CHAPTER 2
BASEBAND DATA TRANSMISSION
PLL & Partial Response
2.1 THE BASEBAND TRANSMISSION CONCEPT

Baseband transmission implies direct transmission of signals between the information source and sink; when baseband transmission applies, it may be assumed that the information source and the information sink are linked together by a pair of wires, or an electrical equivalent of such direct connection. As explained in Chapter 1, the interconnecting link introduces transmission imperfections, such as frequency-dependent attenuation and inter-symbol interference and also adds noise. In order to fight noise and transmission imperfections, the designer can use appropriately shaped filters. The resulting communication link model is shown in Fig. 2.1.

![Diagram of a communication link model](image)

**FIGURE 2.1: Model for baseband communication link**

The simplest type of baseband communication link uses NRZ coding (see Para. 1.3.1), a pulse is transmitted for binary "one", nothing for binary "zero". The decoder has then to search the received signal for intervals having sufficiently high voltage, and these intervals are assumed to carry binary "ones" to the other intervals, the decoder assigns binary "zeroes".

This signalling scheme pertains to the general class of pulse-amplitude modulation (PAM) systems, since the transmission variable is the amplitude of the signal. A simple extension for the two-level, or binary, system described above is the L-level (or generally, multi-level) system, where L is usually some power of 2. Each level (discrete voltage amplitude) carries a binary value equal to its ordering number (ordering number 0 . . . . 0 being assigned to the lowest level), the decoder then analyzes the received signal, to detect the voltage slice in which the signal resides during each interval, and outputs a binary sequence equal to the ordering number of the nominal level contained in that voltage slice. This is shown graphically in Fig. 2.2.

![Examples of PAM systems](image)

**FIGURE 2.2: Examples of PAM systems**
2.2 REGENERATION OF RECEIVED SIGNAL

Referring to Fig. 2.1, it should be obvious that the signal received by the decoder is both distorted and contaminated by noise, and therefore must be processed by some sort of decision device. This decision device must extract from the received signal the information required to output a sequence of standard digital symbols, compatible with the alphabet of the information sink; this sequence must be a reasonable estimate of the transmitted sequence. This decoder function is analyzed in greater detail in para. 1.13. The decoder must also be capable of estimating the rate of information transfer imposed by the transmitter, or, equivalently, of estimating the clock rate. This requirement is especially stringent when dealing with synchronous communication, where the receiver must possess an accurate replica of the transmitter clock. Since the only information made available to the receiver's decoder circuit is the received signal, the decoder must contain circuits for extracting timing information from this signal.

In the following discussion, reference will be made to binary transmission, however, all results can be extended in an obvious way to include multi-level transmission.

2.2.1 EXTRACTION OF TIMING INFORMATION

NRZ-coded signals are distinguished by a waveform whose transitions are synchronous with the transitions of the transmitter clock signal (Fig. 2.3).

![Diagram of NRZ waveform](image)

**FIGURE 2.3: Principle of pulse regeneration**

Transmitting the NRZ signal through a band-limited medium has the effect of smoothing out the step transitions, thereby blurring the timing information contained in its level transitions: uncertainty is introduced as to where the actual transitions were located. In fact, as shown in Figure 2.3, the transitions, as seen at the receiver, become dependent upon the transmitted sequence, due to inter-symbol interference.
The detection of level transitions is the basic instrument for timing extraction; after having detected the approximate location of transitions, we process this information to generate a clock signal for the receiver.

Transitions are usually detected by comparison of the received signal with a fixed reference voltage, and generation of a short pulse each time the received signal passes through the "threshold" set by the reference voltage. The setting of the reference voltage is seen to influence the position of the timing pulses, so its value must be carefully selected.

The technique described above is almost "natural" for binary waveforms; it is less obvious that the same technique is just as useful for extracting timing information from multi-level waveforms. To see why this is true, let us analyze a four-level system.

In a four-level system, each level (symbol) represents a two-bit group (di-bit), and the rate of symbol transmission is half the rate of bit transmission, the latter being equal, by definition, to the transmitter clock rate.

Assuming a "best case" waveform from the point of view of timing recovery, i.e. \ldots 00110011\ldots then transitions occur every second clock cycle; less favorable cases will result in less frequent transitions, but these transitions will still occur at intervals of \( k/2T \), where \( k \) is an arbitrary, varying, integer and \( 2T \) is the duration of a di-bit which equals two periods of the transmitter clock. As a result, the transition detector generates pulses at a rate half that of the transmitter clock; doubling the rate of pulses will allow regeneration of the original clock rate.

The sub-harmonic relationship indicated in the above example can be proved to hold for all multi-level systems, therefore timing information can always be extracted by means of a transition detector; for example, an eight-level system has an average transition rate of one-third that of the transmitter clock.

2.2.2 OPERATION OF PHASE-LOCK LOOPS

Phase-lock loops (PLL) are circuits which synchronize a local variable oscillator to the phase and frequency of an incoming signal.

![Block diagram for phase-lock-loop circuit](image)

\textit{FIGURE 2.4: Block diagram for phase-lock-loop circuit}

The basic form of a PLL is shown in Fig. 2.4; it consists of a variable oscillator, a phase detector and a filter. The frequency of the oscillator is determined by a DC control voltage; this voltage is obtained by filtering the output voltage of the phase detector. The output voltage of the phase detector is proportional to the phase difference between its two input signals: the input signal and the output signal of the voltage-controlled oscillator (VCO). The polarity and absolute magnitude of the phase detector output are such as to bring the VCO to oscillate at a frequency equal to the input frequency. However, a certain phase difference, possibly quite small, must
remain, for the phase detector to generate the control voltage necessary to bring the VCO to the correct operating frequency. Since the phase detector output voltage range is limited, the range of input frequencies over which the VCO frequency can remain synchronized (locked) is also limited; this range is termed the tracking range.

Another important term is the capture range; the range of input frequencies over which the VCO can lock (synchronize), without external help, to the input signal.

Changing the frequency of the input signal causes a change in the phase detector output; the change is, in effect, an AC component.

If the input signal frequency varies rapidly, i.e. because it is frequency-modulated, the AC component generated by the phase detector is greatly attenuated by the low pass filter, and the correction voltage applied to the VCO may not suffice to keep it locked. In this case, the VCO will oscillate at the average input frequency, thus serving as a filter for the input signal; however, large excursions of the input signal frequency may cause loss of lock.

An input signal contaminated by noise is very much like a frequency-modulated signal; up to a certain signal-to-noise ratio, the PLL remains locked, and generates a relatively clean and stable signal having a frequency equal to the average frequency of the input signal, then suddenly breaks lock as the input signal-to-noise ratio is further decreased.

Since the cut-off frequency of the low pass loop filter is normally quite small, the PLL sees an "internal" signal-to-noise ratio far larger than that existing at its input (assuming white input noise). The PLL can therefore maintain lock and generate a stable signal even if the wideband input noise is larger than the desired input signal.

Examination of the output signal generated by a PLL circuit having a noisy input shows that, although its average frequency is stable, its phase changes randomly, i.e. it is phase-modulated by noise. This phase modulation is called jitter. Jitter has a deleterious effect on the performance of circuits driven by the PLL, in comparison with their performance when driven by a stable signal.

Before ending this concise description of phase-lock loops, a modification of the basic circuit is presented (Fig. 2.5).

![Block diagram of modified phase-lock-loop](image-url)
The circuit shown in Fig. 2.5 has a frequency divider in the feedback path. For the loop to lock, the two signals applied to the phase detector inputs must have equal frequencies; this means that \( f_{\text{in}} = f_0 / N \), where \( N \) is the division ratio of the frequency divider. Rearranging terms yields:

\[
f_0 = N \cdot f_{\text{in}}
\]  

(2-1)

This proves that the modified PLL circuit can generate any desired harmonic of the input signal, while maintaining all of the advantages of a regular PLL.

2.2.3 CLOCK REGENERATION BY PLL CIRCUITS

The application of the circuits shown in Figures 2.4 and 2.5 for regeneration of the clock at the receiver is quite straightforward. Referring to Fig. 2.3, the train of pulses generated by the transition detector is applied to the PLL input. The phase detector now updates its output voltage after each transition; its output voltage is filtered and used to bring the VCO to a frequency which minimizes the average timing error between the input transitions and VCO output transitions. Since the input transitions are, in effect, triggered by the transitions of the transmitter clock (however, not all clock transitions generate input transitions for the PLL, see Fig. 2.3), the receiver VCO locks to the transmitter clock frequency.

For a PLL circuit to operate successfully as a clock recovery circuit, its phase detector must be designed so as to maintain its last acquired output voltage when a VCO transition is not followed by an input transition (within the limits of a VCO period). A circuit fulfilling this requirement is shown in Fig. 2.6.

---

**FIGURE 2.6:** Operation of a phase detector
The circuit converts the VCO signal to a linear ramp and samples the ramp voltage whenever an input pulse (transition) is applied. The linear ramp is generated across capacitor C1 (waveform B) by integrating the square wave VCO signal. The sampling of the voltage across C1 is carried out by electronic switch S2, which closes momentarily at the instant an input pulse is applied and transfers the voltage across C1 to the (much smaller) capacitor C2. At the end of each VCO half-period, capacitor C1 is shorted (discharged) by switch S1, so that it becomes prepared for another phase comparison cycle.

A high input impedance buffer applies the voltage across C2 to the loop filter; this ensures that capacitor C2 will retain its last acquired voltage for a long period of time, such as may occur when a data sequence having a low density of transitions is being received.

An analysis of the operation of this phase detector proves that, during steady state operation, that is, constant phase difference between its two input signals, there is no need to update the value of the voltage across capacitor C2, unless it discharges or the circuit conditions change. Therefore, this type of phase detector is quite suitable for use in clock recovery circuits.

2.2.4 REGENERATION OF DATA SEQUENCE
The recovered clock signal is used to regenerate the original transmitted data sequence by supplying timing information for the decision device of the decoder. When analyzing the circuits presented in paragraphs 1.3 and 1.13, it is readily apparent that their correct operation depends upon the existence of a timing signal that must indicate the correct sampling and decision instants and discriminate between adjacent bit intervals. Incorrect operation of the timing circuits introduces inter-symbol interference and degrades the effective signal-to-noise ratio; therefore, it is essential to have a high-performance clock recovery circuit in order to achieve reliable communication. The system used to regenerate the received signal consists of three components:

- Clock recovery circuit,
- Decision device,
- Pulse shaping and retiming circuit.

![Complete block diagram of regeneration circuit](image)

**FIGURE 2.7: Complete block diagram of regeneration circuit**

The clock recovery circuit (Fig. 2.7) is usually comprised of a transition detector followed by a PLL circuit. The output signal of the PLL is processed by logic circuits, contained in the clock recovery block, which provides appropriate timing signals for the other components.
The decision device analyzes the received signal and makes decisions, or estimates, of the symbol most likely to have been transmitted.

The pulse shaping and retiming circuit processes the output of the decision device to obtain a sequence of digital data symbols of standard levels and duration. This sequence conforms to the same standards as the original transmitted sequence, and can be used to drive the information sink (recipient) or another digital communication link with no loss in performance. However, despite the fact that the regenerated sequence looks the same as the original signal, its contents may have been altered by errors made by the decision device, i.e. due to noise, etc. Nevertheless, regeneration does improve performance, as noise and disturbances accompanying the received signal are totally wiped out, insofar as they do not cause the decision device to make an error. In the case of long communication links, the possibility of wiping out noise is a very definite advantage, because digital links can be engineered as a series connection of short sections, each section operating virtually errorless.

In comparison, analog communication links always suffer from noise build-up, and their output signals can never be “cleaned-up” of noise which contaminates the useful signal along the transmission path.

2.3 CALCULATION OF THE ERROR RATE FOR THE BINARY PAM SYSTEM

A binary PAM system transmits pulses for “ones”, and nothing for “zeroes”; assuming that a long random sequence of data is being transmitted, the transmitter output signal will contain a DC component equal to half the peak amplitude. This system is said to be unipolar, because it consists of pulses of one polarity.

Transmission of unipolar data requires a communication channel capable of passing direct current; few practical channels have this capability, therefore the unipolar system must be modified in order to make it more practical.

The simplest modification which permits removal of the DC component is shown in Fig. 2.8; “ones” are represented by positive pulses and “zeroes” by negative pulses of equal amplitude. The modified system is called “bipolar” for obvious reasons.

![Diagram](image-url)

(a) Unipolar PAM sequence   
(b) Equivalent bipolar sequence

**FIGURE 2.8:** Unipolar and bipolar signals
The two systems shown in Fig. 2.8 have identical error rates when noise exists, because the difference between the two levels has been preserved and therefore the systems are equally affected by noise.

We proceed now to calculate the error rate of a bipolar PAM system, assuming it is disturbed by white additive noise, having a Gaussian amplitude distribution, such as thermal noise. This type of noise has been described in paragraph 5.1.

The PAM system consists, in principle, of an encoder/transmitter, a transmission channel and a receiver. The encoder generates a PAM signal of the type described in Fig. 2.8b, and applies this signal, at an appropriate level, to the input of the communication channel.

The channel adds noise to the transmitter signal, attenuates it, but leaves it essentially undistorted. The channel output is applied to the input of the receiver, which consists of a low pass filter, a comparator and a flip-flop driven by the clock recovery circuit.

**FIGURE 2.9: Model for PAM system**

The function of the low pass filter is to improve the signal-to-noise ratio at the decoder input, by attenuating noise components without significantly affecting the useful signal; intuitively, we expect the improvement to stem from the attenuation of those noise components beyond the highest frequency component of interest contained in the useful signal.

The filtered signal is applied to a comparator, which performs the decoding function. If the incoming signal is random, i.e. the probability of receiving a "one" is equal to that of receiving a "zero", and the signal is accompanied by Gaussian noise of zero mean, such as usually encountered in practice, we find that the optimum reference voltage for the comparator is zero volts. Changing the comparator reference to one side or the other will reduce the tolerance to noise and other disturbances for one polarity of the input voltage, while increasing the tolerance for the other polarity of the input voltage, this will result in biased estimation (decoding) of the input signal, normally an undesirable occurrence.
The comparator serves also as a transition detector for the PLL circuit which recovers the clock signal. The PLL provides a square output signal, whose phase is so adjusted as to have the negative transition in the middle of the input bit interval (see also paragraph 1.3.1).

The clock waveform drives the D-flip-flop which performs the shaping and retiming function. The output of the flip-flop is a bipolar PAM signal, synchronized with the recovered clock, whose value is determined by the value of the output voltage of the comparator at the sampling instant (the negative transition in the clock waveform).

From the above description of receiver operation, it is evident that proper decoding depends upon the signal-to-noise ratio at the input of the comparator.

If, at the decision time, the sum of the signal and noise voltages is greater than zero, the comparator would make the decision that a positive pulse had been transmitted, and if the sum of the signal and noise voltages is less than zero, a negative pulse would be assumed.

For a positive pulse, an error occurs if, at the decision time, the noise is more negative than \(-V_p\), where \(V_p\) stands for the peak value of the received signal; similarly, for a negative pulse, an error occurs if the noise is more positive than \(+V_p\). This statement is put in a mathematical format in equation (2–2):

\[ p_e = p_o \cdot \text{prob} (V_n > V_p) + p_t \cdot \text{prob} (V_n < -V_p) = \]

\[ = p_o \cdot \text{prob} (V_n > V_p) + (1 - p_o) \cdot \text{prob} (V_n < -V_p) \]  

(2-2b)

Where:

- \(p_e\) = probability of error,
- \(p_o\) = probability of transmitting “zeroes”,
- \(p_t\) = probability of transmitting “ones”,
- \(V_n\) = instantaneous value of noise voltage,
- \(\text{prob}(x)\) = probability that event \(x\) occurs.

When random data is transmitted, as is usually the case, the probabilities of “ones” and “zeroes” are equal:

\[ p_o = p_t = \frac{1}{2} \]  

(2-3)

Using equation (2–3), the probability of error becomes:

\[ p_e = \frac{1}{2} \text{prob} (V_n > V_p) + \frac{1}{2} \text{prob} (V_n < -V_p) \]  

(2-4)

From equation (2–4) the probability of error is seen to depend on the relative value of noise versus signal and on the statistical distribution of the noise. For Gaussian noise, the distribution
is given in equation (1-6), and the cumulative probability that its instantaneous value is below a certain value is shown graphically in Fig. 1.15. This cumulative probability is called the error function, designated erf(x), where x is the functional variable. Another useful function is the complementary error function, erfc(x); this function is defined as follows:

\[ \text{erfc}(x) = 1 - \text{erf}(x) \]  

(2-5)

Both \( \text{erf}(x) \) and \( \text{erfc}(x) \) have no explicit formulas, but are extensively tabulated.

Using the notation introduced above, the probability of error for the PAM system in Gaussian noise is given by:

\[ p_e = \frac{1}{2} \text{erfc} \left( \frac{V_p}{\sqrt{2} \cdot V_n} \right) = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{S}{N}} \right) \]  

(2-6)

Where:

\( V_n \) = rms value of noise,
\( S/N \) = signal-to-noise power ratio.

The equation is plotted in Fig. 2.10; it can be seen that excellent performance is obtained for modest signal-to-noise ratios, i.e. error rate of \( 10^{-10} \) for a peak signal voltage/rms noise voltage ratio of only 6 (or 16 dB).

![FIGURE 2.10: Probability of error versus (peak signal) to (rms Gaussian noise) for polar PAM system](image)

Analysis of equation (2-6) shows that, for a given signal level, the only parameter available for optimization is the filter bandwidth. Filter bandwidth directly influences the noise voltage value, because of our assumption that white noise is present at the receiver input. On the other hand, reducing the filter passband also alters the signal shape, as explained in paragraph 1.12. An optimum bandwidth must therefore exist, which is narrow enough to suppress a large part of the input noise but remains wide enough to pass the essential components of the useful signal.
2.4 OPTIMIZATION OF PAM RECEIVER BANDWIDTH

In this paragraph, we will develop an argument, enabling the student to determine which filter yields the optimum value for the receiver bandwidth.

Since the noise transmitted by the low pass filter is directly proportional to the cut-off frequency of the filter, we must select the lowest cut-off frequency compatible with the bandwidth of the useful signal. "Bandwidth", in this context, is to be understood as the frequency range which must be transferred.

The frequency range required for transmission of a random bipolar signal starts close to DC and extends to the highest frequency component which is essential for correct transmission of information. To evaluate the highest component, it is sufficient to examine all possible data sequences and isolate those patterns which change fastest; these portions consist of an alternate sequence of "zeroes" and "ones", i.e. ...010101...

The sampling rate, necessary to identify this sequence, is of course, equal to the bit rate, R. The sampling theorem, presented in para. 1.7.1, tells us that the highest allowable frequency component of a signal that can be reconstructed by sampling it at a rate of R samplings/sec is R/2 Hz. It is reasonable to assume that the converse is also true, therefore the highest frequency component which must be transmitted is R/2 Hz.

The reasoning above shows that the filter having the smallest possible bandwidth, that is, the optimum receiver filter, is an ideal low pass filter (see Fig. 1.32) whose cut-off frequency is equal to half the bit rate. Using non-ideal filters will necessitate an increase in bandwidth.

2.5 MATCHED-FILTER RECEIVER

The concept of matched filtering has been described in detail in paragraph 1.13. In sub-paragraph 1.13.1, it was mentioned that it is relatively easy to build a matched-filter receiver for the particular case of rectangular pulse shape. This is exactly the case for PAM systems; the corresponding receiver structure is shown in Fig. 2.11.

![Figure 2.11: Matched filter receiver for PAM system](image-url)
The receiver consists of an ideal integrator, whose output is applied to a comparator and indirectly sampled at the end of each bit period by transferring the comparator output signal into the D flip-flop. This is an equivalent structure for the theoretical. The sampling switch is momentarily closed at the end of each bit interval. The integrator contents are “dumped” after the sampling moment, in order to re-set the proper initial conditions for processing the next bit; this operation is symbolically indicated by the “dump”--switch connected to the integrator output.

Comparing Fig. 2.11 with Fig. 2.9, it is seen that the main difference between the two receiver structures lies in the replacement of the low pass filter with the integrator. The signal-to-noise, \( (S/N)_o \), obtained at the output of the integrator is given by equation (1–52), repeated here for convenience:

\[
\frac{S}{(N_o)} = \frac{\sqrt{2E_s}}{N_o}
\]  
\( (2-7) \)

Where:
- \( E_s \) = total signal energy (Joules)
- \( N_o \) = single-sided noise power density (watt/Hz).

The error rate performance of the receiver is calculated by substituting in equation (2–6) the value calculated by equation (2–7).

2.6 SELECTION OF PAM SYSTEM TRANSMISSION CHARACTERISTICS

In the above discussion of optimal receivers, it was tacitly assumed that the transmission channel has infinite bandwidth and that no filter is used at the output of the transmitter.

In practice, both assumptions may prove false:

- The bandwidth of the transmission channel may be considered “infinite” only when the actual rate of information transfer is much lower than the theoretical capacity. For a binary PAM system, this means that the channel bandwidth, \( W \), must be much larger than the bit rate, \( R \).

- The spectrum of the transmitted signal may have to be limited, because of various engineering considerations, such as the desire to reduce disturbance of other users of the transmission channel. For example, taking into consideration telegraphic communication via telephone cables, the application of steeply-rising square wave signals to a pair of \( -c.c \) contained in the cable may induce strong disturbances (clicks) in adjacent pairs; these clicks may be quite annoying to other telephone users. Low pass filtering of the output signal of the transmitter will reduce the interference to other users, with – hopefully – little degradation of the telegraphic channel.

The effect of filtering on the quality of digital communications must be carefully analyzed. In paragraph 1.12, it was shown that filters generally introduce inter-symbol interference, which may degrade the effective signal-to-noise at the input of the decoder. To avoid undesirable effects and degradation of performance, the design of filters must compromise between the necessity to reduce
bandwidth and the requirement that intersymbol interference be minimized. To this end, the conditions for the minimization of inter-symbol interference must first be found. These conditions are first developed for the overall transfer function between transmitter and receiver, then the optimum division of the transfer function is found.

**FIGURE 2.12: Model for analysis of PAM system transfer functions**

The model for carrying out the analysis is shown in Fig. 2.12; it consists of a transmitter filter, having transfer function \( G_T(\omega) \), a channel of "infinite" bandwidth which adds noise, and a receiver filter, having transfer function \( G_R(\omega) \). The overall transfer function, \( G(\omega) \), of the system is given by:

\[
G(\omega) = G_T(\omega) \cdot G_R(\omega)
\]  

(2.8)

### 2.6.1 OPTIMAL TRANSFER FUNCTION FOR ZERO INTERSYMBOL INTERFERENCE

The condition for zero intersymbol interference can be simply stated as follows: At the sampling instant, the contribution of all previously transmitted symbols, taken over all possible data sequences, to the signal appearing at the input of the decoding circuit, must be null.

The condition for non-interference implies that the filter impulse response, \( h(t) \), must obey the following rules:

\[
\begin{align*}
  h(0) &= h_0, & h_0 &\text{a fixed constant} \\
  h(kT) &= 0, & k &= 1, 2, 3\ldots
\end{align*}
\]  

(2.9a)  

(2.9b)

A transfer function obeying (2–9) is said to belong to Class I of Nyquist functions. (Nyquist, in 1928, was the first to state the conditions for zero inter-symbol interference applying to the transmission of digital signals).

A typical example of a filter having the necessary impulse response is the ideal low pass filter (Fig. 1.32), whose cut-off frequency is set to half the bit rate. However, the ideal low pass filter has two shortcomings:

- Small deviations of the transmission rate from the design value result in significant inter-symbol interference, because the zero crossings of the impulse response are no longer synchronized with the sampling instants (see Figure 1.32b).
The ideal low pass filter has a non-causal response, that is a response which starts before an input signal is applied. This means the ideal filter cannot be physically implemented; moreover, it is exceedingly difficult to even build an approximation to it.

A careful analysis of the conditions set forth in equation (2–9), together with the additional requirement that we restrict our search to transfer functions having a linear phase function, shows that there is a multitude of transfer function shapes which obey equation (2–9). All possible function shapes pass through the point $h_0 T/2$ at frequency $f = 1/2T$ ($T = 1/R$), and all shapes are symmetric about $f = 0$.

The ideal low pass filter is now seen to be the filter with the smallest bandwidth which still obeys equation (2–9). Another class of transfer functions, which has proven particularly useful in communication systems, is the family of cosine roll-off functions.

The transfer functions of this family are defined by the following relationships:

$$H(f) = h_0 T, \quad 0 < |f| < (1 - a) \frac{1}{2T} \quad (2-10a)$$

$$H(f) = \frac{h_0 T}{2} \left[ 1 - \sin \left( \frac{T}{2a} \cdot 2\pi f - \frac{\pi}{2a} \right) \right], \quad (1 - a) \frac{1}{2T} \leq |f| \leq (1 + a) \frac{1}{2T} \quad (2-10b)$$

$$H(f) = D, \quad (1 + a) \frac{a}{2T} < |f| \quad (2-10c)$$

The time domain response corresponding to these transfer functions is given by equation (2–11):

$$h(t) = h_0 \frac{\sin \frac{\pi}{T} t}{\frac{\pi}{T} t} \frac{\cos \frac{\pi a}{T} t}{1 - \frac{2a}{T} t} \quad (2-11)$$

In equation (2–10) and (2–11), the roll-off parameter, $a$, is chosen between 0 and 1; it is also frequently expressed as a percentage, 0 to 100%.

\[ \text{FIGURE 2.13: Cosine roll-off Nyquist Class I transfer function} \]
The variation of the frequency and time response, as a function of \( a \), are shown in Fig. 2.13. It can be seen that for \( a = 0 \), the ideal low pass filter characteristics are obtained and as \( a \) is increased, the frequency response widens and the ripples in the time domain response decrease. Therefore, the roll-off parameter, \( a \), allows a trade-off between bandwidth and time response; increasing the bandwidth makes the filter more tolerant to variations in the transmission rate, or, equivalently, to inaccuracies in the location of the sampling instants.

### 2.6.2 Optimal Transfer Functions for Transmitter and Receiver

The transfer function found in paragraph 2.6.1 specifies the overall system characteristics required in order to avoid inter-symbol interference. However, as stated in the introduction to this paragraph, this leaves open the issue of optimal partition of the overall transfer function, between the transmitter and the receiver. Such a partition must obey equation (2–8).

In practice, the use of two filters, one in the transmitter and another in the receiver, may be dictated by considerations such as occupied bandwidth, interference to adjacent users, etc. In this section, we shall approach the partition issue from the point of view of optimal system performance in white noise. This means that the receiver filter characteristics, \( GR(\omega) \), must be selected so as to minimize the error rate, or, equivalently, to maximize the signal-to-noise ratio at the sampling (decision) instants.

Since the overall system transfer function is selected from amongst the functions obeying equation (2–10), the transmitter filter characteristic will be calculated as follows:

\[
GT(\omega) = \frac{G(\omega)}{GR(\omega)} \tag{2-12}
\]

The result of the optimization, in the case of signals degraded by white noise, is presented in equation (2–13):

\[
GR(\omega) = GT(\omega) = \frac{X(\omega)}{\sqrt{X(\omega)}} \tag{2-13}
\]

Equation (2–13) shows that the overall filtering requirement must be equally distributed between the receiver and transmitter filters; the phase characteristics, although not specified by equation (2–13), may be arbitrarily chosen as long as they compensate each other, and an overall linear phase characteristic is obtained.

### 2.7 Multi-Level Coding

The binary (two-level) PAM system analyzed in the preceding paragraph requires a certain minimal bandwidth for any given transmission rate; the bandwidth cannot be further reduced without introducing errors in the transmitted information. The errors introduced by reducing the system bandwidth are generally not dependent on the signal-to-noise ratio, and, therefore, their number cannot be reduced by increasing the signal-to-noise ratio.
This paragraph and those following it are dedicated to the description of methods and means for reducing the bandwidth necessary for transmission, possibly at the expense of requiring a signal-to-noise ratio higher than that needed in a binary system for a given performance (error rate).

An obvious way to reduce bandwidth is to encode several input bits into one output symbol; then the number of symbols fed to the channel per unit of time decreases in proportion to the number of input bits grouped into one output symbol. 

As an example, let us analyze a four-level system (Fig. 2.14). The transmitter consists of a shift register, in this case having one stage, followed by an encoder, which encodes two bits at a time (the bit stored in the shift register and the bit presently applied to the transmitter input). The encoder outputs symbols with a duration of two input bit intervals, therefore the symbol rate is half the input bit rate. Note that the information transfer rate, as defined in Chapter 1, remains constant.

![Transmitter block diagram](image)

**FIGURE 2.14:**
*Four level transmitter*

Referring to Fig. 2.14b, it can be easily seen that the basic repetition rate of the encoder output is half that of the input (binary) data sequence. Also, the transmitter imposes a delay of one bit between the input and output sequences.

The receiver decodes the sequence generated by the transmitter by detecting the voltage range occupied by the incoming signal plus noise at the sampling instant, and generating a two-bit word which corresponds to the symbol whose nominal level is contained in the previously mentioned voltage range. These voltage ranges can be identified in Fig. 2.14b, waveform D.

To decode the received symbol and send the corresponding two-bit word, the receiver has to wait at least one bit interval. The minimum total delay imposed by the equipment used in the four-level system equals two bit intervals.
We shall evaluate now the probability of error for bipolar PAM systems utilizing multi-level coding, when the received signal is contaminated by white Gaussian noise.

As in the case of binary PAM, a symbol is erroneously decoded when the instantaneous noise voltage is large enough to move the sum voltage (signal plus noise) in another tier (voltage range). To reduce the probability of this event occurring for all possible symbols, given that all symbols are equally probable (as is the case when random information is being transmitted), the levels must be equally spaced. The total voltage swing available is \(-V_p\) to \(+V_p\), therefore the spacing between adjacent levels is \(2V_p/(m-1)\). Symbols located at the extreme levels are erroneously decoded only if the noise voltage has the appropriate polarity. The resulting probability of symbol error is then given by equation (2-14):

\[
p_e = \frac{m-1}{m} \cdot \text{erfc} \left( \frac{V_p}{(m-1) \sqrt{2V_n}} \right)
\]  

(2-14)

where all quantities have the meaning also used in equation (2-6). The probability of symbol error for several multi-level systems, as calculated by equation (2-14), is shown in Fig. 2.10, together with the curve for the binary system. From these curves it can be seen that the signal-to-noise ratio required by the receiver for a specified error rate increases as the number of levels is increased; a coarse “rule-of-thumb” is a 6-dB increase for each doubling of level number.

When comparing systems utilizing different numbers of levels, we must remember that although the signal-to-noise ratio must be increased in proportion to the number of levels utilized, the bandwidth required by the system to carry a given amount of information for a given bit rate is halved each time the number of levels is doubled. Assuming the system is actually designed for the minimum necessary bandwidth, the noise accepted by the receiver filter in its pass band (this is the noise reaching the decoder) is also halved for each doubling of levels, because the noise power is directly proportional to bandwidth (see paragraph 1.5.1).

Assuming that all systems provide the same total voltage swing (\(-V_p\) to \(+V_p\)), the signal-to-noise ratio reaching the decoder is improved by 3 dB for each doubling of the level number. Therefore, the increase in the signal-to-noise ratio actually required by multi-level systems, in order to maintain the error rate when the number of levels is doubled, is only 3 dB. The required increase could be answered, for example, by doubling the transmitter output power.

The last point to be made about equation (2-14) is that this equation gives the symbol error probability, not the bit error probability. Each symbol carries several bits of information, i.e. in a four-level system - two bits, therefore erroneous decoding of one symbol may induce up to \(\log_2 m\) bit errors, where \(m\) is the number of levels. An upper bound for the bit error probability would therefore be \((\log_2 m)\) times the value calculated by means of equation (2-14); a lower bound is given by equation (2-14) itself.

In practice, we may assume that a well engineered system always operates with a relatively high signal-to-noise ratio, so that the probability of having the noise voltage “push” the instantaneous voltage beyond the neighboring voltage tiers is negligibly small. Thus, all errors stem from replacing a given symbol by one of its neighbors, for example, in Fig. 2.14b, the symbol representing 10 may be replaced by the 01 or 11 symbol.
A scheme for reducing the bit error probability to a value close to the lower bound can now be devised; it involves assigning to neighboring levels m-bit binary words differing in only one bit. An appropriate code for this purpose is the Gray code; for the four-level system, the levels would be assigned the values 00, 01, 11, 10, in place of the values 00, 01, 10, 11 assigned in the binary code, i.e. the code used in Fig. 2.14b.

A final point of caution must be made in relation to the multi-level system; such systems are more sensitive than binary systems to imperfections in the transmission and reception equipment, because the difference between adjacent levels is smaller. In practice, systems having more than 16 levels are seldom used.

2.8 THE EYE DIAGRAM

The eye diagram is a convenient technique for the representation and the analysis of the effects of transmission imperfections upon the signal.

As a diagnostic method, the eye diagram can be displayed on an oscilloscope, whose time base (horizontal deflection system) is running at the symbol clock frequency, while the vertical deflection system is fed with the signal.

For equipment design purposes, the eye diagram can be calculated by considering the effect of all system components on the signal, in the time domain.

The power of this technique can be illustrated by a simple example (Fig. 2.15). In Fig. 2.15a, an undistorted sample of a binary bipolar PAM signal is shown, together with the corresponding eye diagram.

In Fig. 2.15b, the same sample is shown after undergoing distortion, such as caused by inter-symbol interference. The eye diagram clearly shows the distortion, as the "eye" becomes distorted and its inner (free) area is reduced.
The effects of noise can also be seen and identified in eye diagrams; noise makes the boundary lines of the diagram appear fuzzy.

![Diagram showing eye characteristics](image)

**FIGURE 2.16:**
*Important characteristics of an eye pattern*

Fig. 2.16 presents the important characteristics of an eye diagram. First, the instant at which the vertical eye opening is maximal is the best sampling instant, because at this time the separation between all possible signal values is maximum and the noise amplitude required to “close the eye” or make an error is the largest.

Secondly, the decision threshold should be set at the level at which horizontal eye opening is maximal.

Imperfections in the transmission characteristics tend to “close the eye”; for example, timing jitter reduces the horizontal eye opening, while inter-symbol interference reduces both the horizontal and vertical eye dimensions.

The rate of eye closure with time indicates how sensitive the system is to timing errors; if a small displacement of the sampling instant from its ideal location causes a large reduction in the vertical eye opening (as measured at the displaced sampling instant), then the system performance may be seriously degraded in actual operation.

The minimum noise margin can also be found from the eye diagram; this is the voltage difference between the decision threshold and the inner boundary line of the eye as measured at the sampling instant.

Since distortion, and especially inter-symbol interference, causes the inner boundary line to come closer to the decision threshold, the effectiveness of the design measures taken to minimize such causes of degradation can be easily evaluated by analyzing the eye diagram.

Most types of clock recovery circuits derive the primary timing information from the zero (threshold) crossings of the signal; this subject is discussed in greater detail in paragraph 2.2.3. For such circuits, the width of the zero-crossing strip, which indicates the amount of waveform distortion at the zero crossing, is very important; a wide strip indicates significant jitter, and this is equivalent to having a noisy input to the clock recovery circuits.
The eye diagrams shown in Fig. 2.15 and 2.16 pertain to binary systems; eye diagrams can also be used to analyze multi-level systems. Fig. 2.17 presents eye diagrams typical for three- and four-level systems.

(a) Eye diagram for three-level PAM system  
(b) Eye diagram for four-level PAM system

FIGURE 2.17: Eye diagrams for multilevel systems

The main difference between multi-level eye diagrams to binary eye diagrams is the appearance of multiple “eyes” — one eye for each level. The eyes must be stacked vertically, along a common sampling instant line.

Eye diagrams permit taking a new, more practical view on the issue of inter-symbol interference. Instead of the classical design principle of zero inter-symbol interference, as described in paragraph 2.6.1, the principle of controlled inter-symbol interference can be applied; this principle states that a controlled amount of inter-symbol interference can be tolerated, or even intentionally introduced into the system, as long as a certain eye opening can be guaranteed. This means that new design trade-offs can be made regarding system bandwidth, which were forbidden under the Nyquist principle of zero inter-symbol interference. A practical design approach, utilizing controlled amounts of inter-symbol interference, is the partial response approach.

2.9 PARTIAL RESPONSE SYSTEMS

The ideal Nyquist channel is capable of transmitting at a rate of 2 symbols per Hz of bandwidth, because it consists of an ideal low pass filter; moreover, this performance is achieved at zero inter-symbol interference.

In practice, however, neither an ideal low pass filter nor zero inter-symbol interference can ever be achieved; nevertheless, channel capacities of 2 symbol/Hz or higher can be achieved, if a carefully controlled amount of inter-symbol interference is introduced by appropriate filters.

Paragraph 1.12, which deals with the effect of filters on digital signals, relates inter-symbol interference to the sluggish response of filters, which extends over periods longer than one bit (symbol) interval. This effect is mathematically equivalent to the effect of an analog memory
which remembers past signal values; therefore, in each symbol interval, the output voltage of a filter is a superposition of several signal voltages, constituting the response of the filter to the present input and to a certain number of past inputs.

Since superposition is in effect simple summation, the statement made above is equivalent to stating that the output of a filter is a linear combination of responses to past signals.

In Chapter 1, paragraph 1.2, such combinations were shown to be, in effect, a code.

If, by a lucky accident or by other, somewhat more mathematical, means, the linear combination appearing at the filter’s output can be decoded and its components be unambiguously identified, then the existence of inter-symbol interference would not hamper transmission of digital signals. Such decoding appears feasible, because, in principle, the system starts at \( t = 0 \) (time of turn-on) from zero initial conditions, thus the first received symbol is not altered by preceding symbols and can be easily decoded. Once the symbol is known, the next symbol can be decoded and so on, until the whole message is decoded.

The scheme proposed above could become quite practical, if the output of the filter is, by design, made dependent on only a few preceding symbols. Indeed, this is the case for actually working systems; these are designed with a memory of one to seven symbols.

Making the output dependent on a small number of symbols overcomes one major flaw of the partial response scheme: error propagation. Error propagation occurs because one decoding error is likely to influence the decoding of the following symbols.

Having a short “memory” stops the chain of errors after a few symbols because the first error is “forgotten” (expelled from the decoder), so that correct operation is resumed automatically. Other methods used to suppress error propagation will be presented in paragraph 2.11.

We proceed now to study several widely used partial response schemes.

### 2.10 THE BITERNARY PARTIAL RESPONSE SYSTEM

The biternary system is a three-level partial response system, whose memory extends over one preceding symbol - one bit. The memory is obtained in analog form, by using the overall system transfer function given in equation (2-15).

\[
G(\omega) = \begin{cases} 
 2 \cos \frac{T}{2} \omega, & \omega \leq \frac{\pi}{T} \\
 0 & \omega > \frac{\pi}{T}
\end{cases} \quad (2-15)
\]

The transfer function described by equation (2-15) can be more clearly visualized when we substitute for \( T \) the value \( 2\pi/R \), where \( R \) is the bit transmission rate. Then the following expression is obtained:
\[
G(f) = \begin{cases} 
2 \cos \frac{\pi f}{R} & f \leq \frac{R}{2} \\
0 & f > \frac{R}{2} 
\end{cases} \tag{2-16}
\]

The required filter transfer function is a cosine shape, having a pass band of 0 to R/2 Hz, the same as that utilized by the ideal Nyquist filter. There is, however, one important difference between these filters: the cosine filter has a very gradual roll-off, while the Nyquist filter ends with a "brick-wall". As a result, the Nyquist filter will pass a .0101. sequence, but the cosine filter will give it an infinite attenuation, because the fundamental frequency of the sequence is R/2 Hz. Moreover, as can be seen from Fig. 2.18, the impulse response of the cosine filter is not identically zero at the sampling instants, so that a significant amount of inter-symbol interference is obtained.

![Figure 2.18: Biternary filter and its impulse response](image)

The impulse response is given by the following expression:

\[
H(t) = \frac{4}{\pi} \cdot \frac{\cos \pi \frac{t}{T}}{1 - 4t^2/T^2} \tag{2-17}
\]

When the sampling instant is set at -T/2, the impulse response scale factor is exactly one; the response due to the preceding symbol (at \( t = +T/2 \)) is also multiplied by unity, while the response due to the other previous symbols is exactly zero. Thus, at the sampling instants indicated by arrowheads in Figure 2.18, the filter output is the sum of the present symbol and of the preceding symbol:

\[
y_k = x_k + x_{k-1} \tag{2-18}
\]

Where:
- \( y_k \) = the value of the filter output at the \( k \)-th sampling instant,
- \( x_k \) = the value of the filter input at the \( k \)-th sampling instant,
- \( x_{k-1} \) = the value of the filter input at \( k-1 \)-th sampling instant.
Assuming bipolar PAM signals, each input, \( x_i \), can assume voltage values \( \pm V_p \), therefore \( y_k \) can assume one of the three following voltages: \( \pm 2 V_p \), 0. This means that the binary input sequence has been converted to a three-level (ternary) sequence, hence the name of this method: "bitermary".

The encoding process carried out by the filter is illustrated in Fig. 2.19 for the simple example of two consecutive "ones".

(a) Bitermary coding for two binary "1"  
(b) Bitermary coding for four consecutive "1"

**FIGURE 2.19: Examples of bitermary coding**

To decode the bitermary sequence, we use again equation (2-18); once \( x_{k-1} \) is found, its value can be subtracted from the present value of \( y_k \), to decode \( x_k \), and so on. A receiver capable of decoding bitermary sequences is shown in Figure 2.20; typical waveforms are shown in Fig. 2.21.

**FIGURE 2.20:**  
Bitermary decoder

**FIGURE 2.21:**  
Typical waveforms for the bitermary system
The receiver analyzes the incoming sequence by means of two comparators, having appropriately selected reference voltages. The comparators are arranged to detect excursions of the input signal beyond the null zone. Examination of the filtered sequence yields the following decoding rules:

- The present symbol is a "zero" if, at the sampling instant, the input lies beneath the lower threshold.
- The present symbol is a "one" if, at the sampling instant, the input is below the lower threshold.
- An input signal falling between the two thresholds indicates that the present symbol is the inverse of the preceding one.

The decoding and sampling of the comparator outputs is carried out by the J-K flip-flop shown in Fig. 2.20.

We proceed now to analyze the error rate of the biteminary system, when disturbed by white noise.

The results presented in paragraph 2.6.2 indicate that, for optimum performance, the filter used in the transmitter and receiver must be identical; under these conditions, the error rate of the biteminary system is given by equation (2−19):

\[ p_e = \frac{3}{2} \cdot \text{erfc} \left( \frac{\pi}{4} \cdot \frac{V_p}{V_n} \right) \]  

(2-19)

where all designations are identical to those used in equation (2−6).

Comparing equation (2−19) to equation (2−6), it is seen that to obtain the same error rate as the ideal binary system, the biteminary system requires a 2.1 dB better signal-to-noise ratio.

In return for this 2.1 dB penalty, versus the unreachable ideal, the biteminary system allows transmission of binary data at the Nyquist rate (2 symbols per Hz of channel bandwidth), using realizable filters. Moreover, whereas the binary system requires unrealizable "brick-wall" filters which have no tolerance to variations in the transmission rate, the biteminary system tolerates large variations in the transmission rate – up to 43% of the nominal design rate. This robustness to variations of the transmission rate is characteristic of partial response systems.

2.11 PRECODING – THE DUOBINARY SYSTEM

An obvious drawback of the biteminary system is that errors, once made, tend to propagate, because of the use of previously decoded symbols in the decoding process for the present symbol.

To remove error propagation, means must be devised to make the decoding of the present symbol independent of preceding symbols. Such a means is the use of precoding.

Precoding involves the transformation of a binary input sequence into another sequence, using the following rule:

\[ b_k = a_k + b_{k-1}, \text{ modulo-2} \]  

(2-20)
Where:
\[ b_k = \text{the output of the precoder during the } k\text{-th bit interval}, \]
\[ b_{k-1} = \text{the output of the precoder during the } k-1\text{-th bit interval}, \]
\[ a_k = \text{the } k\text{-th input symbol.} \]

A circuit implementing the encoding rule given by equation (2–20) is shown in Fig. 2.22.

![Precoder circuit diagram]

**FIGURE 2.22:**
Precoder circuit

Consider now the output of the system filter when the precoded sequence is applied to its input. The filter is a coder by itself, whose operation is described by equation (2–18). The similarity of equation (2–20) and equation (2–18) can be better seen by taking advantage of the properties of modulo-2 addition to bring equation (2–20) to the following form:

\[ a_k = b_k + b_{k-1}, \text{ modulo-2} \]  \hspace{1cm} (2–21)

To prove the correctness of equation (2–21), it suffices to check it by substituting the four possible combinations of variables.

The output signal of the precoder, \( b_k \), is the input signal of the filter, therefore, substituting \( b_i \) for \( x_i \) in equation (2–18) results in the following expression for the filter output in response to the precoded input:

\[ y_k = b_k + b_{k-1} \]  \hspace{1cm} (2–22)

This is an analog equation, dealing with the filter output signal at the sampling instants; this equation is similar to the equation for \( a_k \) (the input sequence) equation (2–21). Therefore when \( a_k \) is even (zero), \( y_k \) is even, and when \( a_k \) is odd (one) \( y_k \) is odd (an extreme level in Fig. 2.21).

Inserting the precoder before the filter makes the filter output independent of the preceding symbol; the instantaneous value of its output voltage (at the sampling instant) gives the present symbol directly, without reliance on correct decoding of other symbols. Error propagation is thus prevented.
The operation of the precoder is illustrated in Fig. 2.23.

\[\begin{array}{ccccccccccc}
\text{INPUT} & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\
\text{SEQUENCE} \\
\hline \\
\text{PRECODED} & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\
\text{SEQUENCE} \\
\end{array}\]

**FIGURE 2.23: Operation of precoder and filter in duobinary system**

Referring to Fig. 2.23, the way to decode the duobinary sequence becomes obvious: rectify the filtered sequence and output a "zero" whenever the rectified signal exceeds a certain threshold. The block diagram of a receiver implementing this decoding rule and its typical waveforms, based upon the filtered signal waveform shown in Fig. 2.23, are shown in Fig. 2.24.

**FIGURE 2.24: Duobinary receiver and associated waveforms**
The error rate of the duobinary system is also given by equation (2–19); however, the actual performance of a duobinary system is better than that of a binary system, because it does not suffer from error propagation. In fact, equation (2–19) is slightly optimistic when applied to binary systems, because it calculates only the primary error rate that is caused by noise alone, and does not account for secondary errors, due to error propagation. The same equation is realistic for duobinary systems, because they do not suffer at all from error propagation.

2.12 PRACTICAL IMPLEMENTATION PROBLEMS

In the preceding paragraphs, the principles of operation of various baseband data transmission systems have been presented, and, where necessary, typical implementations for practical circuits have been given.

This paragraph is devoted to an analysis of deviations from the ideal, observed in communication equipment. These deviations can be classified in two major categories:

Transmission imperfections affecting eye diagrams,

Imperfections in the regeneration of clock and opening data at the receiver.

2.12.1 TRANSMISSION IMPERFECTIONS

All baseband transmission systems require that the decoder receive accurately shaped waveforms; this requirement can also be stated in terms of eye diagram opening. That is, the eye opening must be maximized and the maximal opening be obtained at the sampling instant.

The waveform appearing at the input of the decoder depends upon the characteristics of the transmitter filter and those of the receiver filter; secondary factors are the gain and linearity of the active circuits participating in the transmission process.

If the circuits utilized to process the signal are linear, or very nearly so, the characteristics of the transmit and receiver filters can be represented by their composite transfer functions. The system transfer function must then fulfill the Nyquist requirements or be one of the partial-response functions. In either case, only minor deviations in the amplitude and phase characteristics can be tolerated, because such imperfections distort the waveshape and may reduce the margin between the slicing (decision) levels and the received signal, as measured at the sampling instant. Reducing this margin reduces, by the same amount, the noise that can be tolerated; therefore the error rate is increased relative to that achievable by a perfect system.

In the limit, severe departures from the correct response may introduce a background error rate, called residual error rate, which exists even in the absence of noise.

One of the principal sources of such departures resides in the requirement that the DC and low frequency components of the data spectrum be removed, because of the use of AC coupling in most transmission systems, be they radio or wire (telephone) links.

Although the most part of the baseband signalling waveforms possess significant DC and low frequency components, when truly random data is being transmitted, in practice, this problem
is alleviated by two circumstances:

- The very low frequency components are normally generated by very long strings of "zeroes" or "ones". Most sources of digital information, when active, generate data sequences having a reasonable number of transitions per time interval, so that such long strings are very improbable. When the DC component is not taken into consideration, the ratio of the Nyquist bandwidth (in Hertz) to the lowest frequency which must be transmitted can be as low as 100:1 to 1000:1.

- When the conditions described above are ensured, the DC component stems from the average voltage of the data sequence. Therefore, it can be removed (i.e. by a capacitor, transformer etc.) without degrading performance.

Another factor of importance in the design of baseband data transmission is the stability of the peak-to-peak voltage swing at the input of the decoder. This requirement stems from the use of comparators in most decoder circuits — see for example, Fig. 2.9, 2.11, 2.21, 2.24. These comparators are supplied with constant reference voltages, corresponding to the slicing (decision) levels for the nominal signal level. Increasing or decreasing the signal level reduces the noise margin just as in the case of inter-symbol interference and imperfect system transfer function.

The usual means of ensuring the correct signal level is an automatic gain control (AGC) circuit having a dynamic range adequate for the range of signal levels encountered in practice.

2.12.2 IMPERFECTIONS IN THE REGENERATION OF CLOCK AND DATA
Regeneration of the clock is an essential requirement for any data transmission system, because the clock is used to determine the bit (symbol) intervals, the sampling instants, and many other auxiliary functions.

The recovered clock must have a precise and stable-phase relationship with the input signal applied to the decoder. This may seem a simple task when the input signal is relatively clean of noise and with sharp, well-defined transitions, but these conditions are rarely met in practice.

As a matter of fact, the optimal filters for the receivers described in this chapter have bandwidth in the order of half the bit transmission rate, and, thus, their response times are commensurate with the bit interval duration; as a result, the rise and fall times of the filtered signals take a large percentage of the bit duration (see, for example, Fig. 2.24), and it becomes increasingly difficult to establish the correct timing reference.

The timing reference is usually generated by a transition detector, as explained in paragraph 2.2.1. When the signal possesses slow transition times, the instants at which it crosses the transition detector threshold depend upon the preceding symbols. This effect is illustrated in Fig. 2.22.

The errors made in the detection of the threshold crossings act as a noise signal for the phase-locked loop circuits used to recover the clock signal, just as would be the case if the received signal had noise.

However, there is one important aspect in which the effects of signal-dependent timing errors differ from errors induced by noise: while noise introduces random errors in the threshold
crossings, the signal dependent errors are systematic and every clock recovery circuit would respond in an identical manner, i.e. by an error of the same sign and magnitude, to the transmitted sequence.

Therefore, this effect is far more damaging to the performance of a digital system than white noise. To see why, imagine a long distance digital transmission system composed of a large number of identical sections, each having a data regenerator and driver (together called a "repeater"). White noise is largely removed at each regenerator unit, and its effects are a few errors per million and random jitter added to the signal phase. Since the jitter introduced by noise at one repeater is uncorrelated with that of another repeater, the accumulation of noise induced jitter is relatively slow.

![Diagram showing sequence-dependent timing variations](image)

**FIGURE 2.25: Sequence-dependent timing variations**

Assume now that signal dependent jitter appears since all repeaters respond to the signal in an identical manner; the jitter introduced by each repeater adds in-phase with jitter introduced by other repeaters. Large timing errors may then result, even to the effect that lock is lost in some repeaters. It is seen then that signal dependent errors must be kept as small as possible by design.

Other factors which influence proper timing are the response time of the comparator and the stability of its reference level, relative to the input signal level.

The response time is especially critical in high data rate systems, where bit durations are very short and even response times in the order of nanoseconds may become significant.

The stability of the comparator level and the maintenance of a constant input level have already been analyzed with respect to the noise margin (para. 2.12.1); in this case, their effect is to change the phase of the recovered clock with respect to the pre-established relationship.