Technologies and Services on Digital Broadcasting (8)

MODULATION Systems (part 1)

In the past, digital transmission of video and audio has required a broader frequency bandwidth than analog transmission to transmit the same information. Recent progress in source coding technology for video and audio, however, has made it possible to reduce the bit rate while keeping the quality deterioration to a minimum. The frequency bandwidth required for digital transmission has consequently become about the same as or several fractions that of analog transmission, making digital transmission feasible for broadcasting.

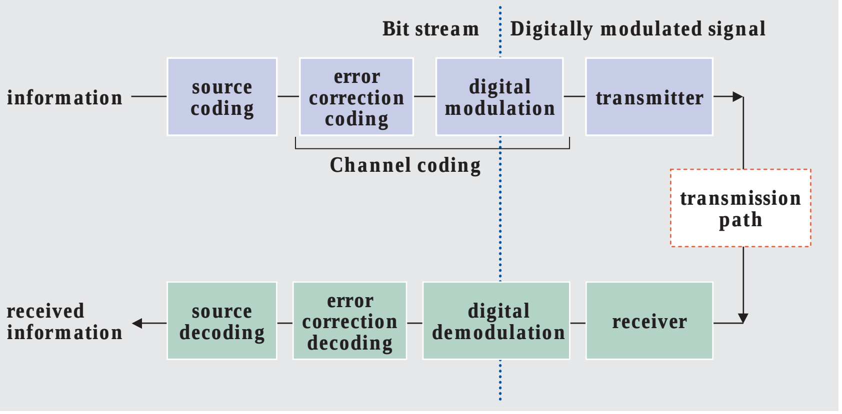


Figure 1: Digital transmission system

Until recently, digital transmission systems for broadcasting have consisted mainly of multiplex broadcasting systems transmitting only audio and text data. Today, digital transmission including video services is possible in the form of satellite and terrestrial digital broadcasting. In this issue, we describe the basics of digital modulation systems.

Figure 1 shows a digital transmission system. The system quantizes and samples analog information such as video or audio data and performs source coding on that data for the purpose of band compression. It then performs channel coding, which consists of adding error correction code to correct the errors generated during transmission, as well as data interleaving and digital modulation. The system converts the digitally modulated signal into one in the transmission frequency band, amplifies that signal at the transmitter, and transmits it by either radio wave or cable.

At the receiver, the system performs inverse processing with respect to the transmission side to restore the original information; that is, it digitally demodulates the received signal, performs de-interleaving and error correction, etc.

1. Digital baseband signal

1.1 Basic code

A data stream consisting of 0's and 1's can be transmitted as a pulse train generated, for example, by the voltage of an electrical signal. Basic binary codes for such a pulse train are shown in Figure 2.

Referring to the figure, "unipolar" means that the input data 0 and 1 correspond to voltages of 0 and +1, respectively, while "bipolar" means that the input data 0 and 1 correspond to voltages of -1 and +1, respectively. Here, RZ (Return to Zero) results in a narrower pulse width, which means that the frequency bandwidth is wider than in NRZ (Non Return to Zero). For this reason, NRZ is more often used as a baseband signal for digital modulation of radio transmissions.

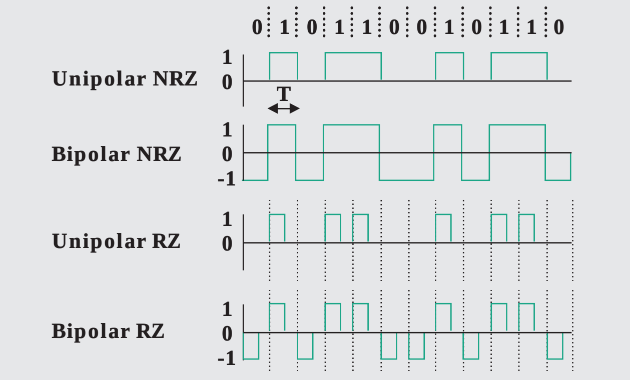


Figure 2: Basic binary codes

1.2 Digital signal spectrum

The power spectrum density (PSD) $S(f)$ of a ± A (Volt) bipolar-NRZ random signal is given by the following expression.

$$S(f) = A^2 T \left(\frac{\sin \pi f T}{\pi f t} \right)^2 \tag{1}$$

Here, T is the NRZ symbol period (sec). A symbol is the minimum unit of code transmission. In this case, one symbol corresponds to one bit. The inverse of the symbol period $1/T$ is called the symbol rate (symbol/sec) or baud rate.

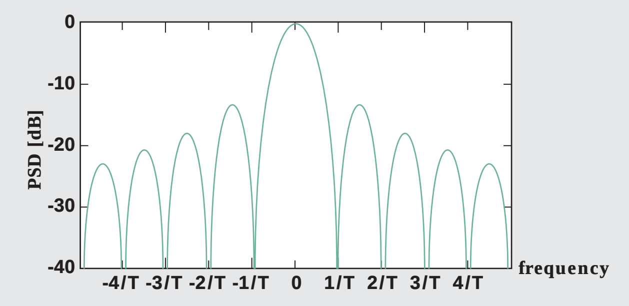


Figure 3: NRZ power spectrum density

The NRZ PSD is shown in Figure 3. The vertical axis is normalized in terms of A^2T. Null points corresponding to zero power appear at frequencies corresponding to various symbol rates for the case of an infinitely large bandwidth.

1.3 Transmission path noise

Noise on a transmission path or in a receiver can be broadly classified into thermal noise, generated from resistors, and artificial noise, from sources such as automobiles that reach the receiving antenna. For relatively clean links like satellite channels and cables, the noise on the transmission path can be considered to be mostly thermal. As shown in Figure 4, the temporal waveform of thermal noise is random in nature; its amplitude distribution follows a Gaussian distribution (normal distribution), whose probability distribution function (PDF) is shown in Figure 5. The average amplitude m of thermal noise is 0. Given a noise power of σ^2, $P(x)$ is the probability that the thermal noise will take on a certain amplitude.

The PSD of thermal noise $N(f)$ is given as follows.

$$N(f) = \frac{N_0}{2} \quad (-\infty < f < \infty) \tag{2}$$

The PSD of thermal noise is thus flat over all frequencies, a condition referred to as "white noise."

The above temporal and frequency characteristics are reflected by thermal noise often being called "white Gaussian noise."

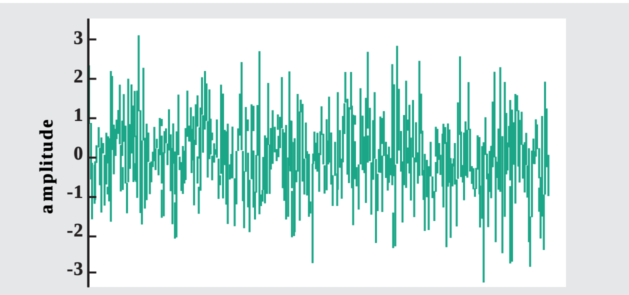
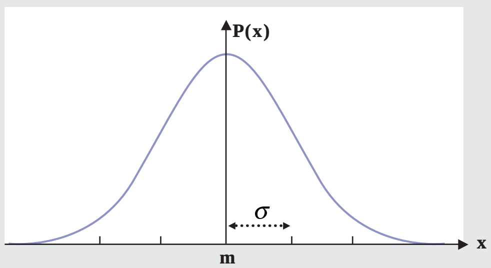


Figure 4: Temporal waveform of thermal noise (σ =1)



$$P_{(x)} = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}$$

When mean = 0, $P_{(x)} = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$

Figure 5: Gaussian probability distribution function

1.4 Errors on a transmission path having noise and distortion

At the receiver, transmitted digital data are judged to be 1 or 0 based on the signal level, which is greater or less than a threshold level, at symbol sample points. The system disregards times other than the sample points, regardless of the type of waveform. Rectangular pulses, such as the NRZ and RZ pulses described earlier, have large bandwidths (infinitely large in theory). Therefore, to make efficient use of finite frequency resources, appropriate band limitations are applied to the transmission, resulting in the smooth waveform shown in Figure 6(b).

If non-linear distortion and reflection occur on the transmission path, prior and subsequent symbols with respect to the symbol in question will be affected and signal levels at sample points will vary. This distortion is called Inter Symbol Interference (ISI). If the ISI is large, the margin of the threshold level with respect to noise drops, and this lesser margin becomes a source of error. If noise is added to the signal, decision levels may exceed the threshold level and errors may be generated as shown in Figure 6(d).

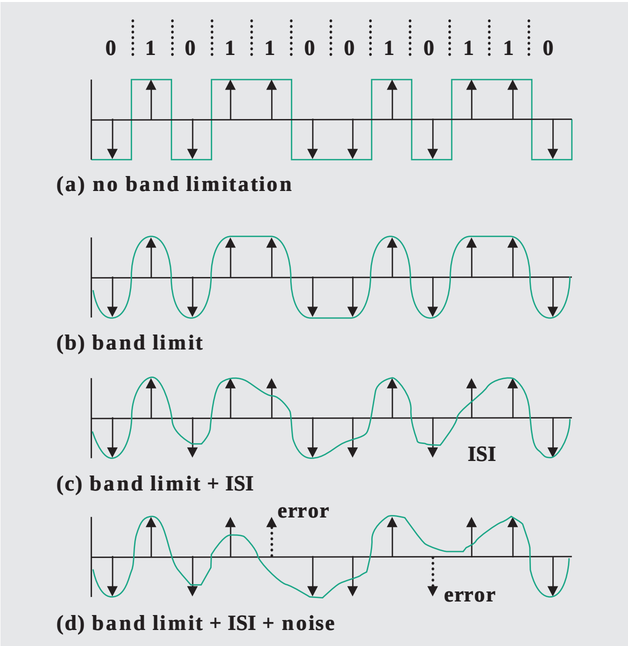


Figure 6: Transmission path having distortion and noise

Transmission errors are evaluated in terms of error rate. For example, one erroneous bit in a transmission of 1000 bits of information corresponds to a bit error rate (BER) of 10^{-3}.

We consider errors caused by noise. If we were to superpose Gaussian noise of variance (noise power) σ^2 (Figure-5) on signals of amplitudes -A and +A corresponding to codes 0 and 1, the amplitude distribution at a sample point would be expressed as $P_0(x)$ and $P_1(x)$ as shown in Figure 7.

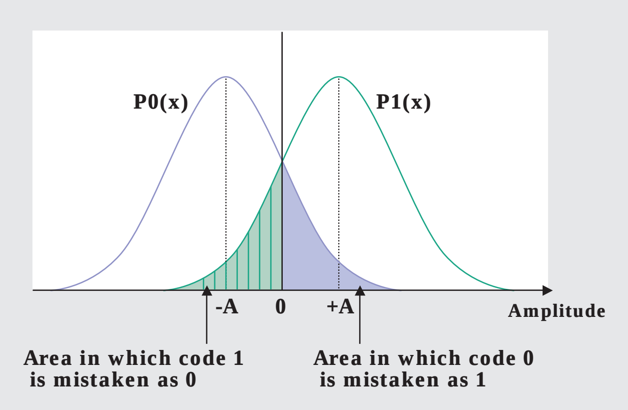


Figure 7: Error caused by Gaussian noise

The error rate *Pe* is the areal percentage of that part exceeding decision level 0. It is computed as follows.

Pe = (probability that code 1 is sent) × (probability that code 1 is mistaken as 0)
+ (probability that code 0 is sent) × (probability that code 0 is mistaken as 1)

Given that the probability of sending code 0 and that of code 1 are both 1/2, we can rewrite the above as follows.

$$P_e = \frac{1}{2} \times \int_{-\infty}^{0} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-A)^2/2\sigma^2} dx + \frac{1}{2} \times \int_{0}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x+A)^2/2\sigma^2} dx$$
$$= \frac{1}{\sqrt{\pi}} \times \int_{A/\sqrt{2}\sigma}^{\infty} e^{-y^2} dy$$
$$= \frac{1}{2} erfc\left(\frac{A}{\sqrt{2}\sigma}\right) \qquad (3)$$

Here, $erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt$ is called the co-error function.

1.5 Condition for undistorted transmission

If band limitations are applied to an NRZ rectangular pulse, ISI will occur in the waveform. Considering, however, that the absence of disturbances from other samples is desirable at a sample point, we note that ISI will not occur (ISI-free transmission) if, for any pulse, the amplitude at the sampling point is non-zero while the amplitude at all other sample points for that pulse is zero. This condition is called the Nyquist criterion.

A band-limited filter that satisfies the Nyquist criterion and its impulse response are shown in Figure 8. Here, the frequency characteristics $H(f)$ are given as

$$H_{(f)} = \begin{cases} 1 & 0 \le f \le \frac{1-\alpha}{2T_s} \\ \cos^2\left\{\frac{T_s}{4\alpha}\left[2\pi f - \frac{\pi(1-\alpha)}{T_s}\right]\right\} & \frac{1-\alpha}{2T_s} \le f \le \frac{1+\alpha}{2T_s} \\ 0 & \frac{1+\alpha}{2T_s} \le f \end{cases} \qquad (4)$$

The above equation can be referred to as a "roll-off filter," where α is the "roll-off rate." The baseband bandwidth B_b is

$$B_b = \frac{1+\alpha}{2T_s} = \frac{(1+\alpha)}{2} f_s \qquad (5)$$

In the above equation, T_s is symbol period and f_s (=1/ T_s) is symbol rate. As α decreases, the bandwidth of the pulse series becomes smaller, and jitter (variance at the sampling point) can easily occur. As α increases, the bandwidth becomes larger. The minimum bandwidth corresponding to α = 0 is called the Nyquist bandwidth.

For example, the following baseband bandwidth is needed when transmitting a 10-Mbps NRZ binary signal at a roll-off rate of 0.5.

B_b = (1 + 0.5) / 2 × 10 Mbps / (1 bit/symbol) = 7.5 MHz

The S/N of the received pulse becomes a maximum on a linear transmission channel if the frequency response of the transmitter and receiver filters behave as if they share a

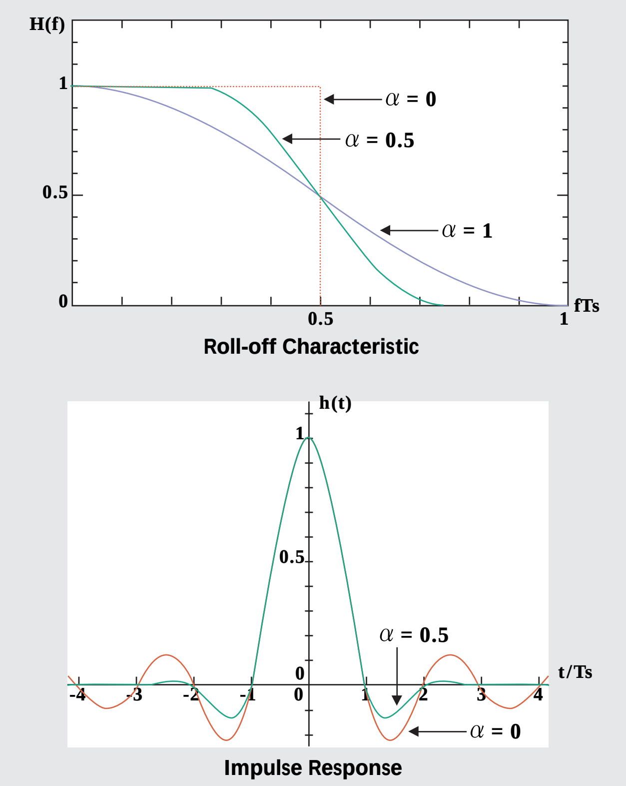


Figure 8: Roll-off characteristics and pulse that satisfies the Nyquist criterion

roll-off filter in the transmission line. It is therefore common to subject the roll-off characteristics to root allocation in the transmitter and receiver as shown in Figure 9. This is called "root roll-off."

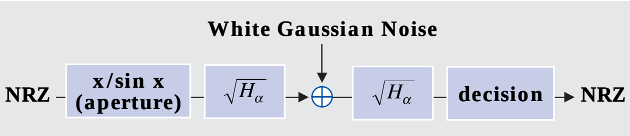


Figure 9: Root roll-off filters in a transmission system

It should be emphasized that the above roll-off characteristics produce an undistorted transmission for an input signal having an ideal impulse. However, as an actual transmission pulse has a width, frequency compensation using reverse characteristics of that spectrum results in a flat spectrum, like the frequency response of an impulse. This is called aperture equalization, and it can be expressed as follows in the case of NRZ.

$$A(f) = \frac{\pi f T_s}{\sin(\pi f T_s)} \quad (6)$$

Aperture equalization is usually done together with roll-off filtering. An example of roll-off characteristics that include NRZ aperture equalization is shown in Figure 10.

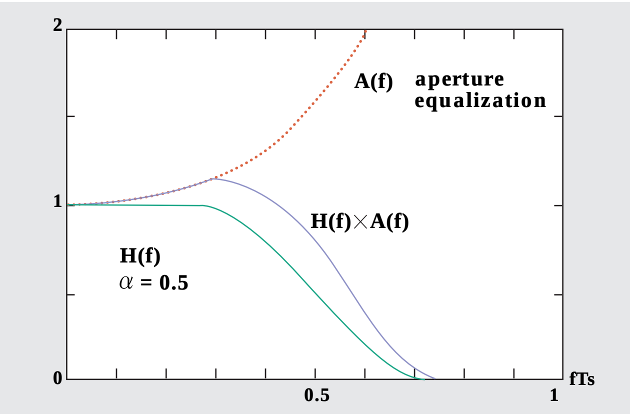


Figure 10: Roll-off characteristics including aperture equalization

The received waveform is in a disturbed state, due to the effects of distortion and noise in the transmission path. This state can be observed in the form of eye patterns. Specifically, if the received signal is displayed on an oscilloscope overlaid by the symbol period, blank areas become visible in the center of the resulting output. These are called eye patterns because of their resemblance to eyes. The size of a single eye here can be used to visually assess the extent of waveform distortion. For example, a "closed" eye means that 1's and 0's cannot be distinguished and the code error is large. Note that there is one eye in a single eye pattern for a binary signal and M-1 eyes for an M-dimensional signal. Figure 11 shows examples of eye patterns.

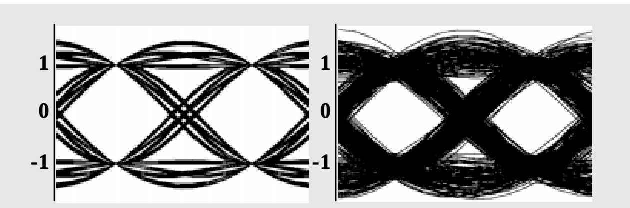


Figure 11: Examples of eye patterns
(α = 0.5; left: no noise, right: with noise)

2. Digital modulation systems

As in analog modulation, the function of "modulation" in digital modulation is to convert the original information signal (baseband signal) into one with a frequency convenient for transmission and to vary the amplitude, frequency, or phase of the carrier. What is different from analog modulation is that the baseband signal in question may be an analog signal or a digital signal.

Varying the amplitude of the carrier (as in AM) according to the 1's and 0's in the baseband signal is called amplitude shift keying (ASK), varying the frequency of the carrier (as in FM) is called frequency shift keying (FSK), and varying the phase is called phase shift keying (PSK). Varying both amplitude and phase, moreover, is called quadrature amplitude modulation (QAM). We here describe several basic examples of such systems, namely, BPSK, QPSK (4PSK), and 16QAM.

2.1 BPSK (Binary Phase Shift Keying)

The most basic form of PSK is BPSK. As shown in Figure 12, BPSK transmits a binary (0 and 1) digital signal by having a carrier phase of π correspond to 0. The configuration of the BPSK modulator is shown in Figure 13. BPSK is a system that sends one bit of information per symbol. Because information is conveyed by phase, BPSK is robust against level fluctuation on the transmission path.

As shown in Figure 14, the BPSK spectrum is equivalent to that of the baseband signal, although it has been shifted (minus side included) to the carrier frequency f_c . The bandwidth of BPSK is therefore twice that of the baseband signal. When inserting a roll-off filter, the BPSK bandwidth for the symbol rate r (=1/T) and roll-off rate α is (1+α)r.

Demodulation by coherent detection simply reverses the process of modulation, as shown in Figure 15. The carrier recovery circuit removes the modulated component and noise component from the received signal and recovers the carrier. The system then multiplies the received signal by this carrier to obtain the baseband signal, which is then input to a clock recovery circuit to recover the same clock as that on the modulator side. With this clock, the system samples the baseband signal, decides whether the symbols are 0 or 1, and restores the digital signal.

On the other hand, demodulation by incoherent detection treats the received signal delayed by one symbol as the reference phase and multiplies the received signal by

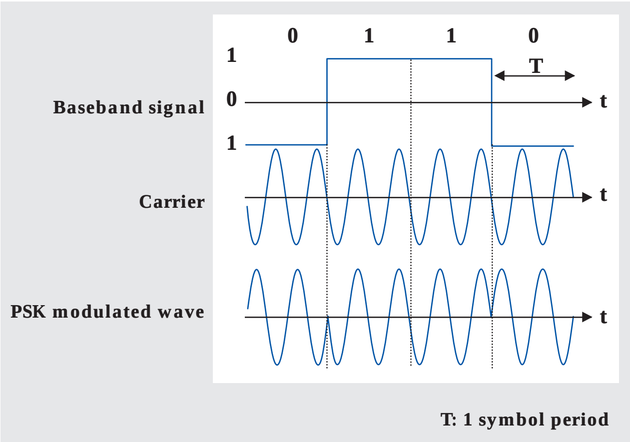


Figure 12: BPSK modulated waveform

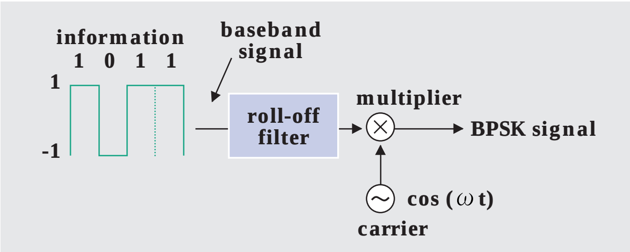


Figure 13: BPSK modulation

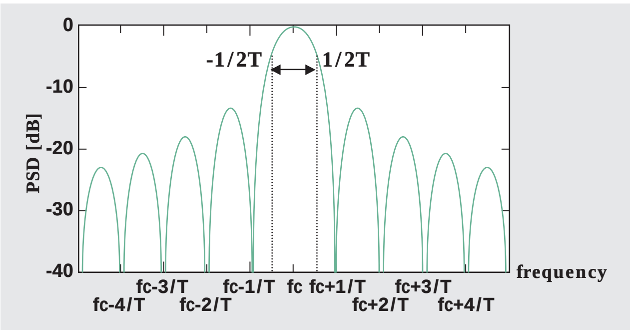


Figure 14: BPSK spectrum

it to obtain the baseband signal, as shown in Figure 16. Incoherent detection is frequently used in terrestrial mobile transmissions since large fluctuations in amplitude due to fading effects make it difficult to recover the carrier. Here, however, because the signal itself, which includes noise and distortion, is used as the reference phase in incoherent detection, the bit error characteristics are worse than those of coherent detection. In addition, as the information is transmitted as a phase difference with respect to the previous symbol, differential conversion of code must be performed on the transmit side before modulation.

BPSK symbol error rate assuming coherent detection of a BPSK signal to which noise has been added to the transmission path can expressed using Eq. (3), since the amplitude probability for the demodulated signal follows the Gaussian distribution shown in Figure 5.

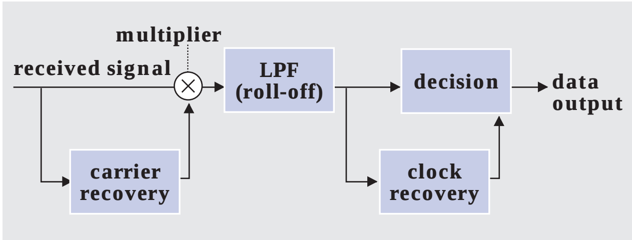


Figure 15: BPSK demodulation by coherent detection

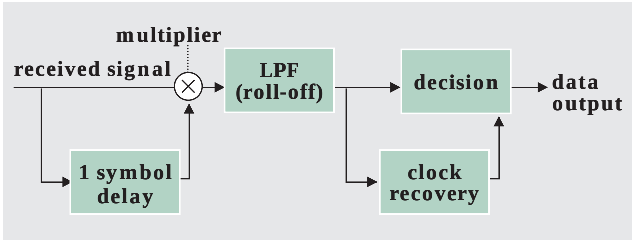


Figure 16: BPSK demodulation by incoherent detection

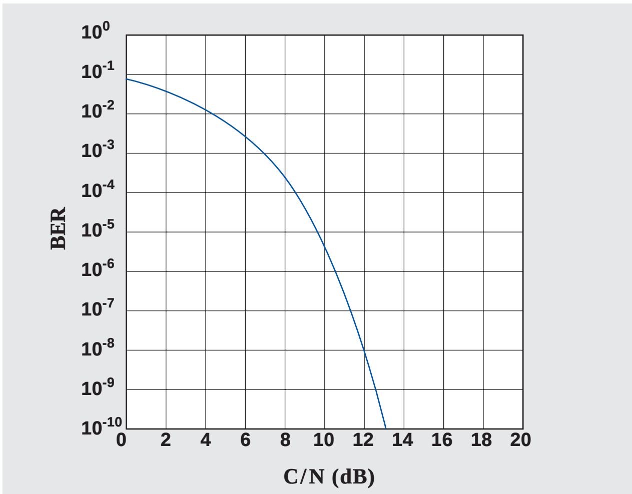


Figure 17: BPSK bit error rate (coherent detection; the noise bandwidth is the Nyquist bandwidth)

$$P_s = \frac{1}{\sqrt{\pi}} \times \int_{A/\sqrt{2}\sigma}^{\infty} e^{-y^2} dy = \frac{1}{2} erfc\left(\frac{A}{\sqrt{2}\sigma}\right) \tag{7}$$

Since signal power C=$A^2/2$ and noise power N=σ^2, the above equation can be rewritten as

$$P_s = \frac{1}{2} erfc\left(\frac{A}{\sqrt{2}\sigma}\right) = \frac{1}{2} erfc\left(\sqrt{\frac{C}{N}}\right) \tag{8}$$

Because one bit corresponds to one symbol in BPSK, the above symbol error rate is the same as the bit error rate P_b. This is shown in Figure 17.

The CN ratio specifies the noise bandwidth, which is usually taken to be the Nyquist bandwidth. The noise bandwidth of a signal with symbol rate *r* (symbol/sec) is *r* (Hz).

2.2 QPSK (Quadrature Phase Shift Keying)

The QPSK system uses carrier phases at 90-degree intervals to send two bits of information per symbol, and for this reason, it is sometimes called 4-phase PSK. It is very efficient in terms of required bandwidth and power and is robust against non-linear distortion on the transmission path. The QPSK system has consequently found wide use in satellite communications in which the CN ratio of the received signal is small.

The configuration of the QPSK modulator is shown in Figure 18. The modulator performs serial-parallel conversion of the input signal, divides the resulting signal into I and Q signals, and performs BPSK modulation on both of the $\cos 2\pi f_c t$ and $\sin 2\pi f_c t$ orthogonal carriers. The system then combines these two modulated signals and outputs a QPSK modulated wave.

Because QPSK combines the BPSK of two orthogonal carriers, the shape of its spectrum will be the same as that of BPSK (Figure 14), provided that the symbol rate is the same as BPSK. However, as QPSK transmits two bits of information for every symbol, it can transmit twice the amount of information (double the bit rate) of BPSK for the same bandwidth.

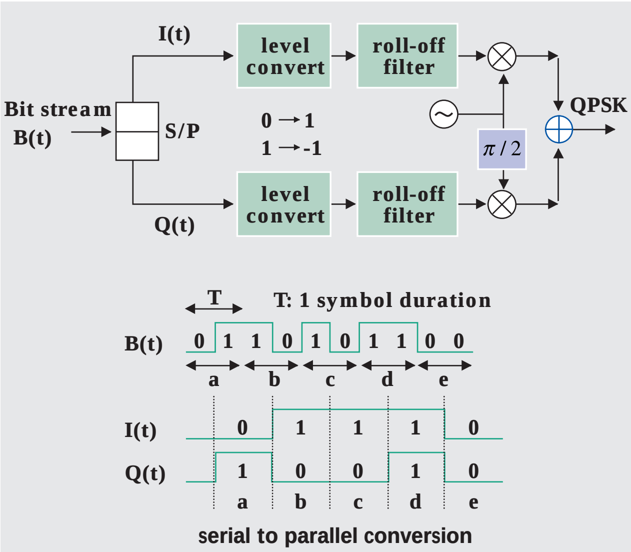


Figure 18: Configuration of the QPSK modulator

A phase map (constellation) in which the horizontal and vertical axes correspond to I and Q signals, respectively, is normally used in digital modulation to represent the relationship between bits transmitted in a symbol and the modulated level or phase. A QPSK constellation is shown in Figure 19. Symbol errors (code errors) occur when noise or distortion cause the signal to enter another quadrant of the phase plane . Symbol errors caused by Gaussian noise can be mostly regarded as errors involving adjacent signal areas. For this reason, multi-dimensional modulations like QPSK employ gray code assignment, whereby adjacent signal points on the phase plane each differ by one bit. In this way, the bit error can be minimized for the symbol error .

As in the case of BPSK, QPSK demodulation can be performed via either coherent detection (Figure 20), which

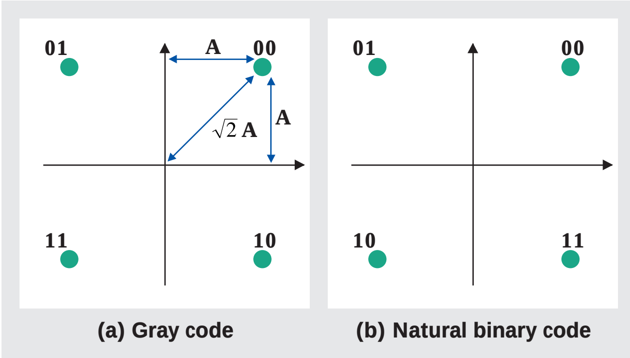


Figure 19: QPSK gray-code-assignment constellation

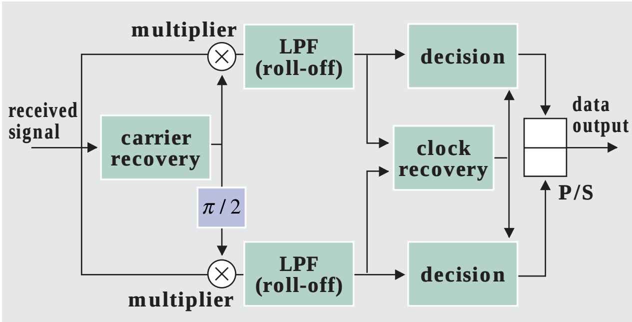


Figure 20: QPSK demodulation by coherent detection

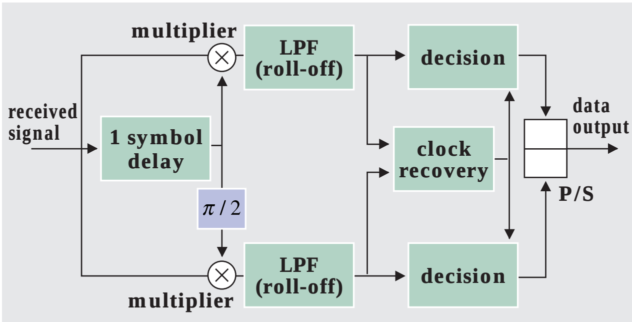


Figure 21: QPSK demodulation by incoherent detection

recovers the carrier in the received signal, or incoherent detection (Figure 21), which does not require carrier recovery.

Denoting the amplitude and phase of the modulated signal by $A(t)$ and $\varphi(t)$, respectively, the transmit signal $s(t)$ is given as follows.

$$s(t) = A(t)\cos\left[2\pi f_c t + \varphi(t)\right]$$
$$= A(t)\cos(\varphi(t))\cos(2\pi f_c t) - A(t)\sin(\varphi(t))\sin(2\pi f_c t) \qquad (9)$$

For BPSK, $\varphi(t) = \pm\pi$, while for QPSK, $\varphi(t) = \pm\pi/4, \pm 3\pi/4$.

In addition, noise added to the transmit signal is expressed as bandpass noise in the following way.

$$n(t) = n_i(t)\cos(2\pi f_c t) - n_q(t)\sin(2\pi f_c t) \qquad (10)$$

In coherent detection, the received signal is detected by using the recovered carrier $L_I(t) = \cos(2\pi f_c t)$ at the receiver, and the I-axis $r_I(t)$ becomes as follows.

$$r_I(t)=[s(t)+n(t)]L_I(t)$$
$$=A(t)\cos(\varphi(t))\cos(2\pi f_c t)\cos(2\pi f_c t)$$
$$-A(t)\sin(\varphi(t))\sin(2\pi f_c t)\sin(2\pi f_c t)$$
$$+n_i(t)\cos(2\pi f_c t)\cos(2\pi f_c t)-n_q(t)\sin(2\pi f_c t)\cos(2\pi f_c t)$$
$$=A(t)\cos(\varphi(t))\frac{[\cos(4\pi f_c t)+1]}{2}-A(t)\sin(\varphi(t))\frac{[\sin(4\pi f_c t)]}{2}$$
$$+n_i(t)\frac{[\cos(4\pi f_c t)+1]}{2}-n_q(t)\frac{[\sin(4\pi f_c t)]}{2} \quad (11)$$

The received signal is also detected by using the recovered carrier $L_Q(t)=-\sin(2\pi f_c t)$, and the Q-axis $r_Q(t)$ becomes as follows.

$$r_Q(t)=[s(t)+n(t)]L_Q(t)$$
$$=-A(t)\cos(\varphi(t))\cos(2\pi f_c t)\sin(2\pi f_c t)$$
$$+A(t)\sin(\varphi(t))\sin(2\pi f_c t)\sin(2\pi f_c t)$$
$$-n_i(t)\cos(2\pi f_c t)\sin(2\pi f_c t)+n_q(t)\sin(2\pi f_c t)\sin(2\pi f_c t)$$
$$=-A(t)\cos(\varphi(t))\frac{\sin(4\pi f_c t)}{2}+A(t)\sin(\varphi(t))\frac{[1-\cos(4\pi f_c t)]}{2}$$
$$-n_i(t)\frac{\sin(4\pi f_c t)}{2}+n_q(t)\frac{[1-\cos(4\pi f_c t)]}{2} \quad (12)$$

After coherent detection, the I-axis and Q-axis signals, $r_I(t)$ and $r_Q(t)$, are passed through low-pass filters (LPFs) to generate baseband signals only, and the transmitted amplitude and phase, $A(t)$ and φ (t), are then recovered from the post-LPF $r_I(t)$ and $r_Q(t)$.

$$r_I(t)=\frac{A(t)}{2}\cos(\varphi(t))+\frac{n_i(t)}{2} \quad (13)$$

$$r_Q(t)=\frac{A(t)}{2}\sin(\varphi(t))+\frac{n_q(t)}{2} \quad (14)$$

In incoherent detection, the received signal delayed by one symbol is multiplied by the current signal, and the I-axis signal $r_I(t)$ becomes as follows.

$$r_I(t)=[s(t)+n(t)][s(t-T)+n(t-T)]$$
$$=s(t)s(t-T)+n(t)s(t-T)+s(t)n(t-T)+n(t)n(t-T) \quad (15)$$

If we extract only the signal component here, we get

$$r_I(t)\cong s(t)s(t-T)$$
$$=A(t)\cos[2\pi f_c t+\varphi(t)]A(t-T)\cos[2\pi f_c(t-T)+\varphi(t-T)]$$
$$=\frac{A(t)A(t-T)}{2}\{\cos[2\pi f_c T+\varphi(t)-\varphi(t-T)]$$
$$+\cos[2(2\pi f_c t)-2\pi f_c T+\varphi(t)+\varphi(t-T)]\} \quad (16)$$

Then, after passing the above through an LPF and extracting only the baseband component, $r_I(t)$ can be approximated as

$$r_I(t)\cong\frac{1}{2}A(t)A(t-T)\cos[2\pi f_c T+\varphi(t)-\varphi(t-T)] \quad (17)$$

Next, with respect to a sample of $2\pi f_c T=2\pi n$ for each symbol, $r_I(t)$ takes on the following form.

$$r_I(t)=\frac{1}{2}A(t)A(t-T)\cos[\varphi(t)-\varphi(t-T)] \quad (18)$$

The phase difference for each symbol on the I-axis is thus demodulated as information.

Incoherent detection of the orthogonal axis (Q-axis) in an orthogonal modulation system such as QPSK results in the following expression for $r_Q(t)$, after shifting the phase of the signal delayed by one symbol by $\pi/2$ and multiplying the result by the current signal.

$$r_Q(t)=[s(t)+n(t)][s(t-T+\pi/2)+n(t-T+\pi/2)] \quad (19)$$

After passing the above through an LPF, we get

$$r_Q(t)\cong\frac{1}{2}A(t)A(t-T)\cos[2\pi f_c T-\pi/2+\varphi(t)-\varphi(t-T)] \quad (20)$$

Then, with respect to a sample of $2\pi f_c T=2\pi n$ for each symbol, $r_Q(t)$ becomes.

$$r_Q(t)=\frac{1}{2}A(t)A(t-T)\sin[\varphi(t)-\varphi(t-T)] \quad (21)$$

The phase difference for each symbol on the Q-axis is thus demodulated as information.

We next consider the QPSK symbol error rate and bit error rate. For coherent detection, QPSK can be viewed as BPSK performed independently on the orthogonal I and Q signals. The QPSK symbol error rate P_{QPSK} is expressed as follows, where the error rates of the I and Q signals are denoted by P_I and P_Q.

$$P_{QPSK}=P_I(1-P_Q)+(1-P_I)P_Q+P_IP_Q \quad (22)$$

Since P_I and P_Q are each the same as the BPSK symbol error rate P_{BPSK}, the above equation can be rewritten using Eq. (7).

$$P_{QPSK}=P_I+P_Q-P_IP_Q$$
$$=2P_{BPSK}-P_{BPSK}{}^2$$
$$=erfc\left(\frac{A}{\sqrt{2}\sigma}\right)-\left[\frac{1}{2}erfc\left(\frac{A}{\sqrt{2}\sigma}\right)\right]^2 \quad (23)$$

Expressing the above in terms of the CN ratio, we get Eq. (24), considering that $C=(\sqrt{2}A)^2/2, N=\sigma^2$ (from Figure 19).

$$P_{QPSK}=erfc\left(\sqrt{\frac{1}{2}\frac{C}{N}}\right)-\left[\frac{1}{2}erfc\left(\sqrt{\frac{1}{2}\frac{C}{N}}\right)\right]^2 \quad (24)$$

Now let us consider the QPSK bit error rate $P_{b\text{-}QPSK}$. In the case of gray code assignment, the 1st and 2nd terms of Eq. (21) correspond to 1-bit error, while the 3rd term corresponds to 2-bit error. Therefore by multiplying the 1st, 2nd, and 3rd terms by 1/2, 1/2, and 2/2, respectively, we get the following expression for $P_{b\text{-}QPSK}$.

$$P_{b\_QPSK} = \frac{1}{2}P_I(1-P_Q) + \frac{1}{2}(1-P_I)P_Q + \frac{2}{2}P_IP_Q$$
$$= \frac{1}{2}P_I + \frac{1}{2}P_Q$$
$$= P_{BPSK}$$
$$= \frac{1}{2}erfc\left(\frac{A}{\sqrt{2}\sigma}\right)$$
$$= \frac{1}{2}erfc\left(\sqrt{\frac{1}{2}\frac{C}{N}}\right) \qquad (25)$$

In other words, for a bit error rate equal to that of BPSK, the CN ratio of QPSK is larger by 3 dB. However, compared with BPSK at the same bit rate, the symbol rate of QPSK is 1/2 that of BPSK, which means that transmission becomes possible at 1/2 the bandwidth. As a result, the signal power C required to obtain a bit error rate identical to that of BPSK is the same power as in BPSK. The bit error rate of QPSK with respect to the CN ratio is shown in Figure 22.

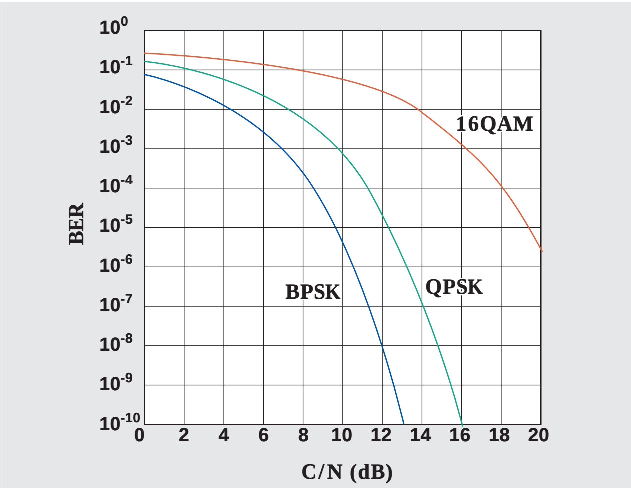


Figure 22: QPSK bit error rate (coherent detection; the noise bandwidth is the Nyquist bandwidth)

2.3 QAM

Reduction of bit rate by using source coding and band compression by using a modulation system are useful when transmitting digital signals over limited bands. In general, the case whereby a signal takes on four or more points of a constellation is called multi-dimensional modulation. A digital modulation that transmits M bits per symbol is called a 2^M-dimensional modulation (for PSK, this becomes 2^M-phase PSK).

The signal obtained by performing orthogonal modulation on two multi-dimensional ASKs is called Quadrature Amplitude Modulation (QAM). The configuration of a 16QAM modulation circuit is shown in Figure 23.

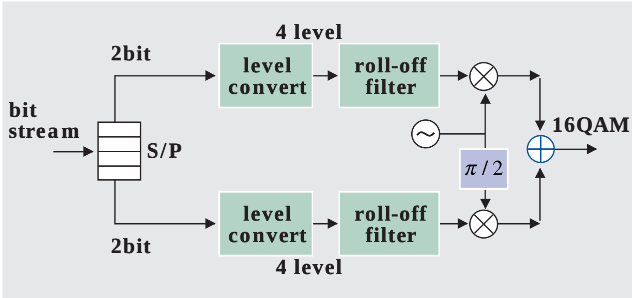


Figure 23: 16QAM modulator configuration

The 16QAM modulation system can transmit four bits of information per symbol. As can be seen from the constellation in Figure 24, however, the distance between each code is small, which means that a large signal power

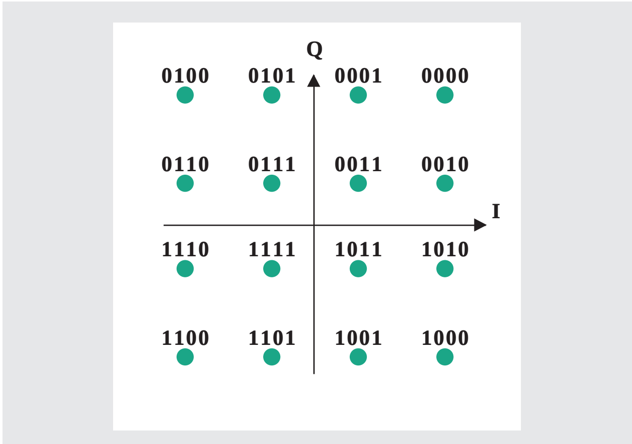


Figure 24: 16QAM constellation

(compared with that of BPSK) is needed to obtain an equivalent error rate. The bit error rate of 16QAM coherent detection is

$$P_b \approx \frac{3}{8}erfc\left(\sqrt{\frac{1}{10}\frac{C}{N}}\right) \qquad (26)$$

Increasing the number of levels on each axis results in higher multi-dimensional modulations such as 64QAM and 256QAM. While frequency-usage efficiency improves with level, a higher receive power is needed to obtain a given bit error rate. The following expressions give the bit error rates for 64QAM and 256QAM.

$$P_b \approx \frac{7}{24}erfc\left(\sqrt{\frac{1}{42}\frac{C}{N}}\right) : 64QAM \qquad (27)$$

$$P_b \approx \frac{15}{64}erfc\left(\sqrt{\frac{1}{170}\frac{C}{N}}\right) : 256QAM \qquad (28)$$

(Shunji Nakahara)