SATELLITE COMMUNICATIONS

TIMOTHY PRATT
CHARLES W. BOSTIAN

PRATT
BOSTIAN

SATELLITE COMMUNICATIONS

WILEY

ISBN 0-471-87837-5

Table 1.1
Operational and Planned Communications Satellite Systems

Program Name	Category Code[a]	Coverage Type[b]	Operational Status[c]	Date of First Operation	Frequency Bands	Operational Satellites
ABC	F	G	IP	1986	Ku	2 planned (+ 1 spare)
AEROSAT	AM	R	IP	198?	C, L, VHF	2 planned
AMERICAN SATELLITE	F	D	UC	1985	C, Ku	3 under construction (incl. spare)
ANIK A (Canada)	F	D	OP	1972	C	1 operational
ANIK B	F	D	OP	1978	C, Ku	1 operational
ANIK C	F	D	OP	1982	Ku	2 operational, 1 under construction
ANIK D	F	D	OP	1982	C	1 operational, 1 under construction
APPLE	F	D	OP	1981	C	1 operational
ARABSAT	E	D	UC	1984	C, L	2 under construction
ASETA (Andean)	F	R	IP	198?	C, Ku	2 or 3 planned
AUSTRALIA	F	D	UC	1985	Ku	3 under construction
AUTOSAT	LM	D	IP	198?	UHF	1 planned
BS-2 (Japan)	B	D	UC	1984	Ku	2 under construction
CBS	B	D	IP	198?	Kc, Ku	4 planned + 1 or 2 spares
CHINA	B	D	IP	198?	C, Ku	2 planned
COMSTAR	F	D	OP	1976	C	4 operational
CS-2 (Japan)	F	D	OP	1983	C, Kc, K	2 operational
DBSC	B	D	IP	1986	Kc, Ku	3 planned + 1 spare
Dominion Video Satellite Network	B	D	IP	198?	Kc, Ku	2 planned + 1 spare
DSCS II	F/MM, Mi	G	OP	1966	X	8 operational
DSCS III	F/MM, Mi	G	OP	1982	X	1 operational
ECS (EUTELSAT)	F	R	OP	1983	Ku	1 operational, 4 more planned

[a] AM = Aeronautical Mobile, B = Broadcast, E = Experimental, F = Fixed, LM = Land Mobile, Mo = Mobile (General), Mi = Military, MM = Maritime Mobile.
[b] G = Global, R = Regional, D = Domestic.
[c] IP = In Planning, OP = Operational, UC = Under Construction.

Source: Wilber L. Pritchard, "The History and Future of Satellite Communications," *IEEE Communications Magazine*, **22**, 22–37 (May 1984) (© 1984 IEEE).

(continued)

Table 1.1 (*continued*)

Program Name	Category Code[a]	Coverage Type[b]	Operational Status[c]	Date of First Operation	Frequency Bands	Operational Satellites
EKRAN (USSR)	B	D	OP	1976	C, UHF	3 operational
FLTSATCOM	F/MM, Mi	G	OP	1978	UHF, X	5 operational
FORDSAT	F	D	IP	1987	C, Ku	3 planned (incl. spare)
GALAXY	F	D	OP	1983	C	1 operational, 2 more planned
GALAXY Ku	F	D	IP	1987	Ku	3 planned
GALS (USSR)	F, Mi	G	IP	198?	X	4 planned
GORIZONT (USSR)	F, Mi	R	OP	1978	C, X	4 operational
GRAPHSAT	B	D	IP	198?	Kc, Ku	2 planned + 1 spare
G-STAR	F	D	UC	1984	Ku	3 under construction, 1 more planned
ILHUICAHUA (Mexico)	F	D	UC	1985	C, Ku	2 under construction
INSAT (India)	F, B	D	OP	1983	C, S	1 operational, 1 more planned
INTELSAT IV	F	G	OP	1965	C	4 operational
INTELSAT IV-A	F	G	OP	1976	C	5 operational
INTELSAT V	F, MM	G	OP	1980	C, L, Ku	6 operational, 3 more under construction
INTELSAT V-A	F	G	UC	1984	C, Ku	6 under construction
INTELSAT VI	F	G	UC	1986	C, Ku	5 under construction, up to 11 more planned
ITALSAT	E	D	IP	1987	EHF, Kc	3 planned (incl. preoperational + spare)
LEASAT	Mo, Mi	G	UC	1984	C, VHF	4 under construction
LES	E/Mo, Mi	G	OP	1976	K, UHF	2 operational
LOUTCH (USSR)	F	R	IP/UC	1983	Ku	8 under construction or planned
LUXSAT (Luxembourg)	B	R	IP	1986	Kc, Ku	2 planned (incl. spare)
MARECS	MM	G	OP	1982	C, L	1 operational, 1 more planned
MARISAT	MM	G	OP	1975	C, L, UHF	3 operational
MOLNIYA 1 (USSR)	F	G	OP	1965	C, UHF	Maybe 6 operational
MOLNIYA 3 (USSR)	F	G	OP	1974	C, UHF	Maybe 7 operational
NATO III	F, Mi	G	OP	1979	X	3 operational, 2 more under construction
NORDSAT (Scandinavia)	B	R	IP	198?	Kc, Ku	2 or 3 planned
OLYMPUS (Europe)	B/E	R	UC	1986	Kc, Ku	1 under construction
OTS (Europe)	E, F	R	OP	1978	Ku	1 operational

PALAPA I (Indonesia)	F	D	OP	1976	C	2 operational
PALAPA II (Indonesia)	F	R	OP	1983	C	1 operational, 1 more planned
POSTSAT (West Germany)	F	D	IP	1986	Kc, Ku	3 planned
RADUGA (USSR)	F, Mi	G	OP	1975	C, X	Maybe 5 operational
RAINBOW	F	D	IP	198?	Ku	2 planned + 1 spare
RCA SATCOM	F	D	OP	1975	C	6 operational, 2 more under construction
RCA SATCOM Ku	F	D	UC	1985	Ku	3 under construction
RCA DBS	B	D	IP	1985	Ku	4 planned + 2 spares
SATCOL (Colombia)	F	D	IP	198?	C or Ku	2 or 3 planned
SARIT (Italy)	B	D	IP	1986	Kc, Ku	1 or 2 planned
SATELLITE SYNDICATED SYSTEMS	B	D	IP	198?	Kc, Ku	4 planned + 1 spare
SBS	F	D	OP	1981	Ku	3 operational, 3 more planned
SBTS (Brazil)	F	D	UC	1985	C	Under construction
SKYNET IV (U.K.)	F, Mi	G	UC	1985	X	2 under construction
SPACENET	F	D	UC	1984	C, Ku	3 under construction, 1 spare planned
STC	B	D	UC	198?	Kc, Ku	2 under construction, up to 4 more planned
SYMPHONIE	F, E	R	OP	1974	C	1 operational
TDF	B	D	UC	1985	Kc, Ku	2 under construction
TELECOM	F	R, D	UC	1984	Ku	3 planned (incl. 2 spares)
TELE-X	E, F/B	R, D	UC	1987	K, Kc, Ku	1 under construction
TELESTAR 3	F	D	OP	1983	C	1 operational, 2 more planned
TVSAT (Germany)	B	D	UC	1985	Kc, Ku	1 or 2 under construction
UNISAT	B	D	IP	1986	Kc, Ku	1 planned + 2 spares
USAT	F	D	IP	1985	Ku	2 planned + 1 spare
USSB	B	D	IP	198?	Kc, Ku	4 planned + 1 spare
VOLNA (USSR)	MM	G	IP	198?	L, VHF/UHF	7 planned
WESTAR	F	D	OP	1974	C	5 operational, 3 more planned
WESTAR Ku	F	D	IP	1985	Ku	3 planned
WESTERN UNION DBS	B	D	IP	198?	Kc, Ku	2 planned + 2 spares

[a] AM = Aeronautical Mobile, B = Broadcast, E = Experimental, F = Fixed, LM = Land Mobile, Mo = Mobile (General), Mi = Military, MM = Maritime Mobile.

[b] G = Global, R = Regional, D = Domestic.

[c] IP = In Planning, OP = Operational, UC = Under Construction.

Source: Wilber L. Pritchard, "The History and Future of Satellite Communications," *IEEE Communications Magazine*, **22**, 22–37 (May 1984) (© 1984 IEEE).

The designer of a satellite communication system is not free to select any frequency and bandwidth he or she chooses. International agreements restrict the frequencies that may be used for particular services, and the regulations are administered by the appropriate agency in each country—the Federal Communication Commission (FCC) in the United States, for example. Frequencies allocated to satellite services are listed in Tables 4.1 and 4.2 in Chapter 4. The bands currently used for the majority of services are 6/4 GHz and 14/11 GHz. The 6/4 GHz band was expanded from 500 MHz bandwidth in each direction to 1000 MHz by a World Administrative Radio Conference in 1979, but other services share the new part of the band and may cause interference to satellite communication links during the 1980s, so most systems designed before 1985 have used a 500-MHz total bandwidth. A similar bandwidth is available at 14/11 GHz, but because of propagation problems in heavy rain, this band is less popular than 6/4 GHz. A different frequency is required for the transmit path (the higher fre-

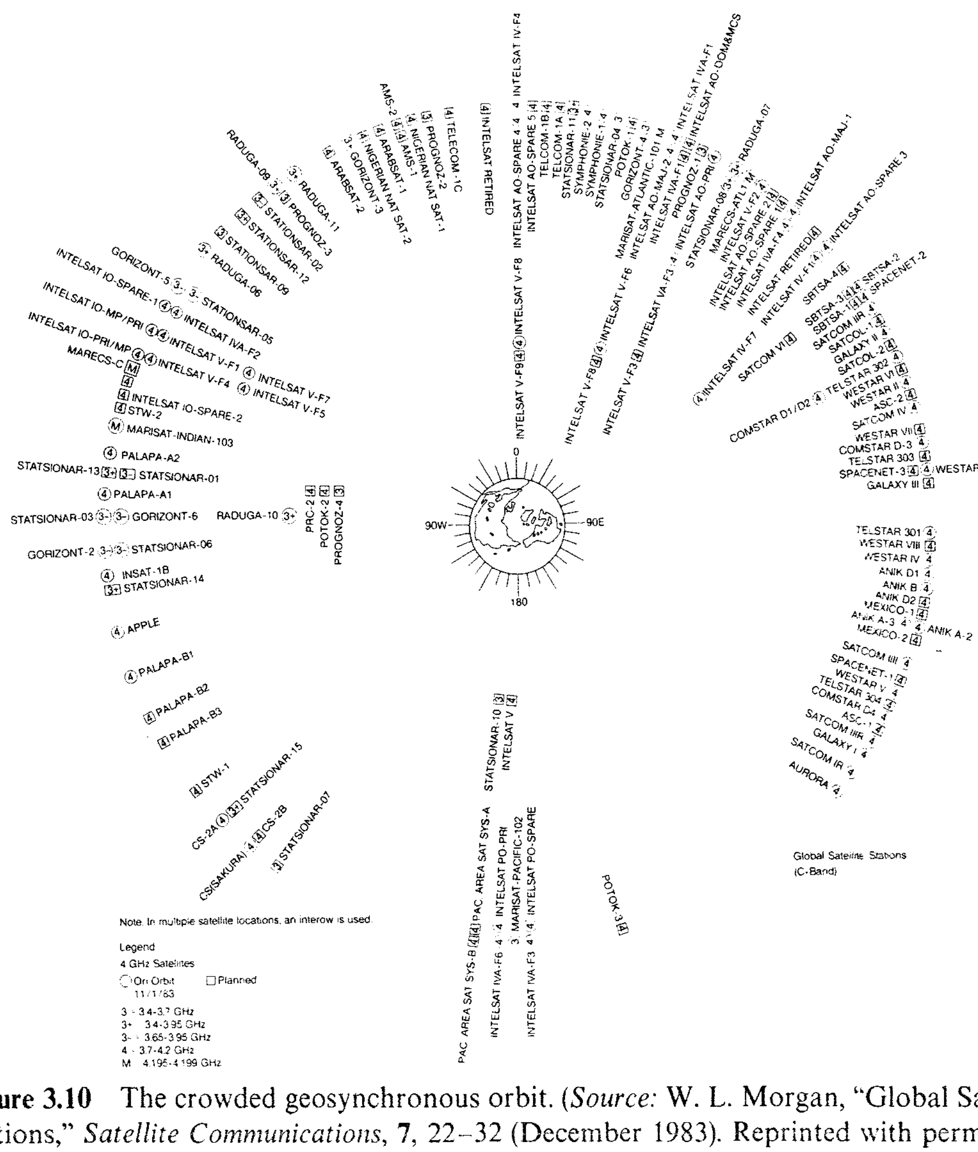


Figure 3.10 The crowded geosynchronous orbit. (*Source:* W. L. Morgan, "Global Satellite Stations," *Satellite Communications*, 7, 22–32 (December 1983). Reprinted with permission of *Satellite Communications*. 6530 So. Yosemite St., Englewood, Colorado, 80111.)

quency) and the receive path in all satellite links because the satellite transmitter generates a signal that would jam its own receiver if both shared the same frequency.

The 500 MHz band originally allocated for 6/4 GHz satellite communications has become very congested and is now completely filled for some segments of the geostationary orbit such as that serving North America. Extension of the bands to 1000 MHz at 6/4 GHz will provide greater capacity as the new frequencies come into use, but many systems are now being designed that will use 14/11 GHz. Figure 3.10 illustrates the crowded nature of the geostationary orbit in 1984. The

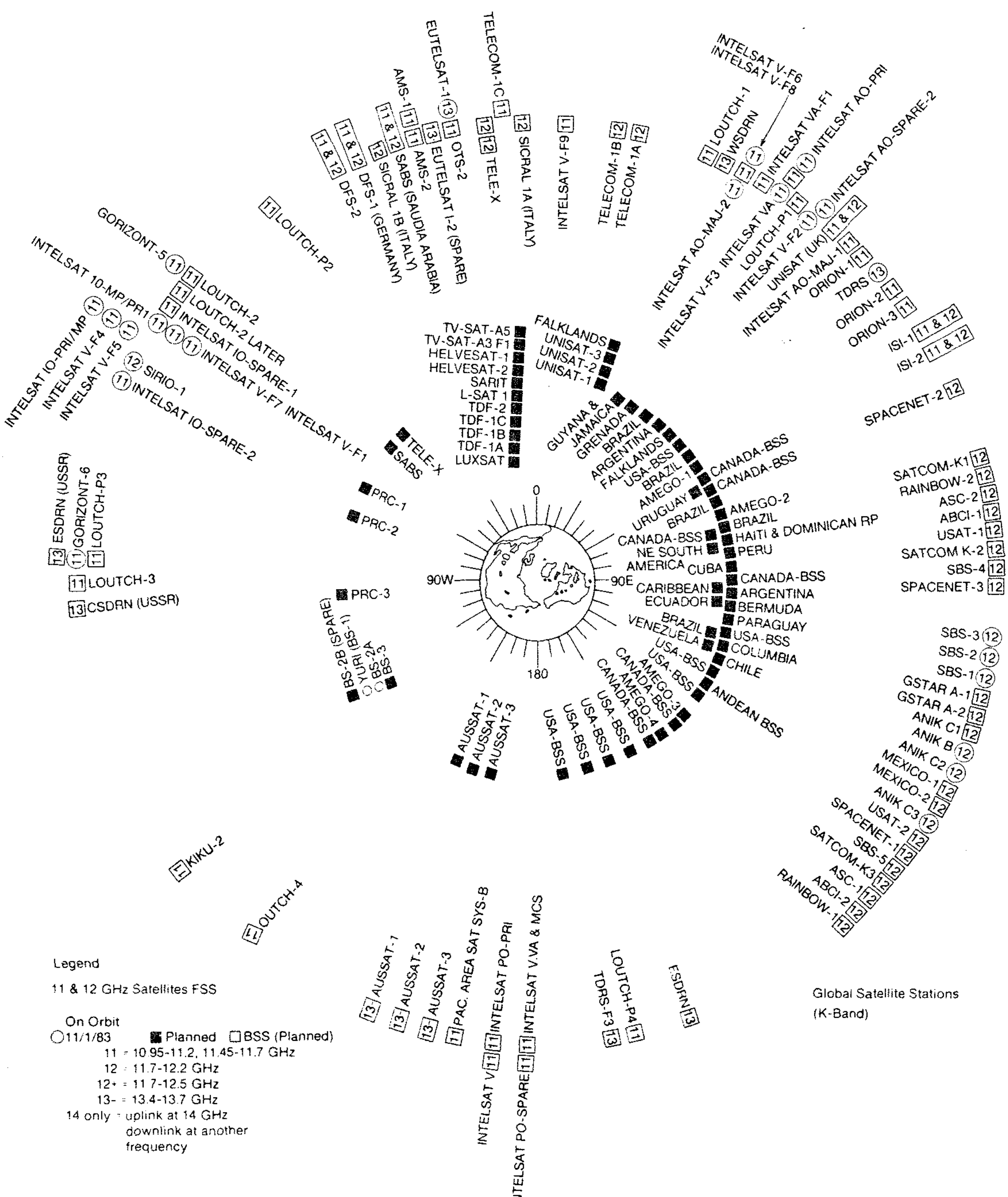


Figure 3.10 *(continued)*

5

MODULATION AND MULTIPLEXING TECHNIQUES FOR SATELLITE LINKS

Communications satellites carry telephone, television (TV), and data signals. Obviously data are always transmitted digitially, but telephone signals may be analog or digital. Digital television is used for teleconferencing, but entertainment TV is still analog. A satellite link will normally relay many signals from a single earth station; these must be separated to avoid interfering with each other. This separation is called *multiplexing*, and its most common forms are *frequency division multiplexing* (FDM) and *time division multiplexing* (TDM). In the first case the signals pass through the transponder on different frequencies; in the second they enter it at different times. Theoretically either multiplexing technique could be used with analog or digital modulation, but TDM is easier to implement with digital modulation and FDM is more convenient with analog modulation. Since choices of multiplexing and modulation techniques cannot be made separately, this chapter treats the two topics together. The next chapter discusses the related problem of multiple access, which is how to allow signals from different earth stations to use the same satellite without interference.

Many books have been written on modulation, demodulation, and multiplexing, and we lack the space to treat these topics in complete detail here. This chapter will review the characteristics of the signals commonly carried by satellites and stress those aspects of modulation and demodulation that are important to satellite link engineering.

5.1 ANALOG TELEPHONE TRANSMISSION

While digital modulation has some inherent advantages over analog frequency modulation (FM) for telephone signals, much of the early investment in the Intelsat

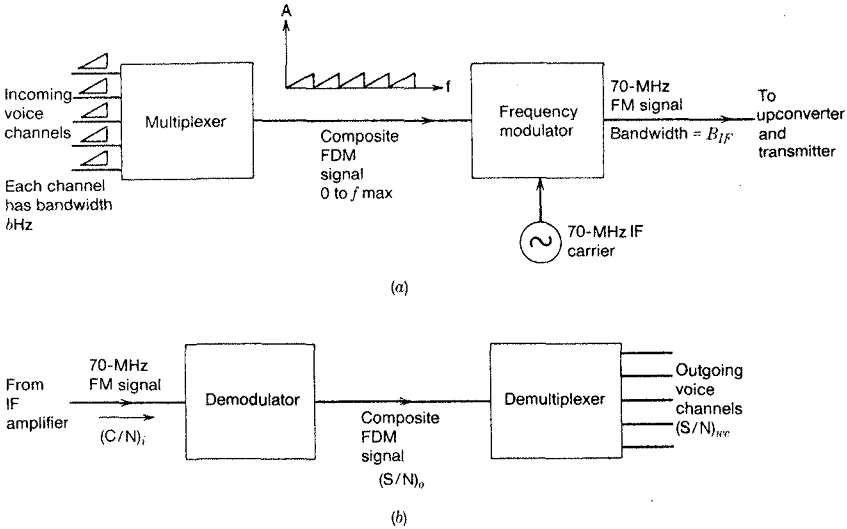


Figure 5.1 Transmitting and receiving ends of a typical FDM system. (*a*) At uplink earth station. (*b*) At downlink earth station.

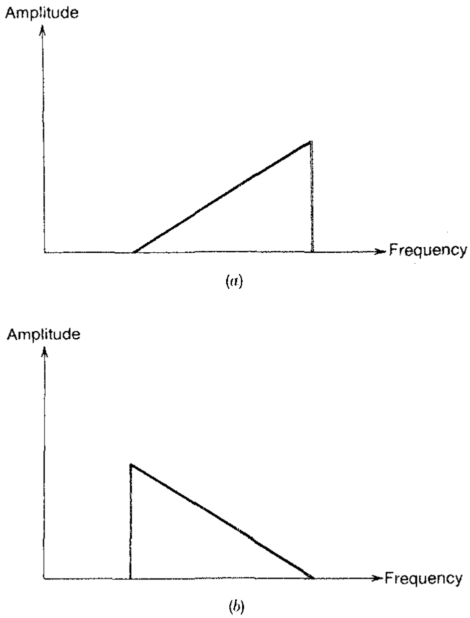


Figure 5.2 Representations of the spectrum of a telephone system baseband voice signal. (*a*) Normal spectrum. (*b*) Inverted spectrum.

network was for FDM/FM telephone systems, and these are still widely used. In this section we will discuss analog multiplex FDM/FM system design and characteristics in detail. We will also present two other analog schemes, *single-channel-per-carrier* (*SCPC*) and *companded single sideband* (*CSSB*), both of which offer advantages over conventional FDM/FM multiplex transmission and over digital modulation under some circumstances.

Satellite FDM/FM analog telephone links resemble the terrestrial microwave point-to-point links that have carried most long distance telephone traffic since the early 1950s. Figure 5.1 sketches a typical system. In it a *multiplexer* takes the *baseband* signals from many individual telephone conversations, translates them to adjacent channels in the RF spectrum, and combines them. Essentially, the multiplexer "stacks" the individual channels in nonoverlapping spectral bands. The resulting composite FDM signal frequency-modulates an IF carrier (usually at 70 MHz) to create an FM (frequency-modulation) multiplex signal. The IF carrier is converted to the appropriate uplink frequency, amplified, and transmitted to the satellite. At the satellite the signal is amplified, downconverted to the downlink band, and retransmitted. At the receiving earth station the downlink signal is amplified and downconverted to IF. The frequency-modulated IF signal drives an FM *demodulator*, which recovers multiplex signals with the voice channels stacked in frequency. Then a *demultiplexer* uses product detectors and filters to translate each channel back to baseband.

Baseband Voice Signals

A baseband voice signal is the voltage generated by an individual telephone set. While its detailed characteristics depend upon the speaker, the Bell System treats it as having a flat spectrum extending from 300 to 3100 Hz. The CCITT recommends 300 to 3400 Hz, but some designs assume a 0 to 3000 Hz spectrum. Here we will follow Intelsat practice and use the CCITT 300 to 3400 Hz spectrum.

Schematically the spectrum of a baseband voice signal is often represented by the triangle shown in Figure 5.2; in a normal spectrum the peak of the triangle is to the reader's right and in an inverted spectrum (one in which the order of frequencies has been reversed) the peak is to the reader's left. The spectrum is not really triangular; this is just a convenient symbol.

The amplitude of a voice signal in a communications link depends on where and how it is measured. In telephone engineering practice (see reference 1, p. 23ff), signal powers are expressed in terms of *transmission level*—that is, their decibel levels with respect to a reference point. At the reference point, the signal power in *dBm* is indicated by the unit *dBm0*; the 0 stands for the zero transmission level point or the test point. Thus, a -2-dBm0 signal is one that produces an average power of -2 dBm at the reference point. A suitable power meter placed at the -5-dB transmission level point would measure the absolute power in the -2-dBm0 signal as -7 dBm.

When telephone engineering began, the test point was accessible and meters could be connected to it. The Bell System standardized the transmission level at the outgoing side of the toll transmission switch as 2 dB below the test level point or −2 dB0 (−2 dB with respect to the zero test level point): With later switchboards, the test point lost its accessibility and disappeared. But the −2 dB standard transmission level at the toll transmission switch remained, and transmission levels are defined from this reference exactly as if the zero test level point still existed.

Under Bell System standards that prevailed for many years, the long-term average power carried by a single voice channel in a telephone system was taken to be −18 dBm0 (CCIR assumes −15 dBm0). The peak instantaneous power in the channel is about 18 dB higher or 0 dBm0. Thus, telephone equipment is often adjusted by applying a 1-kHz tone at 0 dBm0 to the system to simulate peak power on one channel. This is called the *test tone*. We will return to it later in our discussion of multiplexing.

The reader should be aware that the original Bell and CCIR values of −18 dBm0 and −15 dBm0 for the average power level in a single telephone channel are very conservative and that many carriers (including Bell) use other values in some applications. RCA bases its designs on −22 dBm0 while the Bell System uses −19.8 dBm0 for terrestrial FDM/FM links and −22 dBm0 for some satellite links [2]. In general, the number of voice channels that a transponder can carry varies inversely with the average power level per channel [3].

Voice Signal Multiplexing

The process of shifting analog voice channels in frequency and combining them for transmission is called frequency division multiplexing (FDM). The procedure is hierarchial; individual channels are combined into groups; the groups are combined into larger groups; the larger groups are combined into still larger groups, and so on. The names of the groups and their internal channel arrangements vary between administrations and countries. In this section we will use terminology largely drawn from reference 1 for the Bell System and from reference 4 (p.43) for Intelsat.

The first step in voice channel multiplexing is to combine 12 baseband signals into a *basic group* (often called simply a *group*) extending from 60 kHz to 108 kHz. The channels are stacked one above the other at 4-kHz intervals. The stacking is done by double-sideband suppressed-carrier (DSBSC) amplitude modulating each voice channel onto an appropriate carrier, filtering out the upper sideband, and saving and summing the lower sidebands. The result is a single-sideband suppressed-carrier (SSBSC) signal. See reference 5 for a detailed explanation of SSBSC and DSBSC techniques. Figure 5.3 illustrates the process. The carrier frequency in kHz of the nth channel is given by $112 - 4n$; thus channel 1 is at the top of the spectrum and channel 12 is at the bottom. Since each channel occupies only 3.1 kHz, there are 0.9 kHz *guardbands* between channels. These prevent interference and simplify the filtering process when the baseband signals are recovered at the receiver. Selecting the lower sideband in the modulation process inverts the spectra of the channels, but they will be inverted again and put back in the right order at the receiver.

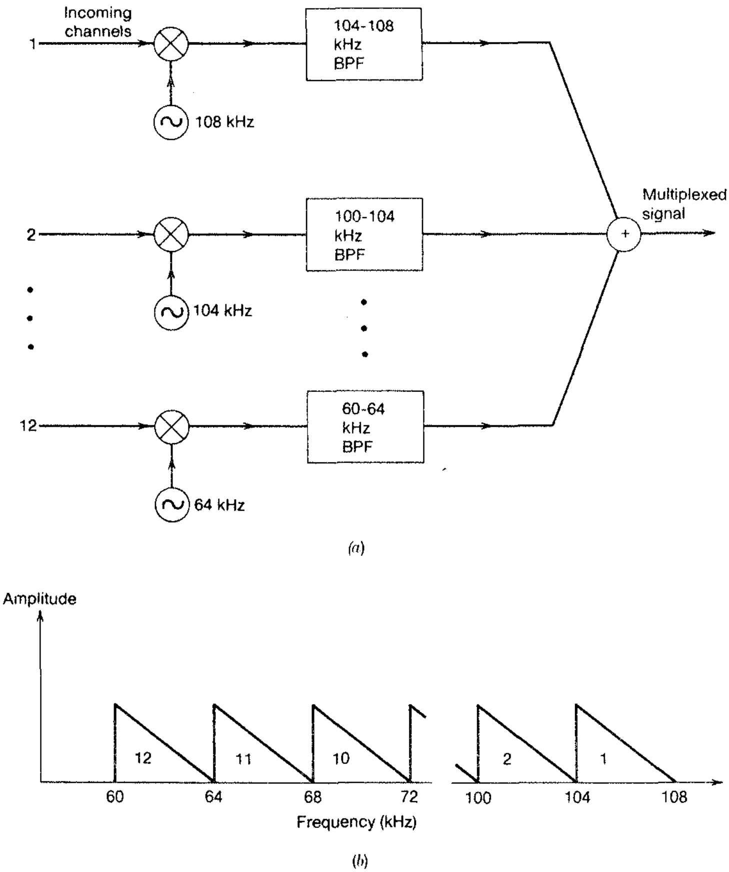


Figure 5.3 Multiplexing 12 telephone channels to form a basic group. (*a*) The basic hardware. (*b*) The spectrum of the multiplex signal showing individual channel spectra.

While a single basic group could be transmitted by itself, most satellite and terrestrial microwave links carry significantly more channels. Bell terrestrial systems use a rigid hierarchy of channel combinations extending from the 12-channel basic group through the 60-channel *basic supergroup* to 600-channel *basic mastergroups* and beyond. The largest named combination in the Intelsat hierarchy is the basic supergroup, five groups stacked in a 240-kHz band. The stacking is done by SSBSC modulating the individual groups onto appropriate carriers and summing the resulting lower sidebands. Figure 5.4 illustrates generation of a 312 to 552 kHz supergroup. The carriers are space 48 kHz apart; that for group 5 is highest at 612 kHz. The spectra of the individual groups are inverted when the basic supergroup is formed. But the individual channel spectra in the groups are themselves

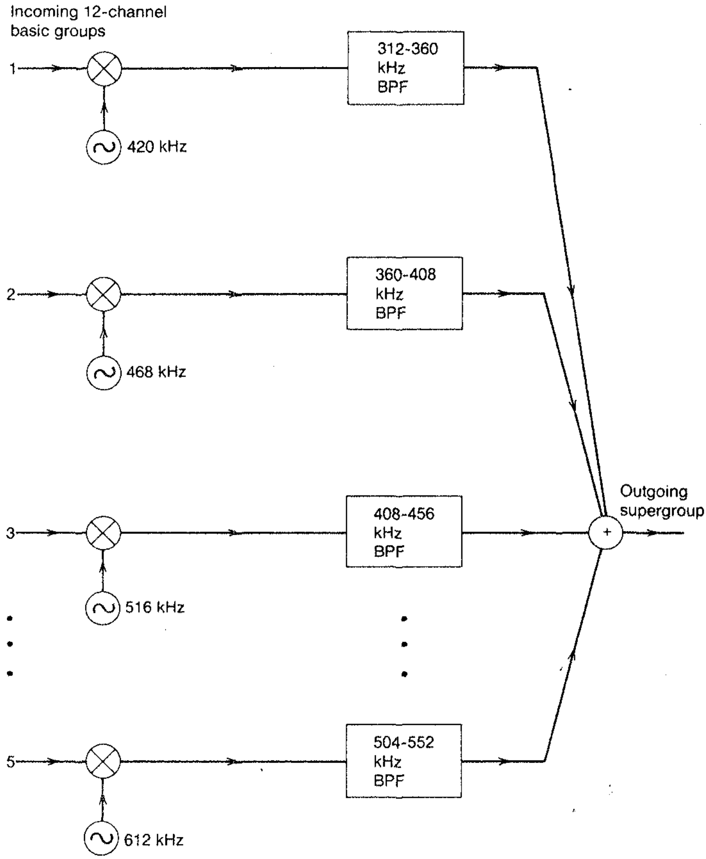


Figure 5.4 Schematic hardware for multiplexing five 12-channel basic groups to form a Bell System basic supergroup.

inverted; the second inversion puts them back in the original order and the individual channel signals in the supergroup are frequency-shifted versions of the original baseband signals. Supergroups are normally separated by 12-kHz guard bands.

Since transponder bandwidth is usually limited, the arrangement of channels in the Intelsat system follows a more flexible format than that used by Bell. In INTELSAT V, for example, earth stations are allowed 12, 24, 36, 48, 60, 72, 96, 132, 192, 252, 312, 432, 492, 552, 612, 792, 972, 1092, 1332, or 1872 voice channels. All of these numbers are divisible by 12 (the number of channels in a basic group), but the channel arrangement varies with the number of channels in order to make efficient use of the transponder. For example, 132 channels are multiplexed by combining a basic group in the 12 to 60 kHz band with one basic supergroup from 60 to 300 kHz and a second basic supergroup from 312 to 552 kHz. Table 5.1 [4:pp. 59ff] summarizes the channel combinations available on the INTELSAT IV through VI spacecraft and lists some of the associated system specifications. The baseband spectrum between 0 and 12 kHz is reserved for an intrasystem channel called the *order wire*, which carries housekeeping information, and for the energy-dispersal signal discussed later.

The amplitude of a multiplexed telephone signal is a random function of time whose characteristics depend upon N, the number of channels. For N greater than or equal to 24, the signal amplitude is usually represented by a Gaussian probability distribution with zero mean and rms value σ. The probability $p(v)$ that the instantaneous voltage has the value v is given by

$$p(v) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(v^2/2\sigma^2)} \tag{5.1}$$

For design purposes, the rms value σ is usually taken as that which will result in 1 mW (0 dBm) total power at the impedance level of the system.

Usually the amplitude of a multiplexed voice signal with $N > 24$ is hard limited to lie between ± 3.16 times the rms value. This causes clipping for something less than 0.2 percent of the time and introduces negligible distortion. Thus it is common practice to equate the peak value of the signal to 3.16 times the rms values. For $N < 24$ the multiplexed spectrum is more "peaky" and an 8.5 peak-to-rms ratio is often used.

Frequency Modulation (FM) Theory

To date frequency modulation is the only form of analog modulation widely used on satellite links. In exchange for a wide bandwidth and poor spectral efficiency, it offers considerable signal-to-noise ratio (S/N) improvement. This means that the signal-to-noise ratio at the output of an FM detector is much larger than the carrier-to-noise ratio (C/N) at the detector input, provided that the input (C/N) is above a threshold value that is characteristic of that detector. Since satellite links have been power limited rather than bandwidth limited and have had to operate with low (C/N) levels, FM has been the analog modulation of choice.

For a detailed discussion of the theory and practice of frequency modulation the reader should consult reference 5. Here we will summarize by quoting from reference 5 that "Frequency modulation results when the deviation Δf of the instantaneous frequency f from the carrier frequency f_c is directly proportional to the instantaneous amplitude of the modulating voltage." An FM modulator is characterized by a maximum frequency deviation $\Delta\omega$, which occurs when the modulating voltage reaches its maximum value. For modulation by a single-frequency sinusoid at radian frequency ω_{mod} whose peak amplitude produces maximum deviation, the expression for an FM waveform $v(t)$ with carrier frequency ω_c is given by

$$v(t) = A\cos\left[\omega_c t + \left(\frac{\Delta\omega}{\omega_{\text{mod}}}\right)\sin(\omega_{\text{mod}} t)\right] \tag{5.2}$$

The ratio $\Delta\omega/\omega_{\text{mod}}$ is called the *FM modulation index m*. This FM waveform can be conveniently represented in terms of m by

$$v(t) = A\cos(\omega_c t + m\sin\omega_{\text{mod}} t) \tag{5.3}$$

Because the FM waveform of Eq. (5.3) takes the form of a cosine of a sine, its spectrum is not obvious. But it can be expanded in an infinite series of discrete components as

$$v(t) = A\{J_0(m)\cos\omega_c t + \sum_{n=1}^{\infty} J_n(m)[\cos(\omega_c + n\omega_{\text{mod}})t + (-1)^n\cos(\omega_c - n\omega_{\text{nod}})t]\} \tag{5.4}$$

where the $J_0, J_1, \ldots J_n$ are Bessel functions of the first kind and order $0, 1, \ldots n$. In theory the spectrum of even a single-frequency sinusoidally modulated FM signal has an infinite number of side frequencies and requires infinite bandwidth; in practice the signal must be filtered to reduce its bandwidth for transmission. An approximate value for the required bandwidth B is given by an equation known as Carson's rule:

$$B = 2f_{\text{mod}}(m + 1) = 2(\Delta f + f_{\text{mod}}) \tag{5.5}$$

where B, Δf (the maximum frequency deviation of the modulator), and f_{mod} (the modulating frequency) are all in Hz. Satellite link designers normally use Carson's rule to calculate bandwidth. The energy associated with sidebands outside the bandwidth B is small, and very little distortion of the modulating waveform occurs when the FM signal is passed through a filter with bandwidth B. An FM signal that is intentionally transmitted through a transponder or to a receiver whose bandwidth is significantly less than the Carson's rule bandwidth is said to be *overdeviated* [2]. We will discuss overdeviation in Chapter 6.

The spectrum of an FM waveform modulated by a real signal is much more complicated than that for a single-frequency sinusoid. But in this case the required bandwidth may still be estimated by Carson's rule if f_{mod} is replaced by f_{max}, the maximum modulating frequency.

$$B = 2(\Delta f + f_{\text{max}}) \tag{5.6}$$

FM Detection Theory: (S/N) Improvement

An FM detector produces an output voltage whose value is proportional to the difference between the instantaneous frequency of the incoming signal and a reference frequency sometimes called the rest frequency. The reference frequency corresponds to the carrier frequency of the previous section. Under normal conditions the detector output is a replica of the modulating waveform that was applied to the carrier before transmission.

As indicated previously, the bandwidth of a frequency-modulated waveform with wideband FM is much greater than the bandwidth of the modulating waveform. Hence the bandwidth of the input signal to an FM detector is much greater than the bandwidth of the output signal. The bandwidth compression provided by the detector is accompanied by an improvement in signal-to-noise ratio (S/N), provided that the input carrier-to-noise ratio is sufficiently large. In other words, the postdetection signal-to-noise ratio $(\mathrm{S/N})_o$ can be considerably larger (by perhaps 20 dB) than the input carrier-to-noise ratio $(\mathrm{C/N})_i$. Since this process is very important to the design and operation of analog FM satellite links and since it is often poorly or incorrectly described in the literature, we will describe it in some detail. For a full development, see pp. 298ff of reference 6.[1]

Assume that an incoming FM signal has an rms amplitude A, occupies an IF bandwidth B_{IF}, and is sinusoidally modulated to have an rms frequency deviation Δf_{rms}. Let the single-sided rms noise power spectral density in the IF bandwidth be η W/Hz so that the noise power at the detector input is ηB_{IF}. For a carrier voltage amplitude A, the average carrier power at the detector input is $A^2/2$ and the incoming carrier-to-noise ratio $(\mathrm{C/N})_i$ is then

$$(\mathrm{C/N})_i = \frac{A^2}{2\eta B_{\mathrm{IF}}} \tag{5.7}$$

Note that $(\mathrm{C/N})_i$ is the same as the overall carrier-to-noise ratio $(\mathrm{C/N})_e$ of the previous chapter.

Let the transfer characteristic of the demodulator be K. This means that a frequency deviation Δf on the incoming carrier produces $K\,\Delta f$ volts at the demodulator output. The rms signal power available at the demodulator output is then proportional to $(K\,\Delta f_{\mathrm{rms}})^2$. If the *output* frequency response of the demodulator extends from f_1 to f_2 Hz, then the noise power output N is given by

$$N = 2\eta(K/A)^2 \int_{f_1}^{f_2} f^2 df = 2\eta\left(\frac{K}{A}\right)^2 \frac{(f_2^3 - f_1^3)}{3} \tag{5.8}$$

Combining Eqs. (5.7) and (5.8) we find that the output signal-to-noise ratio $(\mathrm{S/N})_o$ is

$$(\mathrm{S/N})_o = \frac{K^2(\Delta f_{\mathrm{rms}})^2}{2\eta(K/A)^2 \dfrac{(f_2^3 - f_1^3)}{3}}$$

$$= (\mathrm{C/N})_i \left[\frac{3B_{\mathrm{IF}}(\Delta f_{\mathrm{rms}})^2}{(f_2^3 - f_1^3)}\right] \tag{5.9}$$

[1] The reader who wishes to compare our results with those of reference 6 should note that, in accordance with common satellite communications practice, we compute the noise input noise power in the occupied IF bandwidth while reference 6 computes the noise power in the bandwidth that the modulating waveform would occupy at baseband. Our K is related to the α of reference 6 by $K = 2\pi\alpha$.

Recognizing that for sinusoidal modulation the rms deviation Δf_{rms} is related to the peak deviation Δf_{peak} by a factor of $2^{\frac{1}{2}}$, we may rewrite Eq. (5.9) as

$$(\mathrm{S/N})_o = (\mathrm{C/N})_i \left(\frac{3}{2}\right) \frac{B_{\mathrm{IF}}(\Delta f_{\text{peak}})^2}{(f_2^3 - f_1^3)} \tag{5.10}$$

Remember that Eq. (5.10) is true only for a single-frequency sinusoidally modulated FM signal on a link whose overall $(\mathrm{C/N})_i$ value exceeds a threshold that is typically 10 dB.

For a single (i.e., nonmultiplexed) but nonsinusoidal modulating signal like that used in TV or SCPC telephony, we may use Eq. (5.10) to estimate the FM improvement. To do this, we assume that the spectrum of the modulating waveform extends from 0 to $f_{\max}$ Hz. These limits define the output frequency range of the detector; hence, $f_2 = f_{\max}$ and $f_1 = 0$. Thus

$$(\mathrm{S/N})_o = (\mathrm{C/N})_i \frac{3}{2} \frac{B_{\mathrm{IF}}}{f_{\max}} \left(\frac{\Delta f_{\text{peak}}}{f_{\max}}\right)^2 \tag{5.11}$$

Writing the modulation index m as

$$m = \frac{\Delta f_{\text{peak}}}{f_{\max}} \tag{5.12}$$

and using Carson's rule

$$B_{\mathrm{IF}} = 2f_{\max}(1 + m) \tag{5.13}$$

we can express Eq. (5.11) as

$$(\mathrm{S/N})_o = (\mathrm{C/N})_i \times 3(1 + m)m^2 \tag{5.14}$$

For large m, $3(1 + m)m^2 \simeq 3\,m^3$, while for $m \ll 1$, $3(1 + m)m^2 \simeq 3\,m^2$ and Eq. (5.14) matches the equation given on page 23-11 of reference 7. Both results are correct but they refer to different situations. Such simplified versions of Eq. (5.14) appear frequently in the literature, and the reader should be sure they apply to the problem at hand before using them.

Frequency Modulation with Multiplexed Telephone Signals

The signal-to-noise ratio described by Eq. (5.14) exists at the output of the FM demodulator and describes the ratio of the total power in the multiplexed telephone channels to the total thermal noise power. Let us now consider the FM detector output signal-to-noise ratio for a single telephone channel located at the high-frequency end of a multiplex signal. This is the (S/N) at the output of the demultiplexer. If the channel bandwidth is b Hz, then the noise power output of interest is that between $f_{\max} - b$ and $f_{\max}$. Hence the term $f_2^3 - f_1^3$ in Eq. (5.9) becomes $f_{\max}^3 - (f_{\max} - b)^3$. Since $f_{\max} \gg b$, $(f_{\max} - b)^3 \simeq f_{\max}^3(1 - 3b/f_{\max}) =$

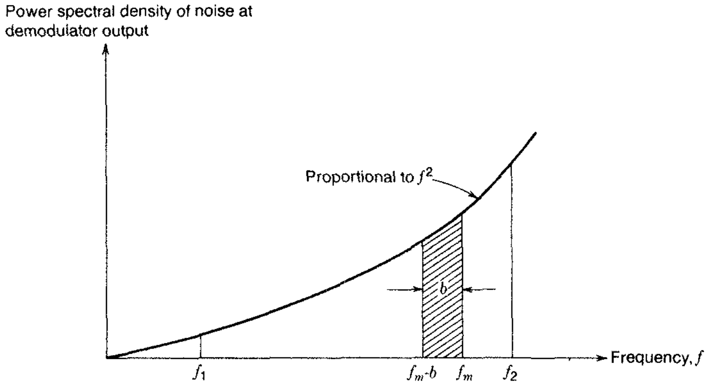


Figure 5.5 Sketch of the noise power spectral density at the output of an FM demodulator. For the narrow band voice channel indicated by the shaded region, the noise power N may be calculated by multiplying the noise spectral density at f_m by b. This leads to Eq. (5.15). For wideband channels the noise power must be calculated by integrating from f_1 to f_m.

$f^3_{\max} - 3bf^2_{\max}$ and Eq. (5.9) becomes

$$(\mathrm{S/N})_{wc} = (\mathrm{C/N})_i \left[\frac{3B_{\mathrm{IF}}(\Delta f_{\mathrm{rms}})^2}{(3bf^2_{\max})} \right] = (\mathrm{C/N})_i \left(\frac{B_{\mathrm{IF}}}{b} \right) \left(\frac{\Delta f_{\mathrm{rms}}}{f_{\max}} \right)^2 \tag{5.15}$$

where the subscript wc means worst channel. See Figure 5.5. Note that b is the bandwidth of one baseband voice channel (nominally 3100 Hz), and $(\mathrm{C/N})_i$ is the overall carrier-to-noise ratio of the link.

Equation (5.15) describes the ratio of signal power to thermal noise power in a telephone channel at the upper end of a multiplexed baseband signal. But the frequency response of neither the human ear nor a telephone receiver is flat, and a telephone listener will respond differently to noise in different parts of the audio spectrum. Some of the noise that is present in bandwidth b will be unnoticed, and the effective signal-to-noise ratio will be higher than that given by Eq. (5.15) by a *weighting factor*. Its value depends on the frequency response of the telephone receiver and of the user's ear. The Bell System uses what is called *C-message weighting*, while CCITT and common satellite practice use *psophometric weighting*. We will adopt the latter and use the symbol p for the psophometric weighting factor. The numerical value of p is 1.78; this corresponds to 2.5 dB [8].

Noise at the high-frequency end of the input spectrum to an FM detector is demodulated with greater output than noise at the low end. (See reference 5, pp. 298 and 322–325.) Figure 5.6*a* sketches this effect. The rising noise at the detector output above some arbitrary frequency f_d can be suppressed as shown in Figure 5.6*a* if a filter with the characteristic of Figure 5.6*b* is installed as indicated in Figure 5.6*c*. This *deemphasis* filter will also reduce the high-frequency content of the modulating signal, but this problem can be eliminated if a *preemphasis* filter is

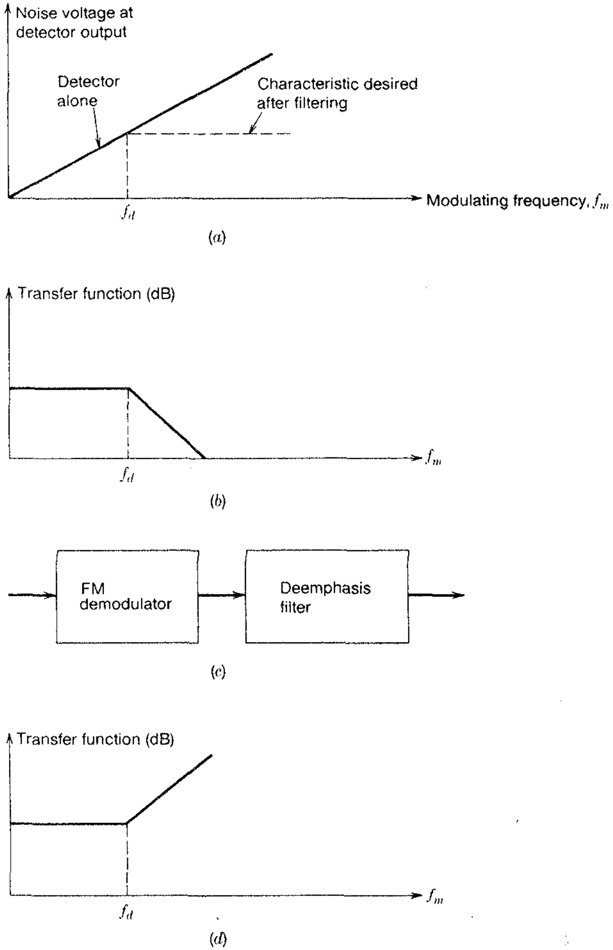


Figure 5.6 Preemphasis. (*a*) Noise voltage at FM detector output. (*b*) Deemphasis filter characteristic. (*c*) Location of deemphasis filter. (*d*) Preemphasis filter characteristic.

inserted at the transmitter ahead of the modulator. The preemphasis and deemphasis filters have inverse characteristics for frequencies above f_d. In practice the characteristics shown need to be maintained only up to the highest baseband frequency present.

This filtering process is called preemphasis although it includes both a preemphasis and a deemphasis filter. It improves the overall (S/N) of the demodulated signal. The degree of improvement depends on the filters and the modulating waveforms used. In CCIR standard satellite and microwave links, preemphasis improves the output signal-to-noise ratio of a telephone system by a factor of 2.5 (4 dB) over that given by Eq. (5.15) [9]. Other values apply to SCPC and TV transmission. Generally the more nonuniform the modulation spectral density (i.e., the more energy below f_d in Figure 5.6), the larger the preemphasis improvement.

CCIR Recommendation 464 [8, pp. 77ff] describes in detail the preemphasis standards for FDM/FM telephone systems. If the highest baseband frequency is f_{max}, the preemphasis filter should provide minimum attenuation at $f_r = 1.25 f_{max}$. Taking attenuation at baseband frequency f_r as the 0 dB reference, the filter attenuation A at any other baseband frequency f should be

$$A = 10 \log_{10}\left\{1 + \frac{6.90}{1 + 5.25/(f_r/f - f/f_r)^2}\right\} \text{dB} \tag{5.16}$$

Reference 8 provides several circuits for preemphasis and deemphasis filters.

Preemphasis improvement and the improvement due to psophometric weighting are independent of each other; hence the right side of Eq. (5.15) may be multiplied by p and by w to yield the psophometrically weighted signal-to-noise ratio on the worst multiplexed telephone channel at the output of an FM link using preemphasis and having an overall carrier-to-noise ratio of $(\text{C/N})_i$.

$$(\text{S/N})_{wc} = (\text{C/N})_i \left(\frac{B_{\text{IF}}}{b}\right)\left(\frac{\Delta f_{\text{rms}}}{f_{\text{max}}}\right)^2 pw \tag{5.17}$$

In decibel form

$$(\text{S/N})_{wc} = (\text{C/N})_i + 10 \log_{10}\left(\frac{B_{\text{IF}}}{b}\right) + 20 \log_{10}\left(\frac{\Delta f_{\text{rms}}}{f_{\text{max}}}\right) + P + W \text{ dB} \tag{5.18}$$

where P is 2.5 dB and W is 4 dB.

Bandwidth Calculation for FDM/FM Telephone Signals

We derived Eq. (5.18) and its predecessors assuming a sinusoidally modulated waveform with an rms frequency deviation Δf_{rms} and requiring a transmission or IF bandwidth B_{IF}. In this section we will relate these quantities to the number of channels N carried by a multiplexed telephone signal and to the available transponder bandwidth.

For link performance calculations, the rms frequency deviation Δf_{rms} that should be used is the *rms test-tone deviation.* This is the rms carrier deviation that a single 1-kHz 0-dBm sinewave called the test tone would produce when supplied to the modulator input, and it represents a standardized test signal in one telephone channel. Putting this another way, the transmitter is designed and adjusted to produce this rms carrier frequency deviation when the modulator input signal is a standard 1-kHz 0-dBm test tone. The rms test-tone deviation is related to the rms deviation that a multiplexed telephone signal will cause by the *loading factor*,

l. For N voice channels, l is given by reference 9 as

$$20 \log_{10}(l) = \mathrm{L} = \begin{array}{ll} -15 + 10 \log_{10}(N), & N > 240 \\ -1 + 4 \log_{10}(N), & 12 \leq N \leq 240 \end{array} \tag{5.19}$$

The product $l\,\Delta f_{\text{rms}}$ is called the *rms multicarrier deviation.*

The ratio of the peak frequency deviation Δf_p to the rms multicarrier deviation $l\,\Delta f_{\text{rms}}$ is given by the *peak factor*, g. For a large number of channels (typically $N > 24$), g is taken as 3.16 (corresponding to 10 dB) and for small numbers of channels (typically $N < 24$), a value of 6.5 (18.8 dB) may be used [9]. If necessary, the incoming voice signals may be amplitude limited to force a true peak-to-rms ratio of 3.16 on the multiplex signal. Thus for an analog FDM/FM telephone link

$$\Delta f_p = lg\,\Delta f_{\text{rms}} \tag{5.20}$$

The use of a high peak-to-mean ratio results in a low rms frequency deviation and consequently a low (S/N) improvement factor for an average baseband level. Amplitude limiting will cause distortion on large signal peaks, but it allows a higher average frequency deviation and a greater (S/N) improvement. In SCPC systems where the peak-to-mean ratio of the single telephone signal is large, companding is often used to reduce the peak-to-mean ratio of the signal before it is applied to an FM modulator. We will discuss companding later.

The 3.16 factor used in calculating the peak deviation of the carrier frequency represents the 0.1 percent extreme of a Gaussian distribution of signal voltages. For 0.2 percent of the time, the signal voltage will exceed 3.16 times the rms value, assuming a Gaussian probability distribution. Because we have to restrict the bandwidth of our RF signal to avoid interference with adjacent channels, the FM modulator at the transmitter will have to be preceded by a limiter that prevents large peaks of signal from overdeviating the carrier frequency. This distorts the multiplexed signal, but for $N > 24$ the effects of limiting are small.

The maximum modulating frequency, $f_{\max}$, depends upon the multiplexing scheme used—that is, on the number of channels multiplexed and how they are organized into basic groups, basic supergroups, and so on. When the standards of a satellite system are established, $f_{\max}$ is tabulated for the allowed values of N. Table 5.1 contains the $f_{\max}$ values for INTELSAT IV through VI. If $f_{\max}$ is not known or if a new satellite link is being designed, a good estimate to use for $f_{\max}$ in kHz is 4.2 N [9].

For a minimum required worst channel $(\mathrm{S/N})_{wc}$ (typically about 50 dB) and an overall $(\mathrm{C/N})_i$ fixed by the link power budget, a satellite systems engineer may trade off values of N, B_{IF}, and Δf_{rms}. The number of channels, N, determines B_{IF} through Carson's rule by

$$B_{\text{IF}} = 2(lg\,\Delta f_{\text{rms}} + f_{\max}) \tag{5.21}$$

where l and $f_{\max}$ depend on N. The minimum $(\mathrm{S/N})_{wc}$ and B_{IF} in turn determine the required value of $(\Delta f_{\text{rms}}/f_{\max})$ in Eq. (5.17). A satisfactory solution requires that B_{IF} not exceed the allocated transponder bandwidth and that the rms test tone deviation Δf_{rms} be achievable by the modulator. After discussing minimum $(S/N)_o$ requirements, we will present an example B_{IF} calculation that illustrates the interdependence of all the variables involved.

Telephone Performance Specifications

While U.S. engineers tend to think of system performance requirements in terms of decibel signal-to-noise ratios, international practice often expresses system specifications in terms of absolute channel noise levels measured in picowatts (psophometrically weighted), abbreviated *pWp*, or in dB above a 1 pWp reference level, abbreviated dBp. (Unfortunately the dBp abbreviation is used both for weighted and for unweighted picowatts.) Picowatts are particularly useful when noise power contributions from several sources must be combined. Decibel powers cannot be added directly.

To convert between picowatts and dBp and milliwatts and dBm, it is necessary first to remember that a psophometric weighting filter reduces the power level of white noise by 2.5 dB and second that 0 dBp (unweighted) corresponds to −90 dBm (unweighted). If P is an absolute power level to be expressed in different units, then

$$P \text{ in dBp (unweighted)} = 10 \log_{10}(P \text{ in pWp}) + 2.5 \tag{5.22}$$

and

$$P \text{ in dBm (unweighted)} = P \text{ in dBp} - 90 \tag{5.23}$$

Assuming a standard 0 dBm signal level, then

$$(S/N) \text{ unweighted} = -P(\text{dBm}) = 87.5 - 10 \log_{10}(P \text{ in pWp}) \tag{5.24}$$

$$(S/N) \text{ weighted} = 90 - 10 \log_{10}(P \text{ in pWp}) \tag{5.25}$$

Thus a 7500-pWp channel noise level corresponds to a weighted (S/N) of 51.25 dB and an unweighted (S/N) of 48.75 dB. Typical satellite link designs allow 7500 to 10,000 pWp total thermal noise for the space segment (up and down links including the intermodulation noise added at the spacecraft). The Intelsat specification for the INTELSAT IV, IV-A, and V space segments is 8000 pWp [4].

Practical Examples

In this section we will apply the equations that describe FDM/FM analog telephone transmission to several examples involving the INTELSAT V spacecraft. The numbers describing the satellite are taken from Table 5.1.

Example 5.1.1

An INTELSAT V transponder using a global beam achieves a 17.8-dB $(C/N)_i$ at an earth station. The transponder carries 972 channels on a single carrier; the FDM/FM signal fully occupies a 36-MHz bandwidth in the transponder.

Table 5.1a
INTELSAT IV-A, V, V-A, and VI Transmission Parameters (Regular FDM/FM Carriers)

Carrier Capacity (Number of Channels)	Top Baseband Frequency (kHz)	Allocated Satellite BW Unit (MHz)	Occupied Bandwidth (MHz)	Deviation (rms) for 0-dBm0 Test Tone (kHz)	Multichannel rms Deviation (kHz)	Carrier-to-Total Noise Temperature Ratio at Operating Point (8000 + 200 $pW0p$) (dBW/K)	Carrier-to-Noise Ratio in Occupied BW (dB)
n	f_m	b_s	b_0	f_r	f_{mc}	(C/T)	(C/N)
12	60	1.25	1.125	109	159	−154.7	13.4
24	108	2.5	2.00	164	275	−153.0	12.7
36	156	2.5	2.25	168	307	−150.0	15.1
48	204	2.5	2.25	151	292	−146.7	18.4
60	252	2.5	2.25	136	276	−144.0	21.1
60	252	5.0	4.0	270	546	−149.9	12.7
72	300	5.0	4.5	294	616	−149.1	13.0
96	408	5.0	4.5	263	584	−145.5	16.6
132	552	5.0	4.4	223	529	−141.4	20.7
96	408	7.5	5.9	360	799	−148.2	12.7
132	552	7.5	6.75	376	891	−145.9	14.4
192	804	7.5	6.4	297	758	−140.6	19.9
132	552	10.0	7.5	430	1020	−147.1	12.7
192	804	10.0	9.0	457	1167	−144.4	14.7
252	1052	10.0	8.5	358	1009	−139.9	19.4

252	1052	15.0	12.4	577	1627	−144.1	13.6
312	1300	15.0	13.5	546	1716	−141.7	15.6
372	1548	15.0	13.5	480	1645	−138.9	18.4
432	1796	15.0	13.0	401	1479	−136.2	21.2
432	1796	17.5	15.75	517	1919	−138.5	18.2
432	1796	20.0	18.0	616	2279	−139.9	16.1
492	2044	20.0	18.0	558	2200	−137.8	18.2
552	2292	20.0	18.0	508	2121	−136.0	20.0
432	1796	25.0	20.7	729	2688	−141.4	14.1
492	2044	25.0	22.5	738	2911	−140.3	14.8
552	2292	25.0	22.5	678	2833	−138.5	16.6
612	2540	25.0	22.5	626	2755	−136.9	18.1
792	3284	36.0	32.4	816	4085	−137.0	16.5
972	4028	36.0	32.4	694	3849	−133.8	19.7
972	4028	36.0	36.0	802	4417	−135.2	17.8
1092	4892	36.0	36.0	701	4118	−132.4	20.7

Source: (Reprinted with permission of the International Telecommunications Satellite Organization from *Standard A Performance Characteristics of Earth Stations in the INTELSAT IV, IV-A, and V Systems Having a G/T of 40.7 dB/K* (BG-28-72E Rev. 1), Intelsat, Washington, DC, December 15, 1982.)

Table 5.1b
INTELSAT IV-A, V, V-A, and VI Transmission Parameters (High-density FDM/FM Carriers)

Carrier Capacity (Number of Channels)	Top Baseband Frequency (kHz)	Allocated Satellite BW Unit (MHz)	Occupied Bandwidth (MHz)	Deviation (rms) for 0-dBm0 Test Tone (kHz)	Multichannel rms Deviation (kHz)	Carrier-to-Total Noise Temperature Ratio at Operating Point (8000 + 200 $pW0p$) (dBW/K)	Carrier-to-Noise Ratio in Occupied BW (dB)
n	f_m	b_s	b_0	f_r	f_{mc}	(C/T)	(C/N)
72	300	2.5	2.25	125	261	−141.7	23.4
192	804	5.0	4.5	180	459	−136.3	25.8
252	1052	7.5	6.75	260	733	−137.1	23.2
312	1300	10.0	9.0	320	1005	−137.1	22.0
492	2044	15.0	13.5	377	1488	−134.4	22.9
612	2540	20.0	17.8	454	1996	−134.2	21.9
792	3284	20.0	18.0	356	1784	−129.9	26.2
792	3284	25.0	22.4	499	2494	−132.8	22.3
972	4028	25.0	22.5	410	2274	−129.4	25.7
1332	5884	36.0	36.0	591	3834	−129.3	23.8

Source: (Reprinted with permission of the International Telecommunications Satellite Organization from *Standard A Performance Characteristics of Earth Stations in the INTELSAT IV, IV-A, and V Systems Having a G/T of 40.7 dB/K* (BG-28-72E Rev. 1), Intelsat, Washington, DC, December 15, 1982.)

If the weighted (S/N) on the top baseband channel is 51.5 dB, find the rms test-tone deviation and the rms multicarrier deviation that must be used. Compare these with the tabulated values.

First we will illustrate the procedure to follow if the multiplexing scheme is not known. Estimating f_{max} as $4200N = 4.082$ MHz and substituting into Eq. (5.18) we have

$$51.0 = 17.8 + 10 \log_{10}\left(\frac{36 \times 10^6}{3.1 \times 10^3}\right) + 20 \log_{10}\left(\frac{\Delta f_{rms}}{4.082 \times 10^6}\right) + 6.5$$

Solving,

$$51.0 - 17.8 - 40.6 - 6.5 = -13.9 = 20 \log_{10}\left(\frac{\Delta f_{rms}}{4.082 \times 10^6}\right)$$

Hence $\Delta f_{rms} = 778$ kHz is the rms test-tone deviation.

Under the loading rule of Eq. (5.19), $L = -15 + 10 \log_{10}(972) = 14.88$ and $l = 10^{(14.88/20)} = 5.55$. Thus the rms multicarrier deviation is $l\,\Delta f_{rms} = 5.55 \times 778$ kHz $= 4.32$ MHz.

We may check this answer by computing the *Carson's rule bandwidth* $B = 2(3.16 \times 4.32$ MHz $+ 4.082$ MHz$) = 35.5$ MHz, which is close to the 36 MHz allowed.

These are slightly different from the published values because the true value of f_{max} (determined by the multiplexing hierarchy) is 4.028 MHz. Using this value of f_{max} we find Δf_{rms} to be 813 kHz and the occupied bandwidth is 36.6 MHz. The published Δf_{rms} is 802 kHz; this leads to an occupied bandwidth of 36.2 MHz and a weighted (S/N) of 50.9 dB.

Example 5.1.2

A single carrier that will occupy (when modulated) 9 MHz of an INTELSAT V transponder can produce a $(C/N)_i$ of 14.7 dB at a standard earth station using the satellite's global beam. Assuming an 8000-pWp space segment noise allocation, how many telephone channels can the transponder carry?

This example illustrates the kind of analysis a systems engineer would perform to determine telephone channel allocations for a proposed spacecraft. It requires an iterative solution.

First, by Eq. (5.24), 8000 pWp corresponds to a weighted (S/N) of 51.0 dB. Substituting this value, the 14.7-dB (C/N), 3.1-kHz channel bandwidth, and IF bandwidth equal to the 9-MHz occupied bandwidth into Eq. (5.18), we obtain

$$51.0 = 14.7 + 10 \log_{10}\left(\frac{9 \times 10^6}{3.1 \times 10^3}\right) + 20 \log_{10}\left(\frac{\Delta f_{rms}}{f_{max}}\right) + 4 + 2.5$$

and

$$20 \log_{10}\left(\frac{\Delta f_{rms}}{f_{max}}\right) = -4.83$$

or

$$\left(\frac{\Delta f_{\text{rms}}}{f_{\text{max}}}\right) = 0.57$$

For N channels f_{max} in Hz is approximately $4200N$. Putting all these numbers, a g value of 3.16, and l for $N < 240$ from Eq. (5.19) into Eq. (5.21), we obtain

$$9 \times 10^6 = 2\{10^{[(-1+4\log_{10} N)/20]} \times 3.16 \times 0.57 \times 4200N + 4200N\}$$

This reduces to

$$1071.43 = N[1.61 \times 10^{(0.2\log_{10} N)} + 1]$$

Substituting a few values we find that $N = 191.2$ solves the equation. The tabulated value for INTELSAT V is 192. Solving this problem required a preliminary assumption that $N < 240$. Suppose instead we had assumed $N > 240$ when getting l from Eq. (5.18). The equation to be solved for N then would have become

$$9 \times 10^6 = 2\{10^{[(-15+10\log_{10} N)/20]} \times 3.16 \times 0.57 \times 4200N + 4200N\}$$

or

$$1071.43 = N[0.320 \times 10^{0.5\log_{10} N} + 1]$$

The solution to this equation is approximately $N = 195.6$ and it violates the hypothesis that $N > 240$. At this point in the process the incorrect initial assumption becomes apparent.

Analog FM SCPC Systems

Single-channel-per-carrier (SCPC) systems avoid the voice signal FDM/FM multiplexing process and instead transmit each telephone channel on its own carrier. Details on the characteristics of FM SCPC systems are given in reference 10, and this section is based on that reference. SCPC reduces the cost of earth terminals on so-called light routes that handle only a few channels because it eliminates expensive multiplexing and demultiplexing equipment. In addition, an SCPC system is easy to reconfigure to meet changing traffic conditions, and thus it is compatible with the demand assignment (DA) schemes discussed in Chapter 6. Further, the carrier for an SCPC channel must be transmitted only when the link is active. In an FDM/FM system the carrier is always present and consuming transponder power. Energy dispersal (Section 5.3) makes this power consumption essentially independent of the channel loading. Since each link in an SCPC system will be active for less than half the time under fully loaded conditions, SCPC offers a saving in transponder power over FDM/FM. But SCPC requires more bandwidth than FM/FDM for the same number of channels, and it is not an economical way to move large amounts of traffic over a fixed route between two earth terminals. Single sideband (SSB) transmission, on the other hand, may be viewed as an SCPC system that uses bandwidth more efficiently than FDM/FM. We will discuss SSB in a later section.

The (S/N) behavior of an SCPC link is described by Eq. (5.14) modified for appropriate preemphasis and weighting. Using 6.3 dB for preemphasis improvement and 2.5 dB for psophometric noise weighting, Ferguson [10] has derived Eq. (5.26) for an analog FM SCPC system.

$$(S/N)_o = (C/N_0)_i - 95.4 + 20 \log_{10}(\Delta f_p) \text{ dB} \tag{5.26}$$

Here N_0 is the noise power density in watts per hertz and Δf_p is the *peak* test-tone deviation in hertz. Note that Eq. (5.26) uses the overall *carrier-to-noise power density ratio* (C/N_0) instead of the overall (C/N). N_0 is the noise power in watts divided by the IF bandwidth in Hz. In decibels

$$(C/N) = (C/N_0) - 10 \log_{10}(B_{IF}) \text{ dB} \tag{5.27}$$

Further improvement is possible through companding, a process in which the dynamic range of speech is compressed before transmission and expanded after detection. Companding distorts a voice waveform in a way that increases the average power by bringing the level of the soft portions close to that of the loud; we will have more to say about it in the following section. Ferguson [10] quotes a companding improvement of 17 dB for SCPC systems, changing Eq. (5.26) to

$$(S/N)_o = (C/N_0)_i - 78.4 + 20 \log_{10}(\Delta f_p) \tag{5.28}$$

RCA achieves a 16-dB companding improvement in its SATCOM system [3].

Companded Single Sideband (CSSB) [2]

Earlier in this chapter we described the process by which individual voice channels are "stacked" in frequency by single sideband suppressed carrier (SSB) modulation and then added to form a multiplexed telephone signal. The multiplexed telephone signal contains a frequency-shifted replica of each outgoing voice channel. With guardbands, each channel occupies 4 kHz, and the individual channels can be recovered independently of each other from the multiplexed signal by an appropriate combination of multipliers and filters.

Now suppose that instead of the 70-MHz frequency modulator used in FDM/FM, another SSB modulator shifted the multiplexed signal up to IF. The resulting uplink signal would be a replica of the incoming voice channels, appropriately stacked in frequency. At the downlink station, each channel would be immediately accessible at RF and at IF. More important, the occupied bandwidth of the uplink and downlink signals would be exactly 4 kHz times the number of channels. Thus this SSB modulation scheme uses bandwidth much more efficiently than FDM/FM (and more efficiently than digital modulation), and it offers the theoretical possibility of sending 9000 channels through a single 36-MHz transponder. Practical considerations, however, reduce this number to about 6000 channels.

Besides bandwidth efficiency, the other advantage that CSSB offers over FM is graceful failure. If the (C/N) at the input to an FM demodulator falls below threshold, the link will immediately stop working. A CSSB receiver has no threshold, and its output (S/N) degrades in proportion to the input (S/N) without any sudden failure.

Since an SSB signal of the type considered here lacks a carrier, it cannot be described in terms of a carrier-to-noise ratio. Instead a signal-to-noise ratio (S/N) is used that is simply the ratio of the received signal power to the received noise power. If all the voice channels feeding a SSB modulator are at their peak levels, then the transmitter output will be at its rated power level, and this power will be divided evenly between the voice channels. Since SSB demodulation (i.e., returning the original voice channels to baseband) provides no improvement in (S/N), the output (S/N) for each channel will be essentially the same as the overall (S/N) at RF, and this will be essentially the same as the overall (C/N) that the same earth stations and transponder could deliver using FDM/FM. This falls far short of the 48 to 50 dB required for telephone channels, and it rules out ordinary SSB for multiplexed voice channel transmission.

If the dynamic range of each voice channel is compressed before multiplexing and expanding after demultiplexing, the channels are said to have been companded. Companding reduces the required (S/N) at RF to a level significantly below that required at baseband and permits satisfactory operation with approximately the same power levels and antenna sizes as for FDM/FM. We will present the equations that govern the capacity and performance of CSSB links in the next chapter.

5.2 ANALOG TELEVISION TRANSMISSION

While telephone signals represent the bulk of communications satellite traffic, satellite technology has had a more dramatic effect on the television industry than on the telephone system. The first commercial satellites made live coverage of international news and sporting events possible, and in the United States these were soon followed by domestic satellites for network program distribution. Since anyone with an earth station and access to a transponder could originate programs, the cost of entering the U.S. television market fell drastically and a number of commercial, educational, and religious organizations began to offer programs via satellite for distribution over cable TV systems in competition with the established networks. Radio amateurs and other electronic experimenters were able to construct equipment capable of receiving these transmissions. Manufacturers began to offer home satellite TV reception equipment for sale to the general public and pressure grew on the U.S. government to allow unrestricted TV broadcasting by satellite. At the time this text was written neither the technology nor the regulatory environment of satellite TV transmission had reached steady state, but clearly TV was one of the most active parts of the satellite communications industry. In this section we will outline television modulation and demodulation; a later chapter will discuss network TV distribution and home satellite reception.

Television Signals

While a number of television transmission standards exist worldwide, the two in most common use are the North American and Japanese 525 line/60 Hz NTSC system and the European 625 line/50 Hz PAL system. These are also called CCIR systems M and B, respectively. In this text we will emphasize the NTSC system.

The video signal of a monochrome (black and white) TV transmission carries an analog representation of the brightness (i.e., the amount of white light) in the picture along a series of horizontal scanning lines. This is called the *luminance* signal. Along with the luminance signal, synchronization pulses are transmitted so that the TV receiver can recreate the scanning process of the camera.

Historically, monochrome TV developed before color, and in the United States color TV was designed so that the color information could be added to monochrome transmissions without degrading the performance of existing black-and-white receivers. Any color may be created by an appropriate combination of red, green, and blue light. Color TV could be transmitted by transmitting the color components of each picture separately, but this scheme would require excessive bandwidth. Instead, three linear combinations of the three components are transmitted and the component values themselves are recovered at the receiver.

The TV camera generates voltage levels corresponding to the red, green, and blue light at each point in the picture. We will identify these voltage levels by the letters R, G, and B. A monochrome receiver would respond to the amount of white light at a point in the picture; this is the *luminance*, Y, and is related to the color voltage levels by

$$Y = 0.30R + 0.59G + 0.11B \tag{5.29}$$

The luminance signal is transmitted so that monochrome receivers can receive a color image in black and white.

For color reconstruction, two other independent linear combinations of R, G, and B must be transmitted along with Y so that all of the color components can be recovered. These are called the I and Q signals, given by

$$I = 0.60R - 0.29G - 0.32B \tag{5.30}$$

$$Q = 0.21R - 0.52G + 0.31B \tag{5.31}$$

The letters I and Q stand for in-phase and quadrature, and together the I and Q signals (when decoded with the luminance signal) carry the *chrominance* information about the color at each point in the picture.

The I and Q signals modulate a color (or chrominance) subcarrier in such a way that the amplitude of the resulting chrominance signal determines the *saturation* (degree of purity) of the color at a point and the phase of the chrominance signal determines the *hue* (perceived shade) of the color. From the amplitude and phase of the chrominance signal, a TV receiver determines the shade of the color and the amount of white light to add. From the luminance signal it determines how bright the color should be.

In terrestrial broadcasting the luminance (Y) signal, filtered to occupy the band from 0 to 4.2 MHz, modulates a "picture" carrier with a vestigial sideband (VSB) modulator. The upper sideband is transmitted in full; the lower sideband is partially removed. The resulting VSB signal is all that needs to be transmitted for the video portion of monochrome television.

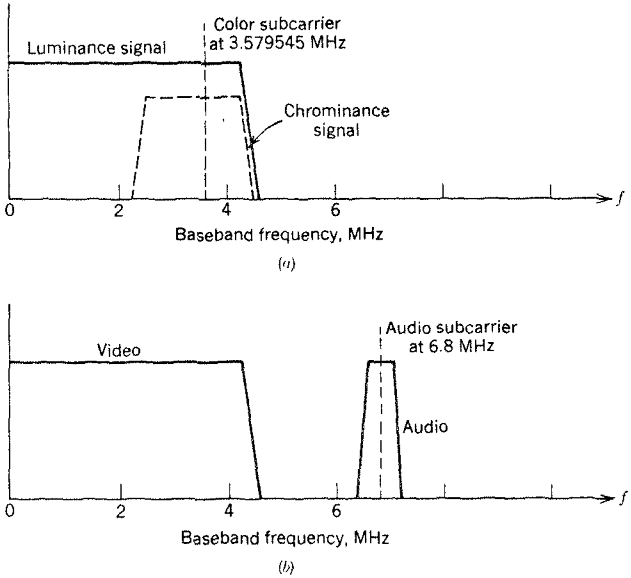


Figure 5.7 Spectra of baseband TV signals. (*a*) Baseband video signal. (*b*) The composite (video plus audio) TV signal as transmitted by US domestic satellites.

The chrominance information is transmitted by a *color subcarrier* at 3.579545 MHz (hereafter abbreviated as 3.58). This value was chosen because it places the chrominance signal at a relatively empty part of the luminance spectrum and minimizes color interference with black-and-white reception. Both the I and Q signals modulate the color subcarrier through double-balanced mixers to generate double sideband suppressed carrier (DSBSC) signals. The subcarrier is phase shifted by 90° before it enters the Q modulator. Thus, both I and Q components may be recovered at the receiver. Figure 5.7*a* presents a sketch of the spectrum of the baseband video signal.

The baseband audio signal extends from 50 Hz to 15 kHz. It frequency modulates an audio subcarrier and the resulting FM waveform is added to the video baseband signal. This leads to the composite TV signal of Figure 5.7*b*; it consists of the baseband video signal below an FM modulated audio subcarrier. In U.S. domestic systems an audio subcarrier frequency of 6.8 MHz is standard; 6.2 MHz is also used.

In terrestrial broadcasting, the audio and video signals are combined and shifted in frequency to an appropriate part of the VHF or UHF band for transmission. The radiated signal is a complex combination of FM (the sound), VSB (the luminance), and quadrature DSBSC (the chrominance). It occupies a 6-MHz bandwidth. For satellite transmission the baseband video signal (luminance and chrominance), frequency modulates a video carrier and the two audio signals frequency modulate two audio carriers. The details of the video modulation depend on the transponder bandwidth available. Typical values for network TV are a peak deviation Δf_p of 10.75 MHz and a maximum video modulating frequency f_V of 4.2 MHz. By Eq. (5.5) this requires a 29.9 MHz transponder bandwidth. Television signals are often overdeviated, trading the larger improvement in video (S/N) that results for the smaller degradation in picture quality associated with truncating some of the sidebands [3].

A TV signal from a satellite is quite different from a broadcast TV signal. Converters that allow reception of satellite television transmission on conventional home receivers must demodulate the incoming FM signals, recover the baseband video and audio channels, and remodulate the audio and video onto a locally generated carrier using the same modulation scheme as a broadcast TV transmitter.

Signal-to-Noise Ratio Calculation for Satellite TV Links

Equations (5.11) and (5.14) are two equivalent formulas that relate the signal-to-noise ratio at the output of an FM demodulator to the overall carrier-to-noise ratio at the input. As derived, they compare total signal power to total noise power and the (S/N) values they predict can be improved by preemphasis and should be weighted to account for the non-uniform response of the eye to white noise in the video bandwidth. The preemphasis factor is called p and the weighting factor is q. Their decibel values, P and Q, add in the decibel versions of Eqs. (5.11) and (5.14):

$$(\mathrm{S/N})_V = (\mathrm{C/N})_i + 1.76 + 10\log_{10}\left(\frac{B_{\mathrm{IF}}}{f_V}\right) + 20\log_{10}\left(\frac{\Delta f_{\mathrm{peak}}}{f_V}\right) + P + Q \text{ dB} \tag{5.32}$$

$$(\mathrm{S/N})_V = (\mathrm{C/N})_i + 10\log_{10}[3m^2(1+m)] + P + Q \text{ dB} \tag{5.33}$$

Here f_V is the maximum video modulating frequency (4.2 MHz by U.S. standards), and the modulation index m is $\Delta f_{\mathrm{peak}}/f_V$. Substituting a typical value of 10.75 MHz for Δf_{peak} into these equations we obtain

$$(\mathrm{S/N})_V = (\mathrm{C/N})_i + 18.5 + P + Q \text{ dB} \tag{5.34}$$

The values used for P and Q depend on the noise characteristics of particular TV systems and on the subjective response of individual viewers to the noise included in a television picture. Numbers for $P + Q$ ranging from roughly 18 to 26 dB are quoted in the literature. These lead to overall improvements in signal-to-noise ratio ranging from 36.5 to 44.5 dB.

5.3 ENERGY DISPERSAL

The satellite telephone and television transmission systems described in the previous two sections both employ frequency modulation. When an input modulating signal is absent, an FM transmitter radiates all of its power at the carrier frequency. With modulation the average carrier power is spread over a large bandwidth. The larger the amplitude of the modulating signal, the smaller is the average transmitted spectral power density in watts per hertz of bandwidth.

To minimize interference with terrestrial microwave systems sharing the same frequencies, most administrations restrict the maximum spectral power density that a satellite may radiate toward the earth. Minimum spectral power density occurs with maximum modulation amplitude; this condition is called *full loading* in telephone practice. Typically for an Intelsat spacecraft the radiated power per 4 kHz of bandwidth must not rise more than 2 dB above its value for full loading.

The process of controlling the radiated spectral density is called *energy dispersal*. It is accomplished at the uplink earth station by adding a symmetric triangular voltage waveform called the *dispersal signal* to the modulating waveform before modulation. The dispersal waveform is removed at the downlink earth station. In television transmission the dispersal signal has a constant amplitude and currently a frequency of 30 Hz for the U.S. NTSC system. The frequency is scheduled to be changed to 60 Hz. The amplitude depends on the spacecraft and modulation used; for example with INTELSAT V the dispersal signal must provide between 1 and 2 MHz peak-to-peak carrier frequency deviation [4].

For telephone transmission the amplitude of the dispersal signal is dynamically adjusted to keep the radiated spectral density within the required bounds. The frequency of the dispersal signal depends on the system; Intelsat requires values between 20 and 150 Hz. In general the amplitude of the dispersal signal for multichannel telephone transmission is determined by finding the frequency shift (peak deviation) ΔF that the dispersal waveform must cause. If ΔF and the FM modulator characteristics are known, then the dispersal signal amplitude may be computed. Reference 11 provides a detailed derivation of the equations for calculating ΔF; we will summarize those results and give a numerical example.

Consider an FDM/FM analog telephone signal modulated onto a carrier whose power is C W and whose loading produces rms multichannel deviation d Hz. Assume that the resulting spectrum is Gaussian and that its power density $W(f)$ may be expressed in terms of frequency f by Eq. (5.35) in which ΔF is the difference between the (unmodulated) carrier frequency f_c and f.

$$W(f) = \left(\frac{C}{d\sqrt{2\pi}}\right) \exp\left[-(\Delta F)^2/(2d^2)\right] \tag{5.35}$$

At full loading the density is $W_{\min}(f)$, given by

$$W_{\min}(f) = \left[\frac{C}{d\sqrt{2\pi}}\right] \exp\left[-(\Delta F)^2/(2d_m^2)\right] \tag{5.36}$$

where d_m is the fully loaded rms multichannel deviation. At no loading the density $W_{\max}$ is determined by the deviation $\Delta F_{\max}$, which the maximum dispersal waveform causes. Assuming that the triangular dispersal signal simply spreads the carrier power C uniformly over a band that extends ΔF Hz on either side of the carrier frequency, then

$$W_{\max} = \frac{C}{2\,\Delta F_{\max}} \tag{5.37}$$

The usual practice is to choose $W_{\max} = W_{\min}(0)$, which leads to the result

$$\Delta F_{\max} = \frac{C}{2W_{\min}(0)} = d_m\left(\frac{\pi}{2}\right)^{\frac{1}{2}} \tag{5.38}$$

The rms multicarrier deviation d_m equals the loading factor of Eq. (5.18) multiplied by the rms test-tone deviation Δf_{rms} of the link. Thus from d_m we may calculate the maximum deviation $F_{\max}$ that the dispersal waveform must be capable of producing.

For less than full loading ($d < d_m$) the dispersal waveform added to the incoming multiplexed telephone signal acting by itself would produce a deviation $\Delta F < F_{\max}$. To calculate ΔF we must solve the integral equation derived in reference 11:

$$\frac{1}{\sqrt{2\pi}}\frac{\Delta F}{d_m} = \frac{1}{\sqrt{2\pi}}\int_0^{\Delta F/d_m\, d_m/d} e^{-x^2/2}\, dx \tag{5.39}$$

Equation (5.39) is written in this form so that the integral may be evaluated from readily available tables. Solving Eq. (5.39) requires a trial-and-error process. The equation actually specifies the maximum ΔF required for a particular loading; Intelsat will permit the spectral density to go 2 dB above the value corresponding to ΔF calculated from Eq. (5.39).

Example 5.3.1

According to Miya's tabulated data [11], a 60-channel telephone link carried by INTELSAT III had an rms test-tone deviation Δf_{rms} of 410 kHz and a fully loaded rms multicarrier deviation d_m of 830 kHz. Find $F_{\max}$ and the value of ΔF that the dispersal signal must produce (i.e., the ΔF the signal would produce if it were acting by itself) at 75 percent loading.

By Eq. (5.39), $\Delta F_{\max} = d_m(\pi/2)^{\frac{1}{2}} = 830(\pi/2)^{\frac{1}{2}} = 1040.25$ kHz. Assume that the deviation d at 75 percent loading is $0.75 d_m$ or 622.5 kHz. Writing $\Delta F/d_m$ as u in Eq. (5.39), we have

$$\frac{u}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}}\int_0^{1.33u} e^{-x^2/2}dx$$

The right-hand side of this equation is the area under the Gaussian probability density function between 0 and $1.33u$. We will represent it by Area ($1.33u$). It is conveniently tabulated in *C.R.C. Standard Mathematical Tables* and similar publications. The equation to be solved then becomes

$$\frac{u}{\sqrt{2\pi}} = \text{Area } (1.33u)$$

A trial-and-error solution yields $u = 1.05$. Thus, $\Delta F/d_m = 1.05$ and $\Delta F = 871.5$ kHz.

5.4 DIGITAL TRANSMISSION

Digital modulation is the obvious choice for satellite transmission of signals that originate in digital form and that are used by digital devices. Familiar examples are data transmissions between computers, printed text, communications between remote terminals and computers, and the like. Such analog signals as telephone channels and television may be put into digital form for transmission and then converted back to analog form for routing to the end user. While this process may be costly in terms of bandwidth, it usually offers improved noise performance and increased immunity to interference. Digital transmission lends itself naturally to time division multiplexing (TDM) and time division multiple access (TDMA); both of these techniques allow one signal to use a transponder at a time and thus avoid intermodulation problems. Finally, analog signals that are transmitted digitally can share channels with digital data; all digital signals are to be handled in the same way, and their content is immaterial. Thus a digital satellite link can carry a mix of telephone and data signals that varies with traffic demand.

There are basically two problems in satellite digital transmission: (1) how to get incoming analog signals into digital form and then back again, and (2) how to transmit and receive digital signals efficiently—whatever their origin and destination. Sections 5.5 and 5.6 will discuss these topics in more detail.

Baseband Digital Signals

We will represent baseband digital signals as serially transmitted logical ones and zeroes. While in computer circuitry a logical zero may be represented by a low voltage (nominally zero) and a logical one may be represented by a high voltage (say 5 V)—or vice versa—this arrangement is inconvenient for transmission over any significant distance and is not used. To understand why, imagine a transmission line carrying a bit stream encoded this way and containing approximately equal numbers of ones and zeroes. About half the time the line voltage will be 5 V and about half the time it will be 0 V; hence the line signal will have a 2.5-V DC component. All circuits that carry this signal must have a frequency response that extends to DC, and this is difficult to achieve since many communication circuits contain transformers. To avoid this problem, digital modulators usually accept their input in a *polar non-return-to-zero* (*NRZ*) format: logical ones and zeroes are transmitted as plus or minus a stated value. Thus a one might be transmitted as +1 V and zero might be transmitted as −1 V. Zero volts is not transmitted except as a transient value. Throughout this text we will assume a polar NRZ format for data signals unless we explicitly state otherwise.

Baseband Transmission of Digital Data

A random sequence of rectangular binary pulses has a power spectral density

$$G(f) = T_b \left[\frac{\sin(\pi f T_b)}{\pi f T_b} \right]^2 \tag{5.40}$$

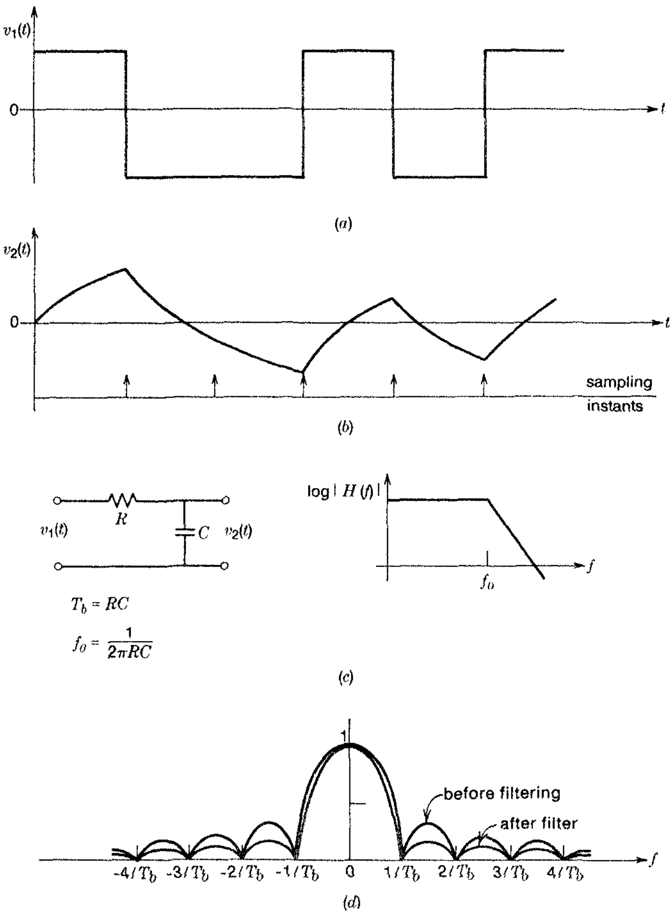


Figure 5.8 Illustration of the effect of low pass filtering on a NRZ signal. (*a*) Random NRZ polar pulse train. (*b*) Waveform output from an RC filter with $T_b = RC$. (*c*) *RC* filter and its transfer function $|H(f)|$. (*d*) Spectrum of bandlimited NRZ pulse train.

where T_b is the duration of the pulse [12]. This spectrum is illustrated in Figure 5.8*d*. The familiar sin x/x shape shows that energy exists at all frequencies; to retain the rectangular pulse shape would require an infinite transmission bandwidth. Practical communication systems are always bandwidth limited; not only is infinite bandwidth not available, interference considerations in radio links dictate that a communication system should use the smallest possible bandwidth, and this is usually one of the design criteria of a communication system.

If we take the random pulse train shown in Figure 5.8*a* and bandlimit it by passing it through a low-pass filter, the pulse shape will be altered. As an example, consider the effect of passing the rectangular pulse train through a single RC section, representing a very simple low pass filter. The resulting waveform, shown in Figure 5.8*b*, has been delayed and pulses are "smeared" in time—the decaying pulse from one transition extends into the next pulse interval. Where the pulse pattern is 10 or 01, the amplitude of the second pulse at the sampling instant shown in Figure 5.8 has been reduced by the presence of a delayed portion of the preceding pulse. This is called *intersymbol interference* (*ISI*) and is likely to occur whenever a digital signal is passed through a bandlimiting filter. When noise is added to the waveform, ISI increases the likelihood that the receiver will detect a bit incorrectly, causing a bit error. In a baseband system, ISI can be avoided by an appropriate

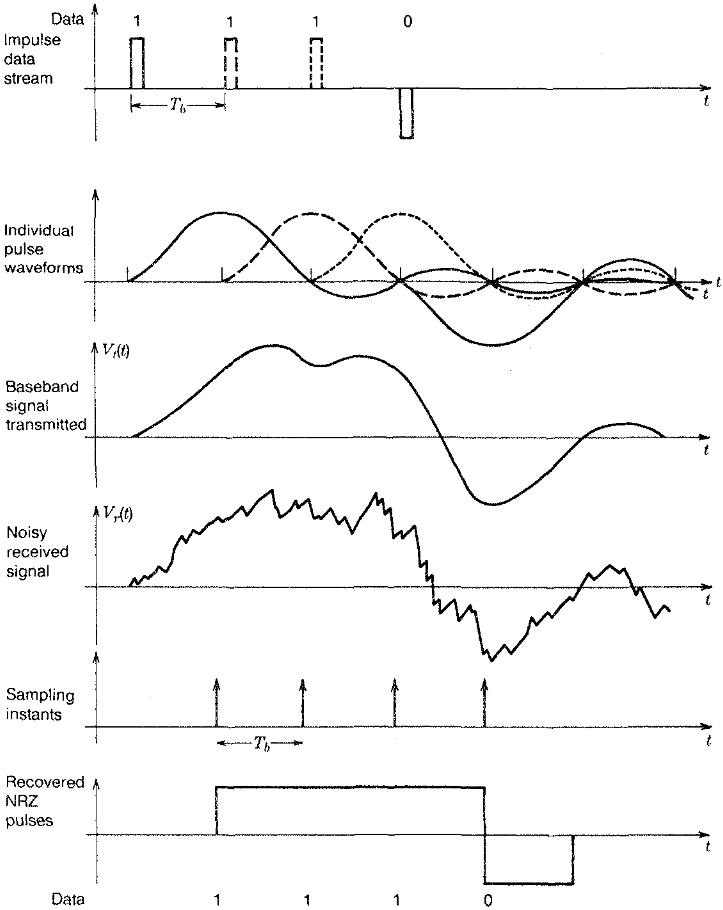


Figure 5.9 Transmission and reception of baseband zero-ISI pulses.

choice of low-pass filter. Nyquist [13] proposed a technique that can theoretically produce zero ISI, now known as the Nyquist criterion. The objective is to create in the receiver a pulse that resembles the sin x/x shape, crossing the axis at intervals of T_b, where T_b is the bit period. The receiver samples the incoming wave at intervals of T_b, as shown in Figure 5.9, so that at the instant one pulse is sampled, the "tails" from all preceding pulses have zero value. Thus previous pulses cause zero intersymbol interference (zero ISI) at each sampling instant.

Filters that produce the required zero ISI waveform in the receiver can be synthesized in several ways. The filter proposed by Nyquist was the "raised cosine" filter, which has a frequency characteristic given by

$$V_r(f) = \begin{cases} T_b, & |f| \leq \dfrac{R_b}{2}(1-\alpha) \\ T_b \cos^2\left\{\dfrac{\pi}{2\alpha R_b}\left[\dfrac{|f| - R_b(1-\alpha)}{2}\right]\right\}, & \dfrac{R_b}{2}(1-\alpha) < |f| < \dfrac{R_b}{2}(1+\alpha) \\ 0, & |f| \geq \dfrac{R_b}{2}(1+\alpha) \end{cases} \tag{5.41}$$

where $0 < \alpha < 1$ and $R_b = 1/T_b$ is the bit rate in bits/second. The pulse shape generated when the filter is driven by an impulse, $\delta(t)$, is $v_r(t)$, the required zero ISI waveform. The waveform $v_r(t)$ is obtained as the inverse Fourier transform of the output from the Nyquist filter, which is simply the spectrum of the input pulse multiplied by the frequency response of the filter.

$$v_r(t) = F^{-1}[V_r(f) \times S(f)] \tag{5.42}$$

where $F^{-1}[\ ]$ indicates the inverse Fourier transform and $S(f)$ is the spectrum of the input pulse. If we use an impulse $s(t)$ as the input signal, $S(f) = 1$ and then

$$v_r(t) = F^{-1}[V_r(f)] \tag{5.43}$$

For the raised cosine filter with $V_r(f)$ given by Eq. (5.43)

$$v_r(t) = \left[\frac{\cos \pi\alpha R_b t}{1 - (2\alpha R_b t)^2}\right]\left[\frac{\sin \pi R_b t}{\pi R_b t}\right] \tag{5.44}$$

Figure 5.10 shows the shape of several raised cosine filter characteristics and the corresponding waveforms generated by the impulse response of these filters. The case of $\alpha = 0$ in Eq. (5.41) yields a filter with a bandwidth of $R_b/2$, the minimum bandwidth through which a bit rate R_b can be transmitted while still satisfying the zero ISI condition. Such a filter is not realizable in practice, since we cannot have an infinitely rapid attenuation slope at one frequency. Practical filters use values of α between 0.2 and 1.0. Figure 5.11 shows a baseband digital link with typical waveforms, and the corresponding spectra.

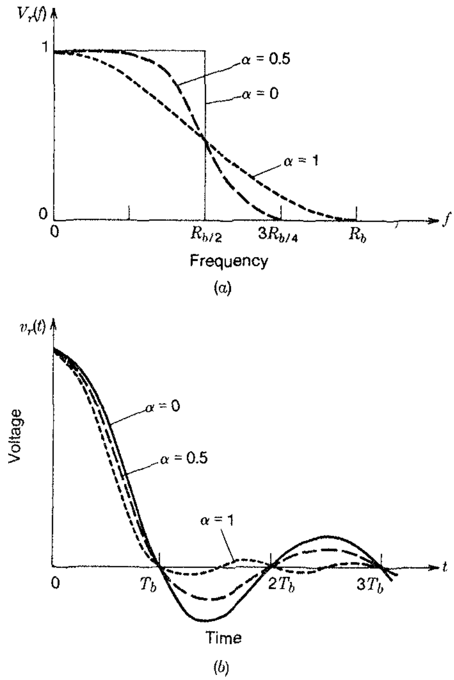


Figure 5.10 Raised-cosine filter frequency characteristic and impulse response. (*a*) Raised-cosine filter transfer characteristics. (*b*) Corresponding impulse responses.

Bandpass Transmission of Digital Data

In a radio frequency communication system that transmits digital data, a parameter of the RF wave must be varied, or modulated, to carry the baseband information. The most popular choice of modulation for a digital satellite communication system is phase shift keying (PSK), as described in the following section. Bandpass (or radio frequency) transmission of digital data differs from baseband transmission only because modulation of an RF wave is required: the receiver demodulates the modulated RF wave to recover the baseband data stream. Thus intersymbol interference will occur at the receiver due to bandlimiting of the modulated waveform unless filters that satisfy the Nyquist criterion are used.

An additional constraint usually exists with radio communication systems. The bandwidth occupied by a transmission is specified to avoid interference with other transmissions at adjacent frequencies: the output of a transmitter must have a carefully controlled spectrum that reduces out-of-band signals to a low level. Figure 5.12 shows the spectrum of a binary PSK (*BPSK*) signal generated from a random train of binary digits. The slow decay of the spectrum beyond $f_c \pm R_b$ results from the sudden phase reversals of the PSK waveform.

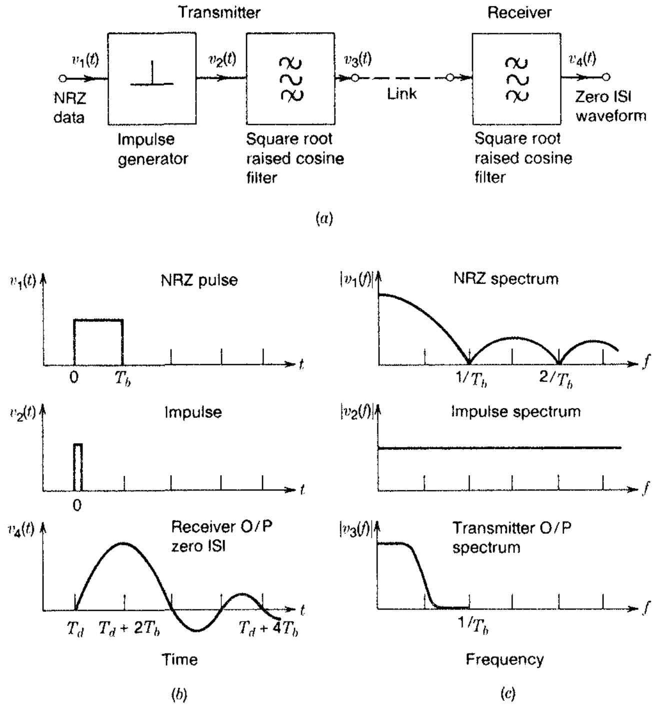


Figure 5.11 Waveforms and spectra in a baseband data system with raised-cosine filters. (*a*) System block diagram. (*b*) Waveforms. (*c*) Spectra.

In many data transmission systems the baseband waveform used has the NRZ format. Nyquist filters produce zero ISI waveforms only when driven by an impulse, as shown by Eqs. (5.42) and (5.43). If the filter is driven by a NRZ waveform, the spectrum of the driving pulse has a sin x/x shape, and the spectrum of the filter output will be $V_r(f)$ sin x/x [14]. To obtain zero ISI at the receiver, we must supply a signal with a spectrum $V_r(f)$, which can be achieved by using an equalizer with a frequency characteristic given by $x/\sin x$. The arrangement is illustrated in Figure 5.12*a*. The raised cosine filter cuts off at $f_c \pm f_0$ where $f_0 \leq 1/T_b$, so the $x/\sin x$ equalizer operates only within the central lobe of the sin x/x function. At $f = 1/T_b$, $x/\sin x$ goes to infinity, so α must be less than 1 for this system to work. In practice, RF filters with raised cosine shaping use α around 0.4, so the maximum gain at the edge of the equalizer band is 8.7 dB in this case.

The discussion of filter characteristics and signal spectra thus far has ignored the phase response of the filters and the resulting phase spectra of the waveforms. It can readily be shown [12] that the phase response of all filters and equalizers must be linear with frequency for the zero ISI condition to be met. Achieving a

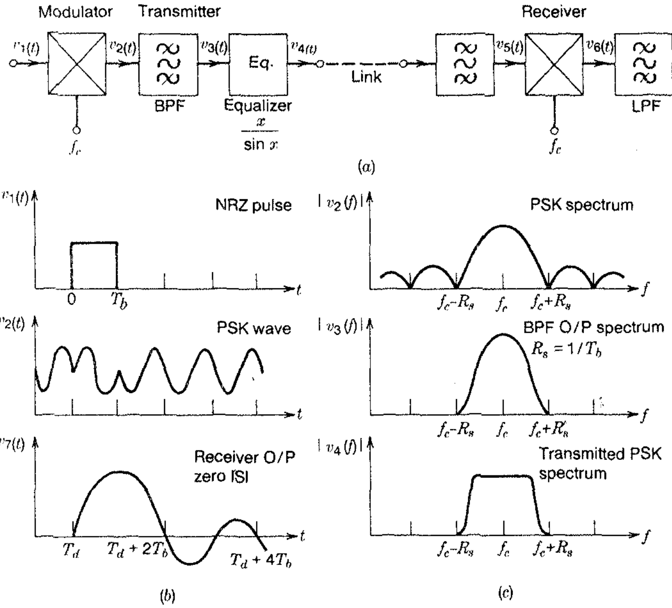


Figure 5.12 Waveforms and spectra in a PSK data system with raised-cosine filters and $x/\sin x$ equalization. (*a*) Block diagram of one channel of a QPSK system. (Equivalent to a BPSK system.) (*b*) Waveforms. (*c*) Spectra.

linear phase response throughout a communication system can be difficult in practice.

Transmission of QPSK Signals Through a Bandlimited Channel

QPSK (quadrature phase shift keying) is the most popular choice of modulation technique for use in satellite communication links carrying digital data. It will be described in more detail in Section 5.5, but basically a digital data stream is taken two bits at a time and used to generate one of four possible phase states of the transmitted carrier. If the data rate is R_b bits/s, the symbol rate for the QPSK carrier is $R_b/2$ bits/s = R_s symbols/s.

In order to recover the symbol stream with zero ISI, we must shape the transmitted spectrum such that after demodulation a single symbol creates a zero ISI waveform at the output of the demodulator. Then sampling of the symbol stream can be achieved with zero intersymbol interference. In practice, a QPSK system has two demodulators, one for each pair of symbols (phase states) in the QPSK carrier. We shall consider only one channel in looking at ISI.

Figure 5.12*a* shows a typical arrangement for one half of a QPSK transmit-receive link. The other half is identical except that the carrier used for modulation and demodulation is shifted in phase by 90°. Since the carriers in the two channels have a 90° phase difference, the channels are identified as *in-phase* (I) and *quadrature* (Q).

The data presented to the QPSK modulators is in NRZ format and causes a jump in carrier phase at each symbol transition. The input data rate to the demodulator is $R_s = R_b/2$, giving the QPSK spectrum shown in Figure 5.13*a*. The

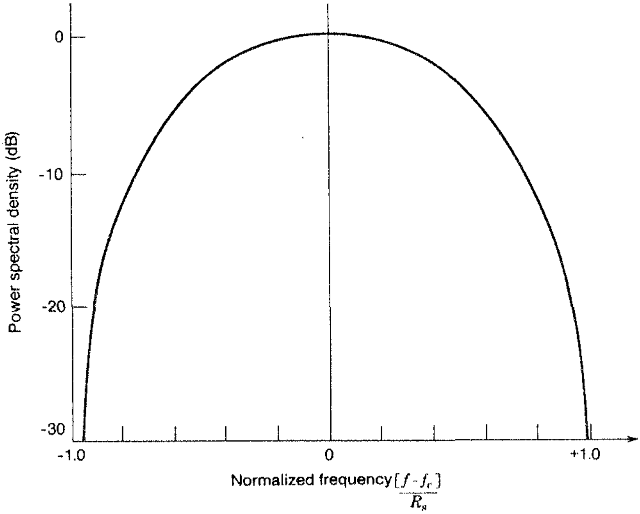


Figure 5.13*a* The frequency spectrum of a QPSK signal. Frequency is relative to the carrier and normalized to symbol rate R_s. Only the central lobe is shown.

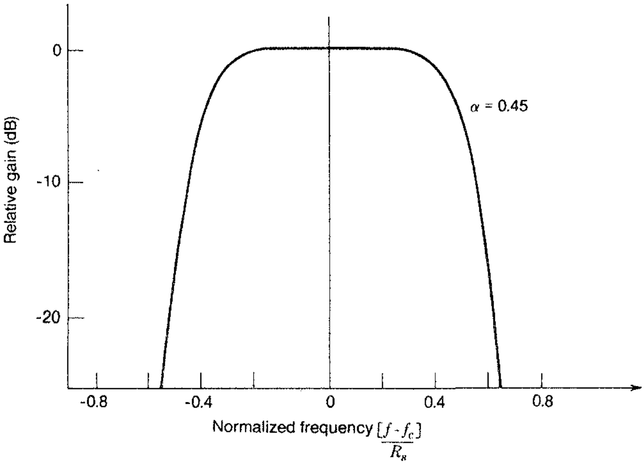


Figure 5.13*b* Frequency response of a bandpass square root raised-cosine filter with $\alpha = 0.45$.

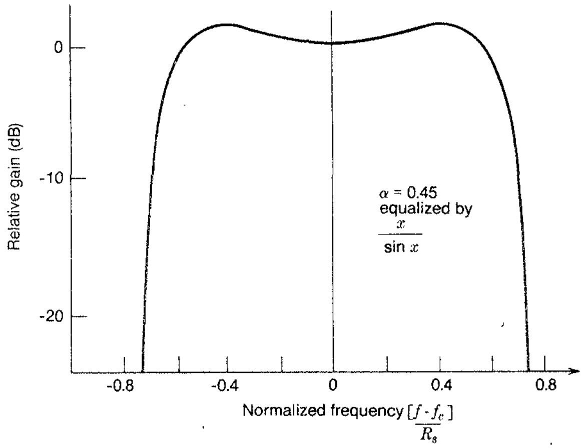


Figure 5.13c Frequency response of a raised-cosine filter with $\alpha = 0.45$, equalized by $x/\sin x$.

central lobe of the spectrum extends from $(f_c - R_s)$ to $(f_c + R_s)$, giving a band occupancy of $2R_s$. The spectrum must be narrowed for transmission via a radio channel, and this is achieved by use of a bandpass filter meeting the zero ISI criterion, for example, a raised cosine filter. The bandpass raised cosine filter is a transformation of the low-pass raised cosine filter and has a response $|H(f)| = 1/2$ at frequencies $(f_c - R_s/2)$ and $(f_c + R_s/2)$. The frequencies at which $H(f)$ falls to zero are determined by the roll-off factor α in Eq. (5.41). Matched filter operation of the link requires that the raised cosine filter response be split between the transmit end and the receive end of the link. Thus a square-root raised cosine response filter is required after the modulator and before the demodulator. Finally, because we are using NRZ pulses rather than impulses, we need an $x/\sin x$ equalizer with $x = [\pi(f_c - f)]/R_s$ to equalize the spectrum from the modulator. The frequency responses of some typical filters are shown in Figures 5.13*b*, *c*.

Example 5.4.1

A data stream at 240 Mbps is to be sent by QPSK on a microwave carrier. The receiver IF frequency is 240 MHz. Find the RF bandwidth needed to transmit the QPSK signal when raised cosine filters with $\alpha = 0.45$ are used.

The 240 Mbps signal is divided into two 120 Msps symbol streams and applied to I and Q channel modulators fed by an IF carrier, with a 90° phase difference. The resulting spectrum from each modulator has a width of 240 MHz between zeros of the central lobe of the PSK spectrum. The I and Q signals are added and applied to an $x/\sin x$ equalizer with

$$x = \frac{\pi(f_c - f)}{R_s}$$

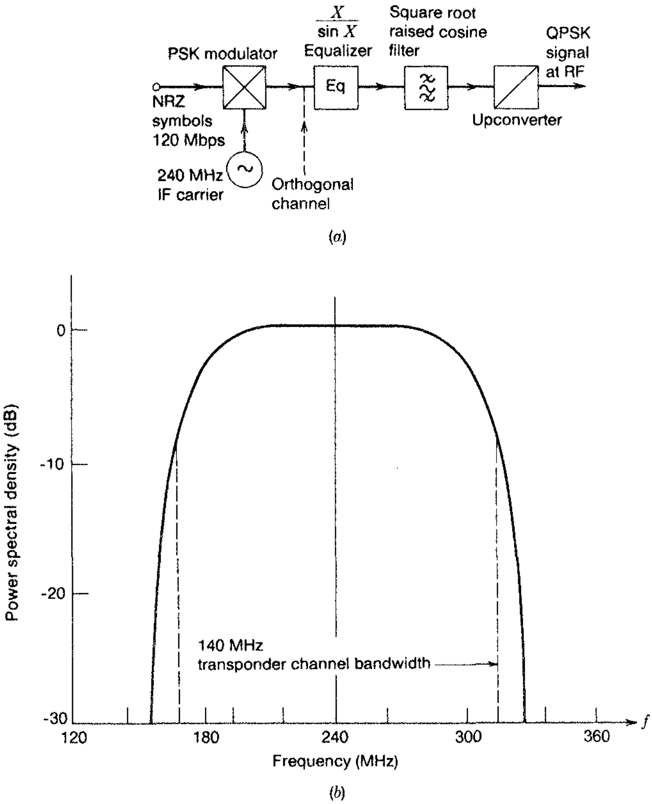


Figure 5.14 A QPSK data system with a NRZ format and $R_s = 120$ Msps symbol rate. IF frequency is 240 MHz, (*a*) Transmit end block diagram for model of 120 Msps symbol rate QPSK link. Only one channel is shown. (*b*) Transmitted QPSK spectrum after processing by square-root raised cosine filter and $x/\sin x$ equalization. Filter roll-off factor α is 0.45.

extending to ± 87 MHz from the carrier. The maximum gain of $x/\sin x$ at ± 87 MHz from the carrier is 9.53 dB. Figure 5.14*a* shows a block diagram of one half of the transmit end of the QPSK link.

The equalized spectrum is applied to the square-root raised cosine filter. The response of this filter is 3 dB down at $f_s \pm R_s/2$, that is, at $f_s \pm 60$ MHz. In practice, one filter combining the square-root raised cosine and $x/\sin x$ responses is used. The combined response of this single filter is shown in Figure 5.13*c*. Thus in the IF amplifier of the receiver, the signal spectrum is 6 dB down at 180 MHz and 300 MHz. The low-pass raised cosine filter with $\alpha = 0.45$ and $R_s = 120$ Msps has $|H(f)| = 0$ for $|f| > R_s/2 + \alpha R_s/2$, so the bandpass filter will have zero response for $f < f_c - (R_s/2)(1 + \alpha)$ and for $f > f_c + (R_s/2)(1 + \alpha)$, that is, below 153 MHz and above 327 MHz. Figure 5.14*b* shows the transmitted QPSK spectrum centered on the IF carrier.

If we examine the spectrum of the QPSK signal at the receiver, we find that the 3-dB bandwidth is 120 MHz and the total frequency band containing all of the signal energy is 174 MHz. A typical satellite transponder for such a signal would have a 3-dB bandwidth of 140 MHz. Beyond 140 MHz the spectrum of the QPSK signal would be attenuated by the transponder filter, leading to some spectral distortion of the receiver signal and consequent ISI in the demodulated waveform. However, the energy contained in the QPSK spectrum beyond ± 70 MHz from the carrier is small, and the ISI caused by the transponder filter is minimal.

Practical filters invariably cause some ISI because it is impossible to realize the raised cosine characteristic exactly. Appendix A.2 presents some spectra for PSK signals using Chebyshev and Butterworth filters. Typically, an extra 2 dB of carrier power must be provided to achieve a 10^{-6} BER, compared to the theoretical power level needed for this error rate, in a carefully filtered QPSK link. The extra power is sometimes called *implementation margin.*

5.5 DIGITAL MODULATION AND DEMODULATION

In this section we will review methods for digital transmission used on or proposed for current satellite links. We will not attempt to summarize the extensive literature of digital communications in general.

Terminology

While any feature of a signal—amplitude, frequency, or phase—may be digitally modulated, phase modulation is almost universally used for satellites. For historical reasons, digital phase modulation is frequently called *phase shift keying*, abbreviated PSK. An M-phase PSK modulator puts the phase of a carrier into one of M states according to the value of a modulating voltage. Two-state or biphase PSK is usually called BPSK, and four-state or quadriphase PSK is termed QPSK. Other numbers of states and some combinations of amplitude and phase modulation are possible and are employed in terrestrial links, but satellite users have been reluctant to adopt anything besides BPSK or QPSK. An important reason for this is the high values of (C/N) required for acceptable bit error rates—typically > 26 dB. Any type of PSK can be *direct* or *differential*, depending on whether it is the state of the modulating voltage or the *change* in state of the modulating voltage that determines the transmitted phase.

Whether direct or differential, a PSK modulator causes the phase of a carrier waveform to go to one of a finite set of values. The transition time plus the time spent at the desired phase constitute a fixed time interval called the *symbol period;* the transmitted waveform during the interval is called a *symbol.* The set of all symbols for a particular modulation type is called its *alphabet*. Thus BPSK has a two-symbol alphabet and QPSK has a four-symbol alphabet.

In the digital modulation process, a stream of incoming binary digits (bits) determines which symbol of the M available in the alphabet will be transmitted. Mathematically, N_b bits are required to specify which of M possible symbols is being transmitted where N_b and M are related by

$$N_b = \log_2(M) \tag{5.45}$$

As defined by this equation, N_b is the number of *bits per symbol* for the M-PSK modulation scheme. Standard practice is to make M a power of 2 so that N_b will be an integer.

Modulation and Coding

The boundary between digital modulation and digital encoding is not well defined. In encoding for forward error correction (*FEC*), redundant bits are added to an incoming bit stream so that errors in transmission may be detected and corrected at the other end of the link. When the redundant bits are added at baseband and the composite (information bits plus redundant bits) bit stream is used to phase modulate a carrier and produce the transmitted symbols, then the division between modulation and encoding is obvious. But the modulator itself may be designed to add redundant bits during the modulation process, making encoding and modulation inseparable. In this section we will ignore encoding for FEC and concentrate strictly on the modulation process for turning an incoming bit stream into RF symbols. We will assume that any FEC encoding is done ahead of the modulator by the methods to be presented in a later chapter.

It is unfortunate that differential phase modulation is frequently called differential encoding, since it is a characteristic of the modulation and demodulation equipment and plays no role in coding as it is usually understood. Differential encoding would more properly be called differential modulation, and we will discuss it after we have presented direct modulation.

Bit and Symbol Error Rates

The figure of merit for a digital radio link is its *bit error rate* (*BER*), also called the *bit error probability* (*PB*). Mathematically this is the probability that a bit sent over the link will be received *incorrectly* (i.e., that a 1 will be read as a 0 or vice versa) or, alternatively, the fraction of a large number of transmitted bits that will be received incorrectly. Like a probability, it is usually stated as a single number—for example 1×10^{-4} or .0001. The BER plays the same role as an indicator of quality in a digital communication system that the signal-to-noise ratio plays in an analog link.

Physically a bit error occurs because a *symbol error* has occurred. At some point in the link noise has corrupted the transmitted symbol so that the decision circuitry at the receiver cannot identify it correctly. For example, the carrier phase may have been transmitted as $+90°$ but additive noise may have changed the received carrier phase to $-90°$. If one symbol carries N_b bits and if differential modulation is not used, then a single symbol error may cause 1, 2, . . . N_b bit errors. With differential modulation, an error on one symbol will cause the symbol that follows to be misinterpreted, and the number of bit errors per symbol error may exceed N_b, the number of bits per symbol.

Symbol errors arise from thermal noise, from external interference, and from intersymbol interference. If only thermal noise is considered, then the *symbol error rate* (*SER*) or *symbol error probability* (*PE*) may be calculated unambiguously from (E_s/N_0), the energy per symbol in joules divided by the noise density in W/Hz, measured in the IF bandwidth at the demodulator input. The higher the value of (E_s/N_0), the lower will be the SER. (E_s/N_0) may be determined from the input value of (C/N), expressed as a ratio.

Assume that C W of carrier power are transmitted during one symbol interval T_s. The energy received during that symbol period is E_s, where

$$E_s = CT_s = \frac{C}{R_s} \tag{5.46}$$

where R_s is the symbol rate in symbols/second. The noise-density N_0 is the received noise power N divided by the IF bandwidth at the demodulator input

$$N_0 = N/B \tag{5.47}$$

Combining the last two equations we have

$$\frac{E_s}{E_0} = \frac{C}{N} \cdot \frac{B}{R_s} \tag{5.48}$$

The square-root Nyquist cosine filter discussed in the preceding section has a noise bandwidth B equal to the symbol rate R_s. Thus a receiver designed with filters of this type to achieve zero ISI also has $BT_s = 1$ and $E_s/N_0 = C/N$. Practical filters such as Butterworth or Chebychev come close to the shape of the square root raised cosine filter, giving BT products close to unity.

While for BPSK bit and symbol errors are the same thing, for modulation schemes with $M > 2$, the relation between the bit error rate and the symbol error rate is not consistently defined in the literature. Equation (5.49), derived in reference 15, is based on the probability that a particular bit carried by a symbol is in error, given that the symbol itself is in error.

$$\text{PB} = \frac{1}{2}\frac{PE}{1 - 2^{-N_b}} \tag{5.49}$$

Here PB and PE are the corresponding values of bit and symbol error probability, and N_b is the number of bits per symbol. For QPSK, $N = 2$ and $PE = 1.5$ PB. Another approach, derived in reference 16 and more frequently quoted in the literature than Eq. (5.49), assumes that a large block of data is to be transmitted either in the form of serial bits or in symbols of M-ary PSK. Each symbol carries

2^M bits. The derivation equates the probability that the word will be received correctly in BPSK with the probability that it will be received correctly in the M-ary system. For PE much less than 1, this equality leads to Eq. (5.50):

$$PB = \frac{\ln 2}{\ln M} PE = \frac{PE}{\log_2 M} \tag{5.50}$$

For QPSK, $M = 4$ and Eq. (5.50) yields $PE = 2\ PB$.

The reader may find it confusing that, while Eqs. (5.49) and (5.50) yield slightly different results for the relation between PB and PE, both show the bit error rate PB to be less than the symbol error rate PE for QPSK. This happens because each QPSK symbol carries two bits. When a given symbol is sent, three symbol errors are possible (i.e., a 00 may be detected as an 01, 10, or 11), but two of these cause only a single bit error, and one will leave a given bit unchanged. Thus if we look at one particular bit out of the two carried by the symbol, the probability that a symbol error will change that bit is about two-thirds—that is, two of the three possible symbol errors that can be made will change the bit. Hence we would expect the bit error rate probability PB to be two-thirds the symbol error probability PE, and this is what Eq. (5.49) yields.

Binary Phase Shift Keying (BPSK)

In binary phase shift keying, an incoming bipolar bit stream $u(t)$ sets the phase of a carrier to plus or minus 90° ($\pi/2$ rad). Thus, if u_i is the ith bit, then the transmitted carrier v_c is given by

$$v_c = V \cos\left(\omega_c t - u_i \frac{\pi}{2}\right) \tag{5.51}$$

where V is an arbitrary amplitude frequently set to 1. Since u_i must have a value of ± 1, a logical one is transmitted by setting the phase to $-\pi/2$ rad and a logical zero (baseband -1) is transmitted with a $+\pi/2$ phase. Using trigonometric identities, we may rewrite Eq. (5.51) as

$$v_c = V u_i \sin(\omega_c t) \tag{5.52}$$

and we see that BPSK resembles amplitude modulation in which the modulating signal has a value $+1$ or -1 only. This causes the BPSK waveform to have a constant amplitude and an envelope AM detector cannot demodulate it.

To recover u_i the receiver must compare the phase of the received signal with that of a reference voltage that has the same phase as the original unmodulated carrier. This may be done with the simple product detector (mixer) of Figure 5.15 where the output voltage v_o is ideally u_i. At the center of each symbol interval, a decision circuit decides whether v_o is positive or negative and thus determines whether u_i was a $+1$ and represented a one or whether it was a -1 and represented a zero. This technique is called coherent detection since it requires a reference voltage that is phase coherent with the transmitter carrier.

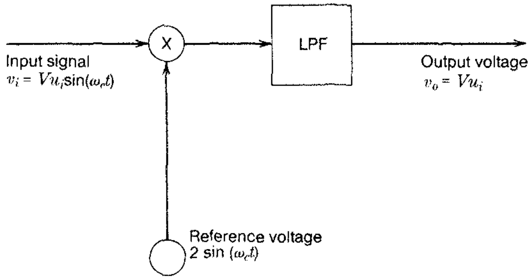


Figure 5.15 A coherent BPSK detector.

The decision circuit will make an error if noise changes the sign of v_o. We may calculate the probability that this will happen and thus the symbol error rate PE by the following argument, based on reference 16. Let the channel noise voltage $n(t)$ be Gaussian distributed with zero mean and rms value σ. Assume that u_i is a -1. At the decision time, v_o is given by

$$v_o = n(t) - V \tag{5.53}$$

If v_o is positive, the decision circuit will interpret u_i as $+1$. That will happen if $n(t) > V$; see Figure 5.16. The probability of $n(t)$ being greater than V is

$$P(N > V) = PE = \tfrac{1}{2}\mathrm{erfc}\left(\frac{V}{\sigma\sqrt{2}}\right) \tag{5.54}$$

The complementary error function, abbreviated erfc, is given by

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-u^2} \cdot du \tag{5.55}$$

where $u = x/\sigma\sqrt{2}$. See reference 6, pp. 60ff., for a discussion of its properties. Equation (5.45) involves the erfc because the noise voltage has a Gaussian distribution.

One symbol lasts for T_s s. The power in the symbol waveform is $V^2/2R$, where R is the input resistance of the decision circuit. Thus the energy per symbol, E_s, is given by

$$E_s = \frac{V^2}{2R} T_s \tag{5.56}$$

assuming that we have a matched filter in our receiver. Then

$$V = \sqrt{\frac{2RE_s}{T_s}} \tag{5.57}$$

and the noise power is given by σ^2/R. The noise density N_0 is the noise power divided by the bandwidth. If we assume that the bandwidth is $1/T_s$, then N_0 is

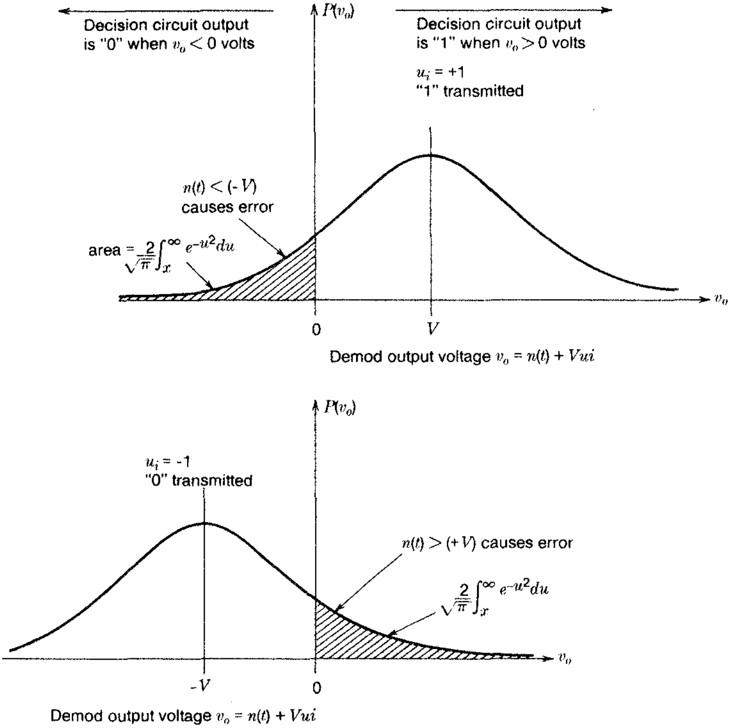


Figure 5.16 Illustration of errors in a binary decision circuit caused by additive Gaussian noise. The threshold is at zero volts.

given by

$$N_0 = \frac{\sigma^2}{R} T_s \tag{5.58}$$

and

$$\sigma = \sqrt{\frac{RN_0}{T_s}} \tag{5.59}$$

Combining Eqs. (5.54), (5.58), and (5.59) yields

$$PE = \tfrac{1}{2}\text{erfc}\left[\frac{\sqrt{2RE_s/T_s}}{\sqrt{RN_0/T_s}} \frac{1}{\sqrt{2}}\right] = \tfrac{1}{2}\text{erfc}\left[\sqrt{\frac{E_s}{N_0}}\right] \tag{5.60}$$

Since for BPSK a bit and a symbol are the same thing, Eq. (5.60) is often written

$$PB = \tfrac{1}{2}\text{erfc}(\sqrt{E_b/N_0}) \tag{5.61}$$

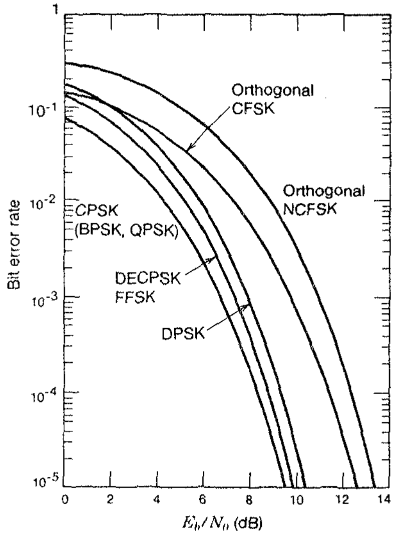


Figure 5.17 Bit error rate versus (E_b/N_0) for a variety of digital modulation and demodulation schemes. CPSK stands for coherent phase shift keying and is the same as the BPSK discussed in the text. DECPSK is differentially encoded PSK, and DPSK is differentially detected PSK. The other curves represent fast frequency shift keying (FFSK), coherently detected frequency shift keying (CFSK), and noncoherently detected frequency shift keying (NCFSK). (Reprinted with permission from V. K. Bhargava, D. Haccoun, R. Matyas, and N. Nuspl, *Digital Communications by Satellite*, John Wiley & Sons, New York, 1981.)

Since coherent detection is the most efficient way of demodulating direct BPSK, Eq. (5.61) is the relation normally used to determine the (E_b/N_0) and hence the (C/N) that a satellite link must maintain to meet a specified bit error rate requirement. Figure 5.17 displays a curve of PB versus (E_b/N_0).

The reference carrier in Figure 5.15 may be generated from the received signal using a carrier recovery circuit, a form of phase locked loop. See Chapter 5 of reference 17 for a detailed discussion of carrier recovery techniques.

Most carrier recovery loops have a 180° phase ambiguity, that is, when the loop is locked the phase of the recovered carrier may differ by 180° from the correct value. This has the effect of interchanging logical ones and zeroes and causes the demodulated bit stream to be the complement of what was transmitted. There are several ways to eliminate the ambiguity; one is to use differential encoding in which adjacent symbols have the same phase if the modulating voltage is a 1 and are 180° out of phase if it is a 0. This may be realized by a binary phase shifter that toggles between 0° phase shift and 180° phase shift each time the modulating bit is a 0. Incoming 1 values have no effect.

Differential modulation is more error prone than direct modulation, since an error on a single bit in a differential system will cause one or more subsequent bits to be interpreted incorrectly. See Section 5.6 of reference 15 for a detailed analysis of errors in differential PSK. Most practical satellite systems avoid differential encoding and check the status of the recovered carrier phase periodically by transmitting a known word. Logic at the receiver looks for this word. If it receives it correctly, then the recovered carrier phase is correct. If it receives the complement of the known word, than the recovered carrier phase is off by 180° and the demodulated data stream should be complemented before it is sent to the end user.

Quadrature Phase Shift Keying (QPSK)

In QPSK the phase ϕ of the carrier is set by the modulator to one of four possible values. We may write the result as

$$v = V\sqrt{2}\cos(\omega_c t - \phi) \tag{5.62}$$

where ϕ takes on the values $\pi/4$, $3\pi/4$, $5\pi/4$, and $7\pi/4$ rad. The factor 2 is for our later convenience. Using trigonometric identities to expand Eq. (5.62) we obtain

$$v = V\sqrt{2}\cos\omega_c t\cos\phi + V\sqrt{2}\sin\omega_c t\sin\phi \tag{5.63}$$

The first term is a BPSK signal in phase with the carrier; it is called the I channel. The second term is a BPSK signal in quadrature with the carrier and is called the Q channel. Thus a QPSK waveform may be generated by combining two BPSK waveforms in quadrature. We may write the result as

$$v = u_I V\cos\omega_c t + u_Q V\sin\omega_c t \tag{5.64}$$

where u_I represents a binary data stream modulating the I channel and u_Q represents a binary data stream modulating the Q channel. On both of these a logical 1 corresponds to u_I or $u_Q = +1$ and a logical 0 corresponds to u_I or $u_Q = -1$. The relationship between u_I, u_Q, and ϕ is given by

$$u_I = \sqrt{2}\cos\phi \tag{5.65}$$

$$u_Q = \sqrt{2}\sin\phi \tag{5.66}$$

and is summarized in Table 5.2. Note that ϕ is conveniently visualized as the phase angle of a phasor whose real component is u_I and whose imaginary component is u_Q. See Figure 5.18.

The bits u_I and u_Q may be selected alternately from one incoming bit stream. For example, u_I may represent the odd-number bits and u_Q the even. In this case one binary data channel enters the QPSK modulator and the outgoing symbol rate is equal to half of the incoming bit rate. Alternatively u_I and u_Q represent binary data channels coming from independent sources, and QPSK may be viewed

Table 5.2
The Relationship Between the Modulating Bit Streams u_I, u_Q and the Phase Angle ϕ of the Modulated QPSK Waveform

u_I	u_Q	ϕ
1	1	$\pi/4$
−1	1	$3\pi/4$
−1	−1	$5\pi/4$
1	−1	$7\pi/4$

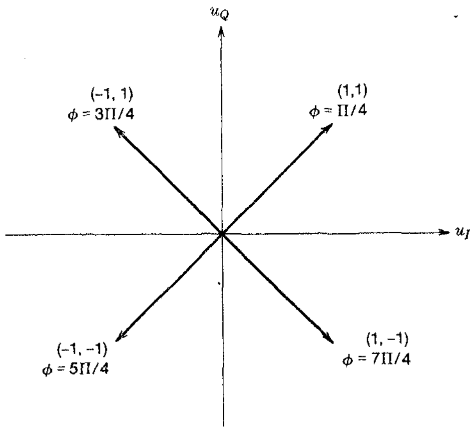


Figure 5.18 Phasor diagram showing the phase of a QPSK waveform modulated with all possible pair combinations of the bits (u_I, u_Q).

as a form of digital multiplexing that combines two BPSK signals with orthogonal carriers. This is the interpretation we first put on Eq. (5.64). When u_I and u_Q come from independent channels, then the incoming bit rate on each of two modulator inputs is equal to the outgoing symbol rate.

QPSK modulators and demodulators are basically dual-channel BPSK modulators and demodulators. One channel processes the u_I bits and uses the reference carrier; the other processes the u_Q bits and uses a 90° phase shifted version of the reference. Figures 5.19 and 5.20 [17] show generalized block diagrams of a QPSK modulator and demodulator. More detailed information is available in reference 18.

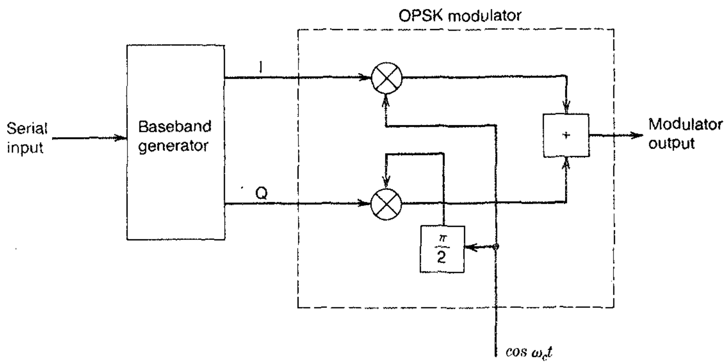


Figure 5.19 A generalized QPSK modulator. (Reprinted with permission from V. K. Bhargava, D. Haccoun, R. Matyas, and N. Nuspl, *Digital Communications by Satellite*, John Wiley & Sons, New York, copyright © 1981.)

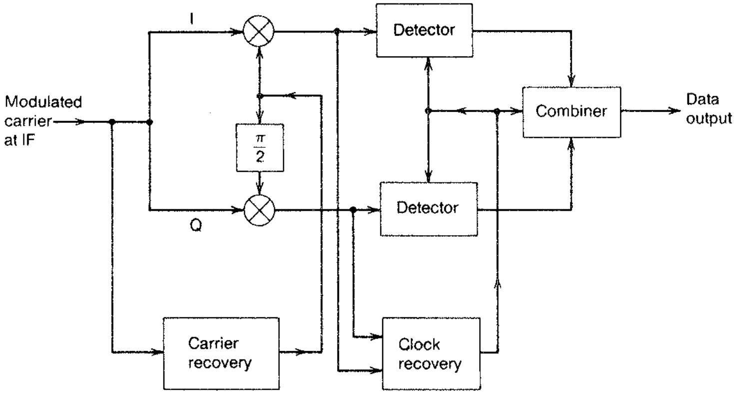


Figure 5.20 A generalized QPSK demodulator. (Reprinted with permission from V. K. Bhargava, D. Haccoun, R. Matyas, and N. Nuspl, *Digital Communications by Satellite*, John Wiley & Sons, New York, copyright © 1981.)

If the transmitted phase angle ϕ takes on the values $\pi/4$, $3\pi/4$, $5\pi/4$ or $7\pi/4$ rad, then a QPSK receiver must simply decide in which quadrant the received phasor signal lies. A decision circuit interprets the phase of all signals that lie in the first quadrant as $\pi/4$; all those in the second quadrant are assumed to have been transmitted with a phase of $3\pi/4$, and so on.

Symbol errors occur when noise pushes the received phasor into the wrong quadrant. Figure 5.21 illustrates the process. In it we assume that the transmitted

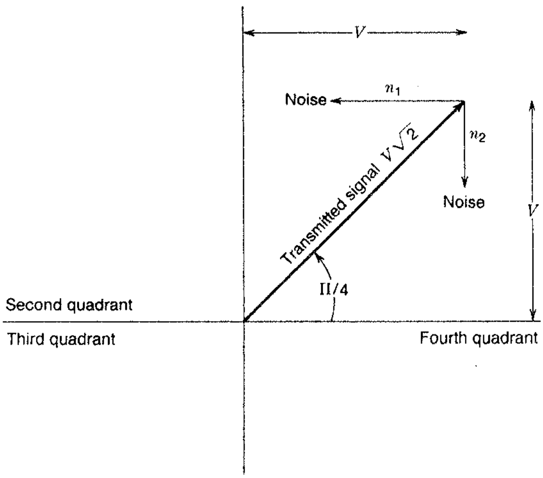


Figure 5.21 Phasor diagram of QPSK signal with phase of $\pi/4$ rad and narrow band noise components n_1 and n_2.

symbol had a phase of $\pi/4$ rad, corresponding to $u_I = 1$ and $u_Q = 1$. Narrow-band white Gaussian noise may be resolved into independent orthogonal components that are respectively in-phase and in quadrature with any arbitrary phasor. We will orient the noise so as to make the component noise phasors n_1 and n_2 point in the directions that are most likely to cause symbol errors.

A symbol error will occur if either n_1 or n_2 exceeds V. The former will place the apparent phase in the second quadrant, while the latter will put it in the fourth quadrant. If both exceed V at the same time, the resultant will go into the third quadrant.

Let the rms noise voltage be σ. This is also the rms value of n_1 and n_2. The probability that n_1 exceeds V is P_A.

$$P_A = \tfrac{1}{2}\operatorname{erfc}\left(\frac{V}{\sigma\sqrt{2}}\right) \tag{5.67}$$

The probability that n_2 exceeds V is P_B. Since n_1 and n_2 have the same rms value, $P_A = P_B$. The probability that n_1 and n_2 simultaneously exceed V is P_A^2. We will assume that this is negligible in comparison to P_A.

An error will occur if either $n_1 > V$ or $n_2 > V$. Since the noise components are independent the probability PE of this happening is

$$PE = 2P_A = \operatorname{erfc}\left(\frac{V}{\sigma\sqrt{2}}\right) \tag{5.68}$$

The rms signal power is proportional to V^2. Following steps similar to those that led to Eq. (5.51), we obtain

$$PE = \operatorname{erfc}\left(\frac{1}{\sqrt{2}}\frac{E_s}{N_0}\right) \tag{5.69}$$

At this point we should emphasize that PE in Eq. (5.68) is the *symbol* error rate for QPSK. Some texts present the same result as if it were a bit error rate, PB. We may calculate PB for QPSK by recognizing that there are two bits per symbol, and hence $E_s = 2E_b$. If we use Eq. (5.50) to relate PB to PE, we find that $PB = 0.5\ PE$. Thus the bit error rate for QPSK is given by

$$PB = \tfrac{1}{2}\operatorname{erfc}(\sqrt{E_b/N_0}) \tag{5.70}$$

Thus QPSK and BPSK have the same bit error rate for the same E_b/N_0 and the error performance of the two modulation systems would seem to be identical; a plot of PB versus (E_b/N_0) for QPSK would be the same as that shown in Figure 5.17.

To compare QPSK and BPSK on an equal basis, assume that we must send R_0 bits per second over a satellite link with fixed bandwidth B and a fixed value of (C/N). For QPSK $R_s = R_0/2$. Hence by Eq. (5.48)

$$\frac{E_s}{N_{0\,\text{QPSK}}} = \frac{C}{N}\frac{R_s}{B_{\text{QPSK}}} = \frac{CR_0}{N2B} \tag{5.71}$$

For BPSK, $R_s = R_0$ and

$$\frac{E_s}{N_{0\,\text{BPSK}}} = \frac{C}{N} \cdot \frac{R_s}{B_{\text{BPSK}}} = \frac{CR_0}{NB} \tag{5.72}$$

By Eq. (5.61)

$$PB_{\text{PSK}} = \tfrac{1}{2}\,\text{erfc}\left(\frac{R_0}{B}\frac{C}{N}\right) \tag{5.73}$$

and by Eqs. (5.69) and (5.50)

$$PB_{\text{QPSK}} = \tfrac{1}{2}\,PE_{\text{QPSK}} = \tfrac{1}{2}\,\text{erfc}\left(\tfrac{1}{2}\frac{R_0}{B}\frac{C}{N}\right) \tag{5.74}$$

Thus QPSK will have a higher BER than BPSK when the two modulation schemes are compared for equal bit rates, bandwidths, and (C/N) values. But QPSK carries twice as much data as BPSK for the same RF bandwidth and using QPSK can double the communications capacity (and revenue-earning power) of a transponder. Some TDMA systems (see Chapter 6) use BPSK in their preamble for rapid and accurate establishment of a link that will subsequently carry QPSK.

Thus far we have discussed direct QPSK. In differential QPSK [18] the carrier phase ϕ of Eq. (5.63) *changes* by an integer multiple of $\pi/2$ rad. The value of the integer depends on the incoming bits u_I and u_Q. Because of the higher error rate associated with differential modulation, this technique has not been widely adopted.

QPSK Variants

We noted after Eq. (5.64) that QPSK may be visualized as the sum of two independent BPSK signals whose carriers are in phase quadrature. In conventional QPSK the bits u_I and u_Q that modulate these carriers both make step changes at the same time. If the bit changes are staggered so that u_I makes step changes at the beginning of each symbol period and u_Q makes step changes at the midpoint of each symbol period, the result is called *staggered QPSK* (*SQPSK*) or *offset QPSK* (*OQPSK*). If instead of steps the bits make sinusoidal transitions between their allowed values of 1 and -1, the result is *minimum shift keying* (MSK) or *fast frequency shift keying* (FFSK). These modulation systems produce spectra that are slightly different from conventional QPSK and that would appear to have some advantages over QPSK for satellite transmission. While they have received considerable academic attention, OQPSK, SQPSK, MSK, and FFSK have not yet been adopted for commercial satellite applications. At the time of writing, the prevailing attitude in the industry seems to be that any theoretical advantages that they might have over conventional QPSK either vanish when these techniques are used over a real transponder or else are so slight as not to justify the added expense that their implementation would require. Because of space limitations and the present lack of practical applications of these QPSK variants, we will not discuss them further in this text. For additional information the reader should consult references 17 and 18.

5.6 DIGITAL TRANSMISSION OF VOICE

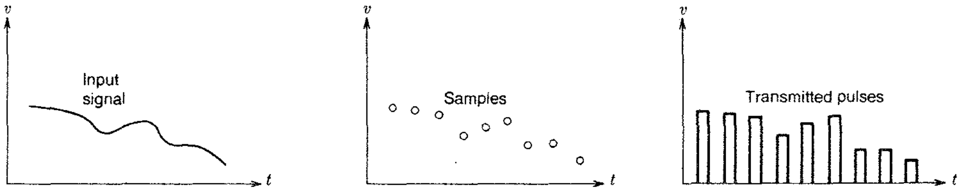
The previous sections have discussed techniques for transmitting and receiving digital information via satellite. Now we will turn our attention to the problem of putting analog voice signals into digital form for transmission and returning them to analog form after reception. While the material presented is generally applicable to all analog signals, we will emphasize baseband voice channels because these are virtually the only analog signals sent over commercial satellite links in digital form.

Sampling and Quantizing

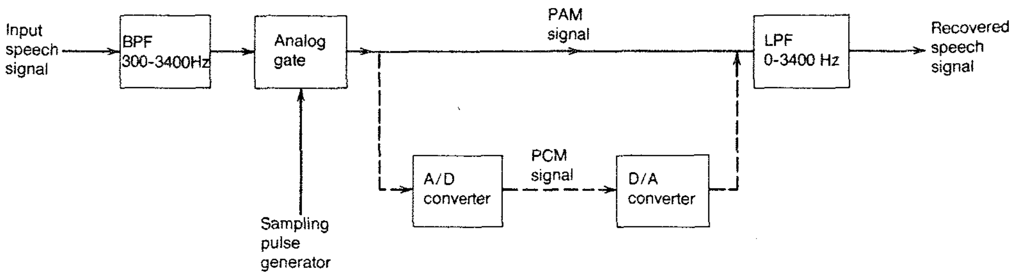
The basic processes in digital transmission of analog information are sampling, quantizing, and encoding. The principles underlying sampling are routinely presented in beginning courses in communications theory, and we will not reproduce them here. See Chapter 5 of reference 6 or Chapter 2 of reference 18 for details. The sampling theorem states that a signal may be reconstructed without error from regularly spaced samples taken at a rate f_s (samples/second), which is at least twice the maximum frequency f_m present in the signal. Instead of transmitting the continuous analog signal, we may transmit the samples. For example, voice signals on satellite links are normally filtered to limit their spectra to the range 300 to 3400 Hz. Thus, one voice channel could be transmitted with samples taken at least 6800 times per second or, as it is usually expressed, with a minimum sampling frequency of 6800 Hz. Common telephone system practice is to use a sampling frequency of 8000 Hz. While transmitting the samples requires more bandwidth than transmitting the original waveform, the time between samples of one signal may be used to transmit samples of other signals. This is *time division multiplexing* (TDM), and we will discuss it later in this chapter.

The samples to which the sampling theorem refers are analog pulses whose amplitudes are equal to that of the original waveform at the time of sampling. The original waveform may be reconstructed without error by passing the samples through an ideal low-pass filter whose transfer function is appropriate to the sampling pulse shape. A communications system that samples an input waveform and transmits analog pulses is said to use *pulse amplitude modulation.* Figure 5.22 sketches this process.

Analog pulses are subject to amplitude distortion, and they are also incompatible with conventional baseband digital signals in which pulses take on only one of two possible values. Hence pulse amplitude modulation is not used over satellite links. Instead, the analog samples are *quantized*—resolved into one of a finite number of possible values—and the quantized values are binarily encoded and transmitted digitally. Thus each sample is converted into a digital word that represents the quantization value closest to the original analog sample. *Quantization* may be *uniform* or *nonuniform* depending on whether or not the quantized voltage levels are uniformly or nonuniformly spaced. At the receiver a digital-to-analog (D/A) converter converts each incoming digital word back into an analog sample; these analog samples are filtered and the original input waveform is reconstructed.



(a) Basic waveforms.



(b) Block diagram of a PAM communications system. Dashed lines show the additional components that would convert it to a PCM system.

Figure 5.22 Pulse amplitude modulation (PAM). (a) Basic waveforms. (b) Block diagram of a PAM communications system. Dashed lines show the additional components that would convert it to a PCM system.

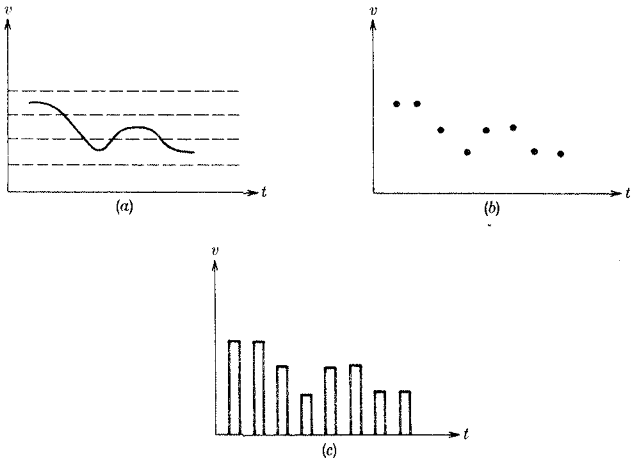


Figure 5.23 The quantizing process. (*a*) The input waveform and the quantization levels. (*b*) Quantized samples. (*c*) Quantized pulses. Their amplitude will be encoded digitally for PCM transmission.

A communications system that transmits digitally encoded quantized values is called a *pulse code modulation* (*PCM*) system.

The quantization process, illustrated in Figure 5.23, prevents exact reconstruction of the digitized waveform. (The sampling theorem requires that analog rather than quantized samples be transmitted.) The error introduced is called *quantization error*; and a person listening to a reconstructed speech signal perceives the quantization error as an added noise called *quantization noise.*

A uniform quantizer (Figure 5.24) operates with L levels spaced Δ volts apart. The input signal is amplitude limited to lie between $-\Delta(L/2)$ and $+\Delta(L/2)$. The quantizer determines in which level an incoming sample falls and puts out the identification number of that level. This identification number is the digital word that represents the sample. Transmitting L levels requires N_l bits where

$$N_l = \log_2 L \tag{5.75}$$

or

$$L = 2^{N_l} \tag{5.76}$$

The levels are normally numbered 0 through $L - 1$. Thus an 8-bit ($N = 8$) PCM system would quantize its incoming samples into one of 256 ($L = 2^8 = 256$) levels numbered 0 through 255. These would be transmitted as binary words ranging from 00000000 (decimal 0) through 11111111 (decimal 255).

If the input signal amplitude is uniformly distributed with an rms value of V_{rms}, the signal-to-noise ratio (S/N) of the reconstructed signal (assuming that *only* quantization noise is present) is given by [19]

$$(\text{S/N}) = 12\left(\frac{V_{\text{rms}}}{\Delta}\right)^2 \tag{5.77}$$

	v	
	4Δ	
Level 7		Signals with $3\Delta < v \leq 4\Delta$ are transmitted as 3.5Δ, encoded 111
	3Δ	
Level 6		Signals with $2\Delta < v \leq 3\Delta$ are transmitted as 2.5Δ, encoded as 110
	2Δ	
Level 5		Signals with $\Delta < v \leq 2\Delta$ are transmitted as 1.5Δ, encoded as 101
	Δ	
Level 4		Signals with $0 < v \leq \Delta$ are transmitted as 0.5Δ, encoded as 100
	0	
Level 3		Signals with $-\Delta < v \leq 0$ are transmitted as -0.5Δ, encoded as 011
	$-\Delta$	
Level 2		Signals with $-2\Delta < v \leq \Delta$ are transmitted as -1.5Δ, encoded as 010
	-2Δ	
Level 1		Signals with $-3\Delta < v \leq -2\Delta$ are transmitted as -2.5Δ, encoded as 001
	-3Δ	
Level 0		Signals with $-4\Delta \leq v \leq -3\Delta$ are transmitted as -3.5Δ, encoded as 000
	-4Δ	

Figure 5.24 Levels and encoding for a uniform 3-bit (8 level) PCM quantizer.

It is common practice (but not a requirement) for uniform quantizers to be designed so that

$$\Delta 2^{N_l} = 8V_{\mathrm{rms}} \tag{5.78}$$

For uniform quantization and a sine wave input

$$(\mathrm{S/N}) = \frac{3}{16} L^2 \tag{5.79}$$

or in decibels

$$(\mathrm{S/N}) = 6N_l - 7.27\,\mathrm{dB} \tag{5.80}$$

For a uniformly distributed input

$$(\mathrm{S/N}) = L^2 - 1 \simeq L^2 \qquad \text{when} \qquad N \geq 5 \text{ bits} \tag{5.81}$$

In Section 5.1 we noted that a common specification for satellite telephone channels is a weighted (S/N) of 51.25 dB (7500 pWp of noise). Achieving this performance with uniformly quantized PCM would require $N_l = 10$ bits per sample [rounded up from the 9.75 calculated using Eq. (5.80)]. At 8000 samples per second, a single telephone channel would require 80,000 bits per second (80 kbps) and a bandwidth of about 80 kHz for transmission by uniform PCM and BPSK. This is much greater than the 4 kHz required for SSB analog transmission. While the 80 kHz can be reduced to 64 kHz by the techniques of the next section, digital transmission is inherently bandwidth inefficient when a small number of signals are to be sent.

Nonuniform Quantization: Compression and Expansion

Uniform quantization introduces more noise when a signal is small and one quantization interval is large in comparison with the signal than it does when the signal is large and one quantization interval is insignificant.. Improved noise performance can be obtained using *nonuniform quantization* in which the size of the quantization intervals increases in proportion to the signal value being quantized. The same effect can be obtained from a uniform quantizer if the input signal is compressed before quantization. The distortion introduced by the compressor must be removed at the receiver by an expander. The transfer functions of the compressor and expander are complementary, that is, their product is a constant and the amplitude distribution of a signal that has passed through both a compressor and an expander is unchanged.

Companding was first employed on terrestrial telephone systems using analog compressors that had logarithmic transfer functions. These were the so-called mu-law and A-law compressors. Later developments in digital technology allowed

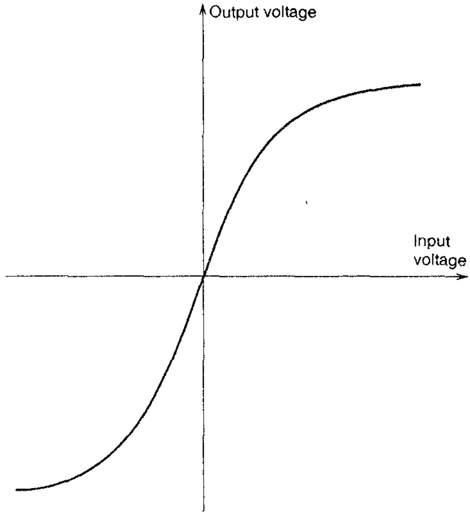


Figure 5.25 The transfer characteristic of a compressor.

digital implementation of the compression and expansion functions and permitted the sampling, compression, quantization, and encoding operations to be combined into one piece of equipment called a *coder*. Common digital compression schemes are the 15-segment coder that provides up to 30 dB improvement over a uniform quantizer with the same number of levels and the 13-segment coder that yields up to 24 dB improvement. (These numbers are taken from reference 1, whose Chapter 28 provides a good discussion of the companding process. See also Chapter 3 of reference 8.) Thus, with a 15-segment coder the 51.25-dB (S/N) requirement of the previous section may be met by the number of bits required by a nonuniform quantizer that delivers a 21.25-dB (S/N) (51.25 minus 30). From Eq. (5.80) this requires $N = 5$ bits. Practical companded systems operate with 7 or 8 bits. Figure 5.25 shows a typical compressor characteristic.

Signal-to-Noise Ratio in PCM Systems

Thermal noise causes bit errors in digital communication links, as discussed in Section 5.5. In a PCM system, the digital data are converted back to a baseband analog signal at the receiver. We need to know the signal-to-noise ratio that corresponds to a given probability of a bit error occurring in the digital data at the receiver. The analysis is straightforward when only one bit error occurs in each PCM word; provided the BER is below 10^{-4} and we have 7 or 8 bits per word, the likelihood of two bit errors occurring in one word is very small. We will assume this to be the case in the analysis that follows.

When a bit is in error in a PCM word, the recovered sample of the baseband analog signal will be at the wrong level. This adds an impulse of amplitude V_n and duration T_s, the period of one sample, to the true analog signal. The bit that is in error may be located in any position in the PCM word. If the least significant bit is in error, V_n is small and equal to Δ, the analog-to-digital converter step size; if it is the most significant bit that is in error, V_n will be large and equal to $2^{N_l-1}\Delta$. Thus the variance of the error in the analog signal, $(\overline{\Delta m})^2$ is given by [6]

$$(\Delta m)^2 = \frac{1}{N_l}\left[\Delta^2 + (2\Delta)^2 + (4\Delta)^2 + \ldots (2^{N_l-1}\Delta)^2\right] \tag{5.82}$$

The sum of the geometric progression in Eq. (5.82) is

$$(\overline{\Delta m})^2 = \frac{2^{2N_l-1}\Delta^2}{3N_l} \approx \frac{2^{2N_l}\Delta^2}{3N_l} \tag{5.83}$$

for $N_l \geq 2$.

Errors occur infrequently for BERs below 10^{-4}, and the mean time between errors is T, where

$$T = \frac{T_s}{N_l PB} \tag{5.84}$$

and PB is the bit error probability.

The two-sided power spectral density of a thermal-noise impulse train is given by [6]

$$G_t(f) = \frac{(\Delta m)^2}{T} = \frac{N_l PB(\Delta m)^2}{T_s} \tag{5.85}$$

To calculate the average noise power in the baseband channel due to noise impulses at the receiver output we must assume a baseband bandwidth. For the square-root raised cosine filter, the equivalent noise bandwidth B is $1/2T_s$, for all values of α. Then the thermal noise after the filter due to random bit errors in the PCM words is

$$N_t = \int_{-B}^{+B} G_t(f)\,df = \int_{-1/2Ts}^{1/2Ts} \left[\frac{2^{2N_l}\Delta^2 N_l PB}{3N_l T_s}\right] df = \frac{2^{2N_l}\Delta^2 PB}{3T_s^2} \tag{5.86}$$

The power in the analog signal, assuming equal probability of any signal voltage, is given by [6]

$$S_0 = \frac{L^2\Delta^2}{T_s^2 12} \tag{5.87}$$

Combining Eqs. (5.86) and (5.87) we find that the signal-to-noise ratio in the baseband channel at the output of the receiver is

$$\frac{S_0}{N_t} = \frac{L^2\Delta^2}{T_s^2 12}\,\frac{3T_s^2}{2^{2N}\Delta^2 PB} = \frac{1}{4PB} \tag{5.88}$$

since $L^2 = 2^{2N_l}$.

The mean quantization noise power for an equiprobable signal is given by [6]

$$N_q = \frac{\Delta^2}{12T_s^2} \tag{5.89}$$

We can combine thermal noise from Eq. (5.77) with quantization noise from Eq. (5.89) to find the overall PCM output signal-to-noise ratio.

$$\frac{S}{N_{\text{PCM}}} = \frac{S_0}{N_t + N_q} = \frac{\dfrac{L^2\Delta^2}{12T_s^2}}{\dfrac{2^{2N_l}\Delta^2 PB}{3T_s^2} + \dfrac{\Delta^2}{12T_s^2}} = \frac{2^{2N_l}}{1 + 4PB2^{2N_l}} \tag{5.90}$$

When PB is small, for example, less than 10^{-8}, the quantization noise will dominate and (S/N) $\simeq 2^{2N_l}$. For $N_l = 8$ bits, this gives S/N = 48 dB. When PB is larger, thermal noise dominates; for example, with $PB = 10^{-4}$ and $N_l = 8$,

$$(\text{S/N}) \simeq \frac{1}{4PB} = 34\,\text{dB}$$

Figure 5.26 shows the transition from thermal noise to quantization noise as the predominant noise source as the probability of a bit error decreases, for PCM systems using 7 and 8 bit words, with linear quantization. Clearly, for BERs

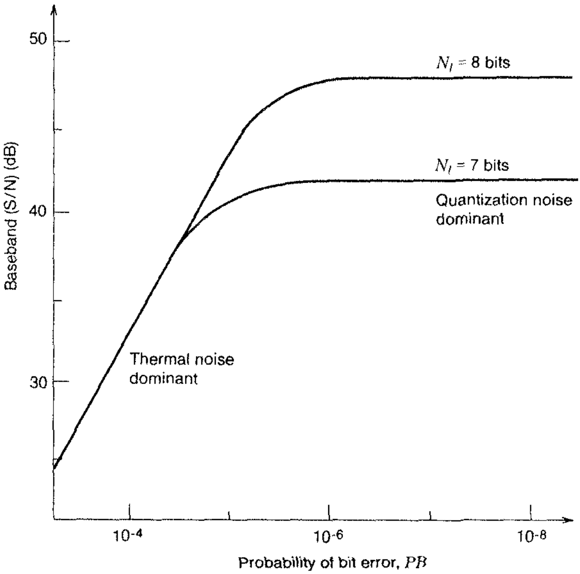


Figure 5.26 Baseband (S/N) in digital speech system using PCM. Signal values between zero and maximum are equally probable.

below 10^{-6}, quantization noise is dominant. Since most PCM links operate with BERs below 10^{-6} most of the time, it is worthwhile using nonlinear encoding (companding) to improve the baseband (S/N) by reducing quantization noise.

Delta Modulation

Delta modulation (*DM*) is a way of digitizing a voice waveform, transmitting the digits, and reconstructing the original analog waveform that avoids the quantizer and the A/D and D/A converters employed in PCM. For details on DM theory and practice the reader should consult reference 20.

In *linear delta modulation* (*LDM*) a circuit like that of Figure 5.27 determines the difference between an incoming waveform $x(t)$ and an estimate $z(t)$. It calculates an error voltage $e(t)$, where

$$e(t) = x(t) - z(t) \tag{5.91}$$

and a sign quantizer determines the sign of $e(t)$. The quantizer output $Q(t)$ is a positive constant when $e(t)$ is positive (i.e., when the signal is greater than the estimate) and a negative constant when $e(t)$ is negative (i.e., when the signal is smaller than the estimate). A sampling circuit samples $Q(t)$ and generates a positive pulse when $Q(t)$ is positive and a negative pulse when $Q(t)$ is negative. These pulses go to a conventional PSK digital modulator for transmission.

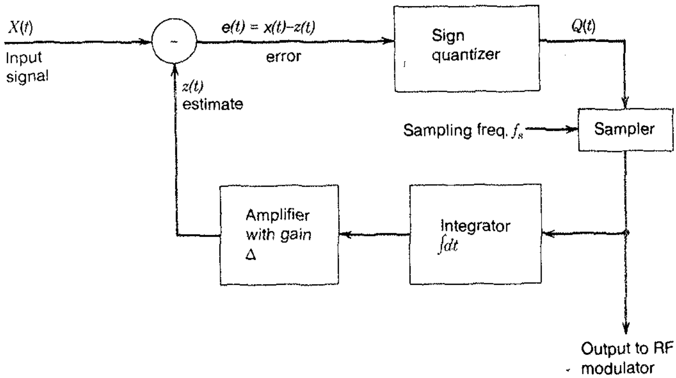


Figure 5.27 A delta modulator.

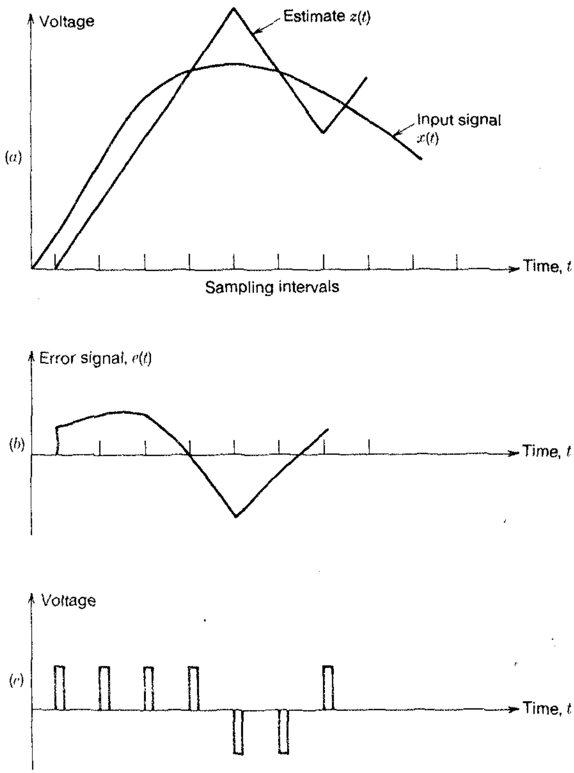


Figure 5.28 Waveforms in the delta modulator of Figure 5.27. (a) The input signal $x(t)$ and its estimate $e(t)$. (b) The error signal $e(t)$. (c) The transmitted pulses. (Based on J. J. Spilker, Jr., *Digital Communications by Satellite*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1977.)

The waveform reconstruction part of an LDM receiver and the estimator portion of the modulator both use the generated pulses to form the estimate $z(t)$ of $x(t)$. The estimate is made by integrating the pulses (which is equivalent to summing them numerically) and multiplying the result by a step size, Δ. If the pulses are of unit area, then the estimate $z(t)$ increases or decreases in value at a rate equal to the sampling frequency f_s if the estimate was smaller than or larger than the input waveform at the time of the last sample. Note that if input waveform is constant, the estimate will "hunt" around the correct value, alternately overestimating and underestimating it by Δ.

Figure 5.28 illustrates the delta modulation process for a hypothetical input signal. The waveforms in the figure are based on a worst-case assumption that the initial value of the estimate $z(t)$ is zero.

The performance of a DM system depends on the step size in two competing ways. The estimate $z(t)$ can change by only Δ volts at each sampling instant; if the input signal $x(t)$ is changing more rapidly, then the estimate cannot keep up and a condition called *slope overload* occurs. Slope overload can be prevented by making Δ large, but this increases *granularity*, the noise that results when the system hunts around a constant input value. Granularity is minimized by making Δ small. For a particular value of the sampling frequency f_s and a particular input signal, the output (S/N) behavior of a LDM system then depends on Δ as indicated schematically in Figure 5.29 [18]. There is an optimum step size; below it slope overload distortion dominates and above it granular noise dominates.

In linear delta modulation, the step size is fixed at a value that provides performance near the peak of Figure 5.28. But better performance may be achieved through a scheme called *adaptive delta modulation* (*ADM*) in which the value of Δ may be varied during the modulation process. Slope overload is characterized by long strings of pulses with the same algebraic sign. Digital logic in the estimator

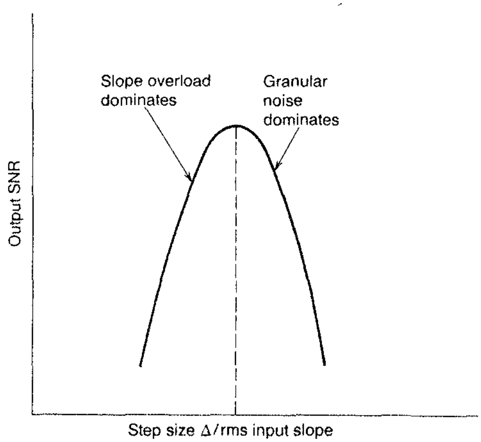


Figure 5.29 Conceptual sketch showing the dependence of (S/N) on step size in a linear delta modulator. (*Source:* J. J. Spilker, Jr., *Digital Communications by Satellite*, copyright © 1977, p. 77. Reprinted by permission of Prentice-Hall, Inc., Englewood Cliffs, NJ.)

circuit watches for this condition and automatically increases the step size after some specified number of same-sign pulses have been sent. Granularity, on the other hand, is associated with long strings of pulses having alternating signs; when this occurs then Δ is made smaller. Thus the step size varies between some specified minimum and maximum; the transmitter and receiver contain identical circuits to determine the step size and they both use the same value for a given sample.

It is difficult to make a definitive comparison between the performance of PCM and DM because DM technology is continually improving. See reference 19 for a detailed discussion of some early comparisons. If the bit rate can be made arbitrarily large, then PCM will outperform DM. But if the systems are compared for successively lower bit rates, then a crossover point will be reached below which DM gives better performance than PCM.

Most commercial satellite links to date use only PCM for digital transmission of speech while DM has been used for military applications or specialized civilian systems like the Space Shuttle in which the simpler equipment of DM and its improved performance advantage at low bit rates (i.e., at narrow bandwidths) over PCM is significant.

5.7 DIGITAL TV AND BANDWIDTH COMPRESSION [21]

While satellite TV for program distribution has retained analog modulation, compressed bandwidth digital television is now offered by at least two carriers (AT&T and SBS) for video teleconferencing. This is a service that offers two-way interactive video and audio for meetings and conferences; see reference 22 for an interesting discussion of the facilities and procedures involved. If the U.S. standard baseband video signal were digitized and transmitted without encoding, it would require a minimum sampling frequency of 8.4 MHz and a bit rate above 50 Mbps. But television pictures are highly redundant, and a much lower bit rate is required to transmit the difference between successive frames than to transmit the frames themselves. Further reduction is possible if the transmitter and receiver use linear predictive encoding [18]. The degree of redundancy in successive TV frames varies inversely with the motion of the main features in the pictures, and a consequence of reducing the bit rate is a deterioration of the system's ability to depict moving objects. This would be objectionable in entertainment TV, but the participants in a teleconference are usually sitting in one place and a technique called motion compensation can be used to improve the presentation of the regular linear motion associated with camera panning. The result is that teleconference TV is now transmitted at bit rates as low as 1.544 Mpbs. (This is the T1 bit rate explained in the next section.)

5.8 TIME DIVISION MULTIPLEXING

In time division multiplexing (TDM) a group of signals take turns using a channel. This contrasts to frequency division multiplexing, presented earlier, where the signals occupy the channel at the same time but on different frequencies. Since digital signals are precisely timed and consist of groups of short pulses with relatively long intervals between them, TDM is the natural way for combining digital signals for transmission.

TDM Terminology: The U.S. T1 24-Channel System

In this section we will describe the U.S. Bell Telephone T1 24-channel TDM system and use it to introduce the terminology of time division multiplexing. While T1 was developed for terrestrial microwave circuits, it appears on FDMA digital satellite links like those operated by RCA Government Communications Services [2]. We present it here as a convenient vehicle for explaining TDM operation. In pure TDMA systems, the multiplexing blends into the multiple access process.

A TDM system transmits a digital word from each channel in turn. Each word is a group of bits that identify the quantization interval of the current sample. The words are organized into *frames*. One frame contains one word from each channel plus some synchronizing information that serves to identify the start of the frame. A frame is then a series of bits numbered sequentially from zero that carry synchronizing information plus the quantized values of one sample from each channel. The bits within a frame are grouped into *slots*. A slot contains all the bits from a common source. The slots within a frame are numbered sequentially from zero. In the Bell T1 system illustrated in Figure 5.30, there are 25 slots, numbered 0 through 24. Slot 0 contains a single bit and carries synchronization information. Slots 1 through 24 each contain 8 bits and carry telephone channels 1 through 24. Thus a T1 frame contains $1 + (8 \times 24)$ or 193 bits.

Standard telephone PCM systems sample at an 8-kHz rate. Said another way, they transmit one sample of each channel every 125 μs. This is the frame interval; the frame rate is the reciprocal of the frame interval and is always equal to the sampling rate. The T1 system must transmit 193 bits in the 125 μs frame interval; hence its bit rate is 193 bits divided by 125 μs or 1.5440 megabits per second (Mbps).

Frame synchronization is established and maintained by transmitting a known bit pattern in slot 0. This bit pattern constitutes what is called the *frame alignment word (FAW)*. The T1 FAW is 100011011100; it contains 12 bits and requires 12 frames for transmission. The group of frames that transmit the FAW make up a *superframe*. Thus in the first frame of a superframe, slot 0 contains a 1. In the second frame slot 0 contains a 0, and so on.

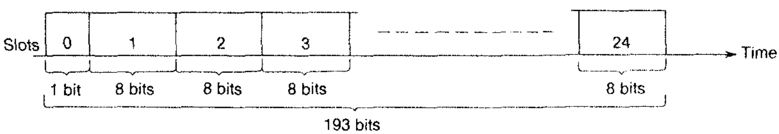


Figure 5.30 Slot organization of one Bell T1 frame.

At this point let us summarize the operation of a Bell T1 digital multiplexer. It receives digitized voice signals from 24 telephone channels, which, for now, we will assume are perfectly synchronized with each other and have exactly the same bit rate. The 8-bit word samples from each channel flow into buffers and wait for the multiplexer to read them out. The multiplexer reads them out and inserts them into outgoing frames as follows. The first frame is transmitted by sending a 1(the first bit of the FAW), then one 8-bit word from channel 1, then one 8 bit word from channel 2, and so forth through one 8-bit word from channel 24. Then a new group of samples flows into the buffers. The multiplexer forms frame 2 by sending a 0 (the second bit of the FAW) followed by the words from each channel. A third group of samples enters and the process continues. When the buffers have been filled and emptied 12 times, one superframe has been sent and the multiplexer begins a new FAW.

At the receiving end of the link a demultiplexer must sort out the bits in each frame and route the appropriate words to each outgoing channel. It must also keep track of the number of the frame (within the superframe) that it is receiving.

We may visualize the demultiplexing process by assuming that the incoming bits are clocked serially into a shift register. At the instant the last bit has entered the register, its contents match the bits in the frame of Figure 5.30. The multiplexer then does a parallel transfer of the bits for each channel into their own individual registers for subsequent serial transmission over their separate paths. At this point the digital channels have been demultiplexed.

The frame alignment bits go into a shift register, which, at the completion of the superframe, should contain the frame alignment word. If it does not, then the multiplexer and demultiplexer are out of sync. When this occurs the demultiplexer seeks to regain alignment; the process is called *reframing*. In it the demultiplexer looks at candidate frame alignment bits until it finds one that is going through the requisite 100011011100 100011011100 100011011100 pattern. Obviously there is a trade-off between the number of frame alignment bits and the time required for reframing. If the entire FAW is transmitted within each frame, then reframing time is much shorter than when the FAW is transmitted with one bit per frame. Typical reframing times are about 50 ms [21], which is sufficiently short not to cause significant degradation of speech in the 24 channels when misalignment occurs.

Along with the information carried in the 24 channels must go the signaling information necessary to route, initiate, and terminate the data channels. In the T1 system this information is transmitted by "robbing" the least significant bit from slots 6 and 12 and using these to form signaling channels A and B, respectively. Thus, channels 6 and 12 are actually carried by a form of 7-bit PCM and channels A and B convey signaling information at an 8 kbps rate.

Other TDM Systems

At the time of writing there seems to be no agreed upon standard international for satellite TDM systems. The CCITT has recommended a standard 1.544 Mbs system (which is slightly different from the T1) and a 2.048 Mbs 30-channel system [23]. For details on their slot and bit organization the reader should consult reference 24.

Channel Synchronization in TDM

Our explanation of the T1 system made the tacit assumption that all 24 incoming PCM channels were synchronized with each other and running at the same bit rate. This condition would hold if the voice channels had reached the originating earth station in analog form and had been digitized by modulators running on a common clock. But if the channels came into the station in digital form, their synchronization would not be guaranteed. They may be resynchronized for TDM transmission by a technique called *pulse stuffing* [1, 18].

In pulse stuffing the incoming words for each channel flow into an elastic buffer. There is one such buffer per channel, and each buffer can hold several words. The multiplexer reads words out of the buffer slightly faster than they come in. Periodically the multiplexer will go to the buffer and find less than a full word remaining. When that happens it inserts a dummy word called a *stuff word* into the frame in place of the word it would have taken from the buffer. At the same time it places a message on the signaling channel that states that a stuff word has been inserted. When the demultiplexer at the other end of the link receives the message it ignores the stuff word. When it is time for the next frame to be sent the buffer will have more than a full word waiting for transmission.

5.9 SUMMARY

Multiplexing is the process of separating the channels transmitted by a single earth station to prevent them from interfering with each other; its most common forms are frequency division multiplexing (FDM) and time division multiplexing (TDM). In the first case the channels are separated in frequency and in the second case they are separated in time.

Most analog telephone channels are transmitted over satellite links using frequency division multiplexing with frequency modulation (FDM/FM). In this method individual voice channels (nominally containing frequencies between 300 and 3400 Hz) are "stacked" in frequency by a multiplexer and the resulting multiplexed telephone signal is used to frequency modulate an uplink carrier. At the downlink earth station, an FM demodulator recovers the multiplexed signal and a demultiplexer recovers the individual channels.

An FM demodulator is characterized by a threshold. Provided that a satellite link's overall carrier-to-noise ratio (C/N) is above this threshold, the signal-to-noise ratio (S/N) of the demultiplexed voice channels will be significantly greater than the incoming (C/N). This effect is called FM improvement. Additional improvement in (S/N) may be obtained through preemphasis and deemphasis. Deemphasis decreases the noise power output of an FM demodulator; preemphasis distorts the multiplexed telephone signal before transmission to compensate for the deemphasis at the downlink earth station.

Analog FDM/FM systems are designed to produce a weighted (S/N) of about 50 dB in the worst telephone channel; this corresponds to about 10,000 picowatts psophometrically weighted (pWp) of noise power. The terms *weighted* and *psophometrically weighted* refer to a procedure for calculating (S/N) that reflects the human ear's nonuniform response to white noise.

Both the bandwidth and the FM improvement of an FDM/FM link depend in a complicated way on the number of voice channels carried, the way in which the voice channels are stacked, and the rms test-tone deviation of the uplink frequency modulator. The last is the rms carrier deviation that a single 1-kHz 0-dBm sinewave called the test tone will produce when supplied to one telephone channel.

Two important alternatives to FDM/FM for analog transmission are FM SCPC (single-channel-per-carrier) and companded single sideband (CSSB) systems. These share a common feature that individual voice channels can be extracted from the downlink radio frequency (RF) signal without demultiplexing all of the unwanted channels. In FM SCPC each voice channel FM modulates its own uplink channel. While requiring more bandwidth per channel than FDM/FM, FM SCPC makes more efficient use of transponder power and is easier to reconfigure to meet changing traffic conditions. It also has economic advantages for low-traffic routes.

In CSSB the individual voice channels are shifted to the uplink frequency by single sideband suppressed carrier modulation. Including guard bands, each signal occupies a 4-kHz bandwidth. Thus CSSB uses bandwidth much more efficiently than competing analog FM and digital modulation schemes, and it provides the same capability as FM SCPC for recovering individual channels at RF. Since SSB demodulation provides no (S/N) improvement, SSB transmission would require prohibitively large uplink and downlink (S/N) values if it were not for companding, a process by which the dynamic range of a voice channel is compressed before transmission and expanded (i.e., restored) after reception. This permits SSB links with (S/N) values on the order of 20 dB at RF to provide subjective (S/N) values of around 50 dB in individual voice channels. Its efficient use of bandwidth makes CSSB more attractive than digital modulation for some applications.

In analog television (TV) transmission by satellite, the baseband video signal and one or two audio subcarriers constitute a composite video signal that frequency modulates an uplink carrier. This system requires a very wide bandwidth for transmission (usually either a full transponder or a half transponder), but it provides the FM improvement necessary to achieve required (S/N) values.

If an uplink FM transmitter is fully modulated (fully "loaded" in the usual terminology), the downlink EIRP is distributed across an entire transponder bandwidth. But if the load decreases and disappears, the downlink EIRP will be concentrated at the center of the transponder bandwidth, and in the absence of uplink modulation it will be radiated at a single frequency. To avoid the interference with terrestrial radio services that this would cause, uplink earth stations are required to add a dispersal waveform to their modulating signal so as to maintain a reasonably constant downlink power spectral density independent of traffic loading. The dispersal waveform is removed by filtering at the downlink earth station.

Digital modulation is obviously the modulation of choice for transmitting digital data. Digitized analog signals may conveniently share a channel with digital data, allowing a link to carry a varying mix of voice and data traffic.

While baseband digital signals are often visualized as rectangular voltage pulses, careful pulse shaping is required to prevent intersymbol interference (ISI) and to permit reasonably distortionless transmission through the limited bandwidth of a transponder. With proper pulse shaping, the symbol rate of a digital link can be made approximately equal to the RF bandwidth.

The common digital modulation schemes used on digital satellite links are binary phase shift keying (BPSK) and quadrature phase shift keying (QPSK). In these an incoming data stream sets the phase of a sinusoidal carrier to one of two (BPSK) or four (QPSK) values. The performance of a BPSK or QPSK link is described by its bit error rate. Digital links are designed to meet bit error rate requirements in the same way as analog links are designed to deliver minimum (S/N) values.

Analog voice signals must be digitized for transmission over a digital link. This involves sampling the signal at a rate that is at least twice the highest frequency present and converting the sample values to digital words. The system that does this is called a quantizer; nonuniform quantizers are analogous to companders in CSSB systems and permit a lower bit rate than uniform quantizers. Standard practice with nonuniform quantization is to sample telephone channels at a 4-kHz rate and transmit each channel at 64 kbps.

Digital signals from different channels are interleaved for transmission through time division multiplexing (TDM). Digitized samples or digitized words from each channel are transmitted in turn; the time interval in which one sample or word from each channel is sent is called a frame. Channels are identified by their position in the frame; individual frames are identified by the presence of synchronization bits that repeat a known pattern.

REFERENCES

1. Technical Staff, Bell Telephone Laboratories, *Transmission Systems for Communications* (5th ed.), Bell Telephone Laboratories, Holmdel, NJ, 1982.
2. K. Jonnalagadda and L. Schiff, "Improvements in Capacity of Analog Voice Mutliplex Systems Carried by Satellite," *Proceedings of the IEEE*, **72**, 1537–1547 (November 1984).
3. W. H. Braun and J. E. Keigler, "RCA Satellite Networks: High Technology and Low User Cost," *Proceedings of the IEEE*, **72**, 1483–1505 (November 1984).
4. *Standard A Performance Characteristics of Earth Stations in the INTELSAT IV, IV-A, and V Systems Having a G/T of 40.7 dB/K* Publication (BG-28-72E Rev. 1), Intelsat, Washington, D.C., December 15, 1982.
5. H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, John Wiley & Sons, New York, 1980.
6. H. Taub and D. L. Schilling, *Principles of Communications* Systems, McGraw-Hill, New York, 1971.

7. *Reference Data for Radio Engineers*, (6th ed.), Howard W. Sams and Co., Indianapolis, 1975.
8. *Recommendations and Reports of the CCIR, 1978*, Vol. IV, International Telecommunication Union, Geneva, Switzerland, 1978.
9. H. L. Van Trees, *Satellite Communications*, IEEE Press, New York, 1979.
10. M. E. Ferguson, "Design of FM Single-Channel-per-Carrier Systems," *IEEE Int. Conf. on Commun.*, **1**, 12-11–12-16 (June 1975). (Reprinted in [9], pp. 336–341.)
11. K. Miya, Ed., *Satellite Communications Engineering*, Lattice Publishing Co., Tokyo, 1975.
12. K. S. Shamnugam, *Digital and Analog Communication Systems*, John Wiley & Sons, New York, 1979.
13. H. Nyquist, "Certain Topics in Telegraph Transmission Theory," *AIEE Transactions*, **47**, 817 (April 1928).
14. K. Feher, *Digital Communications: Satellite/Earth Station Engineering*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
15. W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*, Prentice-Hall, Englewood Cliffs, NJ, 1973.
16. J. R. Pierce and E. C. Posner, *Introduction to Communication Science and Systems*, Plenum Press, New York, 1980.
17. V. K. Bhargava, D. Haccoun, R. Matyas, and N. Nuspl, *Digital Communications by Satellite*, John Wiley & Sons, New York, 1981.
18. J. J. Spilker, Jr., *Digital Communications by Satellite*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
19. N. S. Jayant, "Digital Encoding of Speech Waveforms: PCM, DPCM, and DM Quantizers," *Proceedings of the IEEE*, **62**, 611–632 (May 1974). (Reprinted in [9], pp. 86–107.)
20. R. Steele, *Delta Modulation Systems*, John Wiley & Sons, New York, 1975.
21. T. Ishiguro and K. Iinuma, "Television Bandwidth Compression Transmission by Motion-Compensated Interframe Coding," *IEEE Communications Magazine*, **20**, 24–30 (November 1982).
22. B. A. Wright, "The Design of Picturephone Meeting Service (PMS) Conference Centers for Video Teleconferencing," *IEEE Communications Magazine*, **21**, 30–36 (March 1983).
23. J. Martin, *Communications Satellite Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1978.
24. G. H. Bennett, *Pulse Code Modulation and Digital Transmission*, Marconi Instruments Ltd., St. Albans, Hertfordshire, England, 1978.

PROBLEMS

1. An earth station is to receive 972 multiplexed telephone channels from the INTELSAT V global beam. The rms test-tone deviation specified for the system is 802 kHz and the top baseband frequency is 4028 kHz. Using a peaking factor $g = 3.16$, determine the following.

a. The number of groups and supergroups that are involved.
b. The required receiver bandwidth before FM detection.
c. The overall (C/T) at the receiver input required for operation at a 10-dB (C/N).
d. The overall (C/N) required at the receiver input required for a worst-channel (S/N) of 50 dB.

2. A television signal at baseband extends from 0 to 4.2 MHz. The television signal frequency modulates a 6-GHz transmitter with a peak carrier deviation of 5 MHz. The transmitted signal occupies a bandwidth of 18.40 MHz.

The 6-GHz TV signal will be relayed by a satellite whose antenna noise temperature is 290 K and whose U.S. standard noise figure is 8 dB. Assume that the transponder bandwidth equals the 18.40 MHz signal bandwidth. Successful operation of this system requires a minimum (C/N) of 10 dB at the satellite receiver input.

a. Calculate the required carrier power in dBW at the output of the satellite antenna (i.e., at the input to the transponder).
b. Assuming a satellite antenna gain of 17 dB and a distance from the transmitting station to the satellite of 3.8×10^7 m, calculate the required earth–station EIRP required in dBW.
c. For a 45-dB gain earth station antenna, find the required transmitter power in watts.
d. Assuming that the downlink (C/N) is 11 dB, calculate the overall (C/N) including retransmitted noise.

3. You have rented 36 MHz of bandwidth on an INTELSAT V transponder and your link can deliver an overall carrier-to-noise ratio $(C/N)_o$ to your earth station of 15 dB. You want to run analog telephone FDM/FM/FDMA telephone channels over this downlink. Determine the number of channels that the link can carry if the weighted worst-channel signal-to-noise ratio $(S/N)_{wc}$ is 51.0 dB. In making this calculation use the following equation to relate $(S/N)_{wc}$ to $(C/N)_o$. The equation includes preemphasis and weighting.

$$(S/N)_{wc} = (C/N)_o + 6.5 + 10 \log_{10}(B_{IF}/b) + 20 \log_{10}(\Delta f_{rms}/f_m) \text{ dB}$$

Here b is the bandwidth of a single telephone channel (use 3100 Hz), B_{IF} is the occupied (Carson's rule) bandwidth, $f_m = 4200 \times$ number of channels, and Δf_{rms} is the rms test-tone deviation (you have to find it). If the channels are organized into groups and supergroups, how many can be carried?

4. One of Intelsat's goals is a system of directly relaying messages from one satellite to another without going through an intervening earth station. Assume that they intend to use a frequency in the assigned 55 to 65 GHz band where, at present, 10-W output transmitters and receivers with 750 K noise temperatures are available. In this problem you will do some preliminary analysis for such a system using a *10-W transmitter* and a *750 K* overall noise temperature receiver operating at *60 GHz*. The satellite-to-satellite link will use BPSK (for ease of calculation) with a bit rate of *100 megabits per second* and a bit error rate (*BER*) *of*

b. Calculate the (S/N) for a 0 dBm test tone in the telephone channel at the top of the baseband, including psophometric weighting but excluding preemphasis improvement.
c. How does the capacity of the system compare for FDMA and TDMA?

7. One SCPC voice channel of the SPADE system carries QPSK at a symbol rate of 32,000 symbols per second in a bandwidth of 38 kHz. It achieves a bit error rate of 1×10^{-4}; without coding this requires an E_b/N_0 of 9.4 dB. The single carrier for this SPADE channel is radiated by an INTELSAT V transponder with an EIRP (for this one carrier) of 0 dBW to an earth station 40,000 km away. The channel center frequency is 4095 MHz. Answer questions (a) through (c) below about this channel and its receiver.
a. Tabulated EIRP's for INTELSAT V transponders are about 29 dBW. Why does this single SPADE carrier have a 0-dBW EIRP?
b. What value of (C/N) in decibels must be present at the input to the earth station receiver? Ignore any uplink or intermodulation effects.
c. What is the minimum G/T value in decibels that the earth station must have to achieve the specified bit error rate *with a 3-dB margin*?

8. A digital communication system uses a satellite transponder with a bandwidth of 50 MHz. Several earth stations share the transponder using QPSK modulation. Standard data rates used in the system are 80 kbps and 2 Mbps. The RF bandwidth required to transmit the QPSK signals is 0.75 of the bit rate. The TDMA frame is 125 μs in length, and a 1 μs guard time is required between each access. A preamble of 48 bits must be sent by each earth station at the start of each transmitted data burst.
a. What is the symbol rate of each earth station's transmitted data burst?
b. Calculate the number of earth stations that can be served by the transponder if each station sends data at 80 kbps.
c. Calculate the number of earth stations that can be served if each earth station sends data at 2 Mbps.
d. Repeat the calculations in (b) and (c) for a frame that is 1000 μs in length.

9. The capacity of the system described in Problem 8 can be increased substantially by using satellite switched TDMA. Assume that each earth station sends some data to every other earth station in every frame, and that it takes one microsecond to reposition the satellite antenna beam from one earth station to another. Only the downlink antenna beam is switched; the uplink uses a common zone beam. The frame length to be used is 1000 μs, and the extra antenna gain at the satellite is traded for an increase in the data rate by using 16-phase PSK. In this scheme, the 50-MHz bandwidth transponder can carry a symbol rate of 28.75 Msps corresponding to a data rate of 115 Mbps. Other parameters of the system are unchanged.
a. Find the number of earth stations that can share the transponder when each earth station sends data at 2 Mbps.
b. Find the total data throughput of the transponder after all preamble bits have been removed.

7

ENCODING AND FORWARD ERROR CORRECTION FOR DIGITAL SATELLITE LINKS

7.1 ERROR DETECTION AND CORRECTION

The transmission of information over a satellite communication system always results in some degradation in the quality of the information. In analog links the degradation takes the form of a decrease in signal-to-noise ratio. We saw in Chapter 5 that by using wideband FM we can trade bandwidth for power and achieve a good baseband signal-to-noise ratio with a low carrier-to-noise ratio in the RF signal. In digital links we measure degradation of the information content of a signal in terms of the bit error rate. By using phase shift keying, usually coherent QPSK, we can again trade bandwidth for signal power and achieve good bit error rates with low carrier-to-noise ratios.

A fundamental difference between analog and digital signals is that we can improve the bit error rate of a digital signal by the use of error correction techniques. No such technique is available for analog signals since once the information is contaminated by noise, it is extremely difficult to remove the noise, as we cannot in general distinguish between the signal and the noise electronically. (There are techniques that attempt to distinguish between signal and noise in television pictures, by using the correlation properties of the picture. They have been used successfully to enhance the quality of images of the moon and other planets obtained by the *Voyager* and similar space probes. However, the time taken to process the picture and the computer power needed make such techniques impractical for regular TV and voice transmissions.) In a digital system, we can add extra *redundant bits* to our data stream, which can tell us when an error occurs in the data and can also point to the particular bit or bits that have been corrupted. Systems that can only detect errors use *error detecting codes.* Systems that can detect and correct errors use *forward error correction* (FEC).

Some confusion exists in the literature over the term *coding*, since it is applied to several different processes, not all of which are concerned with error detection and correction. In the popular sense, *coding* is used to describe the rearrangement of information to prevent unauthorized use. This process is known technically as *encryption.* It is widely used on both analog and digital signals that are sent by cable and radio links. Digital signals are much more amenable to encryption, which can be achieved by convolving the data bits with a long pseudonoise (PN) sequence to destroy the intelligibility of the baseband data. To recover the information, the PN sequence used in the encryption process must be known to the recipient; this information is contained in the *key* to the code, which must be changed at frequent intervals to maintain good security. We will not be concerned any further, in this chapter, with encryption. It is, however, an important aspect of communications for commercial and military users who are concerned about interception and improper use of their transmitted data.

Coding is also applied to many processes that change data from one form to another. For example, pulse code modulation (PCM) changes analog data into binary words for transmission over a digital link. It is fundamental to the transmission of voice by digital techniques, and uses a device commonly called a *codec*, short for coder–decoder. The term *coding* is also applied to devices that scramble a digital data stream to prevent the occurrence of long strings of 1s or 0s.

Throughout this chapter we shall use the term *coding* to refer to error detection or error correction. This implies that additional (redundant) bits are added to the data stream to form an error-detecting or error-correcting code. It is possible, in theory, to generate codes that can detect or correct every error in a given data stream. In practice, there is a trade-off between the number of redundant bits added to the information data bits and the rate at which information is sent over the link. The *efficiency* of a coding scheme is a measure of the number of redundant bits that must be added to detect or correct a given number of errors. In some FEC systems the number of redundant bits is equal to the number of data bits, resulting in a halving of the data rate for a given channel transmission rate. (This is called a *rate one-half* scheme.) The loss of communication capacity is traded for a guaranteed low error rate; for example, commercial users transmitting financial data are willing to pay the extra cost in return for guaranteed accuracy in their accounting data, and military users are willing to lower data rates in order to obtain reliable communication in the presence of enemy jamming.

In this chapter we first discuss the techniques of error detection, and how they can be implemented in a satellite communication link, and then consider forward error correction. FEC in some form can be used on digital satellite links to improve bit error rates under low (C/N) conditions. Propagation disturbances, particularly attenuation by rain on links operating at 14/11 GHz and higher frequencies, result in reduced (C/N) for a small percentage of the time. Rather than designing the link with a very high (C/N) in clear weather, it is economic to add FEC to the link and allow the (C/N) to degrade during periods of rain attenuation. In clear air, the FEC system will correct the few errors that do occur; when rain attenuation is present, it maintains the bit error rate at an acceptable level with the degraded (C/N).

Finally, we examine some techniques that can be used to correct errors by requesting retransmission of corrupted data blocks. These are called *ARQ* (automatic repeat request) systems. They are widely used for packet transmission systems, where data are sent in blocks with variable delays, rather than being sent in real time. ARQ systems require a return channel; they cannot be used in one-direction transmission of data, and FEC must be used in such cases.

The operator of a digital satellite communication link has an option of providing FEC as part of the link, or of providing only a basic transmission channel. At 6/4 GHz and on many 14/11 GHz links, the channel is provided without error correction or detection. A minimum BER is guaranteed by the operator for a specified percentage of time, based on the link design and projected performance. The user is then free to add error detection or FEC to the data sent to and received from the link. If digital speech is sent, error detection or correction is rarely applied. With digital data, some measures must be taken to guard against error, and the user will normally provide the necessary equipment.

Links operating at frequencies above 10 GHz are subject to increases in BER during propagation disturbances. The link will be designed with a margin of a few decibels so that the BER falls below an acceptable level, typically 10^{-6}, for only a small percentage of any month or year. The total time for which the margin is exceeded by propagation effects will be less than 0.5 percent of any month in a well designed system. During the remaining 99.5 percent of the month, the E_b/N_0 of the received signal will be well above threshold, and very low BER will result. There may, in fact, be no errors for long periods of time and billions of bits can be transmitted with complete accuracy. Under these conditions, FEC and error-detection systems do nothing for the communication system. However, unless we can detect a falling E_b/N_0, coding may have to be applied all the time to be certain it is available when E_b/N_0 approaches threshold. To that extent, coding is an insurance against the possibility of bit errors; for most of the time it is unnecessary, but when it is needed, it proves invaluable.

Common carriers, who supply communication links to users on a dial-up or leased basis, do not generally apply FEC to their links, nor do they define the *protocols* to be used. (A protocol defines an operating procedure in a link.) These are user-supplied services and must be defined by the user for the data to be sent. In such cases, the error detection and correction equipment will be located at the customers' premises, whereas the earth station may be a long distance away and accessed via terrestrial data links.

The situation may be very different in a single-user network such as a military communication system. The earth station may be located at the data source, and coding and protocol form an integral part of the network operation. A similar situation can arise in carefully controlled systems such as Intelsat's, where the link operator specifies the user's earth station and operating parameters in detail.

Coordination between users is essential in an international system that interconnects the communication systems of different countries. In such cases, all protocols, codes, and FEC techniques must be very carefully defined and standardized. Because of the large number of earth stations in the Intelsat system and their wide geographic distribution, some paths between certain earth stations may have lower than usual E_b/N_0 ratios. It is feasible for those earth stations to employ FEC on the lower quality routes, without FEC being needed by all users.

7.2 CHANNEL CAPACITY

In any communication system operating with a noisy channel, there is an upper limit on the information capacity of the channel. Shannon [1] examined channel capacity in mathematical terms, and his work led to significant developments in information theory and coding.

For an additive white Gaussian noise channel, the capacity H is given by

$$H = B \log_2\left(1 + \frac{P}{N_0 B}\right) \text{bps} \tag{7.1}$$

where B is the channel bandwidth in hertz
P is the received power in watts,
N_0 is the single sided noise power spectral density in watts per hertz

Equation (7.1) is commonly known as the Shannon-Hartley law. We can rewrite Eq. (7.1) specifically for a digital communication link by putting $H = 1/T_b$, where T_b is the bit duration in seconds. The energy per bit is E_b, giving

$$E_b = PT_b = \frac{P}{H} \tag{7.2}$$

Then substituting $E_b/N_0 = P/HN_0$ in Eq. (7.1) yields

$$\frac{H}{B} = \log_2\left(1 + \frac{E_b}{N_0}\frac{H}{B}\right) \tag{7.3}$$

The ratio H/B is the *spectral efficiency* of the communication link, the ratio of bit rate to the bandwidth of the channel. Figure 7.1 shows the ratio $\log_2(H/B)$ plotted against E_b/N_0 in dB for the case when $H < B$ and the link operates at a bit rate H bps. Regardless of the bandwidth used, the E_b/N_0 cannot go below -1.6 dB (ln 2) if we are to operate at capacity. This is known as the Shannon bound. It sets a lower theoretical limit on the E_b/N_0 we can use in any communication link, regardless of the modulation or coding schemes. A link operating with $H < B$ is said to be *power limited* because it does not use its bandwidth efficiently.

Figure 7.2 shows the case for a link with $H > B$, operating at capacity. In this case, we can increase the capacity for a given bandwidth without limit, but only by providing very large E_b/N_0 ratios, implying very high transmitter power. As we saw in Chapter 5, satellite links must operate with E_b/N_0 in the range of 5 to 25 dB, which limits the theoretical spectral efficiency to below 4 bits/Hz. When $H > B$, the link is said to be *bandwidth limited*, implying that we could increase

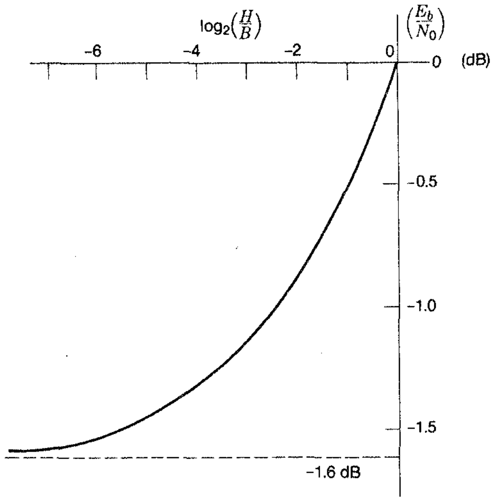


Figure 7.1 Relationship between H/B and E_b/N_0 for power limited case and low E_b/N_0 ratio.

capacity by using the available transmitter power in a wider bandwidth. Practical links using PSK do not achieve capacities anywhere near the Shannon theoretical capacity H. Shannon's theory assumes essentially zero bit errors; to achieve a bit error rate of 10^{-10} in a QPSK link requires a theoretical E_b/N_0 of 13 dB with a spectral efficiency of 2 bits/Hz. Equation (7.3) predicts $E_b/N_0 = 1.77$ dB for this case. What coding, in particular FEC, can do for us is to improve the link performance under conditions of low E_b/N_0, such as during periods of rain attenuation, so that the BER of the link does not rise excessively. This takes us closer to

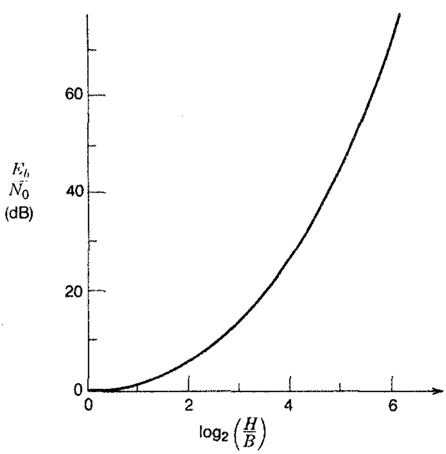


Figure 7.2 Relationship between H/B and E_b/N_0 for bandwidth limited case and high E_b/N_0 ratio.

the Shannon capacity in the region of low E_b/N_0 while not increasing excessively the bandwidth required for transmission.

7.3 ERROR-DETECTION CODING

Error detection coding is a technique for adding redundant bits to a data stream in such a way that an error in the data stream can be detected. Unless one redundant bit is added for every data bit, the exact position of a single bit error cannot be determined. Usually, one redundant bit is added for every N data bits; this allows a single error within that *block* of N bits to be detected. A simple example of an error detecting code system that has been in use for many years for transmission of information is the 8-bit ASCII code [2]. The ASCII code is widely used for transmission of computer and teleprinter data over telephone lines and radio links.

The 8-bit ASCII code consists of 128 characters, each having seven data bits plus a single parity bit. The seven data bits have an internationally agreed-upon interpretation and represent the alphabet in upper- and lowercase letters, numerals 0 to 9, and a set of useful commands, symbols, and punctuation marks. The eighth bit, the parity bit, is used for detection of error in the seven data bits of the character. For example, in a system using *even parity*, the parity bit is 0 when the sum of the data bits is even, and 1 when the sum is odd. Thus the sum of the data bits plus the parity bit is always made even, or 0 in modulo-2 arithmetic. Figure 7.3 shows an example of even and odd parity coding. In *odd parity*, the sum of the data bits plus the parity bit is always odd, or 1 in modulo-2. Single errors in the seven data bits, or the parity bit, are detected at the receiving end of the link by checking the eight received bits of each character for conformity with the

	Data Bits	Parity Bit	Sum (Modulo-2)
Even parity	0101101	0	0
Odd parity	0101101	1	1

Figure 7.3*a* Example of even and odd parity for a 7-bit ASCII word.

Even Parity Example

	Received Codeword	Sum of Bits	Error Detected?
One error	$0101\underline{0}010$	1	Yes
Two errors	$0101\underline{01}10$	0	No
Three errors	$\underline{1}101\underline{01}10$	1	Yes

Figure 7.3*b* Example of error detection in a 7-bit ASCII word with even parity. Error bits are underlined.

parity rule. In modulo-2 arithmetic, $0 \oplus 0 = 0, 0 \oplus 1 = 1, 1 \oplus 0 = 1$, and $1 \oplus 1 = 0$. Similarly, $0 \otimes 0 = 0$, $0 \otimes 1$ or $1 \otimes 0 = 0$, and $1 \otimes 1 = 1$. All the codes that we will be considering are binary, and modulo-2 arithmetic will be used throughout this chapter.

Suppose we have a system using even parity, which transmits the character A in ASCII code, as illustrated in Figure 7.3. At the receiving end of the link we check the eight bits by modulo-2 addition. If the sum is 0, we conclude that the character is correct. If the sum is 1, we detect an error. Should two bits of the character be corrupted, the modulo-2 sum is 0, so we cannot detect this condition. We can easily calculate the improvement in error rate (assuming that we discard corrupted characters) that results with parity in a seven-bit character.

For example, let the probability of a single-bit error occurring in the link be p, and let us suppose that p is not greater than 10^{-1}.

The probability $P_e(k)$ of k bits being in error in a block of n bits is given by the binomial probability function

$$P_e(k) = \binom{n}{k} p^k (1 - p)^{n-k} \tag{7.4}$$

where p is the probability of a single-bit error occurring, and

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{7.5}$$

For example, in the ASCII *codeword* of 8 bits, we have one parity bit, which allows us to detect one error, and seven data bits. Two errors cannot be detected, although three can. The probability that there are two errors is much higher than the probability of four errors, so when parity is used, the probability of an undetected error occurring in an ASCII word is P_{wc} where

$$P_{wc} \simeq P_e(2) = \binom{8}{2} p_c^2 (1 - p_c)^6 \tag{7.6}$$

where p_c is the single-bit error probability for the 8-bit word. When p_c is small, $(1 - p_c)^6 \simeq 1$ so

$$P_{wc} \simeq \binom{8}{2} p_c^2 = 28p_c^2 \tag{7.7}$$

If we had not used parity, the probability P_{wu} of a single error in the 7-bit word with bit error probability p_u is

$$P_{wu} \simeq \binom{7}{1} p_u (1 - p_u)^7 = 7p_u \tag{7.8}$$

Thus the improvement in error rate for the ASCII words is approximately $4p$, provided $p_c \simeq p_u$.

For example, if we have $p_c = p_u = 10^{-3}$, the probability of error for uncoded 7-bit words is 7×10^{-3} and for words with a single parity bit added it is 2.8×10^{-5}.

Codeword		Message
011	011	011
110	110	110
Message bits	Parity bits	

Example of (6, 3) systematic codewords. The parity bits are at the right of the message bits.

Codeword		Message
001	0111	0111
101	1100	1100
Parity bits	Message bits	

Example of (7, 4) systematic codewords. The parity bits are at the left of the message bits.

Figure 7.4 Examples of (6, 3) and (7, 4) systematic codewords.

With a transmission bit error rate of 10^{-6}, the word error rate with parity is approximately 2×10^{-11}. This corresponds to one character error in the transmission of the text of this book 20,000 times over a link with a BER of 10^{-6}.

Some caution is needed in making the assumption that the bit error rate is the same for uncoded and coded transmissions. When we added the single parity bit to a 7-bit ASCII character, the transmission bit rate went up from 7 bits per character to 8 bits per character. The increase in bit rate will result in an increase in BER if the channel remains unchanged, so not all the expected decrease in character error rate will be achieved in practice. This is particularly true when we add several parity check bits to our data bits, to obtain error-correction codes.

The example given above for parity error detection is one case of *block error detection*. We have transmitted our data as *blocks*, in this case eight–bit blocks consisting of seven data bits and one redundant parity bit. In general, we can transmit n bits in a block, made up of k message bits plus r check bits. The n-bit block is called a *codeword* and coding schemes in which the message bits appear at the beginning of the codeword, followed by the check bits, are called *systematic codes*. Figure 7.4 shows an example of some systematic codes.

Linear Block Codes

Linear block codes are codes in which there are 2^k possible message blocks of k bits, to which are added $(n - k)$ redundant check bits generated from the

k message bits by a predetermined rule. In a *systematic linear block code* the first k bits of the codeword are the message and the remaining $(n - k)$ bits are the check bits. A codeword with n bits of which k bits are data is written as (n, k). The (6, 3) and (7, 4) codewords in Figure 7.4 are examples of systematic linear block codes.

The general form of a linear block codeword C is

$$C = DG \tag{7.9}$$

where D is the k-bit message and G is a generator matrix that creates the check bits from the data bits.

Multiplication of matrices can be performed only when the number of columns in the first matrix equals the number of rows in the second matrix. Thus D is a matrix of k columns and one row, and G must be a matrix with k rows and n columns. Multiplication is carried out in modulo-2 arithmetic, which is just like regular multiplication but without a carry. To multiply matrix D by matrix G we first multiply each element in the first column of G by the corresponding elements in D, and then modulo-2 sum the results. Repeating the procedure n times yields the matrix C, a codeword of n bits. The following very simple example for a (3, 2) code illustrates the process.

Let D be a 2-bit data word 10 and G be a generator matrix with

$$G = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Then the codeword $C = DG$ is

$$\begin{aligned} C &= [1 \quad 0]\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\ &= [(1 \otimes 1) \oplus (0 \otimes 0)][(1 \otimes 0) \oplus (0 \otimes 1)][(1 \otimes 1) \oplus (0 \otimes 1)] \\ &= (1 \oplus 0)(0 \oplus 0)(1 \oplus 0) \\ &= 1 \ 0 \ 1 \end{aligned}$$

Thus C is a (3, 2) codeword having two data bits and one parity bit, with even parity. The (3, 2) code in this example could detect one error in the 3-bit codeword.

The reader may wish to verify that the other code words in this set are

$$D = 00, \qquad C = 000$$

$$D = 01, \qquad C = 011$$

$$D = 11, \qquad C = 110$$

The general form of G is

$$G = [I_k \quad P_k]_{k \times n} \tag{7.10}$$

where I_k is the identity matrix of order k and P is an arbitrary k by $n - k$ matrix. No procedures exist for designing the P matrix, but some P matrices lead to codes with desirable properties such as high rate efficiency, ability to detect or correct more than one error, and so on. The subject of code design is involved, and the interested reader is referred to Shamnugam [3] for a review of coding for communication purposes, or to Lin and Costello [4] for a more detailed treatment of the subject. We restrict our survey here to the basic ideas behind the application of coding in communication links.

An example of a generator matrix for a (6, 3) block code is G where

$$G = \begin{bmatrix} 100 & 011 \\ 010 & 101 \\ 001 & 110 \end{bmatrix} \tag{7.11}$$

The message length is 3, the codeword length is 6. The matrix G has the form $[I:P]$ as in Eq. (7.10). There are eight possible codewords (000) through (111). The codewords are generated by Eq. (7.9) as

$$C = DG = [D] \begin{bmatrix} 100 & 011 \\ 010 & 101 \\ 001 & 110 \end{bmatrix} \tag{7.12}$$

For example, if the message word D is 001 the corresponding codeword $C(001)$ is

$$C(001) = [001] \begin{bmatrix} 100 & 011 \\ 010 & 101 \\ 001 & 110 \end{bmatrix} = 001 \quad 110 \tag{7.13}$$

The remaining codewords for this (6, 3) code are given in Table 7.1. The reader should verify these words using Eq. (7.12). Note that the check bits are not unique to a particular message; in fact, they cycle in a symmetrical pattern.

The codewords C are sometimes called *code vectors*. We now need a technique to detect errors in the codewords, which is accomplished by use of a *parity check matrix* H, defined as

$$H = [P^T : I_{n-k}]_{(n-k)\times n} \tag{7.14}$$

Table 7.1
Codewords of the Example (6, 3) Code

Message bits	Codewords
000	000 000
001	001 110
010	010 101
011	011 011
100	100 011
101	101 101
110	110 110
111	111 000

where P^T is the transpose of P_k in Eq. (7.10). P^T is obtained by interchanging the rows and columns of the matrix P. I_{n-k} is an identity matrix for the check bits. Error detection is achieved by multiplying a received codeword R by the transpose of the parity check matrix, H^T. If the codeword R is correct, then $R = C$ and H^T is defined by

$$C \times H^T = 0 \tag{7.15}$$

If the codeword R is in error, such that

$$R = C + E \tag{7.16}$$

where E is an *error vector*, the error will be detectable if E is not zero. We check for errors by finding the *error syndrome* S. The syndrome is a single word of length $n - k$, where $n - k$ is the number of parity check bits in the codeword, formed by multiplication of the received codeword with the parity check matrix. It is always zero if the received codeword has no errors.

$$\begin{aligned} S &= RH^T \\ &= [C + E]H^T = CH^T + EH^T \end{aligned} \tag{7.17}$$

Since CH^T is zero by definition in Eq. (7.15)

$$S = EH^T \tag{7.18}$$

Thus the syndrome S is zero if R is a valid codeword. (It may also be zero for some combinations of multiple errors.) If S is nonzero, we know that an error has occurred in transmission of our codeword.

Example 7.3.1

An example of a code that can detect two errors in a message of length 4 bits is a (7, 4) *Hamming code*. This code has 4 message bits and 3 check bits in a 7-bit codeword.

The generator matrix G for this code is

$$G = \begin{bmatrix} 1000 & 111 \\ 0100 & 110 \\ 0010 & 101 \\ 0001 & 011 \end{bmatrix}$$

$$\quad I_4 \qquad P_3$$

The 16 possible codewords of this (7, 4) code are shown in Table 7.2. Using Eq. (7.14) we can form the parity check matrix H.

$$H = \begin{bmatrix} 1110 & 100 \\ 1101 & 010 \\ 1011 & 001 \end{bmatrix}$$

$$\quad P^T \qquad I_3$$

Table 7.2
Codewords and Weights for the (7, 4) Hamming Code Used in Examples 7.3.1 and 7.3.2

Data, D	Codeword, C	Weight
0000	0000000	—
0001	0001011	3
0010	0010101	3
0011	0011110	4
0100	0100110	3
0101	0101101	4
0110	0110011	4
0111	0111000	3
1000	1000111	4
1001	1001100	3
1010	1010010	3
1011	1011001	4
1100	1100001	3
1101	1101010	4
1110	1110100	4
1111	1111111	7

Now consider a message $D = (1010)$, one of 16 possible messages generated from four message bits. The corresponding codeword is C, which we generate from

$$C = DG = 1010 \quad 010$$

The syndrome $S = CH^T$ is zero. If we now insert an error into the codeword so that the received code vector R is corrupted in the second bit (underlined)

$$R = 1\underline{1}10010$$

Then the syndrome $S = RH^T$ is

$$S = [1110010]\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 110$$

Since S is not zero, we have detected an error.

Example 7.3.2

Let us suppose that 2 bits and then 3 bits in our (7, 4) Hamming codeword are corrupted. Let the received codeword be $R = 101\underline{1}01\underline{1}$ when the message D

is 1010 and the transmitted codeword $C = 1010010$. The syndrome $S = RH^T$ is

$$S = [1011011]\begin{bmatrix} 111 \\ 110 \\ 101 \\ 011 \\ 100 \\ 010 \\ 001 \end{bmatrix} = 010$$

which indicates that the received codeword is incorrect.

If we had three errors, and received, for example, $R = 1\underline{1}10\underline{10}0$, the syndrome is $S = 000$. The errors are not detected, because this is a legitimate codeword in the (7, 4) set. The (7, 4) Hamming code can always detect two errors, but cannot reliably detect three errors. We use a general formula to determine the error-detection capabilities of a linear block code in Section (7.4).

Error Correction with Linear Block Codes

The linear block codes described in the previous section have the capability of locating errors in a received codeword. Generally, a larger number of errors can be detected than can be located, but when an error can be located within a codeword, it can then be corrected. This is a powerful feature of coding and one reason for the widespread use of digital transmission techniques, especially by satellite.

As an example of the error-correction properties of linear block codes, consider again the (7, 4) Hamming code used in Example 7.3.1. The syndrome was generated by using Eq. (7.17), $S = RH^T$. In Example 7.3.1, the received codeword $R = 1110010$ was a corruption of the transmitted codeword $C = 1010010$, with the second bit changed from 0 to 1. The syndrome S was 110, indicating the presence of an error in the received codeword. The syndrome $S = 110$ is identical to the second row of the H^T matrix, pointing to the second bit as the error bit. For single errors, the syndrome is always the row of the H^T matrix corresponding to the bit that is in error. If the error occurs in a message bit, we can correct the message by inverting the appropriate bit in the message part of the received codeword R. If the error occurs in a parity check bit, we may not need to take any action, if only the message bits are transmitted to the receiver output.

7.4 ERROR DETECTION AND CORRECTION CAPABILITIES OF LINEAR BLOCK CODES

Some linear block codes are better for error detection or correction than others. There are some basic theorems that define the capabilities of linear block codes in terms of the *weight*, *distance*, and *minimum distance*. The weight (or Hamming weight), w, of a code vector C is the number of nonzero components of C. The distance (or Hamming distance), d, between two code vectors C_1 and C_2 is the number of components by which they differ. The minimum distance of a block code is the smallest distance between any pairs of codewords in the entire code.

These definitions lead to some useful rules that tell us how many errors a given code can detect or correct. First, the minimum distance of a linear block code is the minimum weight of any nonzero codeword. The minimum distance of the (6, 3) code in Table 7.1 is 3. For the (7, 4) Hamming code considered in Examples 7.3.1 and 7.3.2 and tabulated in Table 7.2, the minimum distance is also 3. Second, the number of errors that can be detected in a code with minimum distance d_{min} is $(d_{min} - 1)$. The number of errors that can be corrected is $\frac{1}{2}(d_{min} - 1)$, rounded to the next lowest integer if the number is fractional.

Example 7.4.1

Consider the (7, 4) Hamming code shown in Table 7.2. The minimum distance of this code is 3, so up to $(3 - 1) = 2$ errors can be detected. The number of errors that can be corrected is $(3 - 1) \times \frac{1}{2} = 1$.

To correct two errors, we must have a code with a minimum distance of at least 5, so that $\frac{1}{2}(5 - 1) = 2$. When the code can correct only one error, the syndrome of the corrupted codeword points to the position in the codeword at which the error occurred. When more than one error exists in a codeword of a code capable of multiple error correction, the syndrome will be nonzero. However, finding the two (or more) bits that are in error becomes more complex.

It is necessary to examine the received codeword to determine which, out of all the codewords that could have been transmitted, the received codeword is closest to. The syndrome is used to indicate the most likely correct codeword. When multiple errors have to be corrected, the procedure becomes extremely cumbersome: as a result, this general approach to error correction is not used in practice.

A subset of linear block codes called *binary cyclic codes* has been developed for which implementation of error-correction logic is much simpler. The codes can be generated using shift registers, and error detection and correction can also be achieved with shift registers and some additional logic gates. A large number of binary cyclic codes have been found, many of which have been named after the people who first proposed them. The best known are the Bose-Chaudhuri-Hocquenghem codes (*BCH codes*), which were independently proposed by three workers at about the same time in 1959–60 [5, 6]. In the following section we look briefly at the techniques used to generate cyclic codes and to correct errors in received codewords.

7.5 BINARY CYCLIC CODES

Algebraic Structure of Cyclic Codes [3]

A binary cyclic code C is made up of (n, k) codewords, V, of the form

$$V = (v_0, v_1, v_2, v_3 \ldots v_{n-1}) \tag{7.19}$$

If the codeword V is shifted to the right by i bits to obtain a new word $V^{(i)}$

$$V^{(i)} = (v_{n-i}, v_{n-i+1}, \ldots v_0, v_1 \ldots v_{n-i-1}) \tag{7.20}$$

then $V^{(i)}$ is also a codeword. Thus there are n possible codewords in an (n, k) binary cyclic code.

The codeword V can be represented by a code polynomial $V(x)$ of the form

$$V(x) = v_0 + v_1 x + v_2 x^2 + \ldots v_{n-1} x^{n-1} \tag{7.21}$$

The coefficients v_0, v_1, and so on, are 0s or 1s obtained by modulo-2 addition and multiplication. The variable x has an index p corresponding to the position of the pth bit in the codeword. In general, we can obtain the code polynomial $V^{(i)}(x)$ for the code vector $V^{(i)}$ as

$$V^{(i)}x = v_{n-i} + v_{n-i+1}x + v_{n-i+2}x^2 + \ldots v_0 x^i + v_1 x^{i+1} + \ldots v_{n-i-1} x^{n-1} \tag{7.22}$$

$V^{(i)}(x)$ is the remainder resulting from dividing $x^i V(x)$ by $x^n + 1$, so

$$x^i V(x) = q(x)(x^n + 1) + V^{(i)}(x) \tag{7.23}$$

where $q(x)$ is a polynomial called the quotient.

We can generate codewords V in two ways. We can form *nonsystematic* codewords by using a *generator polynomial* $g(x)$ such that we form a *code polynomial* $V(x)$ from a *data polynomial* $D(x)$

$$V(x) = D(x)\,g(x) \tag{7.24}$$

where

$$D(x) = d_0 + d_1 x + d_2 x^2 + \ldots d_{k-1} x^{k-1}$$

and

$$g(x) = g_0 + g_1 x + g_2 x^2 + g_3 x^3 + \ldots g_{n-1-k} x^{n-1-k}$$

The coefficients d and g are all 0s or 1s. The codeword V will not necessarily contain the original message D when the codeword is formed using Eq. (7.24).

An alternative method of forming the codeword V is to use the generator polynomial $g(x)$ to make a codeword having the form

$$\begin{gathered} V = (r_0, r_1, r_2 \ldots r_{n-k-1}, d_0, d_1, d_2 \ldots d_{k-1}) \\ n-k \text{ parity check bits} \quad k \text{ message bits} \end{gathered} \tag{7.25}$$

where

$$r(x) = r_0 + r_1 x + r_2 x^2 + \ldots r_{n-k-1} x^{n-k-1}$$

The polynomial $r(x)$ is the *parity check polynomial* of the code and is the remainder obtained when $x^{n-k}D(x)$ is divided by $g(x)$.

$$x^{n-k}D(x) = q(x)g(x) + r(x) \tag{7.26}$$

where $q(x)$ is the quotient.

The code polynomial $V(x)$ is formed by

$$V(x) = r(x) + x^{n-k}D(x) \tag{7.27}$$

An example of the generation of a cyclic code using these two methods will help to illustrate the process [3].

Example 7.5.1

The generator polynomial $g(x) = 1 + x + x^3$ can be used to generate a (7, 4) cyclic code. Let a data word be $D = 1010$. The corresponding message polynomial is $D(x) = 1 + x^2$

From Eq. (7.24), the codeword polynomial is

$$V(x) = D(x)\,g(x)$$

or

$$\begin{aligned} V(x) &= (1 + x^2)(1 + x + x^3) \\ &= 1 + x + x^3 + x^2 + x^3 + x^5 \\ &= 1 + x + x^2 + x^5 \end{aligned}$$

since $x^3 + x^3 = (1 + 1)x^3 = 0x^3 = 0$ in modulo-2 arithmetic.

The codeword V is

$$V = 1 + 1x + 1x^2 + 0x^3 + 0x^4 + 1x^5 + 0x^6 = 1110010$$

Note that the message $D = 1010$ is not evident within this codeword—it does not have the structure given in Eq. (7.25). The systematic codeword form of Eq. (7.25) must be generated using Eqs. (7.26) and (7.27). The last four bits of the required codeword are the data bits $D = 1010$. The first three bits are parity check bits (r_1, r_2, r_3) obtained from

$$\frac{x^{n-k}D(x)}{g(x)} = q(x) + r(x)$$

Using the $g(x)$ generator polynomial and $D(x)$ message polynomial given above, we can find $r(x)$ as the remainder of dividing $x^3(1 + x^2)$ by $1 + x + x^3$. Using algebraic division we obtain

$$\begin{array}{r|l} & x^2 \\ \hline x^3 + x + 1 & x^5 + x^3 \\ & x^5 + x^3 + x^2 \\ \hline & \qquad\qquad x^2 \end{array}$$

Note that there are no minus signs in modulo-2 arithmetic. Subtraction is the same as addition.

The remainder $r(x) = x^2$, so $r = 001$. Hence the codeword is

$$V = 0011010$$

Table 7.3
Systematic and Nonsystematic Codewords of (7, 4) Cyclic Code with Generator Polynomial $g(x) = 1 + x + x^3$

Message Codeword D	Nonsystematic Codeword $V(x) = D(x)g(x)$	Systematic Codeword V
0000	0000000	0000000
0001	0001101	1010001
0010	0011010	1110010
0011	0010111	0100011
0100	0110100	0110100
0101	0111001	1100101
0110	0101110	1000110
0111	0100011	0010111
1000	1101000	1101000
1001	1100101	0111001
1010	1110010	0011010
1011	1111111	1001011
1100	1011100	1011100
1101	1010001	0001101
1110	1000110	0101110
1111	1001011	1111111

Source: Reprinted from K. S. Shamnugam, *Digital and Analog Communication Systems*, John Wiley & Sons, New York, 1979, Table 9.3, p. 466, with permission.

Table 7.3 shows the full 16 codewords for systematic and nonsystematic (7, 4) cyclic codes generated from

$$g(x) = 1 + x + x^3$$

Generation of Cyclic Codes

An encoder to generate the (7, 4) systematic code in Table 7.3 is shown in Figure 7.5. The message input D forms the first 4 bits of the 7-bit codeword and is transmitted when the output switch is set to position 1 (transmitting the most significant bit first). As the four data bits are transmitted, the gate is open allowing the data bits to be shifted into the registers r_0, r_1, r_2. The feedback connection between r_0 and r_1 generates the appropriate $g(x)$ function and is equivalent to a division operation. When the four data bits have been transmitted and fed to the shift register, the switch is set to position 2 and the contents of the register are transmitted. The feedback connections around the shift register correspond to the coefficients of the generator polynomial $g(x)$, and the contents of the register after k data bits have passed though the gate is the remainder, r, which forms the parity check part of the codeword.

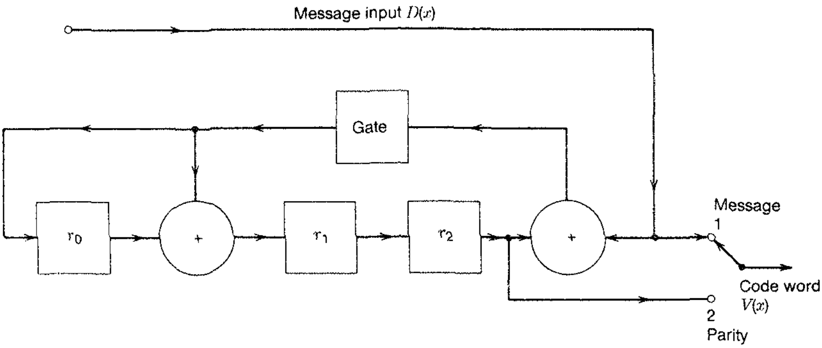


Figure 7.5a A block encoder for a (7, 4) cyclic code. The message bits are read through the switch in position 1 to form the first four bits of the codeword. The message bits are also read into the registers r_0, r_1, and r_2 via the feedback logic to generate three parity bits that are added to the four message bits by setting the switch to position 2.

Codeword

1 1 0	1 0 0 0
Parity bits	Message bits

Figure 7.5b Example of a codeword of a (7, 4) systematic code. See Table 7.3 for a full listing of this code.

For a detailed discussion of cyclic code generation and error correction, the reader should refer to references 4 and 7.

Error Detection and Correction with Cyclic Codes

The detection of errors in a cyclic code is achieved by calculating the syndrome of the received codeword R; if the syndrome is zero, we have a valid codeword. If it is nonzero, one or more errors have occured. The syndrome of a cyclic code is calculated as the remainder obtained when the polynomial of the received codeword, $R(x)$ is divided by the generator polynomial $g(x)$

$$\frac{R(x)}{g(x)} = P(x) + \frac{S(x)}{g(x)} \tag{7.28}$$

where $P(x)$ is the quotient.

The syndrome has order $n - k - 1$ or less. If $E(x)$ is the error that has been added to the transmitted codeword $V(x)$ to get $R(x)$, then

$$R(x) = V(x) + E(x) \tag{7.29}$$

and

$$\frac{R(x)}{g(x)} = \frac{V(x)}{g(x)} + \frac{E(x)}{g(x)} \tag{7.30}$$

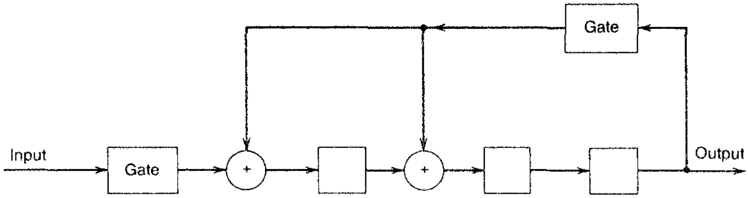


Figure 7.6 Syndrome circuit for the (7, 4) cyclic code generated by $g(x) = 1 + x + x^3$. (*Source:* S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, copyright © 1983, p. 100. Reprinted by permission of Prentice-Hall, Inc., Englewood Cliffs, NJ.)

If we generate $V(x)$ from $V(x) = D(x)g(x)$ we find

$$E(x) = [P(x) + D(x)]g(x) + S(x) \tag{7.31}$$

The error pattern $E(x)$ tells us how to correct the received codeword to get back to the original codeword V.

The strength of cyclic codes is the ease with which the syndrome of Eq. (7.28) may be calculated and the error correction of Eq. (7.31) applied, using digital logic circuits. A circuit for decoding and correcting single errors in the (7, 4) code generated by the circuit of Figure 7.5 is shown in Figure 7.6.

The buffer register holds the received codeword R, which is 7 bits long. As each codeword is read into the buffer, it is entered into the syndrome calculation circuit. The three input AND gate acts as an error detector; if the output of the gate is 0, the current bit in the received codeword is correct. If the gate output is 1, the current codeword bit is incorrect and must be inverted. As the codeword is shifted out of the buffer register, the output gate inverts any bit that the error-detection gate indicates is incorrect.

BCH and Burst Error Correction Codes

The BCH codes form the most powerful group of cyclic error-correction codes yet devised and are readily implemented with shift register and logic circuits. For this reason, they are widely used in systems employing error correction. They have the following properties:

1. Length of shift register: m stages.
2. Block length: $n = 2^m - 1 \qquad (m \geq 3)$.
3. Number of errors that can be corrected: $t \qquad (t < 2^{m-1})$.
4. Number of parity check bits: $n - k \leq mt$.
5. Minimum distance: $d_{\min} \geq 2t + 1$.

For example, a (1023, 923) BCH code has $t = 10$ and $d_{\min} \geq 21$. This code requires the transmission of 100 parity check bits for every 923 data bits, giving a code rate of 923/1023.

The codes discussed thus far have the property of correcting random errors, such as those caused by the peaks of thermal noise. Under some conditions, such as transient interference or cochannel interference caused by multipath, errors may occur in bursts. Special codes have been developed that can correct *burst errors*, that is, errors that occur in adjacent bits. The number of parity check bits needed to correct q sequential bit errors has a lower bound given by

$$n - k \geq 2q \tag{7.32}$$

A code capable of correcting burst errors of length q is called a q *burst-error correcting-code*. Table 7.4 lists some burst-error-correcting cyclic codes.

An alternative technique for burst error correction is to use interleaving of bits from a series of codewords that have good short burst-error-correction properties [4]. The individual bits of a given codeword are spaced l bits apart, so that a burst error of length l corrupts only one bit in each of l codewords. This single error can be corrected at the receiver, rendering the transmission invulnerable to longer error bursts. This technique is more valuable for terrestrial microwave links and troposcatter circuits that suffer short, deep fades due to multipath than for satellite links where the major causes of bit error are rain attenuation and depolarization.

Rain effects change slowly and cause outages lasting for seconds or minutes, so no amount of coding can reasonably correct the resulting loss of thousands or millions of bits.

Table 7.4
Examples of Burst-Error-Correcting Codes

Code (n, k)	No. of Parity Check Bits	Burst Error Correcting Ability q	Code Rate
7,3	4	2	3/7
15,9	6	3	9/15
511,499	12	4	499/511
1023,1010	13	4	1010/1023
131,119	12	5	119/131
290,277	13	5	277/290
34,22	12	6	22/34
169,155	14	6	155/169
103,88	15	7	88/103
96,79	17	8	79/96
56,38	18	9	38/56
59,39	20	10	39/59

Source: Reprinted from K. S. Shamnugam, *Digital and Analog Communication Systems*, John Wiley & Sons, New York, 1979, Table 9.5, p. 475, with permission.

Golay Codes

The Golay code and the single error-correcting Hamming codes are examples of *perfect codes*, in which all possible patterns of a given number of errors are corrected. The Golay code is a (23, 12) cyclic code that corrects all patterns of three or fewer bit errors. It has a minimum distance of 7. A closely related form of the Golay code has one overall parity check bit added to form a (24, 12) code with a minimum distance of 8. The (24, 12) code will detect all patterns of 7-bit errors and correct all patterns of 3-bit errors. It also has the advantage of a coding rate of one-half; rate one-half coding is easier to implement than other rates because the 2:1 ratio between message bits and code bits simplifies clock synchronization between the input data stream and the output coded stream.

7.6 PERFORMANCE OF BLOCK ERROR CORRECTION CODES

Calculation of the improvement of bit error rate with block encoding requires a comparison of the uncoded error rate to that obtained after correction of blocks of encoded data. For the perfect codes given in Section 7.5, the improvement can be determined exactly; when other codes are used, only an upper and lower bound on the error rate after coding can be established.

For the perfect codes, for which t errors can be corrected, the word error probability is

$$P_{we} = \sum_{i=t+1}^{n} \binom{n}{i} q_c^i (1 - q_c)^{n-i} \tag{7.33}$$

where q_c is the bit error rate on the transmission link for the coded data, and n is the codeword length in bits.

If the data were sent uncoded, as k data bits, the word error probability would be

$$P_{wu} = 1 - (1 - q_u)^k \tag{7.34}$$

where q_u is the bit error rate when uncoded data are transmitted on the same link. Because the bit rate for the shorter words of k bits is always lower than for the longer codewords of n bits, q_u is always lower than q_c on a given link. Equation 7.33 is an upper bound for the codeword error probability with other than perfect codes [7].

Figure 7.7 shows a comparison of the performance of a number of cyclic codes when implemented in a coherent PSK link [8]. The curves are for an ideal link with no modem implementation margin and show symbol error probability as a function of E_b/N_0 at the demodulator input.

The (7, 4) code is a single-error-correcting Hamming code. The remainder are BCH codes; the (127, 64) code corrects 10 errors, the (1023, 688) code corrects 36 errors. The (127, 113) BCH code is a double-error-correcting code that has been specified for use in 120 Mbps transmission using the INTELSAT V TDMA system [9].

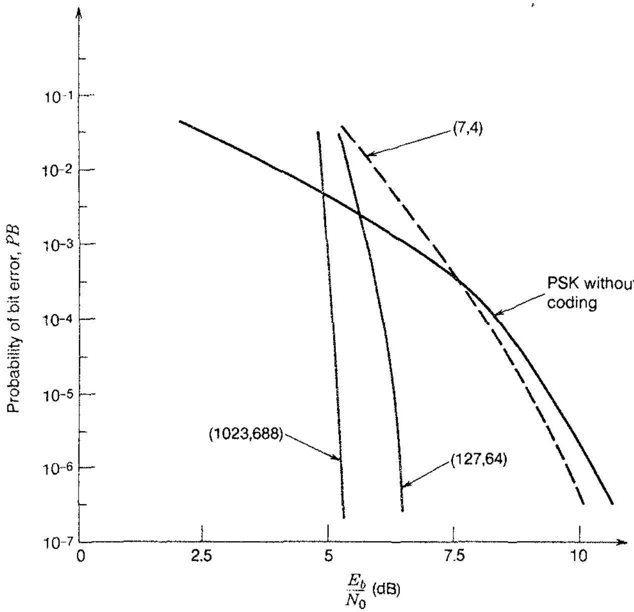


Figure 7.7 Probability of error in a PSK link for various coding schemes. All codes are approximately rate one-half. Note that below 5 dB E_b/N_0, coding does not improve the error rate. (*Source:* K. Feher, *Digital Communications: Satellite/Earth Station Engineering*, copyright © 1983, p. 285. Reprinted by permission of Prentice-Hall, Inc., Englewood Cliffs, NJ.)

Note that by the use of FEC, we can reduce the E_b/N_0 needed to achieve a 10^{-6} error rate from a theoretical 10.5 dB without coding to 7 dB with (127, 64) coding. The symbol error rate now increases very rapidly as the 7 dB E_b/N_0 value is approached; the degradation in error rate is very rapid, falling by a factor of 10^5 for less than 2 dB change in received signal power. The threshold is even more abrupt with (1023, 688) coding.

7.7 CONVOLUTIONAL CODES

Convolutional codes are generated by a tapped shift register and two or more modulo-2 adders wired in a feedback network. The name is given because the output is the convolution of the incoming bit stream and the bit sequence that represents the impulse response of the shift register and its feedback network [10]. Figure 7.8 illustrates an example of a convolutional encoder, which has been widely described in the literature [10, 11, 12, 13]. As each incoming information bit propagates through the shift register, it influences several outgoing bits, spreading the information content of each data bit among several adjacent bits. An error in any

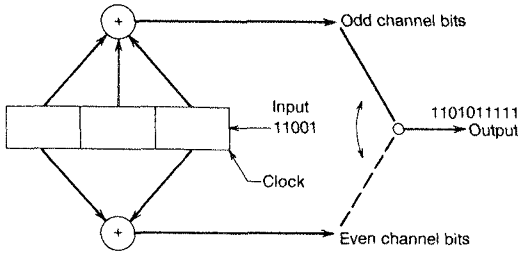


Figure 7.8 The convolutional encoder used in the text example. (Reprinted with permission from Marlin P. Ristenbatt, "Alternatives in Digital Communication," *Proceedings of the IEEE*, **61**, 703–721 (June 1973). Copyright © 1973 IEEE.)

one output bit can be overcome at the receiver without any information being lost. The process is somewhat like forming an image from a hologram, where information is distributed more or less uniformly over a two-dimensional field. The image can be reconstructed from only a portion of the field, although resolution is lost if a significant part of the hologram is discarded.

The *state* of a convolution encoder is defined by the shift register contents that will remain after the next input bits are clocked in. If the shift register is K bits long and input bits enter in groups of k, then the encoder has $2^{(K-k)}$ states. Putting in a group of k input bits causes the encoder to change states; a change of state is called a *state transition*. From a given state, a convolutional encoder can go to only 2^k other states (although one of these 2^k options may be to remain in the starting state). Each state transition is associated with a unique sequence of input bits and is accompanied by a unique sequence of output bits. The quantity K, which measures the length of the shift register, is called the *span* or the *constraint length* of the encoder. If v output bits are transmitted for every k input bits, then the encoding rate is k/v [10].

A decoder keeps track of the encoder's state transitions and reconstructs the input bit stream. Transmission errors are detected because they correspond to a sequence of transitions that could not have been transmitted. When an error is detected, the decoder begins to construct and keep track of all the possible tracks (sequences of state transitions) that the encoder might be transmitting. At some point, which depends on its speed and memory, the decoder selects the most probable track and puts out the input bit sequence corresponding to that track. The other tracks that it had been carrying are discarded. The algorithm used for decoding is named for A. J. Viterbi (see reference 12), and for this reason convolutional codes are sometimes called Viterbi codes.

An Illustrative Example

Consider the encoder of Figure 7.8. We will define the shift register's state by the values of its leftmost two bits before the next input bit is clocked in. After

Table 7.5
State Transitions in the Example Convolutional Encoder

Starting State	Input	Register Contents After Input	Output	Ending State
a:(0, 0)	0	000	00	**a**:(0, 0)
	1	100	11	**b**:(1, 0)
b:(1, 0)	0	010	10	**c**:(0, 1)
	1	110	01	**d**:(1, 1)
c:(0, 1)	0	001	11	**a**:(0, 0)
	1	101	00	**b**:(1, 0)
d:(1, 1)	0	011	01	**c**:(0, 1)
	1	111	10	**d**:(1, 1)

the input bit enters, the two bits that define the state are found in the rightmost two locations. We will call the states **a**, **b**, **c**, and **d**, defined in terms of bits as follows

$$\mathbf{a}:(0,0) \qquad \mathbf{c}:(0,1)$$
$$\mathbf{b}:(1,0) \qquad \mathbf{d}:(1,1)$$

Suppose the register is in state **a**, (0, 0), and a 0 enters. Immediately afterwards the register contents are (0, 0, 0) and the output symbol is $0 \oplus 0 \oplus 0 = 0$ and $0 \oplus 0 = 0$ or (0, 0). Thus when the register is in state **a** and a 0 enters, the output is (0, 0) and the encoder remains in state **a**. Now suppose that the register is again in state **a**, (0, 0), and this time a 1 enters. The register contents become (1, 0, 0) and the output bits are $1 \oplus 0 \oplus 0 = 1$ and $1 \oplus 0 = 1$ or (1, 1). Since the register contents are now (1, 0, 0), the encoder is in state **b**. Thus when the register is in state **a** and a 1 enters, the encoder changes to state **b** and emits the symbol (1, 1) in the process. Table 7.5 illustrates this process and all of the other allowed transitions for this encoder.

Another way to display the operation of this encoder is with the state transition diagram of Figure 7.9. In it a solid line indicates the trajectory that the detector follows when the input bit is a 0 and a dashed line indicates the trajectory that corresponds to a 1. The vertices are labeled with their corresponding states **a**, **b**, **c**, and **d**. The bit pairs at the middle of each trajectory indicate the output corresponding to that transition. Note that from each state it is possible to go to no more than two other states.

To illustrate a process by which the output of a convolutional encoder can be decoded, it is convenient to work with a *trellis diagram* like the one shown in Figure 7.10. Here the states of the encoder are plotted vertically and time is plotted horizontally. The transitions are shown as arrows joining the states occupied by

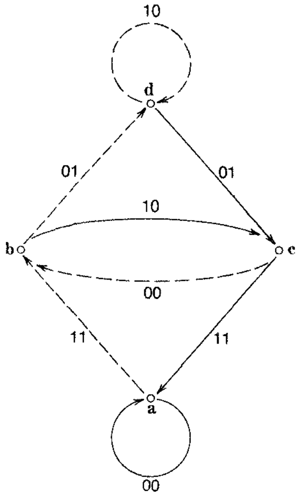


Figure 7.9 State transition diagram for the example convolution encoder. Solid lines indicate trajectories that follow input of a 0 and dashed lines indicate transitions corresponding to 1 inputs. Vertices are the stable states **a**, **b**, **c**, and **d** of the encoder. Output bits are written near the middle of the corresponding transitions.

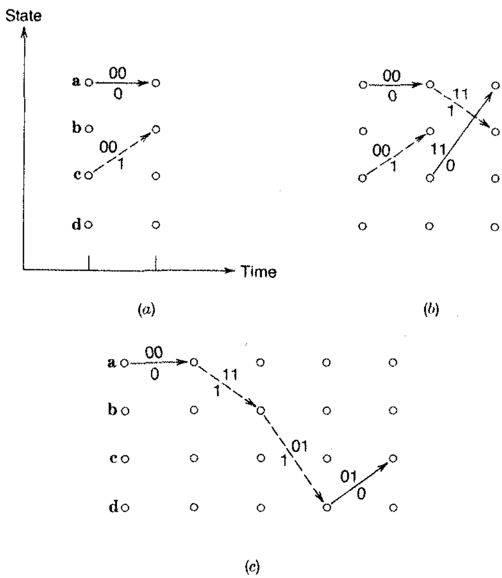


Figure 7.10 Trellis diagrams used in decoding the output of a convolutional encoder. The letters **a**, **b**, **c**, **d**, represent the four stable states of the encoder used in the text example, and the arrows represent transitions. A solid arrow is a transition resulting from an input 0 and a dashed arrow is a transition resulting from an input 1. The input that causes each transition is written below the arrow and the output that is generated by the transition is written above the arrow. (*a*) Two choices for the first transition. (*b*) Alternative choices for the first and second transitions. The top one is correct because the other arrows do not form a continuous path. (*c*) Complete path corresponding to the received bit sequence 00 11 01 01.

the encoder at successive times. If the bit sequence to be decoded was received correctly, then only one set of arrows will form a continuous path through the trellis, and that path traces the sequence of operations that the encoder performed. The corresponding input bits are the decoded version of the received sequence.

As an example, assume that we received an error-free message

00 11 01 01

which we know was encoded on the encoder of Figures 7.8, and 7.9 and Table 7.5. The 00 could represent an input 0 and a transition from state **a** to state **a** or an input 1 and a transition from state **c** to state **b**. We draw both on the trellis diagram as in Figure 7.10*a*. The 11 could represent a transition from **a** to **b** and an input 1 or a transition from **c** to **a** and an input 0. When we draw both as in Figure 7.10*b*, the true path is obviously **a–a–b** since no other combination of arrows is continuous. Using the correct path to continue decoding, we note that the only transition beginning at **b** and putting out 01 ends at **d** and represents an input 1. Likewise the only transition beginning at **d** that puts out an 01 results from an input 0 and ends at **c**. We now have established the complete path as shown in Figure 7.10*c* and decoded the incoming bit sequence as 0110.

In this example we were fortunate in being able to identify the starting state (**a**) of the encoder after only two transitions. In general it may be necessary to follow the first $4K$ transitions of an encoder of span K before the starting point and the subsequent path of the encoder can be identified.

A transmission error is detected when a group of bits is received that does not correspond to an allowed transition. The error may be corrected by following all of the paths that might have been transmitted for a period of time and retaining the path whose bits have the largest correlation coefficient with the bit stream actually received. The correlation coefficient between two bit sequences is found by replacing each 0 by a -1, multiplying corresponding bits in each sequence, and summing the result.

As a simple illustration of the error-detection and -correction process, assume that the bit stream 00 11 01 01 was transmitted but because of a transmission error we received

00 $\underline{1}$1 11 01

where the incorrect bit is underlined. We carry out the steps of Figure 7.11*a* and 7.11*b* and arrive at state **b** when the $\underline{1}$ is received. There is no transition from state **b** that produces a 11 output, so we know that an error was made. What we received as 11 must have been sent as a 01, representing a transition to **d**, or as a 10, representing a transition to **c**. We must draw and keep track of both alternative paths as in Figure 7.11*b*. If the encoder was at state **d** when the last bit pair came in, the 01 represented a transition to **c**. On the other hand, if it started at state **c**, then the 01 represents another error because the only transitions from **c** are to **a** with a 11 output or to **b** with a 00 output. This means that there are two possible paths leaving **c** and we must draw both of them as in Figure 7.11*c*.

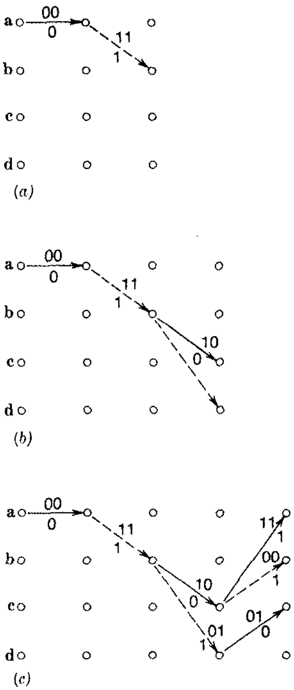


Figure 7.11 The use of alternative paths to decode a bit sequence containing transmission errors. See Figure 7.10. (*a*) Encoder path prior to transmission error. (*b*) Possible encoder paths corresponding to received bit sequence 00 11 11. (*c*) Possible encoder paths corresponding to received bit sequence 00 11 11 01.

At this point we must choose between candidate paths **b–c–a**, **b–c–b**, and **b–d–c**. To make the choice we calculate the correlation coefficient of the bit sequence corresponding to each path with the received sequence 11 01. Here are the calculations.

b–c–a	1	−1	1	1
Received	1	1	−1	1
Product	1	−1	−1	1

Sum of products = 0

b–c–b	1	−1	−1	−1
Received	1	1	−1	1
Product	1	−1	1	−1

Sum of products = 0

b–d–c	−1	1	−1	1
Received	1	1	−1	1
Product	−1	1	1	1

Sum of products = 2

The path with the largest correlation coefficient (sum of products) is **b–d–c**. We choose it as correct and it is. Thus the received sequence is correctly decoded as 0110.

The number of bit errors that can be corrected depends on the number of bits transmitted, the sequence in which the errors occur, and the number and length of the candidate paths that the decoder carries before discarding all but the one(s) with the highest correlation coefficient(s). As an illustration for the particular decoder and bit stream used here, the reader may wish to verify that two bit errors may be corrected if the received stream is 01 11 11 01 but not if it is 00 10 11 01, where again incorrectly received bits are underlined.

In general a Viterbi decoder must be able to retain information on $2^{k(K-1)}$ paths for satisfactory performance [14]. There is of course a trade-off between cost, BER, k, and K in practical convolution codes. Figure 7.12 [14] illustrates the performance of one such code; the reader should consult that reference for a detailed discussion of Viterbi decoder performance.

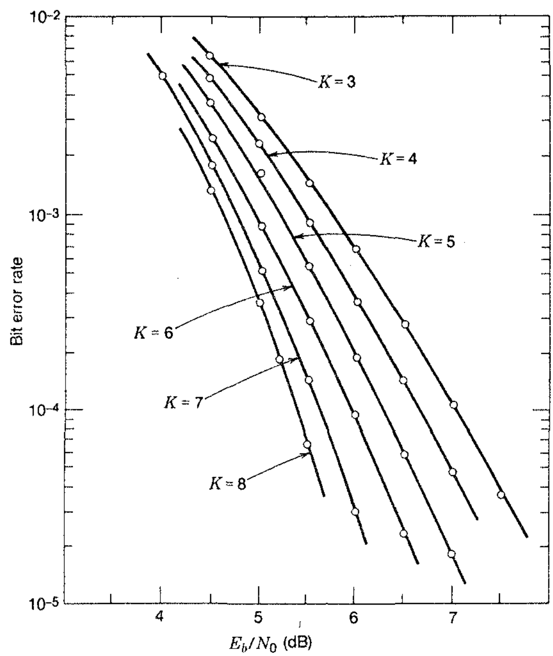


Figure 7.12 Performance of a rate one-half Viterbi codes for constraint length (K) values ranging from 3 to 8 and 32-bit paths. (Reprinted with permission from Jerrold A. Heller and Irwin Mark Jacobs, "Viterbi Decoding for Satellite and Space Communications," *IEEE Transactions on Communications*, **COM-19**, 835–848 (October 1981). Copyright © 1981 IEEE.)

7.8 IMPLEMENTATION OF ERROR DETECTION ON SATELLITE LINKS

Error detection is invariably a user-defined service, forming part of the operating protocol of a communication system in which the earth–satellite–earth segment may only be a part. It allows the user to send and receive data with a greatly reduced probability of error, and a very high probability that uncorrected errors are identified and located within a block of data, so that the existence of an error is known even if the exact bit or word in error cannot be determined. The penalty for the user is a reduced transmission rate, just as in FEC. Implementation of error correction by use of error detection and retransmission requires the use of protocols. A protocol is an agreed-upon set of actions that define how each end of the data link proceeds so that data are transmitted in an accurate and ordered fashion through the link.

Error detection is readily accomplished using the coding techniques described in the previous sections. In most communications systems it is not sufficient simply to detect an error; it must be corrected. When an error-detection code is used, a retransmission of the block of data containing the error must be made so that the correct data are acquired at the receiving terminal. The usual technique for obtaining a retransmission is for the receiving terminal to send an *acknowledge* (*ACK*) signal to the transmitting end when it receives an error-free block of data. If an error is detected in the block, a *not acknowledge* (*NAK*) signal is sent, which triggers a retransmission of the erroneous block of data. This is called an automatic repeat request (ARQ) system.

Such systems work well on terrestrial data links with relatively low data rates and short time delays. Their implementation on satellite links is more difficult due to the long transmission delay; as a result FEC is generally preferred for satellite paths.

There are three basic techniques for retransmission requests, depending on the type of link used. In a one-way, *simplex link*, the ACK or NAK signal must travel on the same path as the data, so the transmitter must stop after each block and wait for the receiver to send back a NAK or ACK before it retransmits the last data block or sends the next one. With a one-way delay of 240 ms on a satellite link, the data rate of such a system will be very low and is suitable only for links in which data are generated slowly, as when someone is typing on a terminal keyboard. Satellite links usually establish two-way communication (duplex channel) by the use of FDM, SCPC, or TDM, as discussed in Chapter 6. The ACK and NAK signals can be sent on the return channel while data are sent on the go channel. However, if the data rate is high, the acknowledgment will arrive long after the block to which it relates was transmitted.

In a *stop-and-wait* system, the transmitting end sends a block of data and waits for the acknowledgment to arrive on the return channel. The delay is the same as in the simplex case, but implementation is simpler. Figure 7.13 shows an example of a stop-and-wait sequence.

In a *continuous transmission* system using the *go-back-N* technique, data are sent in blocks continuously and held in a buffer at the receive end of the link.

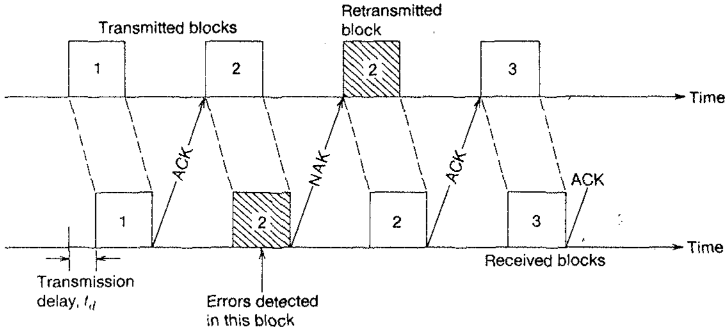


Figure 7.13 Stop-and-wait ARQ system.

Each data block is checked for errors as it arrives, and the appropriate ACK or NAK is sent back to the transmitting end, with the block number appended. When a NAK(N) is received, the transmit end goes back to block N and retransmits all subsequent blocks, as illustrated in Figure 7.14. This requires the transmitter on a satellite link to hold at least 480 ms of data, to allow time for the data to reach the receive end and be checked for errors and for the acknowledgment to be sent back to the transmit end. Since there is a delay in transmission only when a NAK is received, the *throughput* on this system is much greater than with the stop-and-wait method. Throughput is the ratio of the number of bits sent in a given time to the member that could theoretically be sent over an ideal link.

If sufficient buffering is provided at both ends of the link, only the corrupted block need be retransmitted. This system is called *selective repeat* ARQ. In Figure 7.14, block 3 is corrupted, and blocks 4, 5, 6, and 7 are transmitted before the NAK message is received. At this point, we could transmit block 3 only if blocks 4, 5, 6, and 7 are stored at the receive end. On receiving a correct version of

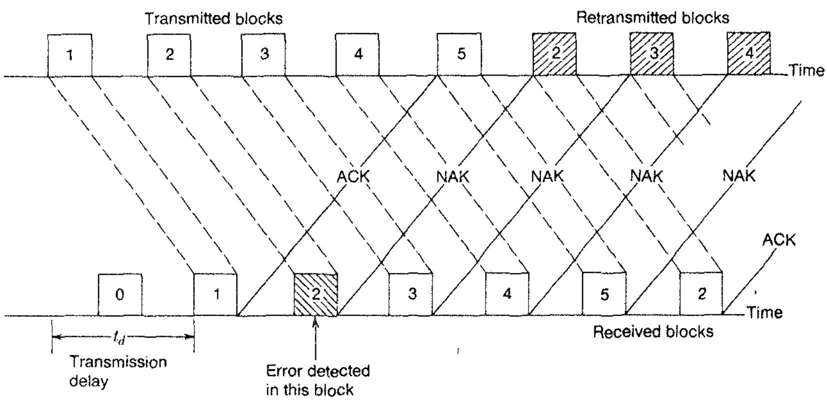


Figure 7.14 Example of a go-back-N-blocks ARQ system. N is three blocks in this example.

block 3, the receive buffer substitutes it for the corrupted version and releases the data for retransmission. In systems handling data rates of megabits per second, the buffer requirements for continuous transmission systems become very large.

Reference 15 contains a good survey of error-detection techniques for use in satellite communication systems and of the various ARQ systems that can be implemented. Some hybrid ARQ systems are described that combine FEC with retransmission of blocks when uncorrected errors are detected. This combines the error-correction properties of the FEC code for a limited number of errors with the error-detection properties of the same code when too many errors are present for all of them to be corrected.

Figures 7.15 and 7.16 [15] show how throughput can be increased as the BER increases for a number of hybrid ARQ systems. In Figure 7.15, curves 1 through 6 are for selective repeat ARQ systems. The improvement in performance over the go-back-N system is clearly shown. Curve 1 is an ideal selective repeat system with infinite buffering, and curves 2 and 3 show the effect of finite buffer size (512 and 1024 bits) when selective repeat ARQ is used. Curves 4, 5, and 6 are for selective repeat ARQ systems with finite buffer (512 bits) and FEC of 3, 5, and 10 bits per block [16]. In each case, the block length n is 1024 bits.

Figure 7.16 shows the effect of using convolutional FEC codes with selective repeat ARQ. The block length is 1024 bits with 1000 data bits. There is a marked

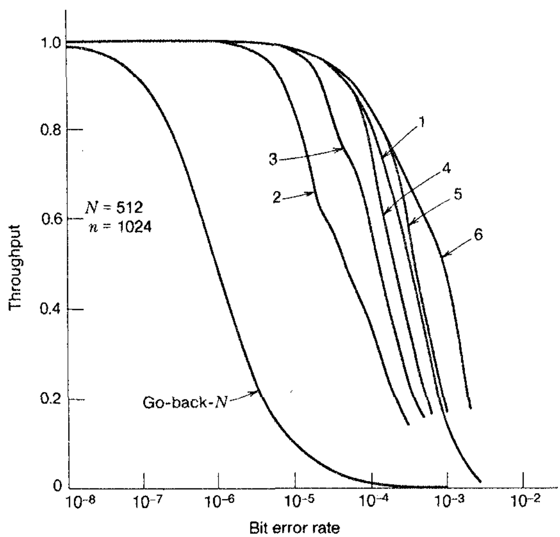


Figure 7.15 Throughput of various ARQ schemes with buffer size $N = 512$ (curves 2, 4, 5, 6), $N = 1024$ (curve 3), and infinite buffer size (curve 1). Curves 1, 2, 3 are all selective repeat ARQ without error correction; curves 4, 5, and 6 are for error correction parameter $t = 3$, 5, and 10. (Reprinted with permission from S. Lin, D. J. Costello, Jr., and M. J. Miller, "Automatic-Repeat-Request Error-Control Schemes," *IEEE Communications Magazine*, **20**, p. 13 (December 1984). Copyright © 1984 IEEE.)

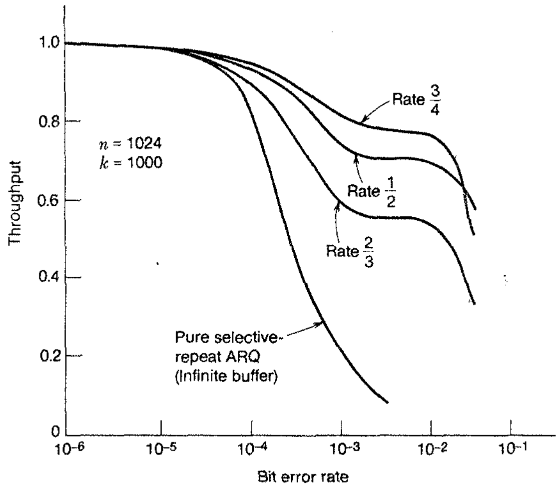


Figure 7.16 Throughput performance of three hybrid ARQ schemes using convolutional codes. (Reprinted with permission from S. Lin, D. J. Costello, Jr., and M. J. Miller, "Automatic-Repeat-Request Error-Control Schemes," *IEEE Communications Magazine*, **20**, p. 14 (December 1984). Copyright © 1984 IEEE.)

improvement over the simpler FEC codes when the BER falls below 10^{-4}, and throughput is maintained above 0.75 for BERs down to 10^{-2} with the rate 3/4 code. A complex decoder is needed with convolutional codes, as discussed in Section 7.7.

Example 7.8.1
Calculate the frequency of retransmission, throughput, and buffer requirements of a satellite link capable of carry data at rates of

(a) 24 kbps (b) 1 Mbps

when a block length of 127 bits is used and the one-way path delay is 240 ms, for a bit error rate of 10^{-4} and a double-error-detecting code (127, 120) using the following ARQ schemes.

1. Stop and wait.
2. Continuous transmission with transmit buffer only (go-back-N).
3. Continuous transmission with buffers at both ends of the link (selective repeat).

For a 127-bit code block, the probability of one or two errors is given by Eq. (7.4)

$$P_e(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

where $k = 2$, $n = 127$, and p is the probability of a single bit error, which is 10^{-4} in this example.

Thus the probability of an error being detected in the block of 127 bits is

$$P_e(1 \text{ or } 2 \text{ errors}) = 127 \times 0.0974 \times 10^{-4} + \frac{127 \times 126}{2} \times 0.095 \times 10^{-10}$$

$$= 1.245 \times 10^{-3}$$

This means that, on average, one in every 803 received blocks has one or two detectable errors.

1. In a stop-and-wait system we must send 127 bits and wait for an acknowledgment. Transmission of 127 bits takes approximately 5 ms at 24 kbps and 127 μs at 1 Mbps. Waiting for the acknowledgment takes 480 ms, so in both systems the transmission rate is approximately two blocks per second, or 254 bps. A block error would be detected, on average, after transmission of 803 blocks, which takes 400 s. Thus the throughput of the system is dominated by the path delay and waiting time when no transmission takes place. The inefficiency of stop and wait methods on satellite links rules out this technique completely for continuous data transmission.
2. Go-back-N system.
   a. The time to send 803 blocks at 24 kbps is 4.25 s. Thus every 4.25 s (on average) we must stop and wait 480 ms for a retransmission of 94 blocks of data. This slows the throughput of the system by about 12 percent. We need a buffer for $94 \times 127 = 11{,}938$ bits at each end of the link.
   b. If the data rate is 1 Mbps, we will detect block errors every 100 ms; since it takes 480 ms to call for a retransmission, the system spends most of the time retransmitting data, giving a throughput of about 172 kbps, on average. This is well below the potential capacity of the 1 Mbps link. We need a buffer for about 480k bits at the transmitter.
3. Selective repeat system. The only time lost in a selective repeat system is in the retransmission of blocks in which an error occurred.For an error in every 803 blocks, the efficiency is $803/804 = 99.87$ percent.
   a. In the 24 kbps system, the average data rate is 23.97 kbps. Transmit and receive buffers must hold 11,520 bits.
   b. In the 1 Mbps system, the average data rate is 998.7 kbps. Transmit and receive buffers must hold 480k bits.

7.9 SUMMARY

The transmission of data over a satellite communication link is likely to result in some errors occurring in the received data, for at least a small percentage of time, because of noise added by the transmission system. Many links guarantee only 10^{-6} bit error rate and may not achieve this accuracy during periods of rain or other propagation disturbances. Bit errors contribute to the baseband (S/N) ratio when digital speech is sent, but it is rarely necessary to correct bit errors in speech; the listener can make such corrections because there is a lot of redundancy in speech. When data are sent over a link, the receiving terminal does not know in advance what form the data take and can only detect or correct errors if extra, redundant bits are added to the transmitted data.

Coding of data provides a means of detecting errors at the receiving terminal. Error-detecting codes allow the presence of one or more errors in a block of data bits to be detected. Error-correcting codes allow the receiving terminal equipment to locate and correct a limited number of errors in a block of data. When error detection is employed, some form of retransmission scheme is needed so that the data block can be sent again when it is found to be in error. Retransmission schemes use ARQ (automatic repeat request) techniques and are easiest to apply in packet switched data networks, where data are not transmitted in real time. The long round-trip delay (480 ms) in a satellite link makes simple stop-and-wait systems unattractive for real-time data transmission. Throughput can be increased by providing data storage at both ends of the link and using continuous transmission in which corrupted data blocks are retransmitted by interleaving them with subsequent data block transmissions.

Forward error correction (FEC) provides a means of both detecting and correcting errors at the receiving terminal without retransmission of data. FEC codes add redundant parity check bits to the data bits in a way that allows errors to be located within a codeword. In general, twice as many errors can be detected by an FEC code as can be corrected. FEC has the advantage over error detection that a single unit at each end of the link (a codec) can insert and remove the FEC code and make corrections as required. Because more errors can be detected than can be corrected, hybrid schemes using both FEC and ARQ have been proposed. ARQ schemes can achieve virtually error-free transmission when the link error rate is not excessive. Provided that all errors introduced by the link are detected, retransmission of the corrupted data block can be repeated until the data are received correctly. In an FEC system, there is always a finite possibility that an error will not be corrected, resulting in incorrect data being received.

Linear block codes are a class of error-detecting and -correcting codes that are easy to implement when the codewords are short. Hamming codes fall into this group. Binary cyclic codes are popular linear block codes for long codewords or data blocks. They can be generated and detected using shift registers and logic gates. The BCH and Golay codes are examples of binary cyclic codes.

Convolutional encoding is a more powerful technique than linear block encoding. The information contained in any one data bit is spread through several bits of the codeword. However, decoding convolutional codes is a complex process that requires decisions on the most likely transmitted data sequence when a codeword is received in error. Convolutional encoding achieves the best improvement in error rate, especially when the link BER is high. It provides good resistance to interference and deliberate jamming, making it popular in military communication systems.

Interference tends to cause burst errors in which many sequential bits are corrupted. Special burst-error correction codes are available with the capability of correcting errors in a number of adjacent bits. Scrambling and interleaving of data bits are other ways in which the effect of burst errors can be reduced.

REFERENCES

1. C. E. Shannon, "A Mathematic Theory of Communications," *Bell Syst. Tech. J.*, Part I, 379–423; Part II, 623–656 (1948).
2. L. W. Couch, *Digital and Analog Communication Systems*, Macmillan, New York, 1983.
3. K. S. Shamnugam, *Digital and Analog Communication Systems*, John Wiley & Sons, New York, 1979.
4. S. Lin, and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
5. A. Hocquenghem, "Codes Corecteurs d'Erreurs," *Chiffres*, **2**, 147–156 (1959).
6. R. C. Bose and D. K. Ray-Chaudhuri, "On a Class of Error Correcting Binary Group Codes," *Info. Control*, **3**, 68–79 (March 1960).
7. W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, 2nd ed., MIT Press, Cambridge, MA, 1970.
8. K. Feher, *Digital Communications: Satellite/Earth Station Engineering*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
9. T. Maratani, H. Saithoh, K. Koga, Y. Mizuno, and Y. J. S. Snyder, "Application of FEC Coding to the Intelsat TDMA Systems," *Proc. 4th Intl. Conf. on Digital Satellite Comm.*, Montreal (October 1978).
10. Marlin P. Ristenbatt, "Alternatives in Digital Communication," *Proceedings of the IEEE*, **61**, 703–721 (June 1973). (Reprinted in [11], pp. 212–230.)
11. Harry L. VanTrees, Ed., *Satellite Communications*, IEEE Press, New York, 1979.
12. J. J. Spilker, Jr., *Digital Communications by Satellite*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
13. G. David Forney, Jr., "The Viterbi Algorithm," *Proceedings of the IEEE*, **61**, 268–278 (March, 1973). (Reprinted in [11], pp. 286–296.)
14. Jerrold A. Heller and Irwin Mark Jacobs, "Viterbi Decoding for Satellite and Space Communications," *IEEE Transactions on Communications* COM-19, pp. 835–848, October 1971. (Reprinted in [11], pp. 273–286.)
15. S. Lin, D. J. Costello, Jr., M. J. Miller, "Automatic-Repeat-Request Error-Control Schemes," *IEEE Communications Magazine*, **20**, 5–7 (December 1984).
16. S. Lin and P. S. Yu, "A Hybrid ARQ System with Parity Retransmission for Error Control of Satellite Channels," *IEEE Transactions on Communications, COM-30*, pp. 689–694, July 1982.

PROBLEMS

1. Alphanumeric characters are transmitted as 7-bit ASCII words, with a single parity bit added, over a link with a transmission rate of 9.6 kbps.

   a. How many characters are transmitted each second?

   b. If a typical page of text contains 500 words with an average of five characters per word and a space between words, how long does it take to transmit a page?

   c. If the bit error rate on the link is 10^{-5}, how many characters per page are detected as having errors? How many undetected errors are there?

   d. On average how many pages can be transmitted before
      (i) a detected error occurs
      (ii) an undetected error occurs?

   e. If the BER increases to 10^{-3}, how many detected and undetected errors are there in a page of the text?

2. A (6, 3) block code is formed from a generator matrix

$$G = \begin{bmatrix} 100 & 110 \\ 010 & 101 \\ 001 & 011 \end{bmatrix}$$

   a. Draw up a table of the eight codewords in this code.

   b. A message 111 000 100 001 110 011 is to be sent using this code. Convert the message to its coded form.

   c. What is the minimum distance of this code? How many errors can be detected in a codeword? How many errors can be corrected in a codeword?

3. The following series of codewords is received on a link using the (7, 4) Hamming code given in Table 7.2.

1100001 1001100 1110011 1110100 0011101

   a. Are there any errors in the received codewords? (Check against Table 7.2.) Which words contain errors?

   b. Calculate the syndrome for any code words that have detected errors.

   c. Find the correct codewords using the parity check matrix in Example 7.1. Check your own answer by comparison with Table 7.2.

4. Find the weight and minimum distance of the systematic (7, 4) cyclic code in Table 7.3.

   a. How many errors can this code detect?

   b. How many errors can it correct?

5. Analysis of a 56-kbps data link shows that it suffers burst errors that corrupt several adjacent bits. The statistics for burst errors on this link are given in Table P.5.

Table P.5
Statistics for Burst Errors on a Link in Problem 5

No. Adjacent Bits Corrupted	Probability of Occurrence
2	4×10^{-2}
3	2×10^{-3}
4	3×10^{-4}
5	1×10^{-6}
6	2×10^{-9}
7	5×10^{-11}
8	1×10^{-12}
9	3×10^{-14}
10	2×10^{-17}

a. Using Table 7.4, select a burst error correcting code that will reduce the probability of an uncorrected burst error below 10^{-10}.
b. Calculate the data rate for messages sent over the link using the code you selected.
c. Estimate the average bit error rate for the coded transmission.

6. A rate $\frac{1}{2}$ convolutional encoder is connnected as shown in Figure P7.6

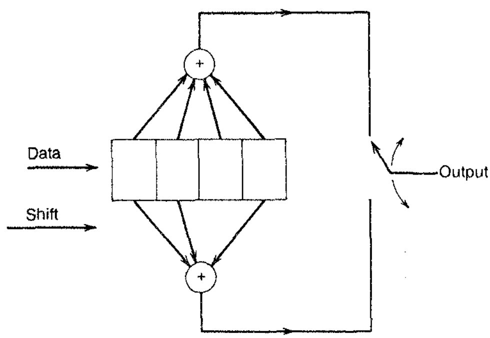


Figure P7.6

a. Determine and draw the state-diagram representation of the encoder.
b. Determine how the encoder would encode the binary sequence 10110. Assume that the left-most bit is transmitted first.
c. Assuming that the encoder started with 0000 in the shift register prior to accepting the first data bit, find the most probable decoding of a received message 00 11 10 10 00. Assume that the left-most bit was received first. How many bit errors does the received encoded message contain?

7. A satellite link carries packet data at a rate of 256 kbps. The data are sent in 255-bit blocks using a (255, 247) code that can detect three errors. The

probability of a single bit error, p_b, varies from 10^{-6} under good conditions to 10^{-3} under poor conditions. The one way link delay is 250 ms.

a. If no error detection is used, what is the message data rate for the link?
b. For a link BER of 10^{-6}, find the probability of detecting an error in a block of 255 bits. Hence find how often an error is detected.
c. Estimate the probability that a block of 255 bits contains an undetected error when the link BER is 10^{-3}.
d. Find the message data throughput when the link BER is 10^{-6} and a stop-and-wait ARQ system is used, assuming one retransmission always corrects the block.

8. Repeat Problem 7 using a block length of 1024 bits and a (1024, 923) code that can detect 22 errors in a block.

The (1024, 923) code can correct 10 errors. Find the average number of blocks that can be transmitted before an uncorrected error occurs when the BER is 10^{-3}. Repeat the analysis for a BER of 10^{-2}. *Note:* The probability of an unlikely event (11 or more errors in a block of 1024 data bits with $p_b = 10^{-3}$ in this case) can be calculated from the Poisson distribution more easily than from the binomial distribution. The Poisson distribution is given by

$$P(x = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

where $\lambda = Np_b$. N is the block length, and k is the number of bits in error.

8 PROPAGATION ON SATELLITE–EARTH PATHS AND ITS INFLUENCE ON LINK DESIGN

In Chapter 4 we introduced the concept of a link power budget. The key equation in that development was Eq. (4.11), repeated here as Eq. (8.1) and by this time presumably familiar to the reader.

$$P_r = \text{EIRP} + G_r - L_p - L_a - L_{ra}\ \text{dBW} \tag{8.1}$$

This equation indicates how the received power in dBW depends on the transmitter EIRP, the receiving antenna gain, and the various losses in the system. Everything on the right-hand side is independent of time except the atmospheric loss, L_a.

Usually L_a is written as the sum of a constant term called the *atmospheric absorption* and a variable term called the *attenuation*, A. At most frequencies of commercial interest, the atmospheric absorption is relatively unimportant (a few tenths of a decibel), and it disappears into the general uncertainty about the other terms in Eq. (8.1). The attenuation is zero in clear weather, but it can increase to large values during unfavorable propagation conditions. Rapid fluctuations in attenuation are called *scintillations*, while longer term increases are called *fades*. Rain fades are a particular problem above 10 GHz.

Attenuation can affect all types of satellite links; those that employ orthogonal polarizations to transmit two different channels on a common or overlapping frequency band are also degraded by *depolarization*. This is a conversion of some of the energy in a transmitted signal from one polarization to another. Absent under ideal propagation conditions, it can cause co-channel interference and crosstalk in dual-polarized satellite links. Rain is a primary cause of depolarization.

earth stations, and 6 spacecraft. By 1987 the number of coast earth stations is expected to grow to 29 [8]. Inmarsat headquarters are in London.

Inmarsat leases transponders from Marisat, ESA, and Intelsat. Marisat is a U.S. system that serves the U.S. Navy and merchant marine; Comsat General is the system manager and majority stockholder [3]. The ESA spacecraft that Inmarsat leases is MARECS A; MARECS B was lost during launch and will be replaced in 1985 by MARECS B2. The INTELSAT V spacecraft that carry Inmarsat traffic are equipped with a maritime communications subsystem (MCS).

REFERENCES

1. Joseph N. Pelton, "INTELSAT: Making the Future Happen," *Space Communication and Broadcasting*, **1**, 33–52 (April 1983).
2. Jose L. Alegrett, "U.S. Role in International Satellite Communications II: Leadership Continues," *Satellite Communications*, **8**, 20–23 (March 1984).
3. *COMSAT Guide to the INTELSAT, MARISAT, and COMSTAR Satellite Systems*, Communications Satellite Corporation, Washington, DC.
4. *Standard A Performance Characteristics of Earth Stations in the Intelsat IV, IVA, and V Systems Having a G/T of 40.7 dB/K*(BG-28-72E Rev. 1), Intelsat, Washington, DC, December 15, 1982.
5. K. Miya, Ed., *Satellite Communications Engineering*, Lattice Publishing Co., Tokyo, Japan, 1975.
6. *Standard "C" Performance Characteristics of Earth Stations in the INTELSAT V System (14 and 11 GHz Frequency Bands)*, (BG-28-73E Rev. 1), Intelsat, Washington, DC, December 15, 1982.
7. Paul Branch and Alex Da Silva Curiel, "Inmarsat and Mobile Satellite Communications," Part I, *Telecommunications*, pp. 30–32, 38, 40, 65 (March 1984).
8. "Inmarsat Data," *Ocean Voice*, **4**, 14–15 (January 1984).

11

SATELLITE TELEVISION: NETWORK DISTRIBUTION AND DIRECT BROADCASTING

Television is probably more strongly associated by the public with satellite communications than any other aspect of the industry. Quite rare in 1980, home satellite TV receivers are now common in the United States, and a whole new industry has grown up to sell and service them. Originally intended as relays for network TV and cable TV systems, current communications satellites are becoming de facto direct broadcast satellites (DBS) with their signals received by large numbers of private individuals. This has accelerated movement toward the construction and launch of the first true DBSs, intended to broadcast TV and possibly audio programs to the general population.

Some authorities feel that TV broadcasting will be one of the most important future uses for communications satellites in that it is difficult to foresee how any competing technologies (e.g., optical fiber cables) can offer the accessibility of satellite transmissions. This is particularly important to those users who want their television receivers to be easily portable. Satellites also offer the possibility of broadcasting high-definition TV, a service for which terrestrial television stations lack the necessary bandwidth.

In this chapter we will attempt to describe the current state of satellite TV distribution and home satellite TV reception. We will also summarize the current plans for DBS systems. The reader who would like more details on how to construct or assemble his or her own home earth station should consult reference 1 or 2 and *Satellite TV Magazine*.

11.1 TRANSPONDER FREQUENCIES AND DESIGNATIONS

The cable TV industry has developed a standard numbering system for satellite TV channels that has become widely accepted and that is used to label the tuning controls of most commercially available satellite TV receivers. As summarized in Table 11.1, it is based on 24 downlink center frequencies spaced 20 MHz apart with channel 1 centered at 3720 MHz and channel 24 centered at 4180 MHz. The channels themselves are 40-MHz wide. The spectra of adjacent channels radiated by a "24-channel satellite" thus overlap; interference is avoided by using alternating horizontal and vertical polarization on adjacent channels. The common practice is to use horizontal polarization for the odd-numbered channels and vertical polarization for the even-numbered channels. But this scheme is not universal, and WESTAR IV, for example, reverses it. Some satellites lack dual polarization and transmit only 12 40-MHz channels corresponding to the odd-numbered channels of the standard numbering system. The "12-channel satellites" are usually horizontally polarized [3]. In a 24-channel satellite the transponder numbers correspond to the channel numbers, while in a 12-channel satellite transponder numbers 1 through 12 correspond to odd-numbered channels 1 through 23. Reflecting this, some home satellite receivers have a tuning control with 12 positions. The first is marked with something like 1/2 for channels 1 and 2; the second is marked with 3/4 for channels 3 and 4, and so on. A toggle switch or pushbutton controls the antenna polarization and provides the center frequency offset necessary to receive the desired channel.

Published directories provide detailed listings of what is carried on each channel of each commercial satellite. One of the most complete appears in reference 4.

Table 11.1
Satellite TV Downlink Channel Numbering System

Channel	Center Frequency (MHz)	Channel	Center Frequency (MHz)
1	3720	13	3960
2	3740	14	3980
3	3760	15	4000
4	3780	16	4020
5	3800	17	4040
6	3820	18	4060
7	3840	19	4080
8	3860	20	4100
9	3880	21	4120
10	3900	22	4140
11	3920	23	4160
12	3940	24	4180

11.2 SATELLITE TELEVISION RECEIVERS

While there is considerable variation between manufacturers, the typical TV receive-only (TVRO) earth terminal consists of a small antenna (typically 8 to 15 ft in diameter) a low noise amplifier (LNA), a downconverter, and an IF receiver (simply called a receiver). The market closely resembles that for audio equipment, with some vendors offering integrated systems and others specializing in one or two components. In this section we will briefly survey the characteristics of the equipment that was available at the time of writing. Some of the specifications quoted will quickly become outdated, and the reader who is interested in assembling a personal TVRO terminal should consult the popular electronics press. We will begin with the receiver and work our way back up the chain to the antenna.

In Chapter 5 we summarized the modulation standards for television signals distributed by satellite. The composite video signal that frequency modulates the uplink carrier consists of a baseband video waveform extending from 0 to 4.2 MHz plus one or more frequency-modulated audio subcarriers, which are typically at 6.8 or 6.2 MHz. The first is now the most common, and the less-expensive receivers typically provide selection between only these two. More expensive receivers offer continuously tunable subcarrier reception at frequencies as high as 8 MHz.

While as much as 40 MHz of transponder bandwidth may be available, typical satellite TV signals require 23 to 30 MHz IF bandwidth in the receiver. The industry standard is supposed to be 30 MHz, but a random survey of current receiver specifications shows that 28 and 27 MHz are common, and at least one unit offers selectable IF bandwidths of 23, 25, and 27 MHz. There is an obvious advantage to being able to minimize noise by choosing the smallest bandwidth consistent with desired video quality. As we indicated in Chapter 5, satellite TV signals are frequently overdeviated (sent through a transponder bandwidth narrower than their Carson's rule bandwidth) to enhance video (S/N) in exchange for slightly lower picture quality.

The standard input frequency for TVRO IF receivers is 70 MHz. Most units have a threshold (C/N) at the IF input of 7 or 8 dB, and an FM improvement of about 34 dB can be expected. There is considerable difference between receivers in the detailed video and audio specifications for such things as distortion, differential gain, and the like; these may be important to an electronic hobbyist and unimportant to someone whose main interest in a TVRO terminal is to watch TV. Noise figures are usually not quoted for IF receivers since these make a negligible contribution to the overall TVRO terminal noise performance.

Downconverters translate a selected channel in the 3.7 to 4.2 GHz downlink band to IF. The downconverter local oscillator must be controlled by the IF receiver so that the local oscillator frequency corresponds to the channel for which the receiver is tuned. Typical noise figures for downconverters range from 15 dB down to a nominal 12 dB typical, 10 dB minimum, depending on the manufacturer and cost. Some manufacturers integrate the downconverter and the low noise amplifier into one unit, while others provide separate assemblies.

Low noise amplifiers are available at noise temperatures that range upward from 30 K, but those priced for home TVRO stations typically offer values between 90 and 120 K. Gains range from 30 to 50 dB with the higher figure being typical of units that combine an LNA and a downconverter in one package.

Feed systems for home TVRO antennas are primarily circular horns with rectangular waveguide outputs. The simplest provide only a single polarization and must be physically rotated to change between the horizontal and vertical polarizations of a 24-channel satellite. More sophisticated feeds provide remote polarization adjustment by means of a servo motor or a switch. The feed polarization control system must be compatible with the receiver, since some receivers combine polarization selection with channel selection.

Reflectors range upward in size from 7.5-ft diameters. With present transponders even a 9-ft antenna is marginal for most users, and 10- to 12-ft reflectors are more common. Except for zoning laws, the only limits to home TVRO antenna size are economic, and 15- and 16-footers are available.

Antenna positioners range from the "armstrong rotator" (an old radio amateur term for a hand-operated crank) to sophisticated servo systems that will point the antenna to the desired satellite at the touch of a pushbutton. The owner of a home TVRO system can have the degree of convenience that he or she is willing to pay for!

As in all problems of earth satellite design, it is not possible to give a simple answer to the question "Which system should I buy?" often asked of satellite communications engineers. The problem is exactly like that of assembling a home audio system in that it involves trading dollars for performance. One user may be very happy with a snowy picture while another may require network quality. See reference 5 for a discussion of the relationship between perceived picture quality and weighted (S/N) values. According to the results presented there, about 90 percent of the population considers a 35-dB signal-to-noise ratio "passable" and a 50-dB signal-to-noise ratio "excellent." Rounding the 36.5 dB minimum improvement factor of Chapter 5 up to 37 dB, these numbers correspond to (C/N) values ranging from -2 to 13 dB. The first is below threshold and unrealistically low, but these numbers indicate that (C/N) values only a few decibels above the demodulator threshold will be acceptable to most viewers. A recent article for prospective TVRO owners describes (C/N) values in the range 8 to 10 dB as offering good video quality and 10 to 14 dB as providing excellent quality [6]. These numbers are consistent with the foregoing discussion.

Construction details for those readers who want to build their own TVRO system from scratch are beyond the scope of this text. The primary problems are the reflector, feed, and LNA, three areas in which it is difficult to come close to the price and performance offered by industry. Do-it-yourself articles are available from mail order sources and in the popular electronics press; see the April 1980 *Radio Electronics* magazine for a particularly good set.

11.3 LEGAL MATTERS

At the time this book was written, many legal questions about home satellite TV reception were unresolved—at least in the minds of the interested parties. The originators of most of the television programs currently distributed by satellite (Home Box Office, for example) earn their revenues by selling their programs to cable TV systems. They believe that people who receive the programs directly from a satellite without permission or payment are violating their property rights. At the other extreme are electronic experimenters who believe that U.S. law has always permitted anyone to receive any available signal so long as (1) it is for a noncommercial purpose and (2) it violates no one's rights of privacy. The situation is exactly analogous to that involving copying of books and records. When technology makes it easy to copy something and when copying is cheaper than buying the original, many people feel that they have a right to copy. When it is as cheap to receive TV signals on your own terminal as it is to join the local CATV system, most people will want their own terminals. How they can have them without violating the property rights of the program originators is still unresolved. At the present time it is legal to own and operate a TVRO terminal, and no license is required. Whether permission to receive a particular program is required from the program originator is, in our opinion, still unresolved. See reference 2 for further discussion.

11.4 DIRECT BROADCAST SATELLITES

The satellite systems described in the previous sections of this chapter were designed to distribute television programs to TV broadcast stations and to CATV systems. Their widespread reception by private individuals and the current popularity of home TVRO terminals were probably not foreseen. But while most distributors of satellite TV may be looking for ways to discourage home TVRO viewers, a large number of companies are preparing to launch direct broadcast satellites (DBS) designed for convenient home reception. Some will probably provide encoded transmissions, which only paying subscribers can unscramble, while others will presumably be paid for by advertisers in the same way that broadcast TV is now.

This book was written before widespread DBS service began, and the material we will present about it may be outdated in a short time. Most of it was taken from the special satellite communications issue of *IEEE Communications Magazine* published in March 1984.

As approved by the FCC in 1983, the primary allocation for the so-called broadcast satellite service (*BSS*) is an uplink band from 17.3 to 17.8 GHz and a downlink from 12.2 to 12.7 GHz. In addition, that portion of the "normal" fixed satellite service (*FSS*) allocation from 11.7 to 12.2 GHz may be used for direct broadcast satellites on a noninterfering basis [7].

Currently there are 15 potential operators of DBS systems. Of these, Satellite Television Corporation (STC) is probably closest to operation. As an interim measure it expects to provide DBS service to a limited area of the northeast United States beginning in 1984 with several transponders of the Satellite Business Systems SBS IV spacecraft. This will offer a 53-dBW EIRP [8]. The FCC has authorized STC to launch two DBS satellites to share an orbital location of

100.8° W and broadcast a total of six TV channels to the eastern United States [9]. Dedicated DBS spacecraft will be launched by a number of operators in the 1985–86 time frame.

Proposed DBS transmitters will offer EIRP values in the 54 to 60 dBW range—10 to 20 dB better than what is now available from U.S. domestic satellites. These will provide clear-weather (C/N) values on the order of 14 dB in a 24 MHz bandwidth using home TVRO terminals with G/T values on the order of 12 dBK^{-1}. These may be achieved, for example, by employing 0.9-m dishes and receivers with overall noise temperatures of 480 K [10]. This is possible because small offset-fed antennas are available with aperture efficiencies of 75 percent and because GaAsFET amplifiers are in production with noise figures in the 2.8 to 3.2 dB range [8].

The high transmitter powers required by DBS spacecraft—typically 200 W per transponder—will limit the number of channels that a single satellite can carry to something between three and six [8]. Exactly how many DBS channels will ultimately be available is not yet clear, since it will depend on the bandwidths, polarizations, and frequency plans used. The most commonly accepted numbers are in the 32 to 40 channel range [11].

11.5 SUMMARY

Communications satellites are widely used to distribute television programming to terrestrial TV stations and to CATV systems. Many individuals receive these signals as if they were broadcast to the general public, and a number of companies are seeking to provide true direct broadcast satellite (DBS) service.

Most satellites currently used for TV distribution offer either 12 or 24 40-MHz bandwidth transponders. If only 12 are used, they are all on one polarization (usually horizontal). If 24 are used, they overlap in frequency and avoid interference by putting overlapping channels on orthogonal polarizations.

Home TVRO (television receive-only) terminals are available with antenna diameters from 7.5 ft up. For satisfactory performance, they must deliver (C/N) values of at least 8 dB to their FM demodulators. If proposed high-power DBS spacecraft are launched, terminals to receive them are expected to be significantly smaller and cheaper than those now required for conventional satellites.

REFERENCES

1. Anthony T. Easton, *The Home Satellite TV Book*, Putman Publishing Group, New York, 1982.
2. Robert J. Traister, *Build a Personal Earth Station for Worldwide Satellite TV Reception*, TAB Books, Blue Ridge Summit, PA, 1982.
3. C. J. Schultheiss, "The Ultimate Receiver," *Satellite TV Magazine*, **1**, 30–34 (March–April 1983).
4. Cheryl R. Carpinello, Lucinda Weindling, and Pat Hoyos, "Satellite Tran-

sponder Chart: Who's Carrying What on Which Satellite," *Satellite Communications*, 7, (un-numbered foldout section) (June 1983).
5. *Reference Data for Radio Engineers (6th ed.)*, Howard W. Sams and Co., Indianapolis, 1975.
6. C. J. Schultheiss, "Understanding TVRO," *Satellite TV Magazine*, **1**, 8–13, 84–85 (March–April 1983).
7. Trudy E. Bell, "Technology '84 Communications," *IEEE Spectrum*, **21**, 53–57 (January 1984).
8. D. K. Dement, "United States Direct Broadcast Satellite System Development," *IEEE Communications Magazine*, **22**, 6–10 (March 1984).
9. Hale Montgomery, "First In, First Out," *Satellite Communications*, **8**, 12 (July 1984).
10. Wilbur L. Pritchard and Harley W. Radin, "Direct Broadcast Satellite Service by Direct Broadcast Satellite Corporation," *IEEE Communications Magazine*, **22**, 19–25 (March 1984).
11. Richard G. Gould, "Transmission Standards for Direct Broadcast Satellites," *IEEE Communications Magazine*, **22**, 26–34 (March 1984).

PROBLEMS

1. Calculate the video-weighted (S/N) that can be expected from a "super-cheap" home TVRO earth terminal that uses a 7.5-ft diameter dish with a noise temperature of 50 K and an LNA with a noise temperature of 120 K. Assume that all of the receiver noise is contributed by the LNA, the operating frequency is 4 GHz, the satellite EIRP is 34 dBW, the distance from the satellite to the earth station is 38,000 km, the earth station antenna's aperture efficiency is 0.55, and the FM demodulator has an 8-dB (C/N) threshold.

2. Direct broadcast satellites (DBS) will probably be launched in 1985 to provide low-cost home TV via inexpensive terminals. The allocated frequencies are 17.3 to 17.8 GHz uplink and 12.2 to 12.7 GHz downlink; these provide a total bandwidth of 500 MHz. Transmission standards for DBS are not yet fully developed; obviously every prospective operator wants to maximize the number of channels that can be carried in the available bandwidth, minimize the number of satellites required to cover this bandwidth, and minimize the cost of the required earth terminals.

In this problem you will examine some of the trade-offs involved in planning a DBS system. Your goal is to design a system that meets the following specifications so that (1) the number of channels (transponders) carried by an individual satellite is as large as you can make it but (2) if more than one satellite is required to cover the allocated 500-MHz band, each satellite carries the same number of transponders.

a. Spacecraft Specifications:
   (i) Receive G/T: $+1$ dBK^{-1}
   (ii) Saturation uplink flux density: -90 dBW/m^2

(iii) Total RF power: 1400-W end of life. The sum of the total transponder output powers on any one spacecraft cannot exceed 1400 W.
(iv) Downlink antenna pattern: Circularly symmetric with 3° minimum half-power beamwidth across downlink band.
(v) Number of transponders: One for each TV channel carried.
(vi) Transponder bandwidth: TV channel bandwidth multiplied by 1.10
(vii) Total bandwidth (sum of transponder bandwidths): 500 MHz maximum.
(viii) Backoff required: None. Only one carrier is present and the transponder is assumed to be linear.
(ix) Orbital location: 101° W.

b. Uplink Earth Station Specifications:
(i) Location: 80.438° W, 37.229° N.
(ii) Transmitting antenna: 20 ft in diameter with a 55 percent aperture efficiency.

c. Downlink Earth Station Specifications:
(i) Location: Anywhere in eastern half of the United States.
(ii) Antenna: 1 m in diameter with 75 percent aperture efficiency.
(iii) Antenna noise temperature: 90 K in clear weather.
(iv) Receiver U.S. standard noise figure: 3 dB (includes all noise contributions except those from the antenna).
(v) Demodulator threshold: 7.5 dB.
(vi) Receiver noise bandwidth: Negotiable. It can be made equal to the specified TV channel bandwidth.
(vii) Video (S/N) required: 45 dB at beam edge, calculated by $(S/N) = (C/N) + 18.8 + 10 \log_{10}[3\, m^2(m + 1)]$ dB. This includes all preemphasis and weighting factors. The modulation index m is given by $m = f_p/f_v$ where f_p is the peak deviation (you choose its value) and f_v is 4.2 MHz.

In your design do not worry about rain margin. Do not assume overdeviation; calculate occupied bandwidths from Carson's rule.

APPENDIX

A.1 DECIBELS IN COMMUNICATIONS ENGINEERING

Most readers of this book will be familiar with the practice of expressing power ratios in decibels, abbreviated dB. The dB ratio A of two power levels p_1 and p_2 is given by

$$A = 10 \log_{10}\left(\frac{p_1}{p_2}\right) \text{dB} \tag{A.1}$$

provided that p_1 and p_2 are expressed in the same units. Although the decibel is formally defined only for a power ratio, p_1 and p_2 and A can also be expressed in terms of many combinations of voltage, current, resistance, electric field strength, magnetic field strength, and so on.

It is common practice in communications engineering to use decibels and the mathematical properties of the logarithm to transform multiplicative equations to additive equations, to manipulate the additive equations into particularly convenient forms, and to define new logarithmic units with dB in their names for some of the quantities that appear. When first presented, this practice is confusing to many people, and we hope to clarify it here.

Consider the simple voltage divider circuit with resistors R_S and R_L shown in Figure A.1. The rms voltages across the source and the load resistance are V_S and V_L, respectively; the rms power supplied by the source is p_S W and the rms power delivered to the load is p_L W. From elementary circuit theory, these quantities are related by

$$p_S = \frac{V_S^2}{R_S + R_L} \tag{A.2}$$

$$p_L = \frac{V_L^2}{R_L} \tag{A.3}$$

GLOSSARY

AC Assignment channel. A channel carrying assignment information in the Intelsat TDMA system.

ACK Acknowledge signal. This is sent when data are received correctly in an ARQ system.

ADM Adaptive delta modulation. Delta modulation that varies its step size in response to the behavior of the input signal.

AKM Apogee kick motor. A rocket engine that boosts a satellite from the Space Shuttle orbit into a transfer orbit.

ALOHA A random access system.

Antenna gain An increase in flux density in a given direction produced by an antenna relative to flux density produced by an isotropic antenna transmitting the same total power.

Aperture The area across which an antenna radiates or receives energy.

Aperture antenna A microwave antenna using a horn or a feed and reflector.

Aperture efficiency The ratio of effective aperture of an antenna to its physical area, typically in range 50 to 75 percent. A lossless, uniformly illuminated aperture has an aperture efficiency of 100 percent.

Apogee The point in an orbit of greatest distance from the earth.

APT Automatic picture transmission. A format for sending weather data from a satellite.

ARQ Automatic repeat request.

ASCII code A standardized binary code to represent alphabetic letters, numbers, and symbols. ASCII stands for American Society for Computer Information Interchange.

AT&T American Telephone and Telegraph Company.

Attenuation Time-varying fading.

Attitude and Orbit Control System (AOCS) A system that keeps a satellite in the correct orbit and pointing in the correct direction. Made up of gas jets and/or momentum wheels.

Audio subcarrier A carrier that is modulated by the audio component of a television signal.

Autotrack Automatic tracking of satellite motion by earth station antenna.

Azimuth The horizontal angle measured east from north to the line from an observer to a satellite.

Backoff The process of reducing the input power level of a traveling wave tube to obtain more linear operation.

Baseband The frequency band that a signal occupies when initially generated.

BCH codes Bose-Chaudhuri-Hocquenghem codes, named after discoverers. These are linear block codes.

BER Bit error rate. The fraction of a sequence of message bits that are in error.

Binary cyclic codes Block codes with cyclic properties. These can be generated with shift registers.

Block code A binary data transmission format in which message bits and parity check bits are formed into blocks, each with a predetermined number of bits.

Blockage The loss of energy in a reflector antenna caused by the presence of obstacles in the aperture (feed, feed supports, etc.).

Boresight The center of an antenna beam, usually the direction of maximum gain.

Boltzmann's constant The constant used to obtain noise power from noise temperature. Value is 1.38×10^{-23} J/K, or -228.6dBW/K/Hz.

BPSK Binary (or bipolar) phase shift keying. A digital modulation technique in which the carrier phase takes on one of two possible values.

BSS Broadcast satellite service.

Burst errors Errors that occur in adjacent bits.

C band Frequency range 4 GHz to 8 GHz.

Carson's rule bandwidth The theoretical bandwidth needed to transmit an FM signal without distortion.

Cassegrain antenna A two-reflector antenna with a configuration originally used by William Cassegrain for an optical telescope.

CATV Originally, this stood for community antenna television, but now it is used as a general abbreviation for cable television.

CCIR International Radio Consultative Committee.

CCITT International Telegraph and Telephone Consultative Committee.

CDMA Code division multiple access. A multiple access scheme in which stations use spread-spectrum modulation and orthogonal codes to avoid interfering with one another.

CES Coastal earth station in the Inmarsat system.

Circumscribed circle A circle that is tangential to an orbit and that encloses the orbit.

Clear sky Conditions under which atmosphere does not cause excess attenuation or depolarization of RF signals.

Clip The time delay between the onset of sound and the completion of a channel assignment and connection in a demand-assignment system.

(C/N) Carrier-to-noise ratio.

Code vector A codeword.

Codec A coder–decoder.

Codeword A combination of message bits and parity check bits into a word of fixed length.

Coma Distortion of plane wave in antenna aperture caused by movement of the feed away from the focus in the focal plane.

Companding The process by which the dynamic range of speech is compressed before transmission and expanded after detection. The name is a contraction of compression and expanding.

Comsat The Communications Satellite Corporation.

Conical scan Autotrack technique that derives angular error of a satellite in two planes simultaneously by rotation of the antenna beam around the boresight axis.

Convolutional code A code in which there is not a one-to-one correspondence between data bits and coded bits.

Corrugated horn A microwave antenna with high efficiency. It is used as a feed for reflector antennas because of its symmetrical beam shape and low cross-polarization.

CR/BTR Carrier recovery/bit timing recovery.

Crosstalk Transfer of signal between two separate channels or circuits.

CSC A common signaling channel in the SPADE system.

CSSB Companded single sideband. An analog FDM technique using companding and single sideband amplitude modulation.

dBK Noise temperature unit expressed in decibels greater than 1 K.

dBK^{-1} Units of G/T ratio, decibels per Kelvin.

dBm Decibels relative to one milliwatt.

dBp Power in dB above 1 picowatt. It may be weighted or unweighted, depending on context.

DBS Direct broadcast satellite.

dBW Decibels relative to one watt.

Deemphasis The removal of some high-frequency noise from a demodulated FM signal. See *preemphasis.*

Delta modulation A process for digitally encoding and transmitting an analog waveform that avoids the use of analog-to-digital and digital-to-analog converters.

Demultiplexer A device that recovers individual signals that had been multiplexed for transmission.

Despun antenna An antenna, mounted on a satellite with a spinning body, which is rotated in the opposite direction to the body rotation so that the antenna beam points in a fixed direction.

Differential modulation Digital modulation in which the change in the carrier phase or amplitude from one symbol to the next is determined by the modulating signal. This is in contrast to direct modulation, in which the state of the carrier phase (or amplitude) is determined by the modulating signal.

Direct modulation A digital modulation scheme in which the phase (or amplitude) of the carrier is determined by the modulating signal. This is in contrast to differential modulation, where the change in carrier amplitude (or phase) is determined by the modulating signal.

DM Delta modulation.

Doppler shift The change in radio frequency that results from motion of the transmitter or receiver.

Downlink The communications channel from a satellite to an earth station.

DSI Digital speech interpolation. The digital implementation of TASI, a technique for reassigning channels during speech pauses.

Duplex link A link capable of simultaneous two-way transmission of data.

Eccentricity A measure of the ellipticity of an orbit. If the eccentricity is zero, the orbit is circular.

Eclipse A satellite in the shadow of the earth.

EIRP Effective isotropically radiated power. This is equal to antenna gain multiplied by transmitted power.

Elevation The angle measured up from local horizontal to the line from an observer to a satellite.

ELV Expendable launch vehicle.

Encryption Systematic modification of a signal to prevent unauthorized use.

Energy dispersal The process of modulating a lightly loaded link to maintain a low power spectral density and avoid interference with terrestrial links.

Equalizer A filter with an amplitude or phase characteristic designed to correct for amplitude or phase distortion in another part of a communication channel.

Erlang The unit of traffic flow in a communication system.

Erlang B model The preferred equation for calculating the number of channels required to carry a given amount of traffic.

Error correction The correction of bits in a digital data stream that have been corrupted during transmission.

Error detection The detection of bits in a digital data stream that have been corrupted during transmission.

Error syndrome A word that indicates whether a codeword is correct or in error.

ESA European Space Agency.

Even (odd) parity Addition of one or more bits to a data signal to permit error detection or correction. In even parity, the binary sum of the bits in a character or block is even. In odd parity, the binary sum is odd.

Far field A region sufficiently distant from a transmitting antenna that the observed wave is locally plane.

FAW Frame alignment word. A pattern of bits that identifies the start of a frame.

FCC Federal Communications Commission. The government agency in the United States that regulates radio communications and allocates frequencies.

FDM Frequency division multiplexing. A technique whereby several signals from the same earth station share a transponder by using different frequencies.

FDMA Frequency division multiple access. A technique whereby signals from several earth stations share a satellite or transponder by using different frequencies.

FEC Forward error correction.

FH Frequency hopping. A spread-spectrum technique in which the transmitter frequency is "hopped" over a wide bandwidth.

Flux density Power per unit area when an electromagnetic wave is incident on a surface.

FM Frequency modulation.

FM improvement Increase in (S/N) at the output of an FM demodulator relative to the (C/N) at its input.

FPD Focal plane distribution. Spatial variation of energy in the focal plane of an antenna.

Frame (TDM) A portion of a digital transmission that contains one word from each channel and sufficient synchronization information to identify the start of the frame.

Frame (TDMA) A portion of a digital transmission that contains one burst from each earth station plus synchronization and housekeeping information.

Frequency coordination A process by which interference is avoided when planning an earth station installation.

Frequency reuse Transmitting to or receiving from a satellite two independent data channels at the same radio frequency. The channels are separated by the spacecraft antenna using directional beams in spatial frequency reuse and by orthogonal polarizations in polarization frequency reuse.

Front-fed antenna An antenna with a single reflector and a radiating feed at the focus of the reflector.

FSK Frequency shift keying.

FSS Fixed satellite service. An FCC term describing satellite communications with fixed (i.e., nonmobile) earth stations.

GaAsFET Gallium Arsenide Field Effect Transistor.

GCE Ground control equipment (at earth station).

Geostationary arc A circle of radius 42,242 km lying in the plane of the equator and having its center at the center of the earth.

Geostationary orbit A circular orbit of radius 42,242 km lying in the plane of the equator. To an observer on the ground, a satellite in geostationary orbit remains at the same fixed position in the sky.

Geosynchronous orbit A circular orbit of radius 42,242 km that is not in the earth's equatorial plane. A satellite in geosynchronous orbit has an orbital period equal to the earth's rotational period, but its inclination with respect to the equatorial plane makes its position with respect to an observer on the ground change with time.

Global beam A satellite antenna beam covering the whole of earth, as seen from the satellite.

Grating lobe An unwanted sidelobe caused by two radiating or receiving antennas or elements that are separated by more than one half wavelength.

Gregorian antenna A two-reflector antenna with a configuration originally used by James Gregory for an optical telescope.

G/T Antenna gain to noise temperature ratio. Used to characterize earth stations.

Guardband An empty frequency band used to separate FDM signals.

Guard time An empty time interval used to separate TDMA bursts.

Hamming code A simple block code that is useful for correcting or detecting a small number of errors per block.

Housekeeping This term refers to systems used on a satellite to keep it operating, but that do not form part of the payload.

HPA High-power amplifier.

I channel The bit stream that modulates the in-phase carrier of a QPSK system.

IF Intermediate frequency.

Illumination efficiency A measure of the efficiency with which an aperture collects energy from an incident plane wave. Uniform illumination of the aperture gives 100 percent efficiency when the antenna transmits or receives.

Inmarsat The International Maritime Satellite Organization.

Intelsat The International Telecommunications Satellite Organization.

Intermodulation noise Noise generated by the interaction of signals in a nonlinear device (usually a TWT in the context of this text).

ISI Intersymbol interference. This is interference between different symbols in a serial digital transmission.

Isotropic source (antenna) An antenna that radiates equal power in all directions.

Julian date The number assigned to a day in a standard astronomical dating system.

Kelvin Unit of noise temperature.

Key The information needed to successfully decode an encrypted data sequence.

Ku-band Frequency range of 12 GHz to 18 GHz

L-band Frequency range of 1 GHz to 2 GHz

LDM Linear delta modulation. Delta modulation that uses a fixed step size.

LHCP Left-hand circular polarization.

Link power budget Method for calculation of received power in a satellite link.

LNA Low noise amplifier.

Loading A measure of the traffic carried by a multiplex telephone link.

Loading factor A factor that is multiplied by the rms test-tone deviation to yield the rms deviation of an FDM/FM multiplexed telephone signal.

Look angles The coordinates to which an earth station antenna must be pointed to communicate with a satellite.

Luminance The amount of white light at a point in a television picture.

Margin The amount (usually in decibels) by which a received signal exceeds a predetermined lower limit.

Modem A modulator/demodulator.

Momentum wheel A heavy wheel, mounted within a satellite, which rotates at high speed to provide angular momentum in one direction. It is used in the attitude control system.

Monopulse An autotrack technique that derives angular error of a satellite in two planes simultaneously.

MTBF Mean time before failure.

Multiple access The sharing of a transponder or satellite by signals from several earth stations.

Multiplexer A device that combines signals for transmission.

Multiplexing The sharing of a transponder by several signals from the same earth station.

(n, k)code An arrangement of binary data such that in a block of n bits, k bits are the message and $(n - k)$ are parity check bits.

NAK Not acknowledge signal. It is sent when data are received incorrectly in an ARQ system.

NASA The National Aeronautics and Space Administration.

NOAA National Oceanographic and Atmospheric Agency. U.S. Government body that operates weather satellites.

Noise figure A measure of the noise power generated by a device or system.

Noise power budget A method for calculation of total noise power in a receiving system.

Noise temperature A measure of noise power that is independent of measurement bandwidth.

North–South, East–West, stationkeeping maneuver Movement of the satellite perpendicular to (N–S) and along (E–W) the geostationary orbit to correct orbital errors.

NPR Noise power ratio. A measure of intermodulation noise in a multiplexed telephone channel.

NRZ Non-return-to-zero. A baseband digital transmission scheme in which the logical 1 and 0 signals correspond to equal-amplitude positive and negative voltages.

NTSC The 525-line/60-Hz television transmission standard used in North America and Japan.

Nyquist filter A filter that produces zero ISI waveforms when driven by an impulse.

Offset reflector A reflector that is not a rotationally symmetric conic section.

OMT Orthogonal mode transducer. This separates polarizations in an antenna feed.

OQPSK Offset QPSK modulation.

Orbital elements A set of six constants that are sufficient to specify an orbit.

Osculating orbit The orbit that a spacecraft would follow if all perturbing forces were removed.

OTS Orbital Test Satellite, built by the European Space Agency.

Outage Loss of communication.

Overdeviation Intentionally modulating an FM transmitter so that the Carson's rule bandwidth of the uplink signal will be significantly greater than the transponder (and possibly the IF receiver) bandwidth.

P channel The bit stream that modulates the in-phase carrier of a QPSK system.

PAL The European 625-line/50-Hz television transmission system.

PAM Payload assist module. A rocket engine used in Space Shuttle launches to move a satellite from low orbit to geostationary orbit.

Parabolic torus reflector A reflector that is circular in one plane and parabolic in the orthogonal plane. Beam can be scanned by feed movement in the plane of circular curvature.

Paramp Parametric amplifier.

Parity check matrix A method for calculating the error syndrome in a block codeword.

Path loss The apparent loss of power caused by spherical expansion of a transmitted wave. Equal to $(\lambda/4\pi R)^2$.

PCM Pulse code modulation. Any modulation scheme that transmits digitally encoded quantized samples of an incoming signal.

Peak factor A factor that is multiplied by the rms deviation to yield the true deviation of an FDM/FM multiplexed telephone signal.

Perigee The point in the orbit where a satellite is closest to the earth.

PN Pseudonoise. This is also the abbreviation for a spread-spectrum scheme in which a transmitted signal is spread through multiplying it by a PN sequence.

PN sequence A sequence of apparently random ones and zeros that actually repeats after a finite time. The spectrum of a PN sequence is similar to that of white noise.

Polarizer A device to generate wanted polarization in an antenna feed.

Preamble The part of a TDMA burst that contains synchronization and other housekeeping information.

Preemphasis A process for increasing the signal-to-noise ratio of a demodulated FM signal by filtering out some of the high-frequency noise. The signal is distorted before modulation by a preemphasis filter, and the distortion is removed along with some of the noise by a deemphasis filter.

Principal planes The planes in an antenna that align with reference axes such as the direction of linear polarization and its orthogonal plane.

PSK Phase shift keying. A modulation technique in which the phase of the RF carrier is varied.

Psophometric weighting A correction factor used in signal-to-noise ratio calculation that accounts for the nonuniform response of the human ear to white noise.

Pulse stuffing Insertion of dummy words into a multiplexed digital transmission. It is done to allow an input channel with a low transmission rate to catch up with the rest of the system.

pWp Picowatts psophometrically weighted. Power in picowatts corrected for the nonuniform response of the human ear to white noise for use in signal-to-noise ratio calculation.

Q channel The bit stream that modulates the quadrature channel of a QPSK system.

QPSK Quadrature phase shift keying. A digital modulation technique in which the carrier phase takes on one of four possible values.

Quantization The process of resolving analog samples of a signal into one of a finite number of possible values.

RARC Regional Administrative Radio Conference. Sets frequency allocations for a region of the world.

RCA Radio Corporation of America.

Redundancy The provision of spare components in a satellite or earth station to ensure continuous operation when one component fails.

Redundant bits Extra bits inserted in a digital data stream to enable error detection or correction to be carried out.

Reference burst The burst that marks the start of a TDMA frame.

Reliability Measure of the likelihood of a component or system continuing to operate correctly after a given time.

RF Radio frequency.

RHCP Right-hand circular polarization.

rms multicarrier deviation The rms deviation of an FDM/FM telephone signal.

Roll, Pitch, Yaw Rotational motion about three Cartesian axes centered in the spacecraft.

S band Frequency range of 2 GHz to 4 GHz.

SAM Simple attenuation model.

SBS Satellite Business Systems. A U.S. company that operates satellite networks, primarily digital.

SC Satellite channel. The bits representing 16 terrestrial channels (TC) in the Intelsat TDMA system.

Scalar feed Feed for front-fed reflector antennas with high efficiency and symmetry.

Scintillation Rapid fluctuations in signal strength.

SCPC Single channel per carrier. An FDM technique in which each signal is modulated onto its own carrier for transmission.

SER Symbol error rate. The fraction of a sequence of message symbols that are in error.

SES Ship earth station in the Inmarsat system.

Shannon capacity The maximum possible data rate of a communication channel for a given signal-to-noise ratio, with no bit errors.

Shaped reflector Antenna reflector with modified profile for improved efficiency.

SIC Station identification code.

Simplex link A link that is capable of only one-way transmission of data.

Site-diversity A scheme in which two or more redundant earth stations located a few kilometers apart communicate with the same satellite.

Sky noise Electromagnetic radiation from galactic sources and thermal agitation of atmospheric gases and particles.

Slot A grouping of bits from one channel (in TDM) or from one earth station (in TDMA).

(S/N) Signal-to-noise ratio.

SOF Start of frame in a TDMA system.

Solar sails Long flat panels covered with solar cells, projecting from a satellite. The panels are rotated to face the sun and generate the spacecraft's electrical power from incident sunlight.

SORF Start of receive frame. The reference time for an earth station's TDMA downlink.

SOTF Start of transmit frame. The reference time for an earth station's TDMA uplink.

Space qualification Special test procedures to ensure high reliability of spacecraft parts.

Spacecraft Any vehicle in orbit. All communication satellites are spacecraft.

SPADE Single channel per carrier PCM multiple-access demand assignment equipment. An Intelsat demand-access system.

Sparklies White flashes or dots appearing on a TV screen in FM TV systems. They are caused by noise impulses from the FM demodulator.

Spherical reflector A reflector having a spherical surface, which can employ a movable feed to scan the beam.

Spillover Energy lost in an antenna that does not contribute to the illumination efficiency.

Spinner A satellite with a body that rotates to provide gyroscopic stabilization of attitude.

Spot beam Very narrow beam covering a small part of the earth's surface.

SQPSK Staggered QPSK modulation.

SS Spread-spectrum. A modulation technique in which a signal's energy is spread over a bandwidth much wider than that required for transmission.

SS-TDMA Satellite switched TDMA. A TDMA system in which the satellite can interconnect different uplink and downlink spot beams.

SSPA Solid-state power amplifier.

STS Space Transportation System. The Space Shuttle.

Subdish Secondary reflector in a Cassegrain or Gregorian antenna.

Subsatellite point The place where a line drawn from the center of the earth to a satellite passes through the earth's surface.

Superframe The number of TDM or TDMA frames required to transmit all housekeeping and synchronization information once.

Surface errors Departure of reflector surface from ideal curvature due to manufacturing tolerance, alignment errors, etc.

Symbol One of the unique states of the carrier in digital modulation. A symbol may carry one or more bits.

System noise temperature The total apparent noise temperature at a reference point in a receiving system.

Taper Distribution of field strength across an aperture or reflector.

TASI Time-assigned speech interpolation. A technique for reassigning channels during speech pauses.

TC Terrestrial channel. The bits containing 16 samples of one voice channel carried in the Intelsat TDMA system.

TDM Time division multiplexing. A technique whereby several signals from the same earth station share a transponder by using it at different times.

TDMA Time division multiple access. A technique whereby signals from several earth stations share a satellite or transponder by using it at different times.

Telemetry, Tracking, and Command (TT&C) A system that provides two-way communication between an earth station and the satellite to monitor spacecraft systems and to send instructions for changes.

Test tone A 1-mW (0-dBm) 1-kHz signal used to adjust telephone equipment.

Three-axis stabilization (body stabilization) A technique to maintain the body of a satellite in the same orientation relative to earth at all times. The body of the satellite does not rotate.

3-dB beamwidth A measure of the width of an antenna beam. The power level observed by a receiving antenna placed at the 3-dB point of an antenna pattern is one-half of the power received when the antenna is at boresight.

Threshold A level at which the operation of a system changes suddenly.

Thruster A small rocket used to provide fine control of the velocity or attitude of a spacecraft.

T1 carrier A 24-channel TDM system developed by the Bell System for telephony.

Transfer orbit An intermediate orbit used in the process of launching a geostationary satellite.

Transparent transponder A transponder that does not modify the signal, other than by amplification and frequency conversion.

Transponder Basically a receiver followed by a transmitter. A transponder receives an uplink signal at one frequency, amplifies it, and retransmits it on another frequency.

Transponder hopping One TDMA earth station working several transponders.

Trellis diagram A technique for locating errors in convolutional code codewords.

TRT The timing reference transponder for an Intelsat TDMA network.

TVRO Television receive-only terminal.

TWT Traveling wave tube.

TWTA Traveling wave tube amplifier.

Universal time (UT) Standard time for space operations and scientific observations.

Uplink The communications channel from an earth station to a satellite.

UW Unique word. A sequence of bits used in TDMA to mark each frame or superframe.

Viterbi codes Convolution codes named for A. J. Viterbi.

WARC World Administrative Radio Conference. This sets frequency allocations for the whole world.

Windowing A process in which a TDMA earth station searches for the unique word only during some time window.

X band Frequency range of 8 GHz to 12 GHz.

XPD Cross-polarization discrimination. A measure of the conversion of power from one wave polarization to the orthogonal polarization.

XPI Cross-polarization isolation. A measure of the coupling between nominally orthogonally polarized channels.

Z-axis intercept Intersection of the satellite Z axis and Earth's surface. Defines the antenna pointing direction.

Zone beam Satellite antenna beam covering a part of the earth's surface.

LIST OF SYMBOLS

The following is a list of the symbols used in most of the equations of this text. We have tried to adopt the symbols used in the relevant literature, recognizing that in some cases this causes the same symbol to be used for more than one quantity, while in other cases the same quantity is represented by more than one symbol.

a	Arc length used in look angle calculation.
a	Height of antenna aperture.
a	Semimajor axis of orbit.
$\mathbf{a}$	Electric field vector of a vertically polarized transmission.
$\mathbf{a}_c$	Co-polarized electric field vector that is received when $\mathbf{a}$ is transmitted.
a_n	Coefficients used in the infinite-series representation of a general nonlinear device.
$\mathbf{a}_x$	Cross-polarized electric field vector that is received when $\mathbf{a}$ is transmitted.
aR^b	Specific attenuation at rain rate R.
A	Amplitude of a general sinewave.
A	Traffic intensity in erlangs.
A	Attenuation, dB.
A	Subjective signal-to-noise ratio improvement provided by a compandor.
A	Area of an antenna aperture.
A_c	Total attenuation due to the clear atmosphere on a vertical path.
A_e	Effective area of an antenna aperture.
A_j	Joint attenuation in a diversity reception system.
$A(P)$	Attenuation exceeded for P percent of the time.
A_S	Average single-site attenuation in a diversity system.
Az	Azimuth angle.
$A_{0.01}$	Attenuation value exceeded for 0.01 percent of a year.
b	Width of an antenna aperture.
b	Semiminor axis of orbit.

b	Voice channel bandwidth in hertz.
$\mathbf{b}$	Electric field vector of a horizontally polarized transmission.
$\mathbf{b}_c$	Co-polarized electric field vector that is received when $\mathbf{b}$ is transmitted.
$\mathbf{b}_x$	Cross-polarized electric field vector that is received when $\mathbf{b}$ is transmitted.
B	Bandwidth in hertz.
B_0	Geomagnetic flux density in Teslas.
B_{IF}	IF bandwidth in hertz.
BO_i	Input backoff in decibels.
BO_o	Output backoff in decibels.
c	Arc length used in look angle calculation.
C	Codeword
C	Polar angle between earth station and subsatellite point.
C	Received carrier power.
C	Undetermined constant in differential equation solution.
C	Total number of available channels.
$(\mathrm{C/N})_i$	Carrier-to-noise ratio at input of an FM demodulator.
$(\mathrm{C/N})_t$	Overall carrier-to-noise ratio on a satellite link.
$(\mathrm{C/N_0})$	Carrier-to-noise power density ratio.
$(\mathrm{C/N})_D$	The value that the downlink carrier-to-noise ratio would have if the transponder did not retransmit incoming noise.
$(\mathrm{C/N})_I$	Term introduced into (C/N) equations to account for the effects of intermodulation noise.
$(\mathrm{C/N})_{\mathrm{IS}}$	The value of $(\mathrm{C/N})_I$ when a TWT is operating at saturation.
$(\mathrm{C/N})_U$	Uplink carrier-to-noise ratio.
dD	Incremental raindrop diameter.
d_m	rms multichannel deviation at full loading.
dt	Differential time.
D	Antenna aperture diameter.
D	Data message
D_m	Median raindrop diameter.
e	Eccentricity of an orbit.
e_{in}	Input voltage of a general nonlinear device.
e_{out}	Output voltage of a general nonlinear device.
E	Electric field.
E	Orbital eccentric anomaly.
E	Number of bits by which a received sequence can differ from the unique word and still be counted as correct.
E_b	Energy per bit.
El	Elevation angle.
E_s	Energy per symbol.
EIRP_S	Satellite effective isotropic radiated power.
f_c	Carrier frequency in hertz.
$f_{\max}$	Maximum modulating frequency in hertz.
f_{mod}	Modulating frequency in hertz.
f_r	frequency of minimum attenuation in a preemphasis filter.
f_N	IF carrier frequency of Nth channel in SPADE.
f_R	Received frequency.
f_T	Frequency that a moving transmitter would radiate when at rest.
f_V	Maximum video-modulating frequency.

F	Focal length of a parabolic reflector.
F	Probability of a false alarm (confusing an extraneous bit sequence with a unique word).
F	Flux density.
F^{-1}	Inverse Fourier transform.
F_{max}	Deviation in hertz that the maximum dispersal waveform causes.
F_S	Single-carrier saturation flux density at a satellite in dBW/m^2.
g	Peak factor for a multiplexed telephone signal.
$g(f)$	Factor used in scaling attenuation values with frequency.
G	Antenna gain.
G	Amplifier gain.
G	Green signal in color TV transmission.
G	Universal gravitational constant (6.67×10^{-11} Nm^2/kg).
G	Channel utilization in ALOHA.
G_l	Gain (less than unity) of a lossy device.
G_p	Spread-spectrum processing gain.
G_r	Receiving antenna gain.
G_t	Transmitting antenna gain.
$G_u(f)$	Power spectral density of random sequence of rectangular pulses.
G_D	Diversity gain.
$(G/T)_E$	Figure of merit of an earth station.
$(G/T)_S$	Figure of merit of a transponder's uplink antenna and receiver.
$G(\theta)$	Antenna gain as a function of angle θ.
h	Quantity of traffic (total holding time).
$\mathbf{h}$	Orbital angular momentum.
h_0	Earth station elevation above sea level in kilometers.
h_R	Rain height used in CCIR rain-attenuation model.
H	Parity check matrix.
H	Magnetic field.
H_e	Effective height of the top of a rainstorm in kilometers.
H_i	Height of zero-degree isotherm in kilometers.
H_0	Earth station elevation above sea level in kilometers.
H^T	Transpose of parity check matrix.
i	Orbital inclination.
I	Identity matrix
I	In-phase component of a TV chrominance signal.
I	Number of received incorrect bits.
I_D	Diversity improvement.
J_n	Bessel function of order n.
J_n	Legendre polynomial of order n.
k	Boltzmann's constant, 1.38×10^{-23} JK^{-1}.
k	Propagation constant, $2\pi/\lambda$.
k	Number of message bits in a codeword.
k	Number of stages in a shift register.
$\mathbf{K}$	Surface current density.
K	Transfer characteristic of FM demodulator.
l	Loading factor.
l_A	West longitude of point A.
l_B	West longitude of point B.
l_e	West longitude of earth station.

l_s	West longitude of the subsatellite point.
L	Mean time duration of a speech spurt.
L	Path length in kilometers.
L_a	Loss due to atmospheric effects.
L_A	North latitude of point A.
L_B	North latitude of point B.
L_e	North latitude of earth station.
L_{eff}	Effective path length in kilometers.
L_G	Horizontal projection of the portion of a satellite path length that is in rain.
L_p	Total downlink path loss.
L_{ra}	Loss in receiving antenna system.
L_s	North latitude of the subsatellite point.
L_S	Path length in rain in CCIR attenuation model.
m	FM modulation index.
m	Mean time before failure.
M	Orbital mean anomaly.
M_E	Mass of the earth.
M_r	Required fade margin for link reliability r.
n	Number of bits in a codeword.
n	The number of PN sequences that a shift register can generate.
$\mathbf{n}$	Unit vector normal to reflector surface.
N	Number of channels carried by a multiplexed telephone system.
N	Received noise power.
N	Channel number in SPADE-unprimed numbering system.
N	Mean number of times that ALOHA packets must be retransmitted.
N	Electron density in electrons/m^3.
N	Number of chips used to spread each bit in a PN spread-spectrum system.
N'	Channel number in SPADE-primed numbering system.
N_b	Number of bits required to transmit an alphabet.
$N(D)$	Number of raindrops.
N_f	Number of failed components in a life test.
N_s	Number of components at start of life test.
N_0	Noise power density.
N_0	*A constant describing the size distribution of raindrops.*
N_0	Number of surviving components in a life test.
p	Psophometric weighting factor expressed as a ratio.
p	Semilatus rectum of an orbit.
p	Power level in watts.
p	Bit error probability.
$p_r(t)$	Output voltage of a Nyquist filter when excited by a unit impulse.
$p(v)$	Probability that an instantaneous voltage has value v.
$p(x)$	Probability that a new call will be made in time interval dt.
p_L	Power delivered to a load.
p_S	Power supplied by a source.
P	Psophometric weighting factor, 2.5 dB.
P	Percent of time.
P	Number of bits in a preamble.
P	Probability of all combinations of I errors and $N - I$ errors in N consecutive bits.

P	Probability of failure.
P	Probability that a unique word will be received correctly.
P_n	Noise power.
P_r	Received power.
$P_{r_{clear\_weather}}$	Received power under clear-weather conditions.
$P_r(f)$	Amplitude response of a Nyquist filter.
P_t	Transmitted power.
P_0	Total power.
PB	Bit error probability.
PE	Probability of a symbol error.
$P(C, A)$	Probability that the number of required channels will exceed the number of available channels when the traffic intensity is A.
$P_J(A)$	Probability that the joint attenuation exceeds A.
P_S	Average power in a voice channel before compression.
$P_S(A)$	Probability that single-site attenuation exceeds A.
$P(\theta)$	Power in a direction θ.
Q	Weighting factor in video signal-to-noise ratio calculation.
Q	Probability of a miss (failure to recognize a correct unique word).
Q	Probability of failure.
r	Link reliability, the percentage of time that a link performs satisfactorily.
r	Distance between two points.
r	Number of parity bits in a codeword.
r	Radial distance in a circular aperture.
r_e	Radius of earth, 6370 km.
(r_o, ϕ_o)	Satellite polar coordinates in the orbital plane.
r_p	Factor used in CCIR rain-attenuation model to account for spatial non-uniformity of rain.
r_s	Orbital radius.
r_s	Symbol rate.
r_0	Radius of a circular aperture.
R	Distance from a satellite to an earth station.
R	Red signal in color TV transmission.
R	Rainfall rate in millimeters per hour.
R	Transmission bit rate.
R	Distance from the center of an antenna aperture to an observation point.
R	Reliability.
RA	Right ascension.
R_b	Bit rate.
R_c	Chip rate of a PN spread-spectrum system.
R_L	Load resistance.
$R(P)$	Rainfall rate exceeded for P percent of the time.
R_s	Symbol rate.
R_S	Internal resistance of source.
R_T	Total resistance in circuit.
s	Half-perimeter of spherical triangle used in look angle calculations.
S	Demodulated signal power.
S	Error syndrome of a block code.
$S(f)$	Spectrum of a general pulse.
(S/I)	Voice channel signal to co-polarized interference ratio.

$(S/N)_D$	The value that the downlink signal-to-noise ratio would have if the transponder did not retransmit incoming noise.
$(S/N)_i$	Overall signal-to-noise ratio of a CSSB link.
$(S/N)_I$	Term introduced into (S/N) equations to account for the effects of intermodulation noise.
$(S/N)_o$	Signal-to-noise ratio at output of FM demodulator.
$(S/N)_U$	Uplink signal-to-noise ratio.
$(S/N)_V$	Weighted video signal-to-noise ratio.
$(S/N)_{wc}$	Signal-to-noise ratio in worst channel at demultiplexer output.
(S/X)	Signal to cross-polarized interference ratio.
t	Time.
t_1	Time at which satellite position is to be determined.
t_o	Time at which satellite position is known.
t_p	Time of perigee.
t_x	The total time that x out of N channels are in use during time interval T.
T	Noise temperature.
T	Orbital period.
T	Reference time period for traffic calculations.
T_A	Antenna noise temperature.
T_b	Duration of a rectangular baseband pulse.
T_b	Increase in antenna noise temperature in Kelvins caused by rain.
T_c	Elapsed time in Julian centuries between 0 hours UT Julian day JD and noon UT on January 1, 1900.
T_e	Elapsed time since the x_r axis coincided with the x_i axis.
T_F	Frame period.
T_l	Noise temperature of a lossy device.
T_n	Noise temperature of a device.
T_p	Physical temperature.
T_s	Duration of one symbol.
T_s	System noise temperature.
(T_T/N)	Test-tone-to-noise ratio.
u	$1/r_o$ where r_o is the satellite radial coordinate in the orbital plane.
U	Coefficient used in calculating XPD from attenuation.
v	Instantaneous value of voltage.
v	Orbital linear velocity.
V	Coefficient used in calculating XPD from attenuation.
V	Bit rate for a single voice channel.
$V(D)$	Terminal velocity of raindrop with diameter D.
$V(f)$	Voltage spectral density.
V_P	Phase velocity of light in free space.
V_T	Component of transmitter velocity directed toward the receiver.
w	Preemphasis improvement factor expressed as a ratio.
w_b	Bandwidth required to transmit symbols at bit rate R_b.
W	Preemphasis improvement factor in decibels.
W	Bandwidth required for a spread-spectrum system with chip rate R_c.
$W(f)$	Power spectral density of an FDM/FM signal at RF.
x	Distance in aperture plane.
x	Number of channels in use.
(x_i, y_i, z_i)	Satellite rectangular coordinates in the geocentric equatorial system.

(x_o, y_o, z_o)	Satellite rectangular coordinates in the orbital plane.
xpd_V	Numerical cross-polarization discrimination on V-polarization channel.
xpi_V	Numerical cross-polarization isolation on V-polarization channel.
(x_r, y_r, z_r)	Rotating rectangular coordinate system.
X	Vertex angle in look angle calculation.
X	Decibel increase in average power provided by a compressor.
XP	Logical variable used in unique word identification.
XPD_V	Decibel cross-polarization discrimination when V-polarized waves are transmitted.
XPI_V	Decibel cross-polarization isolation on V-polarized channel.
XQ	Logical variable used in unique word identification.
y	Distance in aperture plane.
Y	Luminance signal of a television transmission.
Y	Vertex angle in look angle calculation.
Z	Path length in meters.
α	Vertex angle used in look angle calculation.
α	Rolloff factor of a Nyquist filter.
$\alpha_{g,o}$	Right ascension of the Greenwich meridian at 0 hours UT.
γ	Central angle between subsatellite point and earth station.
γ	Empirical constant used in the simple attenuation model.
$\delta(t)$	Unit impulse function.
Δf	Frequency deviation in hertz.
Δf	Frequency shift.
Δf_{rms}	rms test-tone deviation.
ΔN	Increase in received thermal noise power resulting from rain.
$\Delta\phi$	Change in polarization angle of a linearly polarized wave.
η	Average orbital angular velocity.
η	Single-sided rms noise power spectral density.
η	Aperture efficiency.
θ	Angle between antenna axis and a line to an observation point.
λ	Wavelength in meters.
λ	Packet transmission rate in ALOHA.
λ	Reciprocal of MTBF.
λ'	Packet reception rate in ALOHA.
Λ_e	Earth station latitude.
μ	Kepler's constant (3.9861352×10^5 km^3/s^2).
σ	rms value of a signal.
τ_L	Telephone load activity factor.
ϕ	Arbitrary phase angle.
ϕ	Angle between received electric field and local horizontal.
ϕ	Rotation angle in Faraday rotation.
ϕ	Angle between geomagnetic field and direction of propagation.
ω	Argument of perigee.
ω	Radian frequency.
ω_c	Radian carrier frequency.
ω_{mod}	Radian modulating frequency.
Ω	Right ascension of the ascending node.
Ω_e	Orbital angular velocity of the rotating coordinate system.

INDEX