INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION

INTERNATIONAL CENTRE FOR SCIENCE AND HIGH TECHNOLOGY
c/o INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS 34100 TRIESTE (ITALY) VIA GRIGNANO, 9 (ADRIATICO PALACE) P.O. BOX 586 TELEPHONE 040-224572 TELEFAX 040-224575 TELEX 460449 APH I

H4.SMR/537-21

SECOND COLLEGE ON THEORETICAL AND EXPERIMENTAL RADIOPROPAGATION PHYSICS
(7 January - 1 February 1991)

Co-sponsered by ICTP, URSI, ICSU and with the participation of ICS

PULSE MODULATION

S. C. Dutta Roy
Indian Institute of Technology, Delhi
New Delhi, India

PULSE MODULATION

S C DUTTA ROY
Department of Electrical Engineering
Indian Institute of Technology
New Delhi 110 016, India

Text of a series of lectures to be delivered at the URSI/ICTP sponsored second college on Radiopropagation physics, International Centre for Theoretical Physics, Trieste, Italy, January 7 - February 1, 1991

INTRODUCTION

In CW modulation, some parameter of a sinusoidal carrier wave is varied in accordance with the message. In pulse modulation, some parameter of a pulse train is varied in accordance with the message. Two types of pulse modulation need to be distinguished, viz. pulse analog modulation and pulse code modulation (PCM). In the former, a periodic pulse train is the carrier, and some characteristic of each pulse (eg amplitude, duration or position) is varied in a continuous manner in accordance with the pertinent sample value of the message. In PCM, on the other hand, a discrete time, discrete amplitude representation is used for the signal, and as such it has no counterpart in CW modulation.

SAMPLING

Sampling is the first basic operation in the implementation of all forms of pulse modulation systems. Consider the finite energy signal $g(t)$, shown in Fig. 1(a), and let it be sampled once every T_s seconds to obtain $g_\delta(t)$, as shown in Fig. 1(b). T_s is called the sampling period and $1/T_s$ is

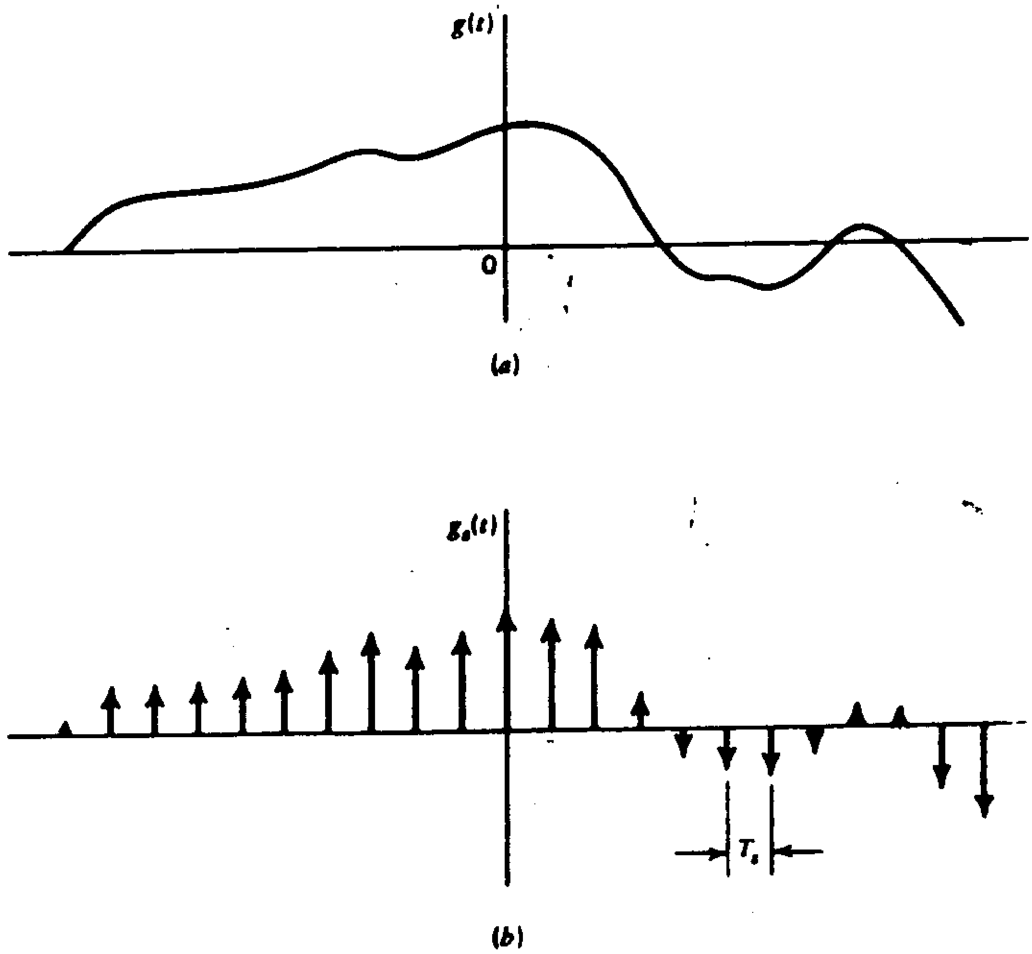


Fig. 1

(a) continuous-time signal and (b) its sampled version the sampling rate or sampling frequency f_s. $g_\delta(t)$ may be viewed as an infinite series of delta functions, weighted by the corresponding instantaneous value of $g(t)$ i.e

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} g(nT_s)\, \delta(t - nT_s) \qquad (1)$$

$$= g(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \qquad (2)$$

$$= g(t) \quad \delta_{T_s}(t) \qquad (3)$$

(Dirac comb)

Taking the Fourier transform of both sides, we have

$$G_\delta(f) = G(f) * \frac{1}{T_s} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_s}\right) \qquad (4)$$

$$= \frac{1}{T_s} \sum_{n=-\infty}^{\infty} G(f) * \delta\left(f - \frac{n}{T_s}\right) \qquad (5)$$

(by interchanging Σ and $*$)

$$= \frac{1}{T_s} \sum_{n=-\infty}^{\infty} G\left(f - \frac{n}{T_s}\right) \qquad (6)$$

Thus uniform sampling in the time domain results in a periodic spectrum in the frequency domain with a period equal to the sampling rate

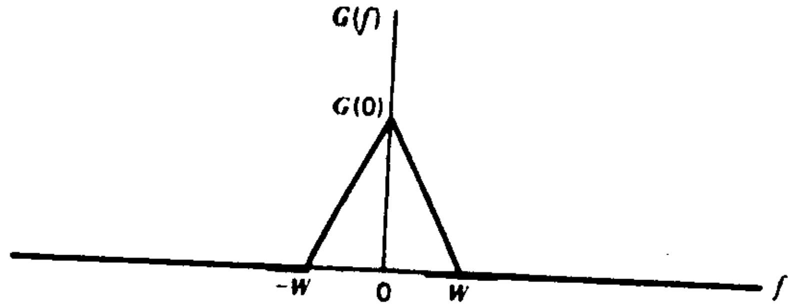
Another useful expression for ~~the~~ $G_\delta(f)$ is obtained from (1) by taking the FT of both sides. The result is

$$G_\delta(f) = \sum_{n=-\infty}^{\infty} g(nT_s)\, e^{-j2\pi n f T_s} \qquad (7)$$

This is a complex Fourier series representation of the function $G_s(f)$ with $g(nT_s)$ defining the coefficients of expansion. Thus

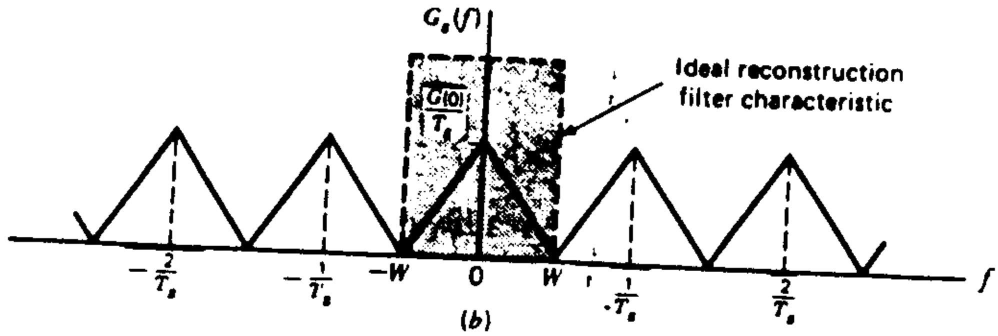
$$g(nT_s) = T_s \int_0^{1/T_s} G_s(f)\, e^{j2\pi n f T_s}\, df \qquad (8)$$

Now suppose $G(f)$ is bandlimited i.e. $G(f) = 0$ for $|f| > W$ and in addition, let $T_s = \frac{1}{2W}$. Then, typically, the spectra of $G(f)$ and $G_s(f)$ will look like those shown in Fig. 2(a) and Fig. 2(b) respectively. For this situation, (7) becomes



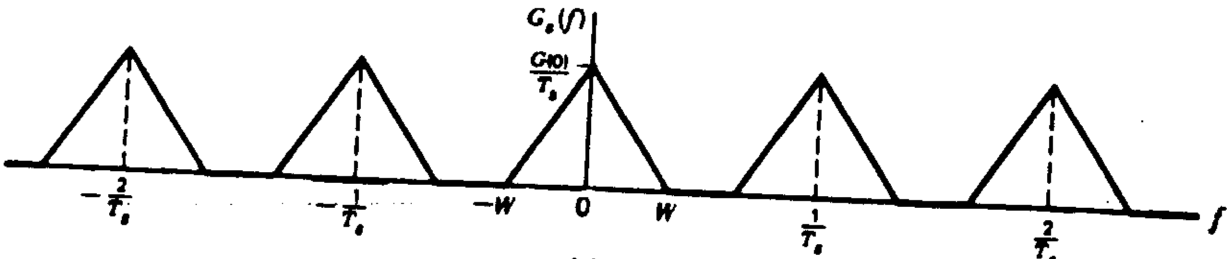
(a)

Spectrum of original signal



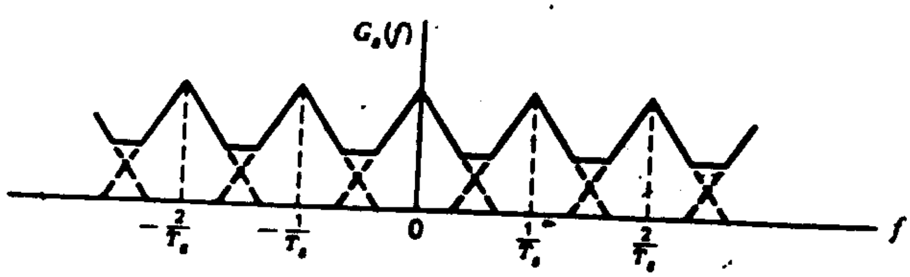
(b)

Spectrum of sampled signal: $f_s = 2W$



(c)

Spectrum of sampled signal: $f_s > 2W$



(d)

Spectrum of sampled signal: $f_s < 2W$

Fig. 2

$$G_s(f) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) e^{-j\pi n f/W} \tag{9}$$

Also, from (6), we have

$$G(f) = \frac{1}{2W} G_s(f) \qquad -W \le f \le W \tag{10}$$

Hence

$$G(f) = \frac{1}{2W} \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) e^{-j\pi n f/W} \; ; \; -W \le f \le W \tag{11}$$

Thus if $g\left(\frac{n}{2W}\right)$ is known for all n, then $G(f)$ can be determined from (11), and hence by inverse FT of $G(f)$, $g(t)$ is uniquely determined. In other words, $g\left(\frac{n}{2W}\right)$ contains all the information of $g(t)$.

For reconstructing $g(t)$ from sample values $g\left(\frac{n}{2W}\right)$, note that

$$g(t) = \int_{-\infty}^{\infty} G(f) e^{j2\pi f t} df$$

$$= \int_{-W}^{W} \frac{1}{2W} \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) e^{-j\pi n f/W} e^{j2\pi f t} df \tag{12}$$

$$= \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \frac{1}{2W} \int_{-W}^{W} e^{j2\pi f\left(t - \frac{n}{2W}\right)} df \tag{13}$$

$$= \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \frac{\sin(2\pi W t - n\pi)}{2\pi W t - n\pi}$$

$$= \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \operatorname{sinc}(2Wt - n) \tag{14}$$

Thus is the interpolation formula and the sinc function $\text{sinc}(2Wt-n)$ plays the role of an interpolation function. Note that (14) also represents the response of an ideal LPF of bandwidth W and zero transmission delay, which is produced by an input signal consisting of a sequence of samples $g(\frac{n}{2W})$ for $-\infty \le n \le \infty$. This is intuitively satisfying since from Fig. 2(b), $G(f)$ may be recovered from $G_s(f)$ by passing the latter through an ideal LPF ~~of BW~~ of bandwidth W.

We may now state the sampling theorem for bandlimited signals of finite energy:

A bandlimited signal of finite energy which has no frequency components higher than W Hz, is completely described by specifying values of the signal at instants of time separated by $\frac{1}{2W}$ seconds

or

A bandlimited signal of finite energy having no frequency components higher than W Hz, may be completely recovered from a knowledge of its samples taken at the rate of $2W$ per second.

The sampling rate $f_s = 2W$ is called the Nyquist rate. In practice, it is customary to use $f_s > 2W$ so as to ensure physical realizability of the reconstruction filter. See Fig. 2(c); the band $W \le f \le \frac{1}{T_s} - W$ may be used to accommodate the transition from the passband to the stopband, which is a requirement for a realizable filter.

One can see from Fig. 2(d) what happens when $\frac{1}{T_s} < 2W$. The distortion is said to arise out of aliasing effect i.e effect of high frequencies taking on the identity of a lower frequency. Hence $g(t)$ cannot be recovered from its samples.

SAMPLING OF BANDPASS SIGNALS

In the case of a bandpass signal with bandwidth small compared to the highest frequency component (f_h) it is possible to use $\frac{1}{T_s} <$ ~~[illegible]~~ $2f_h$. Consider the signal shown in Fig. 3. The minimum sampling rate required for this signal is $2(f_c + W)/m$, where m is the integer ~~[illegible]~~ part of $(f_c+W)/(2W)$. When the frequency band occupied by the signal is located between two adjacent multiples of $2W$, then $m = (f_c + W)/(2W)$ in which case the minimum sampling rate is $4W$ irrespective of the highest frequency component. If this is not the case, then the sampling rate lies between $8W$ and $4W$

TIME DIVISION MULTIPLEXING

Because of sampling, a message occupies the transmission channel for only a fraction of the sampling interval on a periodic basis; at other times, other independent messages can be sent. We thereby obtain

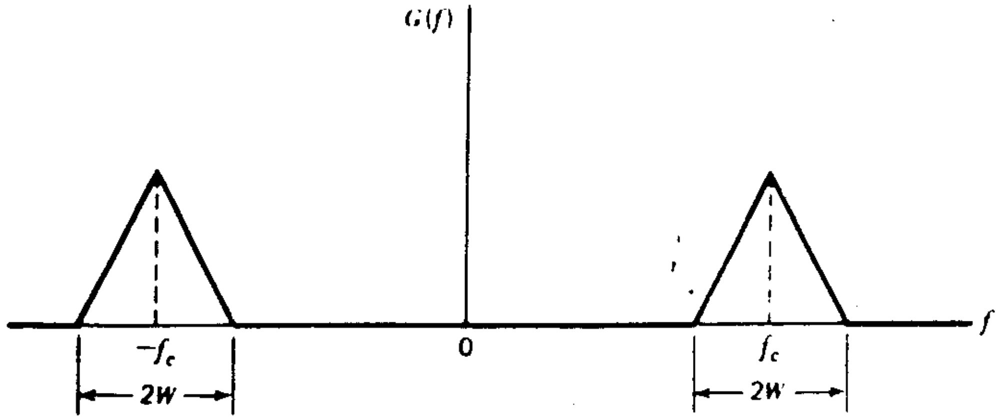


Fig. 3
Spectrum of a bandpass signal

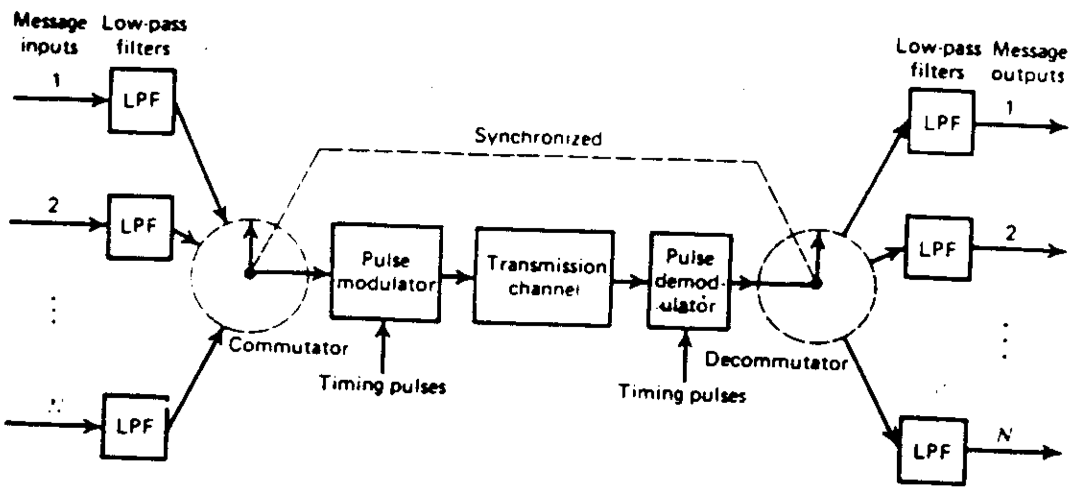


Fig. 4

Block diagram of a TDM system

what is known as time division multiplexing ~~system~~ (TDM). A block diagram of a typical TDM system is shown in Fig-4.

Obviously TDM introduces a bandwidth expansion factor N because of squeezing N samples into one sampling interval. Method of synchronization of commutator and decommutator ~~is~~ ~~affected~~ depends upon the method of pulse modulation used.

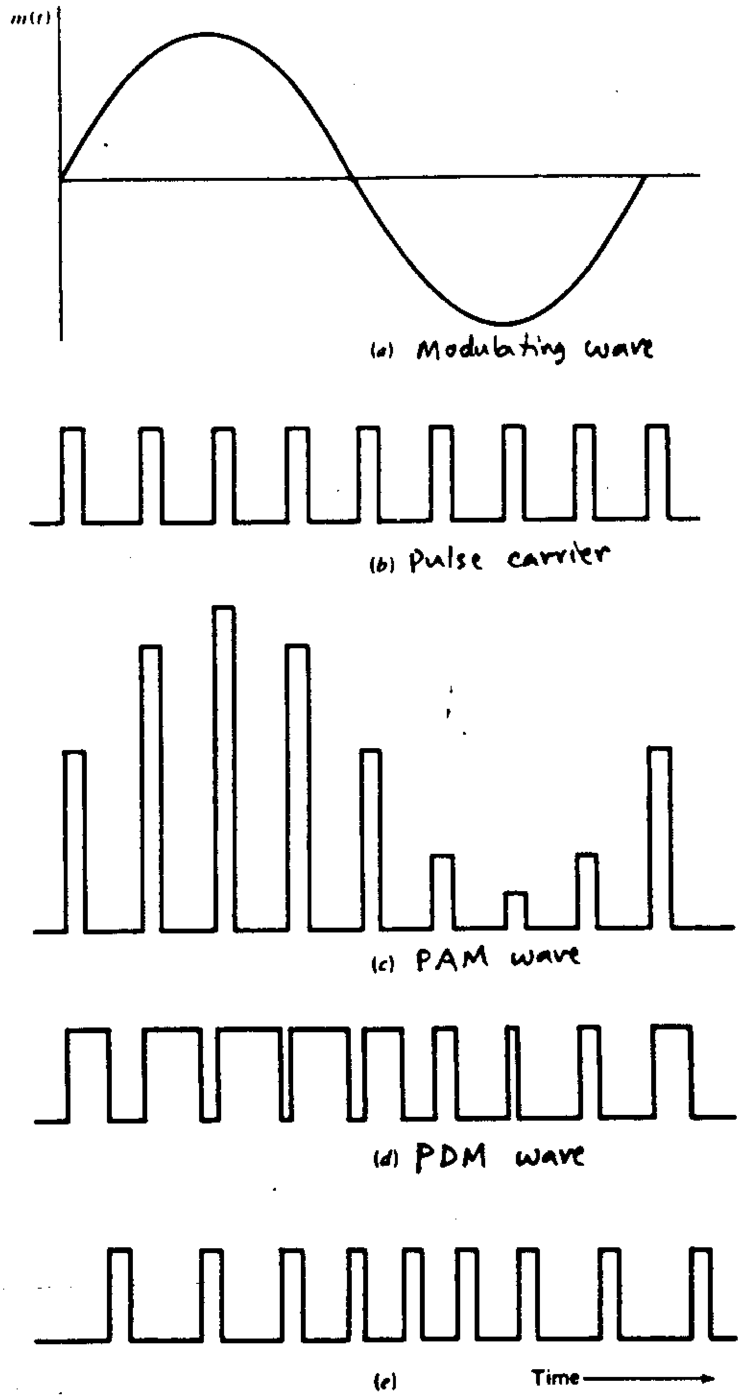
TDM is highly sensitive to linear channel distortion (amplitude of TF not a constant, phase not linear), but is relatively insensitive to nonlinear distortion in the channel.

PULSE AMPLITUDE MODULATION (PAM)

In PAM, the amplitude of a regularly spaced rectangular pulses vary in direct proportion to the instantaneous sample values of a continuous message signal, as shown in Fig.5. Thus a PAM wave is defined by

$$s(t) = \sum_{n=-\infty}^{\infty} \left[1 + k_a m(nT_s)\right] g(t - nT_s) \qquad (15)$$

where $m(t)$ is the message signal and k_a is a constant so chosen that $1 + k_a m(nT_s) > 0 \;\forall n$. $1/T_s$ is chosen to be no less than twice the highest frequency component in $m(t)$.



PPM wave

Fig.5

Various form of pulse modulation

PULSE TIME MODULATION

In pulse duration modulation (PDM), the samples of message signal are used to vary the duration of the individual pulses. This form of modulation is also referred to as pulse width modulation or pulse length modulation. The modulating wave may vary the leading edge, the trailing edge or both edges of the pulse. Fig. 5 shows the essence of PDM.

In PDM, long pulses expend considerable power without bearing any additional information. If only time transitions are preserved, we obtain a more efficient type of pulse modulation known as pulse position modulation (PPM)

PULSE-CODE MODULATION (PCM)

In pulse analog modulation, only time is discretized while the respective modulation parameter (pulse amplitude, duration, position) are varied in accordance with the message in a continuous manner. In a PCM system, both time and amplitude are quantized so that message can be transmitted as coded electrical signals. The block diagram of a PCM system is shown in Fig. 6.

When multiplexing is used, it becomes necessary to synchronize the receiver to the transmitter for the overall system to operate satisfactorily.

Sampling must be done at a frequency $\geq 2W$; the LPF ensures exclusion of frequencies greater than W in its output

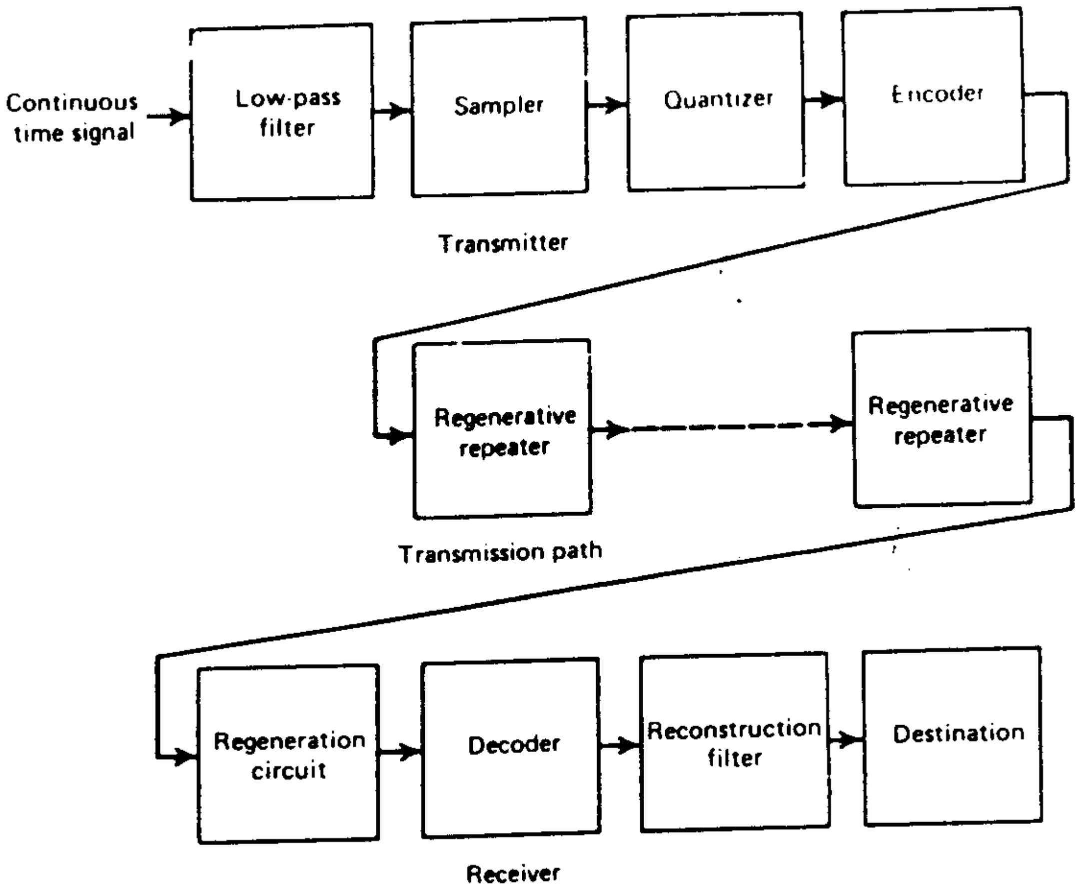


Fig. 6
Block diagram of a PCM system

Quantizing operation transforms an analog signal into a digital one. Graphically, the quantizing process replaces the straight line relation between input and output of a linear continuous system by a staircase characteristic as shown in Fig. 7(a). The difference between two adjacent discrete values is called a quantum. The quantizing error consists of the difference between input and output signals. It is obvious that the maximum instantaneous value of this error is one half of quantum step, as shown in Fig. 7(b).

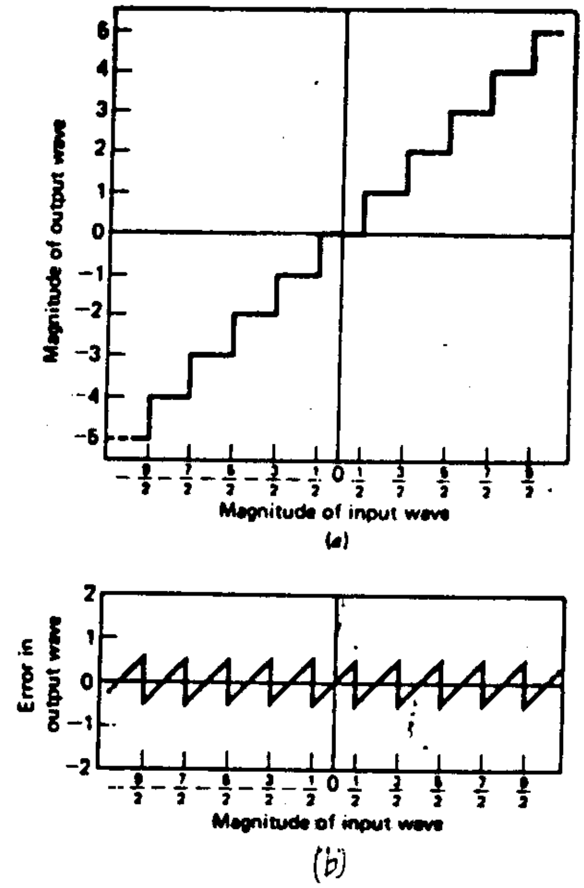


Fig. 7

(a) Quantizer characteristics

(b) Error charactristics

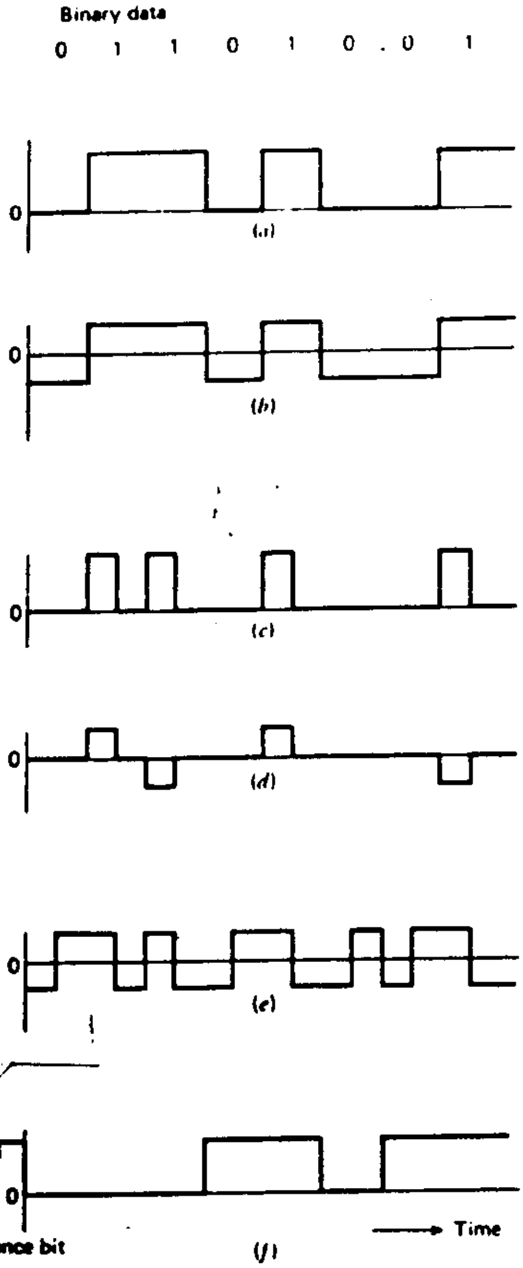
the quantum size may be uniform as shown in Fig. 7 or nonuniform. Voice signals are preferably quantized in a nonuniform manner, so that weak passages are favoured at the expense of loud ones. The use of a nonuniform quantizer is equivalent to passing the baseband signal through a compressor and then applying the compressed signal to a quantizer. In the receiver, one has to use an expander to restore the signals to their original values

ENCODING

The discretized signal obtained by sampling and quantizing is not suited to transmission over a line or radio path. This signal has to be <u>coded</u> i.e. each of these discrete values has to be represented as a particular arrangement of discrete events, which is called a <u>code</u>. One discrete event in a code is called a <u>code element</u> or <u>symbol</u>. For example, the presence or absence of a pulse is a symbol. A particular arrangement of symbols used in a code to represent a single value of the discrete set is called a <u>code word</u> or <u>character</u>.

In a binary code, the two symbols are 0 and 1 (corresponding to absence or presence of a pulse). One can think of ternary and other codes, but the binary code is most immune to noise and is easy to regenerate. If a binary code consists of n bits (binary digit), then a total of 2^n distinct levels can be represented. The ordinal number of the level may be represented in binary to form a code.

Various ways in which binary symbols 0 and 1 can be represented by electrical signals are shown in Fig. 8



(a) ON-OFF (NRZ)

(b) BIPOLAR (NRZ)

(C) RETURN TO ZERO (RZ)

(d) PSEUDO-TERNARY
+ve and -ve pulses alternately for 1, no pulse for 0 (avoids the necessity of transmitting d.c. and l.f.)

(e) SPLIT-PHASE OR MANCHESTER CODE
1 → +ve pulse followed by a -ve pulse of equal amplitude and half duration.
0 → -ve pulse followed by ...

(f) DIFFERENTIAL ENCODING
0 → Transition
1 → No transition

Fig. 8
Various representations of Binary Data

REGENERATION

The most important feature of a PCM system lies in its ability to control the effects of distortion and noise introduced by the channel. This is done by means of a chain of regenerative repeaters (RR) located at sufficiently close spacing along the channel. Each RR performs the functions of equalization, timing and decision making, as illustrated by the block diagram of Fig. 9. The amplifier equalizer

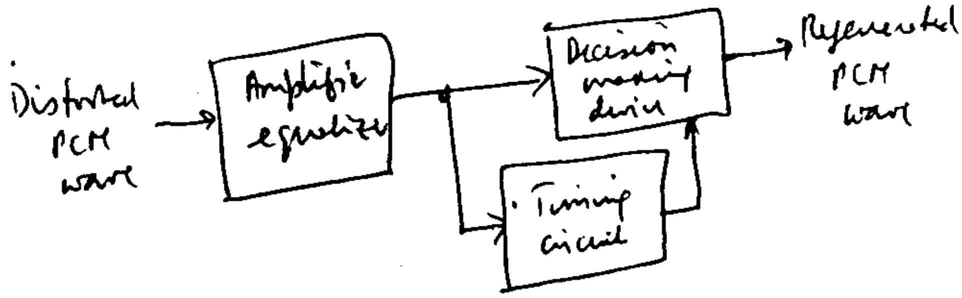


Fig. 9

compensates for the amplitude and phase distortions occurring in the channel. The timing circuit provides a periodic pulse train, derived from the received pulses, for sampling the equalized pulses at the instants of time where the S/N ratio is a maximum. The decision making device is enabled, when at the sampling time determined by the timing circuitry, the amplitude of the equalized pulse plus noise exceeds a predetermined voltage level. With on-off signalling, for example, the repeater makes a decision in each interval as to whether or not a pulse is present. If the decision is "yes", a clean new pulse is transmitted to the next repeater. If, on the other hand, the decision is "no", a clean base line is transmitted. The regenerated signal, in practice, may deviate from the original signal due to two reasons:

1. If, due to noise and interference, a wrong decision is made, then bit errors are introduced into the regenerated signal
2. The spacing between received pulses may deviate from the assigned value, thereby introducing a jitter in the regenerated pulse position.

DECODING AND FILTERING

The received pulses are first regenerated i.e. reshaped and cleaned up in the receiver and then the clean pulses are regrouped into code words and decoded (mapped back) into a quantized PAM signal. Essentially, a pulse whose amplitude is the linear sum of all the pulses in the code word is generated, each pulse in the code being weighted by its place value ($2^0, 2^1, \ldots$)

The signal wave is then recovered by passing the decoder output through a low-pass reconstruction filter whose cutoff frequency is the message bandwidth W.

MULTIPLEXING

TDM is used to transmit a number of different messages, whereby each source maintains its individuality throughout its journey. Because of this individuality, message sources may be dropped or re-inserted with ease. In practice, the number of independent messages has to be limited, because of the difficulty of generating and transmitting narrow pulses.

SYNCHRONIZATION

For PCM with TDM, it is necessary to synchronize the timing operations in the receiver with those of the transmitter, after accounting for the delay in the channel. One possible procedure is to set aside a code element or pulse at the end of a frame (consisting of a code word derived from each of the independent message sources in succession) and to transmit this pulse at every other frame only. The receiver includes a circuit that searches for the pattern of 1's and 0's alternating at half the frame rate.

BANDWIDTH, PULSE SHAPING AND INTERSYMBOL INTERFERENCE

Bandwidth limitations and nonlinear phase characteristics of the channel may cause successive pulses to overlap in time. This gives rise to intersymbol interference (ISI), which is particularly objectionable in TDM, when adjacent pulses may represent symbols of code words derived from different message sources.

The basic elements of a baseband system for transmission of PCM waves is shown in Fig 10.

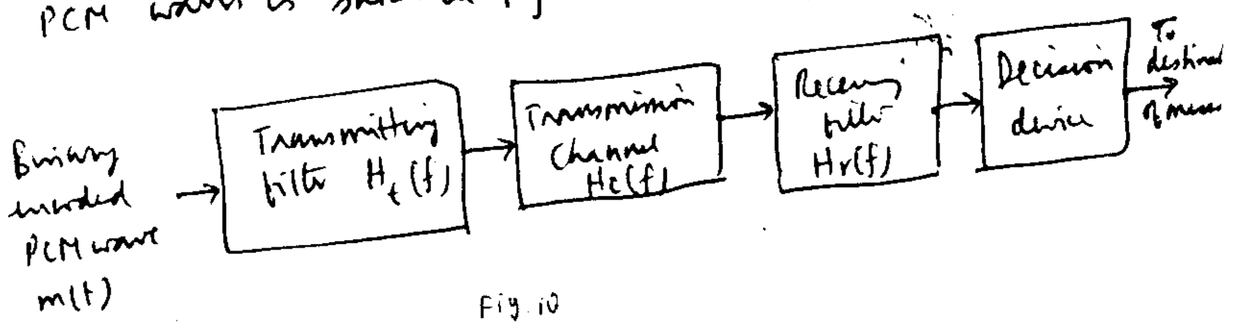


Fig. 10

Consider a PCM system having a single alphabet with $M = 2^q$ symbols or discrete amplitude levels with uniform spacings. For a given average power, the spacing is maximized by using equal +ve and −ve range of values for the amplitude levels. Thus if A_n denotes the n-th amplitude level, ~~then~~ $A_n = \pm A/2, \pm 3A/2, \ldots, \pm (M-1)A/2$ and the PCM wave may be denoted by

$$m(t) = \sum_n A_n \, g(t - nT) \tag{16}$$

where $g(t)$ is a specified pulse signal of duration equal to the symbol duration T. The signalling rate $1/T$ is referred to as the baud rate, expressed in units of bauds. Consider the binary digit symbol (1 or 0); a PCM system with M symbols and a symbol duration T is equivalent to a binary system with a bit duration T_b, where

$$T_b = T / \log_2 M \tag{17}$$

$1/T_b$ is called the bit rate of the system, in bits/sec. Since a baud = one symbol/sec. and in an M level system, each symbol is equivalent to $\log_2 M$ bits, it follows that a baud is equal to $\log_2 M$ bits/second. In the special case of a binary system $M = 2$ and $T = T_b$, so that bit rate and baud rate are the same.

From (16), taking the Fourier transform, we get

$$M(f) = \sum_n A_n \, G(f) \, e^{-j2\pi n f T} \tag{18}$$

The input of the decision device in Fig. 12 receives a signal $y(t)$ whose Fourier transform is

$$\begin{aligned} Y(f) &= H_r(f) \, H_c(f) \, H_t(f) \, M(f) \\ &= \sum_n A_n \, P(f) \, e^{-j2\pi n f T} \end{aligned} \tag{19}$$

where

$$P(f) = H_r(f)\, H_c(f)\, H_t(f)\, C(f) \qquad (20)$$

then

$$y(t) = \sum_n A_n\, p(t - nT) \qquad (21)$$

In the function $P(f)$, typically $H_c(f)$ and $C(f)$ are specified, and the problem is to determine $H_t(f)$ and $H_r(f)$ so that the receiver can recognize the sequence of values of A_n in the received signal wave.

In solving the problem, we have to overcome the ISI caused by the overlapping tails of other pulses adding to $A_n\, p(t-nT)$. One signal waveform that produces zero ISI is

$$p(t) = \text{sinc}(2B_T t) = \frac{\sin(2\pi B_T t)}{2\pi B_T t} \qquad (22)$$

where $B_T = 1/(2T)$. The corresponding $P(f) = 1/(2B_T)$ for $|f| < B_T$ and zero for $|f| > B_T$ i.e. no frequencies exceeding half the Baud rate are needed. The function $p(t)$ is the impulse response of an ideal LPF with gain $1/(2B_T)$ and bandwidth B_T. The function $p(t)$ has its peak value at $t = 0$, and goes through zero at $t = \pm T, \pm 2T, \ldots$; hence pulses defined by $A_n p(t-nT)$ with arbitrary amplitude A_n and $n = 0, \pm 1, \pm 2, \ldots$ will not interfere with each other.

An ideal LPF is of course not physically realizable. Further, $p(t)$ decreases as $1/|t|$ for large $|t|$, resulting in a slow rate of decay. Accordingly, there is no margin of error in sampling times in the receiver, and timing error may cause ISI and wrong decisions.

An alternative solution, given by Nyquist is the so-called cosine rolloff:

$$P(f) = \begin{cases} \dfrac{1}{2B_T} & |f| < f_1 \\ \dfrac{1}{4B_T}\left\{1 + \cos\left[\dfrac{\pi(|f| - f_1)}{2B_T - 2f_1}\right]\right\} & f_1 < |f| < 2B_T - f_1 \\ 0 & |f| > 2B_T - f_1 \end{cases} \quad (23)$$

where

$$\rho = 1 - \frac{f_1}{B_T} \quad (24)$$

is the rolloff factor. For $\rho = 0$ i.e. $f_1 = B_T$, we get the minimum bandwidth solution of the ideal LPF. The inverse FT of (23), $p(t)$ is given by

$$p(t) = \operatorname{Sinc}(2B_T t)\,\frac{\cos(2\pi\rho B_T t)}{1 - 16\rho^2 B_T^2 t^2} \quad (25)$$

Plots of $P(f)$ and $p(t)$ are shown in Fig. 11. Note that $\rho = 0$ corresponds to the ideal LPF. $p(t)$ consists of two factors — the first corresponds to ideal LPF and the second that decreases as $1/|t|^2$ for large $|t|$. The first factor causes zero crossings of $p(t)$ at the desired sampling instants $t = nT$. The second factor reduces the timing errors and hence ISI; obviously, ISI decreases with increasing ρ.

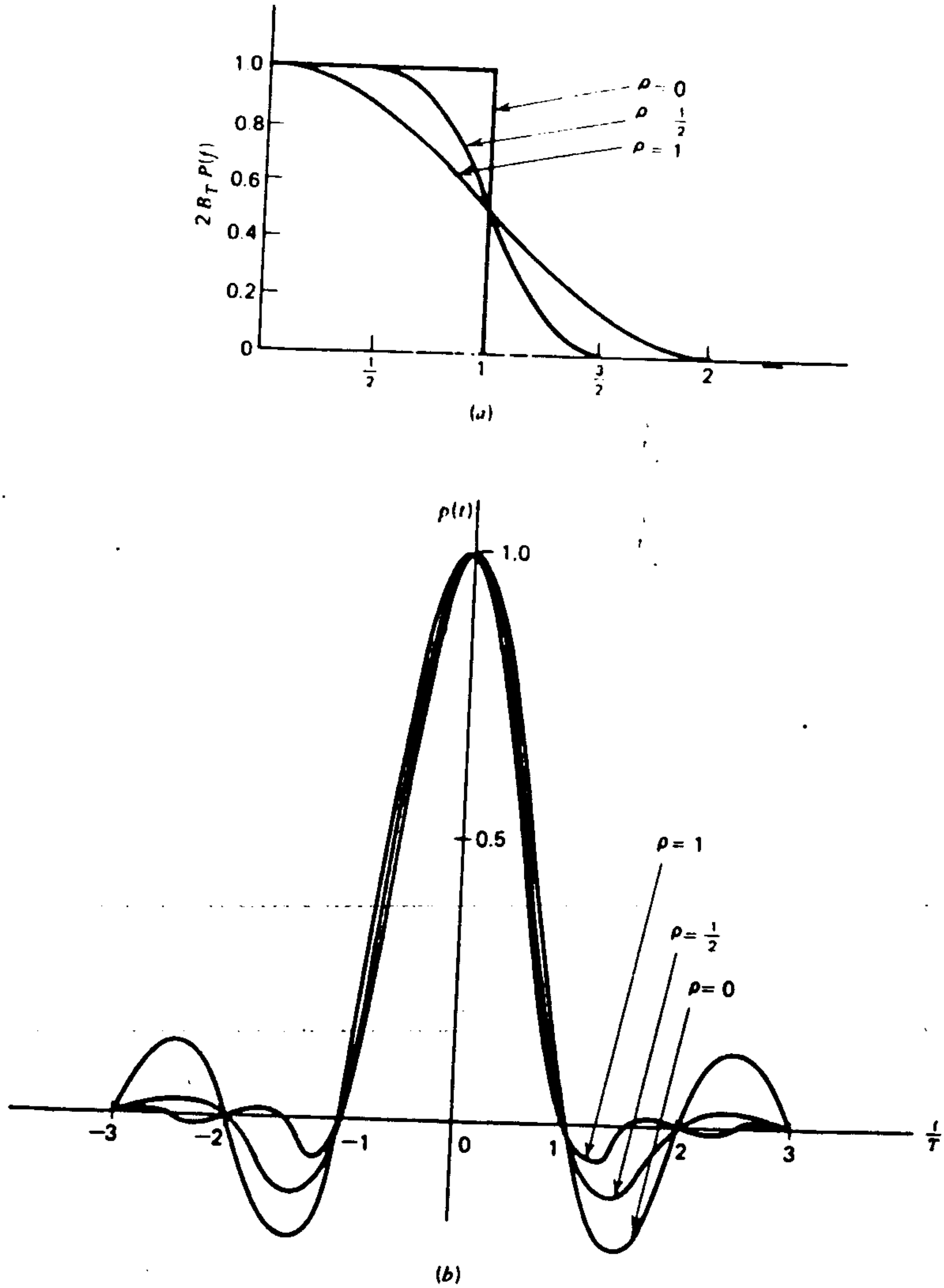


Fig. 11
Responses for different rolloff factors: (a) Frequency response,
(b) Time response

For $\rho = 1$,

$$p(t) = \frac{\operatorname{sinc} 4(B_T t)}{1 - 16 B_T^2 t^2} \tag{26}$$

which exhibits two interesting properties viz. (1) $p(t) = 1/2$ at $t = \pm T/2 = \pm 1/(4B_T)$ i.e. the pulse width measured at half amplitude is exactly equal to the symbol duration T, and (2) there are additional zero crossings at $t = \pm 3T/2, \pm 5T/2, \ldots$ that can be used in the receiver to generate a timing signal for the purpose of synchronization. However, the price to be paid for these desirable features is that the transmission bandwidth is double of that for the case $\rho = 0$.

As an example, consider a TDM system for 24 independent voice sources using PCM. Since voice frequency signal occupies the band 300 – 3400 Hz, a sampling rate of 8 KHz is sufficient for realizable filtering. Also, listening tests confirm that 128 quantizing levels are needed for good quality of voice transmission. Hence 7 bits are needed with binary coding. In addition to voice, telephone transmission system must also pass a d.c. for supervision and addressing, and 20 Hz a.c. for ringing. To provide for this signalling information, an extra bit is multiplexed to each code word associated with a voice signal. Furthermore, another bit is to be added at the end of the frame for synchronization.

For N message sources, with each contributing a code word of n bits to a frame, the total number of bits in a single frame would therefore be $nN + N + 1$. Thus the bit duration T_b is

$$T_b = \frac{T_s}{nN + N + 1} \qquad (27)$$

where T_s is the sampling period. In this case, we have $T_s = 125\ \mu s$, $n = 7$ and $N = 24$. Thus a frame consists of 193 bits, and

$$T_b = 0.647\ \mu s$$

Assuming ideal LPF characteristic for the transmission channel, the minimum transmission bandwidth B_T is

$$B_T = \frac{1}{2T_b} = 772\ \text{KHz}$$

Using cosine rolloff with $\rho = 1$, we need $B_T = \frac{1}{T_b} = 1.544$ MHz

In contrast, the minimum bandwidth requirement for a corresponding FDM system is only 96 KHz!

NOISE IN PCM SYSTEMS

Two major sources of noise are i) transmission noise introduced in the channel, and (ii) quantizing noise introduced in the transmitter. We consider them separately. The effect of transmission noise is to introduce bit errors, causing the receiver occasionally to make wrong decisions by reading a digit 1 when actually 0 was transmitted, and vice versa. The average probability of error is given by

$$P_e = \frac{1}{2}\,\mathrm{erfc}\left(\frac{1}{2}\sqrt{\frac{\gamma}{2}}\right) \qquad (28)$$

where erfc is the complementary error function

$$\mathrm{erfc}(u) = \frac{2}{\sqrt{\pi}}\int_u^{\infty} \exp(-z^2)\,dz \qquad (29)$$

and γ is the peak pulse-to-noise ratio. There is a fairly defined error threshold of about 20 dB, corresponding to an average error of 1 in 10^6 bits. Above threshold, the effect of transmission noise is practically negligible.

Assuming a quantizer of uniform step size, and all quantizing levels to be equally likely, we find that the mean square value of quantizing noise is equal to $\delta^2/12$, where δ is the step size. The spectral density of quantizing noise is approximately flat. When the channel is noise free, and the step size is small enough (no. of levels ≥ 64), the output SNR is given by

$$10\log_{10}(\mathrm{SNR})_o = 6n - 7.2 \text{ dB} \qquad (30)$$

where n is the number of bits in a code word representing a sample of the signal.

Quantizing noise limits the information capacity of a PCM system to

$$C = B_T \log_2\left(1 + \frac{12P}{K^2 N}\right) \text{ bits/sec} \qquad (31)$$

where K is a constant typically equal to 10, B_T is the transmission bandwidth, P is the mean transmitted signal power, and N is the mean noise power. The relation shows that in a PCM system, power and bandwidth are exchanged on a logarithmic basis

DIFFERENTIAL PULSE CODE MODULATION

When a voice or video signal is sampled at a rate slightly higher than the Nyquist rate, successive samples are found to be significantly correlated. When these are encoded in a PCM system, the resultant signal contains redundant information. By removing this redundancy before encoding, we obtain a more efficient coded signal

For this purpose, we use a process called <u>prediction</u>. Let the baseband signal $m(t)$ be sampled at the rate $1/T_s$ and that at $t = nT_s$, we have available the sequence

$$m(nT_s - NT_s), \ \ldots \ , m(nT_s - 2T_s), \ m(nT_s - T_s), \ m(nT_s)$$

A linear prediction of $m(T_s)$, based on the past N samples, is

$$\hat{m}(nT_s) = \sum_{i=1}^{N} w_i \, m(nT_s - iT_s) \qquad (32)$$

where the weights w_i are constant. A realization of (32) is shown in Fig. 12 where the upper sequence of delays forms a tapped delay line and $\{w_i\}$ defines the impulse response of the tapped delay line filter. Thus as each sampling time approaches, the filter output yields the likely value of the next sample. In general, of course $\hat{m}(nT_s)$ will differ from $m(nT_s)$

In differential pulse code modulation (DPCM), the difference between $m(nT_s)$ and $\hat{m}(nT_s)$ is quantized, encoded and transmitted. To construct the original signal, the receiver must make the same prediction as the transmitter, and then add the same correction. Thus a DPCM wave can be generated by using the scheme shown in Fig. 13.

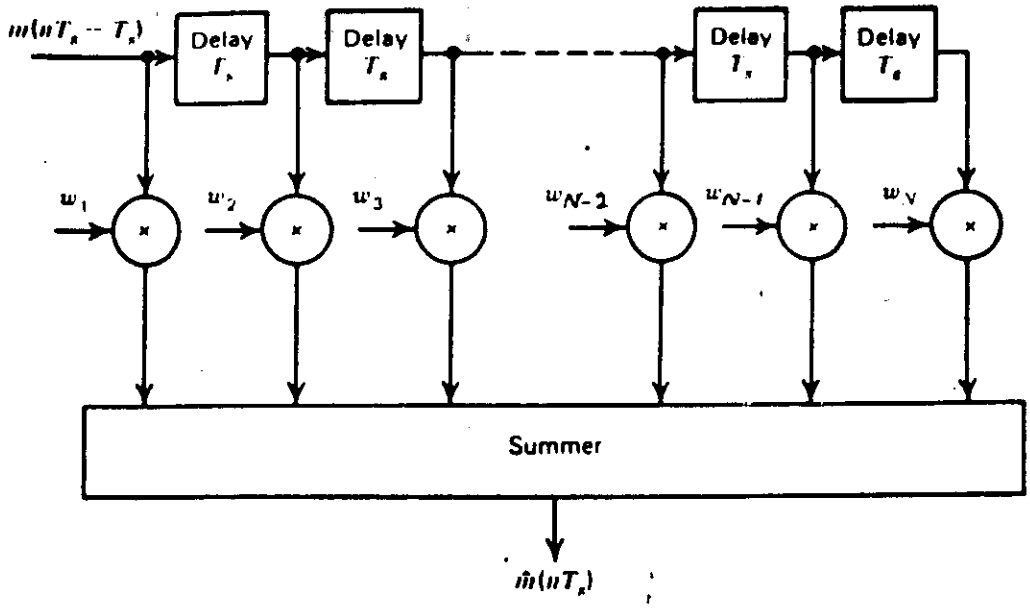


Fig. 12
Tapped delay line filter

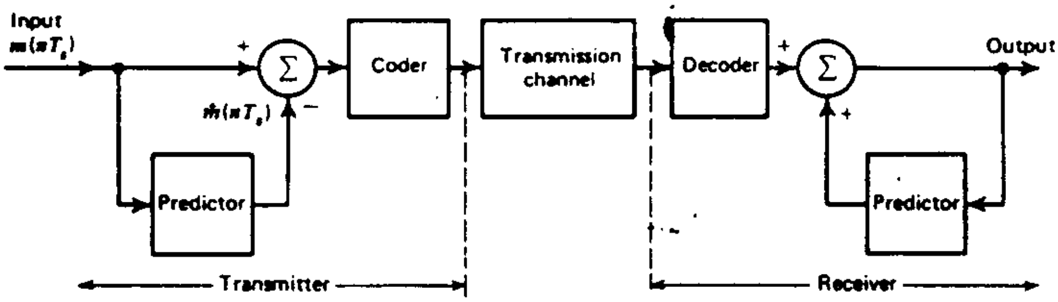


Fig. 13
DPCM system using prediction from input signal

DELTA MODULATION

Delta modulation is a 1-bit (2 level) version of DPCM, in which the baseband signal is highly oversampled — so as to purposely increase the correlation between adjacent samples. In its simplest form, DM provides a staircase approximation to the oversampled signal, as shown in Fig. 14. The difference

between the input and the approximation is quantized into only two levels viz. $\pm\delta$, corresponding to +ve and −ve differences, respectively. If the approximation falls below (lies above) the signal at any sampling epoch, it is increased (decreased) by δ.

With the symbols as in Fig. 14, DM can be formalized by the following equations:

$$b_n = \operatorname{sgn}\left[m(nT_s) - m_a(nT_s - T_s) \right] \qquad (33)$$

$$m_a(nT_s) = m(nT_s - T_s) + \delta b_n \qquad (34)$$

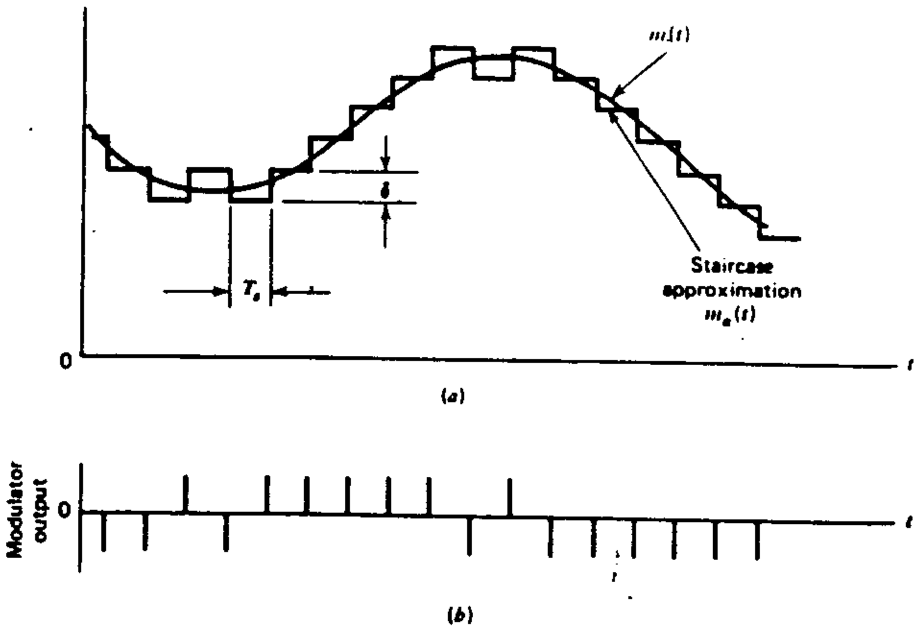


Fig. 14
Delta Modulation

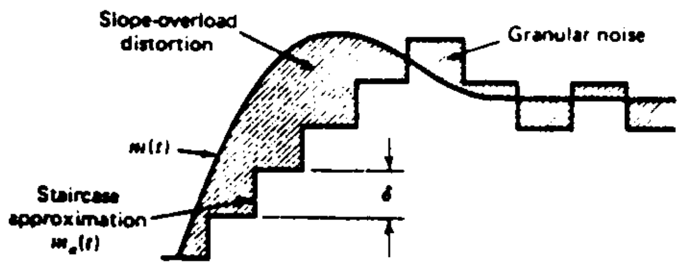


Fig. 15
Quantizing error in DM

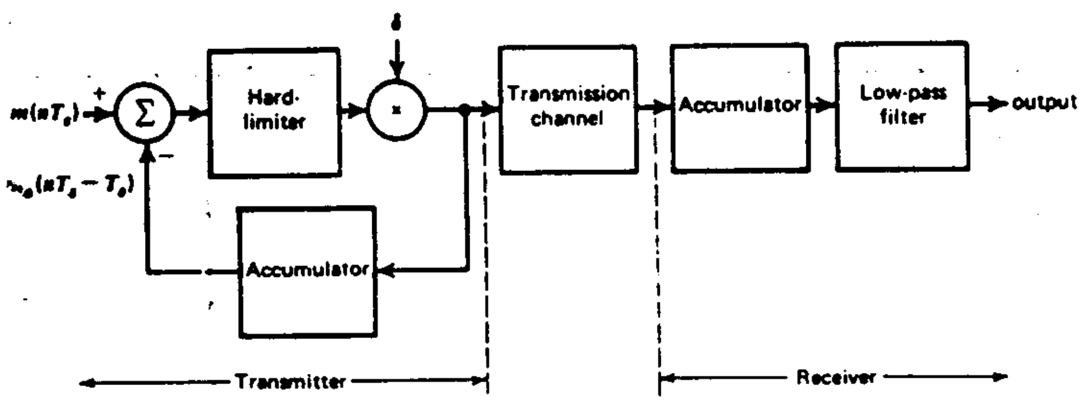


Fig. 16
Delta Modulation System

For each signal sample, the transmitted signal is the single bit b_n, and the rate of information transmission is simply the sampling rate $1/T_s$, as shown in Fig 14(b).

DM is subject to two types of quantizing error — slope overload distortion and granular noise, both of which are illustrated in Fig. 15. The former occurs when δ is too small to follow the high rate of change of $m(t)$, while the latter occurs if δ is too large for the small rate of change of $m(t)$.

The principal virtue of a DM is its simplicity, as demonstrated in the block diagram of Fig. 16.

From Fig. 17, we have

$$m_q(nT_s - T_s) = \sum_{i=1}^{n-1} \delta b_i \qquad (35)$$

At every sampling instant, the accumulator increments the approximation by a step δ in the direction of b_i.

In the receiver, $m_q(t)$ is reconstructed by passing the received sequence of +ve and -ve pulses through an accumulator in a manner similar to that used in the transmitter. The out of band quantizing noise in the high frequency staircase waveform $m_q(t)$ is rejected by passing it through a LPF with a BW equal to the original signal BW.

SOURCE

S. Haykin, Communication Systems, John Wiley, 1983.