TECHNICAL INSTRUCTION

GP.2

PULSE CODE MODULATION

CONTENTS

SECTION 1

INTRODUCTION TO PULSE MODULATION

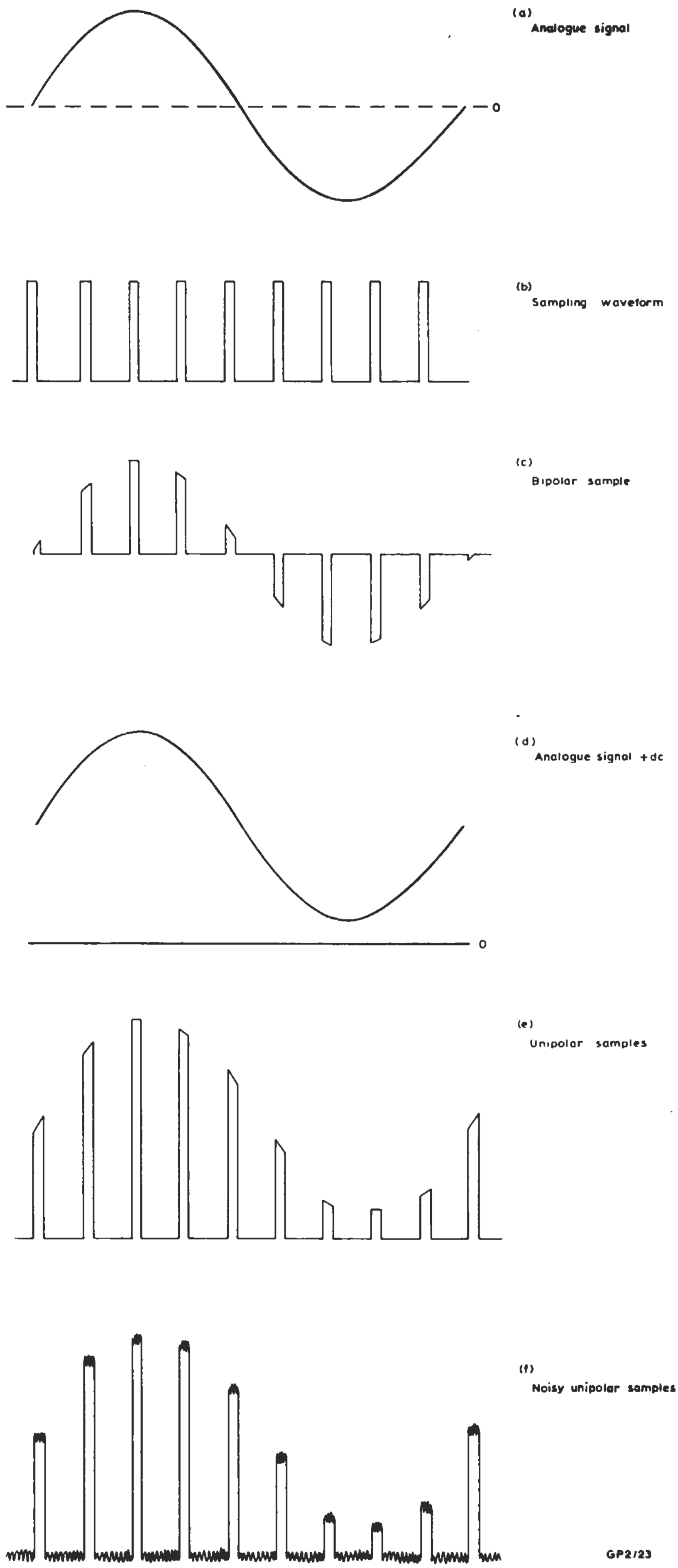

Fig. 1.1. Pulse Amplitude Modulation Waveforms

1.1 General

Modulation is the process by which the characteristics of one electrical signal are impressed upon those of another. In broadcasting, for example, the amplitude of a sound or video signal is used to control the amplitude or frequency of the r.f. carrier wave radiated from the transmitter.

Similarly a train of regular pulses may be regarded as a carrier wave and modulation is possible by varying one of the characteristics of the pulses. Some of the modulation systems so obtained have important advantages over a.m. and f.m.

There are in general three types of pulse modulation:

Pulse amplitude modulation (p.a.m.)
Pulse time modulation (p.t.m.)
Pulse code modulation (p.c.m.)

1.2 Pulse Amplitude Modulation

As the name implies, the amplitude of the carrier pulses is varied in accordance with the amplitude variations of the analogue signal. This is shown in Fig. 1.1 where f_a, waveform (a), is the analogue signal and f_s, waveform (b) is the carrier. The result of modulating f_s with f_a is a train of bipolar p.a.m. pulses as shown in waveform (c). If a d.c. component is added to the analogue signal before modulation takes place, as shown in waveform (d), the result is a train of unipolar p.a.m. pulses as shown in waveform (e). The effect of a small amount of noise in a transmission path on the bipolar p.a.m. signal of waveform (e) is shown in waveform (f).

The only advantage of p.a.m. (an advantage shared with other forms of pulse modulation) is that it enables several signals to be interlaced for transmission over a common path. Because each signal has exclusive use of the path for short periods of time, this process is known as *time-division multiplex (t.d.m.).*

Using p.a.m. gives no improvement in the signal-to-noise ratio of the received signal over that obtained when the analogue signal itself is sent along the same path. In fact the signal-to-noise ratio of the p.a.m. signal will be worse than that of the unmodulated analogue signal unless the noise at the zero voltage level (i.e. between pulses) is gated out at the receiver.

Any limitation in the high-frequency response of a transmission path which is carrying p.a.m. time-division-multiplex signals causes distortion of the pulse shape and, if this distortion causes adjacent pulses to overlap, cross-talk results. Cross-talk can also be caused by low-frequency distortion on the transmission path. The bandwidth required to avoid cross-talk in a p.a.m. time-division-multiplex system is greater than that required for good-quality p.c.m.

Note that p.a.m. is an intermediate step in the production of p.c.m.

1.3 Pulse Time Modulation

There are two variations of pulse time modulation. These are:

(a) *pulse duration modulation (p.d.m.)* also known as *pulse width modulation (p.w.m.)*
(b) *pulse position modulation (p.p.m.)*

In both systems the pulse timing is varied but the pulse amplitude remains constant.

(a) In p.d.m. (see Fig. 1.2) the duration of the pulses changes in accordance with the amplitude variations of the analogue signal; the timing of the leading edges of the pulses remains constant.

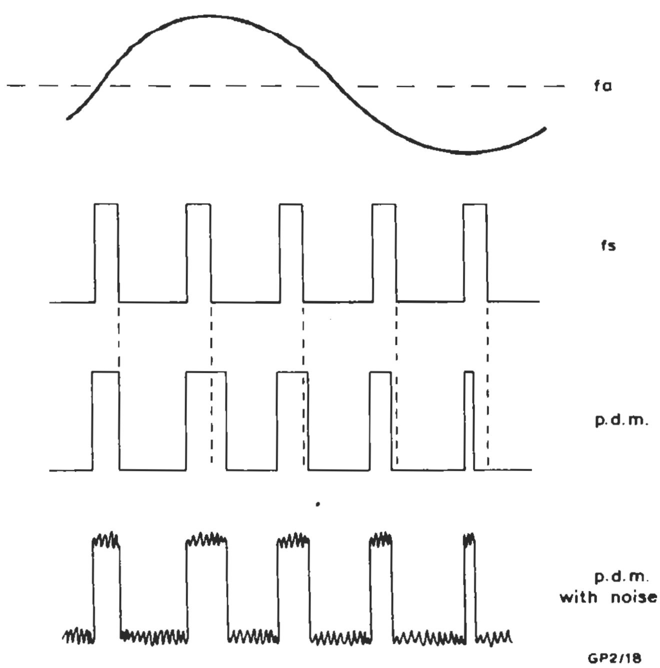

Fig. 1.2. Pulse Duration Modulation Waveforms

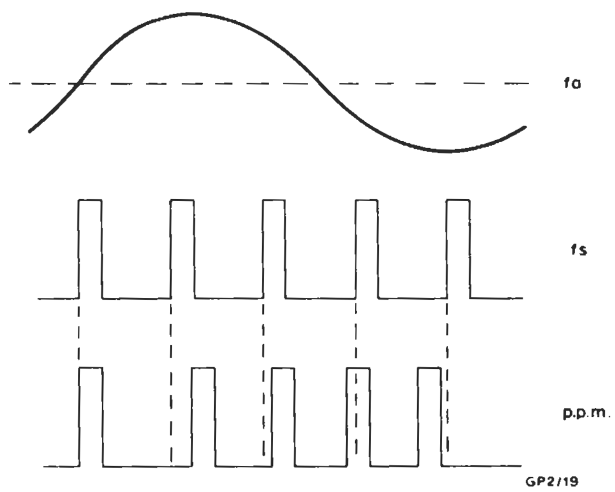

Fig. 1.3. Pulse Position Modulation Waveforms

(b) In p.p.m. (see Fig. 1.3) the pulse duration remains constant but the position (timing) of the leading edges of the pulses changes in accordance with the amplitude variations of the analogue signal.

An advantage which p.t.m. possesses over p.a.m. is its immunity to noise on the transmission path provided that the bandwidth is wide enough to preserve the shape of the pulses. This advantage is illustrated (for p.d.m.) in the last waveform of Fig. 1.2 which shows that noise has no effect on the pulse duration if the pulse edges are vertical. Unfortunately, preserving vertical edges requires an infinite bandwidth and so the noise immunity of the system is only of academic interest. When the bandwidth is restricted the pulse edges do not remain vertical and any noise effects the pulse duration. Thus the signal-to-noise ratio is decreased as the bandwidth is decreased.

SECTION 2

PRODUCTION OF A PULSE–CODE–MODULATED SIGNAL

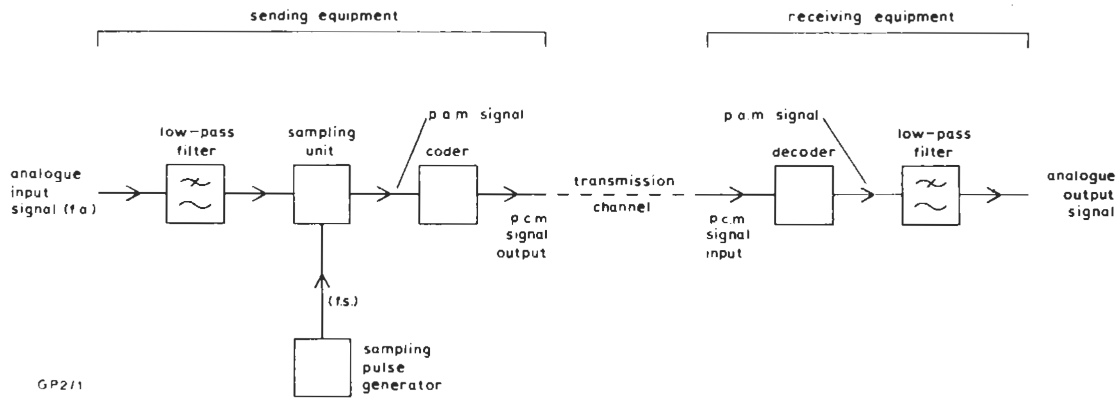

Fig. 2.1. Block Diagram of a P.C.M. System

There are, basically, three processes in the production of a p.c.m. signal. These are:

Sampling
Quantising
Coding

The sampling process turns the analogue signal into a pulse-amplitude-modulated (p.a.m.) signal; the quantising process gives each sample a discrete amplitude and the coding process turns the quantised levels into a p.c.m. signal. The processes are shown in block diagram form in Fig. 2.1 and are described in the remainder of this section.

2.1 Sampling

2.1.1 Sampling Theory

Sampling is the name given to the process of converting an analogue signal into a sequence of pulses, the amplitude of which vary in accordance with the magnitude of the analogue signal at discrete intervals of time. The sampling process is illustrated in Fig. 2.2 where a signal source V_1, waveform (a), is connected through a switch S to a load R. The switch closes periodically, under the stimulus of sampling waveform (b) and when it does so the signal source is connected to the load to produce a train of pulses as shown by waveform (c). The p.a.m. signal shown by waveform (c) is a bipolar signal; i.e. the pulses are ranged on both sides of the zero line. To produce a unipolar p.a.m. signal, a d.c. component must be added to the analogue signal as shown in waveform (d); the result of sampling (d) with (b) is shown in (e).

If the analogue signal is a sine-wave of frequency f_a and the sampling frequency is f_s then the sampling process produces a spectrum of sine waves consisting of:

1. zero frequency (d.c.)
2. analogue frequency (f_a)
3. sampling frequency plus upper and lower sidebands ($f_s - f_a$, f_s, $f_s + f_a$)
4. twice sampling frequency plus upper and lower sidebands ($2f_s - f_a$, $2f_s$, $2f_s + f_a$) etc.

The sine waves produced by the sampling process are represented on a frequency scale in Fig. 2.3(a). The relative amplitudes of the sine waves depend on the width of the sampling pulses. The greater the pulse width the larger the amplitude of the f_a component.

Text continued on page 2.4

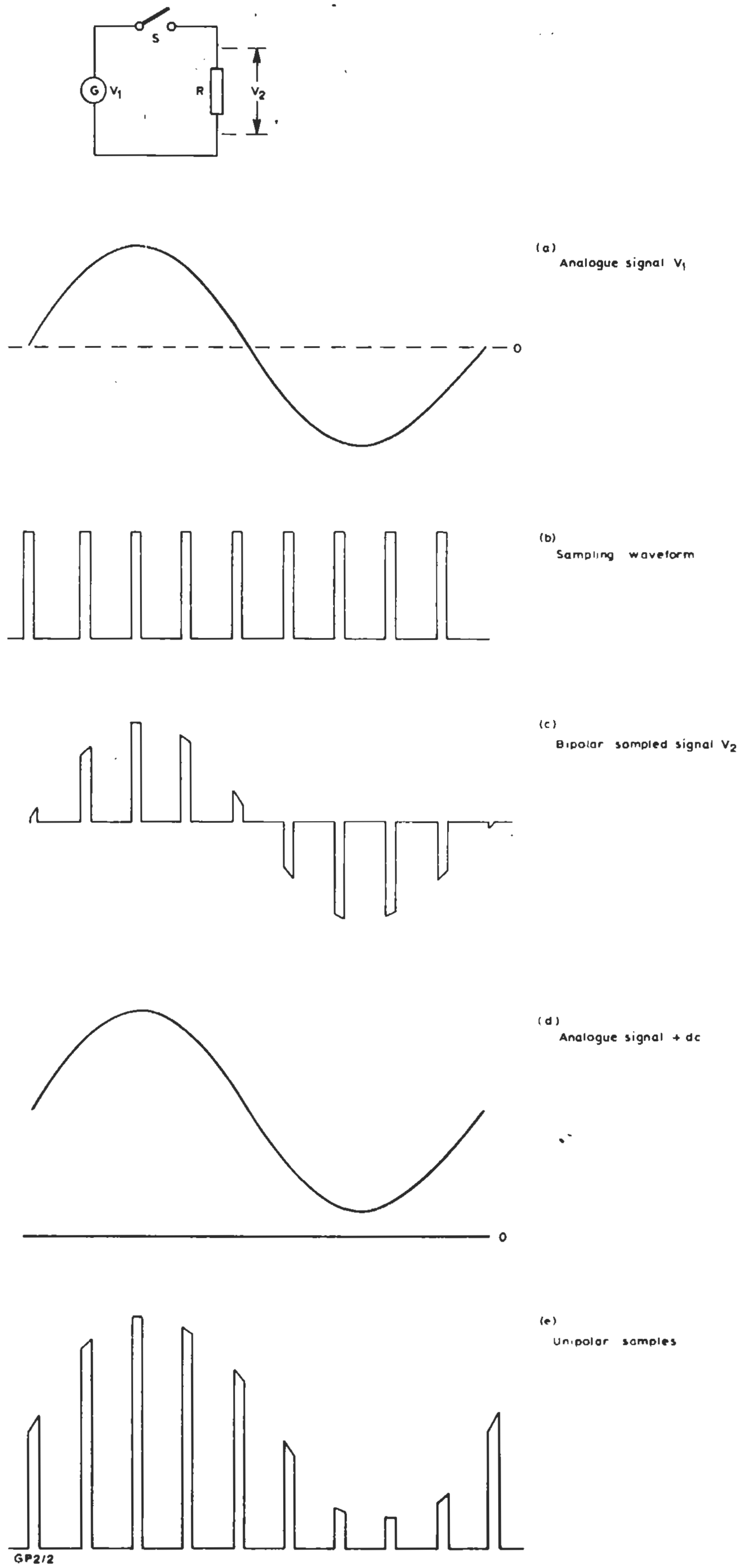

Fig. 2.2. The Sampling Process

If the sampling frequency for a given analogue signal is arranged to be twice the highest frequency contained in the analogue signal, the resulting frequency spectrum is as shown in Fig. 2.3(b).* The highest analogue frequency and the extremity of the lower sideband of f_s coincide when f_s equals twice f_a max. If f_s is less than twice f_a max., the lower sideband of f_s overlaps the higher frequencies of f_a and it is impossible to extract an undistorted analogue signal at the demodulator. In the demodulator, the analogue signal is recovered by means of a low-pass filter which rejects all frequencies above f_a max. For f_s to be twice f_a max, the filter characteristic must have an ideal shape. However, because of the limitations of practical filters, it is necessary to make f_s from 2·2 to 2·3 times f_a max. The resulting frequency spectrum is shown in Fig. 2.3(c).

Thus f_s must be at least twice f_a. It could be much higher than this but a high value offers no advantages and indeed makes the production of a p.c.m. signal more difficult by reducing the time between samples and increasing the bit rate of the system.

The relationship $f_s = 2{\cdot}3 f_a$ applies regardless of whether f_a is a telephone signal, a high-quality sound signal or a video signal. In practical systems the following figures have been used:

System	f_a *(max.)*	f_s
British Post Office	3·4 kHz	8 kHz
BBC sound-in-syncs	14 kHz	31·25 kHz
BBC video p.c.m.	5·5 MHz	13 MHz

To summarise: If the analogue signal is sampled at a rate which is slightly greater than twice the highest frequency component of that signal, the original signal can be completely recovered by means of a low-pass filter.

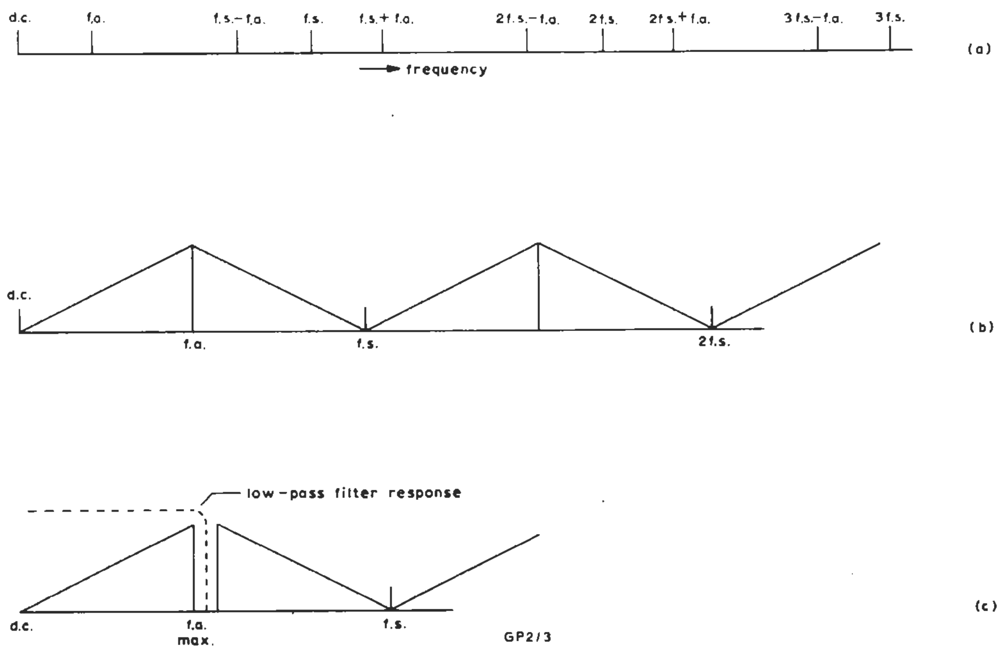

Fig. 2.3. Frequencies Produced by Sampling

2.1.2 Sample and Hold

Coding the individual samples of a p.a.m. signal into digital form cannot be done instantaneously; therefore the initial amplitude of each sample must be held constant whilst coding takes place. The sample-and-hold process is illustrated in Fig. 2.4. When the switch is closed (i.e. during the sample-pulse period) the capacitor is charged to a potential which represents the sample amplitude; when the switch is open the charge is retained except for a slow leakage through the very high input impedance of the second amplifier. Thus the amplitude of the analogue signal

*The convention used in Figs. 2.3(b) and 2.3(c) is that the height of a triangle at any point is a measure of the analogue frequency concerned. For instance, the triangle that starts at d.c. represents the whole range of modulating frequencies up to the maximum value of f_a. See BS 3939 Section 25.

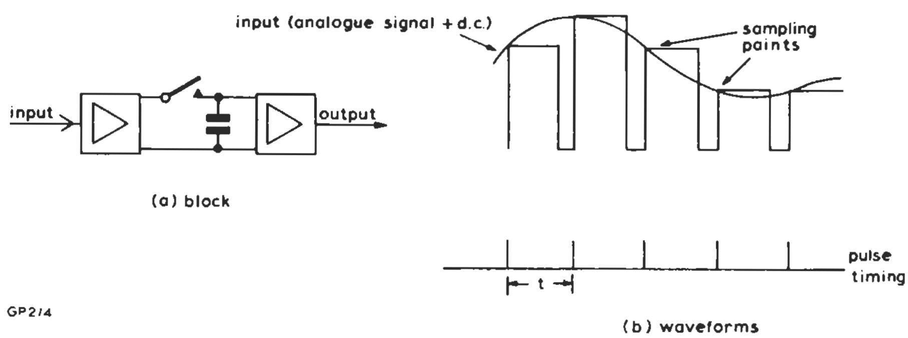

Fig. 2.4. The Sample-and-hold Process

at each sampling period is held constant for most of the time interval (t) between sampling periods and it is during this period that the assessment of amplitude is made. Immediately prior to each sample-pulse period the capacitor is discharged

2.2 Quantising

2.2.1 Quantising Level

The full amplitude range of an analogue signal can be divided into a number of equal divisions, or levels, and each level can be assigned a reference number (e.g. a number in binary form). These discrete levels are known as *quantising levels.*

In the four-bit binary system shown in Fig. 2.5 the number of possible quantising levels is $2^4 = 16$. The binary numbers corresponding to each level are shown on the diagram. Only rarely are samples exactly equal to one of the quantising levels and so the level that is used for coding is the level below the sample amplitude. For example, in Fig. 2.5 the first sample is just above level 1000 and so it is coded as 1000; the last sample just fails to reach level 0100 and so it is coded as 0011.

The differences between the sample amplitudes and the quantising levels used for coding cause errors in the reproduced signal. The maximum error possible is equal to one quantising step (i.e. the distance between adjacent quantising levels) and so is proportionally smaller for a large-amplitude signal than for a small-amplitude one. Thus quantising errors are subjectively more noticeable on small signals than on large ones.

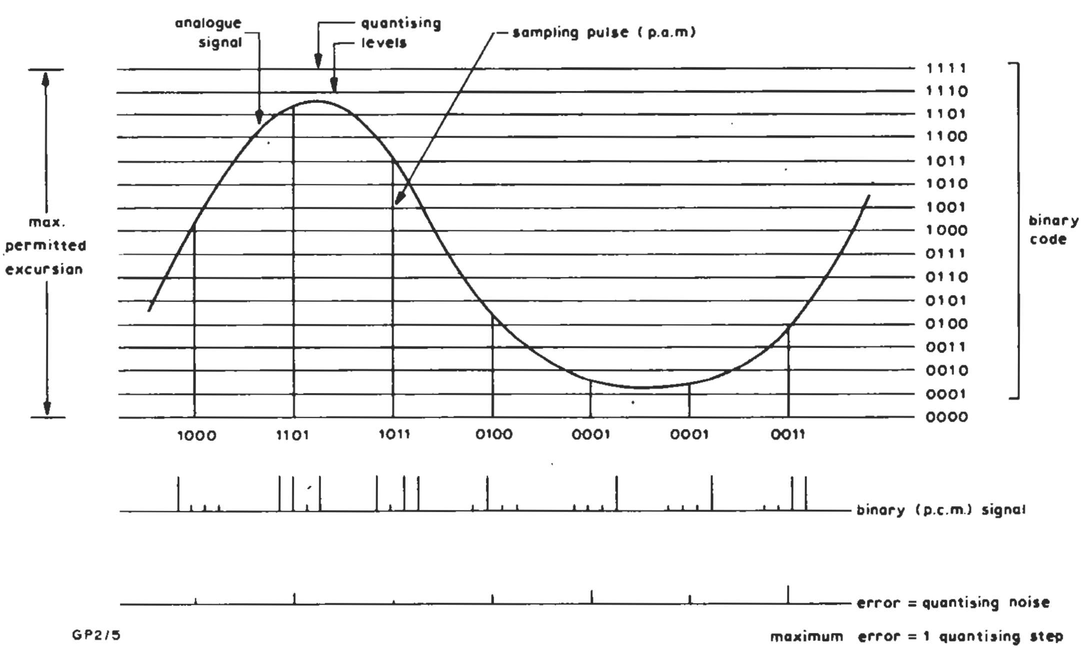

Fig. 2.5. Binary Coding for a Four-bit P.C.M. System

2.2.2 Quantising Noise

The error so introduced by the quantising process is known as *quantising noise.* For large amplitude signals (i.e. signals which cover a large number of quantising levels) the subjective effect of quantising noise resembles that of white noise. However, if the signal amplitude falls so low that only a few of the available quantising levels are used, the subjective effect of quantising noise resembles non-linear distortion and is known as *granular distortion.*

The greater the number of quantising levels used, the smaller the quantising error and the lower the quantising noise level. If the number of levels is doubled the quantising noise is halved, and to double the number of levels the binary code used must be increased by one bit. For example; a five-bit code gives $2^5 = 32$ levels, a six-bit code gives $2^6 = 64$ levels, a seven bit code gives $2^7 = 128$ levels and so on.

The maximum quantising error is one quantising step and the maximum signal amplitude covers all quantising levels; therefore the peak signal to peak quantising noise ratio is 2^n (where n is the number of bits in the binary code) because 2^n is the total number of levels. Table 2.1 gives the peak signal to peak quantising noise ratios for p.c.m. systems with from 10 to 13 bits.* When the quantised signal covers a large number of levels, quantising noise is subjectively similar to white noise. This assumption makes it possible to specify p.c.m. signal-to-noise ratio in terms of peak-signal to peak-weighted noise and so obtain figures (also given in Table 2.1) which can be easily measured on standard BBC equipment.

*This range covers most broadcast-quality sound systems. Telephony p.c.m. systems use only 7 bits (for a 3-kHz audio bandwidth) but employ non-linear quantising to improve the signal-to-quantising-noise ratio.

In the absence of signals quantising noise ceases. In practice a small amount of noise, such as hum, is likely to be present at the input to a p.c.m. system when signals are absent; if this is sufficient to cause the analogue-to-digital converter to jump intermittently from one quantising level to the next, corresponding voltage steps will appear at the receiving end of the system and produce what is known as *idling noise.* Because of its discontinuous nature idling noise is more objectionable to the ear than the continuous background hiss produced by quantising noise. However, if quantising noise can be reduced so that it is barely audible, idling noise is negligible.

Fig. 2.5 shows linear quantising; i.e. quantising in which all the steps are of equal amplitude. However, some p.c.m. systems use non-linear quantising in which the quantising levels are spaced so that small-amplitude signals cover more levels than they would for linear quantising. The purpose of non-linear quantising is to improve the signal-to-noise ratio of small-amplitude signals. For details see Section 3.

2.3 Coding

Coding enables information on the quantised samples to be conveyed over a transmission path without degradation. The coding process turns each sample into a digital number which is transmitted as a group of pulses.

TABLE 2.1

Signal to quantising noise Ratios in Binary P.C.M. Systems

Number of digits	*Number of quantising levels*	*Peak signal/peak quantising noise in dB*	*Peak signal/peak weighted quantising noise in dB*
10	1024	60	53
11	2048	66	59
12	4096	72	65
13	8192	78	71

2.3.1 Coding Systems

In any given p.c.m. system the number of different states that each pulse can assume is referred to as the *base* of the code for that system. The code is characterised by the number of pulses *n* corresponding to each signal sample and the number of states *m* which each pulse can assume.

The simplest form of p.c.m. uses a binary code in which the base *m* is equal to 2 and information is conveyed by the presence or absence of a pulse at a given instant of time; the presence of a pulse corresponds to 1 and the absence of a pulse corresponds to 0. Another form of p.c.m. uses a ternary code with a base of 3 in which the pulses can have three possible amplitudes; 1, 1/2 or 0.

Binary-based systems can tolerate higher noise levels than ternary-based systems because, with binary-coded signals, the decoder at the receiving end of the system has only to distinguish between the presence and absence of a pulse. On the other hand, with the ternary system, the decoder must be able to distinguish three separate levels and is thus more susceptible to noise. The only advantage that the ternary system has over the binary system is that it requires less bandwidth, because pulses are transmitted at a slower rate than that required for the binary system. An example which compares the binary and ternary systems for signals with approximately the same number of quantising levels is given below.

The number of quantising levels in a binary system is 2^n and in a ternary system it is 3^n, where n is the number of digits per sample.

If n = 8 for the binary system and
n = 5 for the ternary system

then $2^8 = 256$ and $3^5 = 243$; i.e. the number of levels is nearly the same. Bandwidth is proportional to digit-rate (the rate at which binary digits are transmitted) and so both the digit-rate and the bandwidth for the ternary system are 5/8 of that required for the binary system. Because the number of levels used in each system is nearly the same the quantising noise is similar for both systems. Thus the reduced bandwidth of the ternary system is accompanied by a reduction in the noise level at which decoding errors occur; in other words bandwidth is exchanged for noise immunity.

The comparisons made between the binary and ternary based systems apply equally to codes with higher bases than 3. In fact, because of their increased susceptibility to noise, higher based codes are seldom used.

Both the BBC and the British Post Office* use binary coded p.c.m. systems, therefore all subsequent description deals with the binary system only.

2.3.2 Binary P.C.M.

In a binary p.c.m. system the amplitude of each quantised sample is coded into a succession of 1's and 0's, and so information is conveyed by the presence or absence of a pulse. The advantage of using only two levels in the coded signal is that a large amount of noise can be tolerated before errors in decoding occur. This tolerance is illustrated by Fig. 2.6 which shows a four-bit word (1101) both before and after its transmission through a noisy band-limited path. To determine the binary content of the received word the waveform is inspected at the appropriate time to determine the voltage levels, at the intersections at the dotted lines of Fig. 2.6; in this way the decoder can unambiguously determine whether a 1 or an 0 is present. If the received waveform is sampled at its mean height, errors in decoding occur only if the peak noise equals half the pulse amplitude. Note that the shape of the waveform does not matter provided that the presence of a 1 or 0 can be recognised without error.

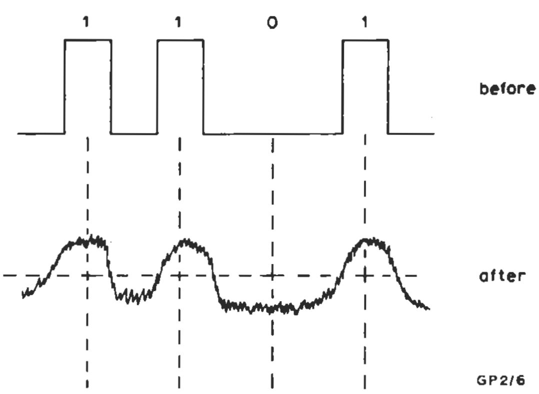

Fig. 2.6. The Effect on a Binary Signal of a Noisy Transmission Path

If the transmission path is so long that the p.c.m. signal is likely to be degraded to the point where errors occur, the signal can be regenerated at an intermediate point and a new set of digits produced which are completely free from extraneous noise. By using this technique the p.c.m. signal can be sent over any path length without being affected by the path imperfections. If, however, the signal-to-noise ratio is degraded to the point where frequent errors occur in the decoding process the system fails catastrophically; there is no gradual rise in the error rate as the noise increases.

*In the British Post Office p.c.m. system alternate 1's are inverted to give a bipolar waveform.
Two advantages of the bipolar system are:
(a) no d.c. component
(b) the pulse train has its principal energy content at half the bit rate with a null at the bit rate; this minimisation of energy at high frequencies helps to reduce cross-talk from adjacent p.c.m. circuits.

2.3.3 P.C.M. Bandwidth

The bandwidth required for a p.c.m. signal depends on the highest frequency component of the analogue signal (f_a max.) and the bit-rate of the p.c.m. signal. The bit-rate depends on:

(a) the sampling frequency f_s

(b) the number of bits required for each sample.

The sampling frequency f_s is related to the highest analogue frequency (f_a max.) by the expression $f_s = 2{\cdot}2\ 2f_a$ max. (see subsection 2.1.1) and the number of bits used is determined by the amount of quantising noise that can be tolerated. For example, a 13-bit system is required for a high-quality p.c.m. sound-distribution system and with this number of bits quantising noise is imperceptible. For high-quality sound f_a max is about 15 kHz; thus the required bit rate is 13 x 2·2 x 15 kHz = 429,000 bits per second.

The bandwidth required for the transmission of this bit rate must be decided bearing in mind that:

(a) the pulses are distorted if the bandwidth is restricted,

(b) a p.c.m. signal can suffer a lot of distortion without the accuracy of the decoding process being affected.

One way of mitigating the effects of limited bandwidth is to pre-form the p.c.m. pulses e.g. to sine squared shape which have negligible energy content outside the passband of the system. No degradation of the pulse shape then occurs during transmission.

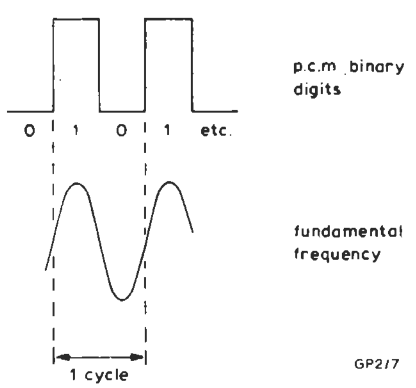

Fig. 2.7. Fundamental Frequency of Alternate 1's and 0's

The highest fundamental frequency arises from a pattern consisting of a sequence of alternate 1's and 0's. This frequency is half the bit rate, as shown in Fig. 2.7. In general terms the bit rate is n x 2·2 x f_a max., where n is the number of bits in the p.c.m. code and f_a max. is the maximum analogue frequency. Thus an analogue signal with a bandwidth of 15 kHz requires a p.c.m. bandwidth of 429 kHz for transmission by means of a 13-bit system. Because the highest frequency pattern obtainable is half the bit rate, the theoretical minimum bandwidth equals n x 1·1 x f_a max; i.e. the p.c.m. bandwidth is just over n times the analogue bandwidth. In practice a bandwidth which is equal to the bit rate is often used e.g. in the sound-in-syncs systems.

Economies in the bandwidth can be obtained by reducing the number of bits per sample and by using special techniques to reduce the effect of quantising noise; these techniques are described in Section 3.

2.3.4 Methods of Coding

There are many ways of coding p.c.m. signals, but all are derived from the basic methods outlined below.

(a) Successive-approximation Coders

Successive-approximation coders, sometimes called serial coders, compare the voltage of the sample to be coded with a number of precise reference voltage levels which are presented successively in descending order of magnitude; each level corresponds to one digit position in the code. A block diagram of a successive-approximation coder is given in Fig. 2.8.

In the first stage of the coder the sample voltage and the maximum reference voltage, V, are both applied to a voltage comparator; if V is less than the sample voltage the comparator gives a logic 1 output for the most significant digit. The reference voltage is then subtracted from the sample voltage and the difference is passed to the second stage of the coder for similar treatment. (If the reference voltage on the first stage is greater than the sample voltage the comparator gives a logic 0 output and the whole of the sample is applied to the second stage.) The reference voltage for the second stage has a magnitude of V/2; subsequent stages have reference voltages of V/4, V/8 and so on. The procedure applied to the first stage is repeated at each subsequent stage until all digit positions have been coded.

A practical coder which uses a variation of the successive-approximation principle is shown in block diagram form in Fig. 2.9. The operation is described below for a sample input with a magnitude of 9·2 units; a four-bit system is used and so the coder can handle a signal with a magnitude of 16 units.

When coding commences, clock pulses are fed to the logic circuits and the maximum reference voltage of 8 units is applied to the comparator. Because the reference voltage is smaller than the sample input the reference voltage is retained and the generator produces a logic 1 output. The retained 8-unit reference voltage is added to the next reference voltage and the sum (8 + 4 = 12) is applied to the comparator. Because 12 is greater than 9·2 a logic 0 is produced by the generator, the 4-unit reference voltage is rejected and the 2-unit reference voltage selected.

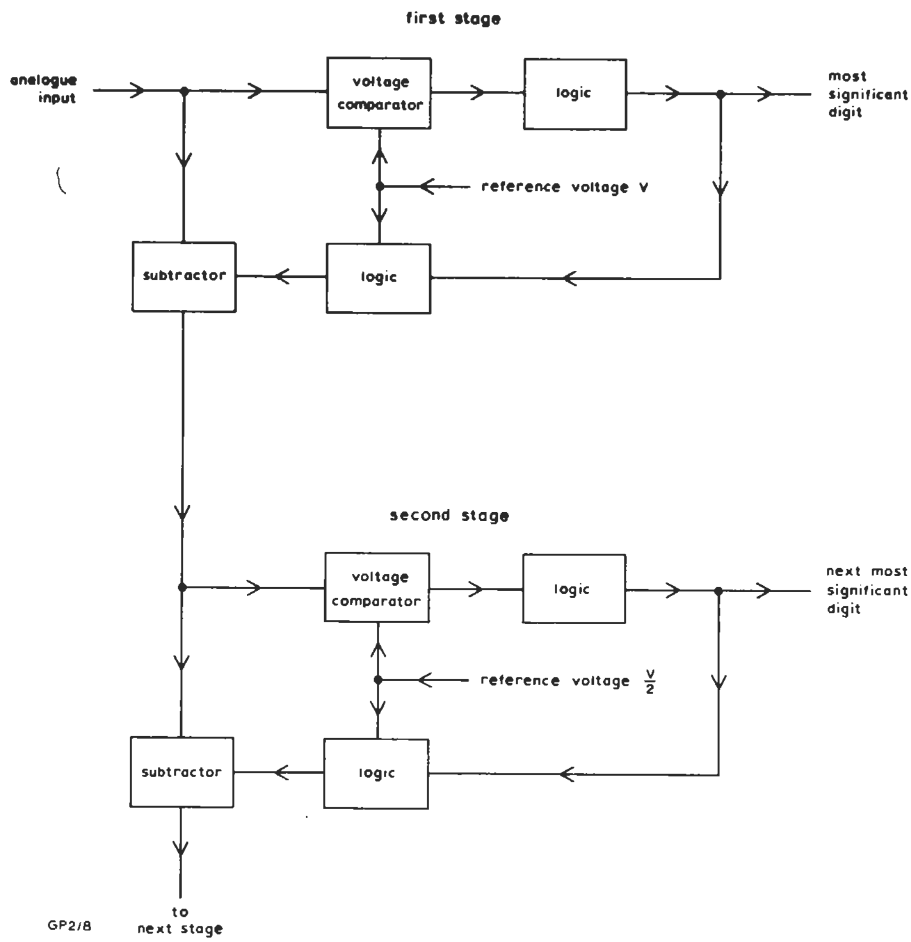

Fig. 2.8. Simplified Block Diagram of a Successive-approximation Coder

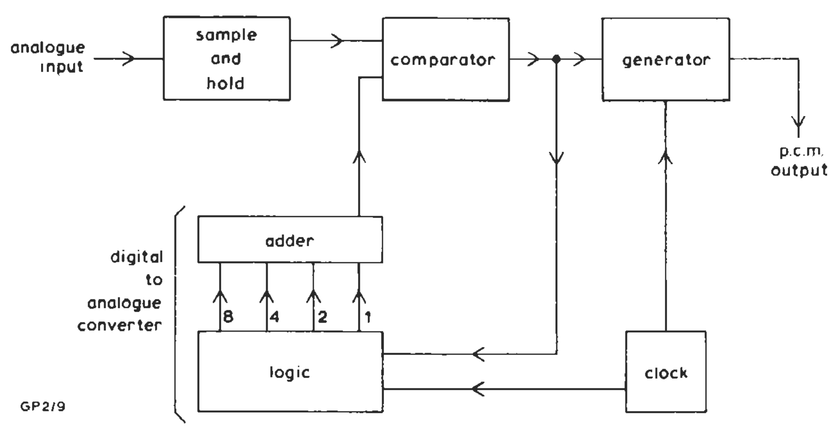

Fig. 2.9. Block Diagram of Practical Successive-approximation Coder showing reference levels

The sum of 8 plus 2 is applied to the comparator and because it is still greater than 9·2 another logic 0 is produced by the generator and the 2 is rejected and replaced by the 1-unit reference voltage. The sum of 8 and 1 is less than 9·2, a logic 1 is produced by the generator and the coding process is completed. The logic output for a 9·2-unit sample is thus 1001; the reference levels for the coding sequence are shown in Fig. 2.10.

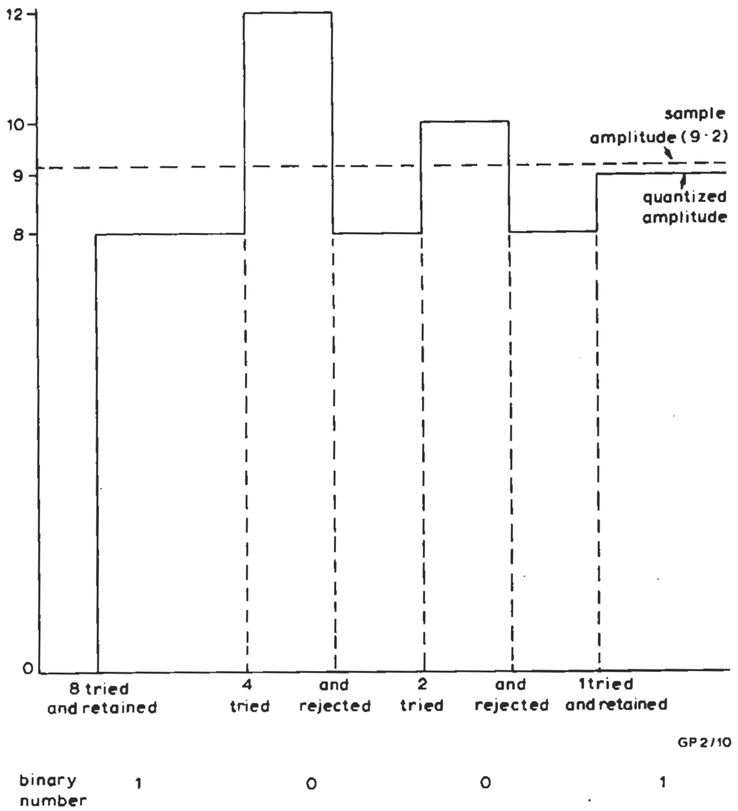

Fig. 2.10. Illustrating the operation of the Coder Shown in Fig. 2.9 for an input of 9.2 V, showing reference levels

(b) Parallel Coders

In this type of coder the sample amplitude is compared simultaneously with a complete set of reference voltages which correspond to all the quantising levels. Parallel coders are not used for high-quality sound signals because of the high degree of instrumental accuracy required and because the large number of comparators needed would make the circuit too complex and expensive.

(c) Hybrid Coders

It is possible to make a hybrid coder which is a combination of the parallel and series types. The hybrid coder needs less equipment than a parallel coder and is faster in operation than a serial (i.e. successive-approximation) one; it produces two or more digits at each comparison instead of the one digit produced by the serial coder. A simplified block diagram of an 8-bit hybrid coder is given in Fig. 2.11.

The coder has two similar stages each of which handles 4 bits in parallel. The first stage determines the state of the four most significant digits and the second stage determines the state of the four least significant digits. In the first stage the sample is fed to 15 voltage-level comparators which are connected in parallel. A logic network converts the comparator outputs into a 4-bit binary code and this represents the four most significant digits of the signal.

To obtain the lower order digits it is first necessary to obtain a 16-level quantised version of the signal and this is done by decoding the four most significant digits. The resulting signal is subtracted from a suitably delayed version of the input sample and the difference signal obtained indicates the amount by which the input sample exceeds the level given by the first four bits. The difference signal (A – B) in Fig. 2.11 is applied to a second set of comparators the outputs of which are converted by logic circuits into a four-bit number which represents the four least significant digits.

Finally, both sets of digits are passed through digit

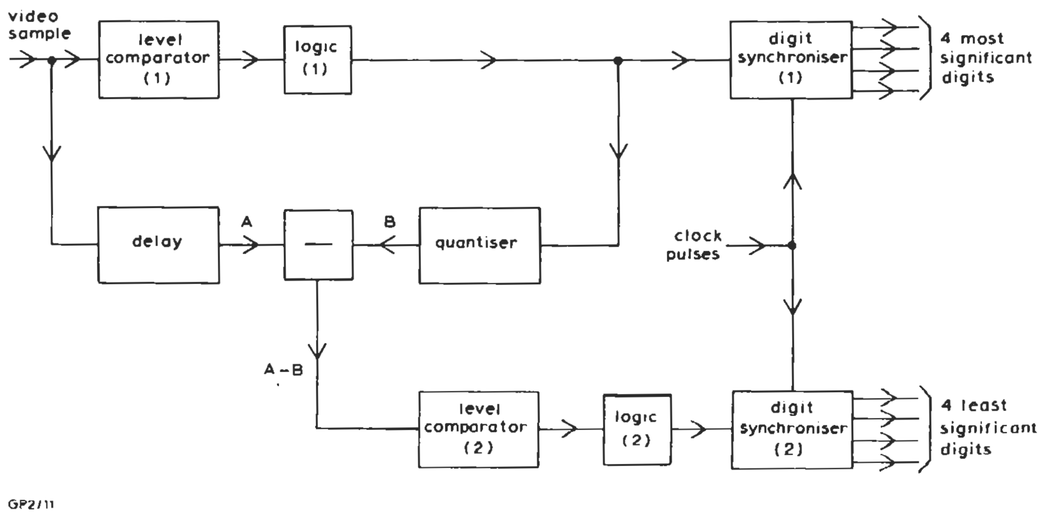

Fig. 2.11. Block Diagram of a Hybrid (Parallel/Serial) Coder

synchronisers which retime them so that all 8 digits corresponding to a given sample appear simultaneously at the output of the coder.

(d) Counter Coders

A simplified block diagram of a counter-type coder is shown in Fig. 2.12. The sample is compared with the voltage developed across a capacitor which is charged from a constant-current source; the combination of capacitor and constant-current generator is known as a ramp generator.

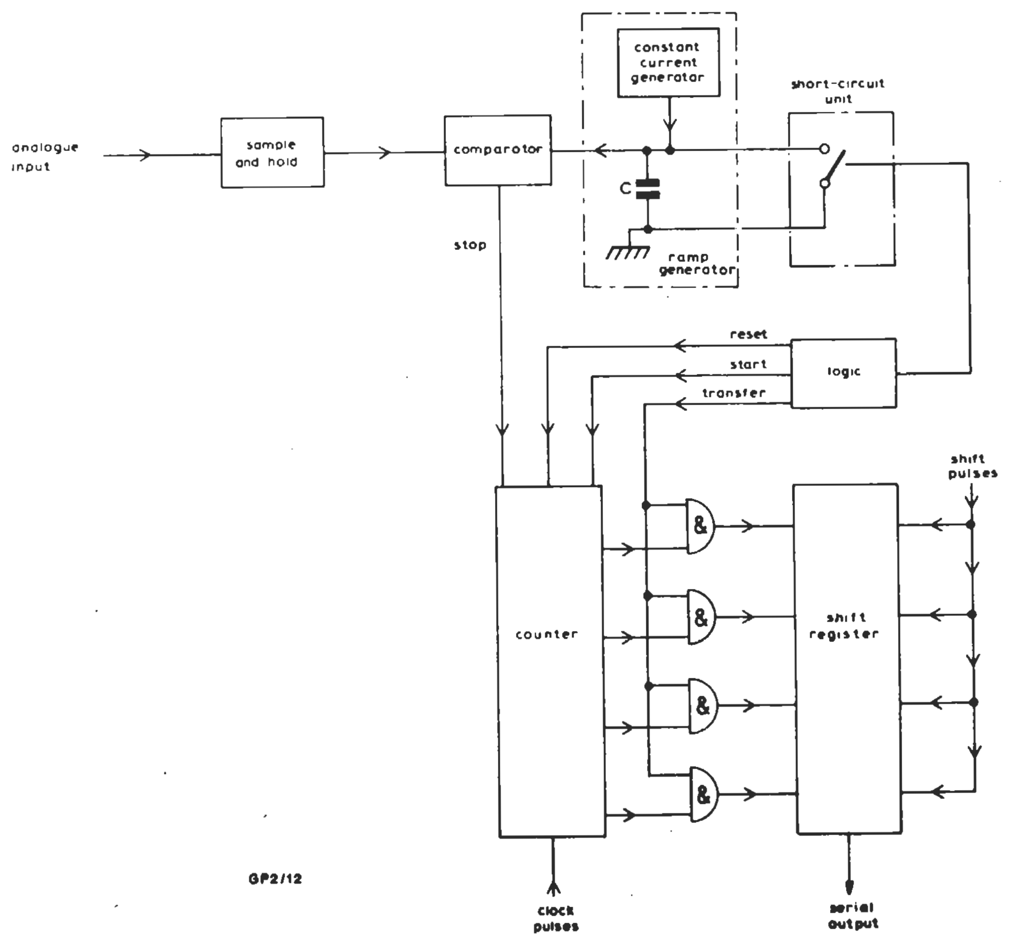

Fig. 2.12. Block Diagram of a Counter Coder

Initially the capacitor is short-circuited and the counter is held in the reset condition (all outputs at logic level 0). When a sample is applied to the comparator, signals from the logic circuits remove the capacitor short-circuit and, simultaneously, allow clock pulses to be fed to the counter. The capacitor charges linearly and the counter accumulates a number which is proportional to the charging time. During the counting process the sample-and-hold circuit ensures that the initial sample amplitude is held constant.

When the capacitor voltage equals the sample voltage the comparator produces a pulse which stops the counter; the number held in the counter when it stops is the quantised binary representation of the sample amplitude. The binary digits held by the counter are fed to the shift register by applying transfer pulses to the *And* gates which link the counter to the register. The counter is then cleared and the capacitor short-circuited in preparation for coding the next sample. The shift register performs a parallel/series conversion and the application of shift pulses to the register causes the digital information to be fed to the output in time sequence.

With counter type coders, current or voltage references requiring accurate adjustment are not used. This advantage is shared by no other coding method.

2.3.5 Methods of Decoding

Two different methods of decoding are in use.

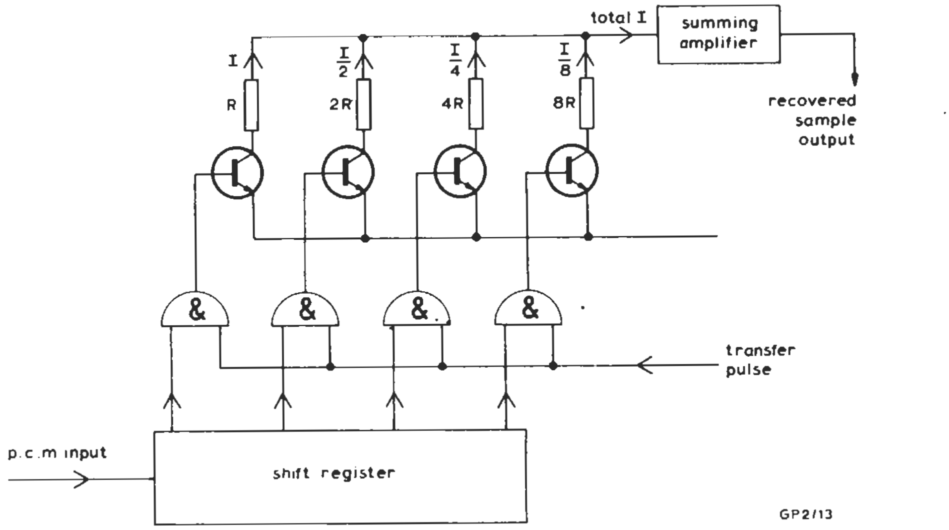

Fig. 2.13. Block Diagram of a Parallel Decoder

(a) Parallel Decoders

The digital number to be decoded is first assembled in a store or shift register. When the number is complete a command pulse applied to the store causes each digit representing 1 to turn on a current proportional to the value signified by that digit. The sum of the currents produced is proportional to the digital number.

The principle of operation is illustrated, for a four-bit code, in Fig. 2.13. The p.c.m. signal is fed serially into the shift register and the four shift register outputs are thereby set to the logic 1 or logic 0 states. A transfer pulse is applied to the associated *And* gates and those transistors which have a 1 applied to their base are switched on. The current passed by each transistor is determined by the value of the collector resistor and the values of these resistors are chosen to give a current proportional to the significance of the digits applied to the associated shift-register output. The most significant digit, if a 1, switches on a current of magnitude I, the next digit a current of magnitude I/2, the third digit a current of magnitude I/4 and the last digit a current of magnitude I/8. The transistors conduct only if a 1 is present. The sum of the currents passed by the transistor stages corresponds to the original sample amplitude and is obtained by use of a summing amplifier.

(b) Counter Decoders

The counter decoder is the reverse of the counter coder and is shown in block diagram form in Fig. 2.14.

The received p.c.m. signal is assembled in a shift register and the binary information at the output of each section of the register is transferred to a counter. The counter counts down towards zero under the action of clock pulses and, at the same time, a capacitor is charged from a constant-current source (i.e. a ramp generator). When the counter registers a sequence of logic 0's the charging circuit is broken and the voltage across the capacitor is proportional to the quantised value of the original sample.

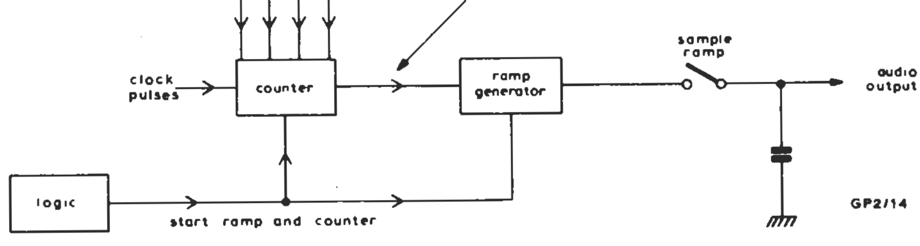

Fig. 2.14. Block Diagram of a Counter Decoder

SECTION 3

IMPROVING THE SIGNAL TO QUANTISING NOISE RATIO OF A P.C.M. SIGNAL

The quantising noise can be reduced by increasing the number of quantising levels, and hence increasing the bit rate and the bandwidth. However, other methods of improving the signal-to-noise ratio are available which do not require an increase in the bit rate (bandwidth). These methods are:

Pre-emphasis and de-emphasis
Companding
Adding a dither waveform and white noise.

3.1 Pre-emphasis and De-emphasis

The high-frequency components of the analogue signal are pre-emphasised prior to coding and de-emphasised after decoding with a consequent reduction in noise. This method of improvement depends on the assumption that in normal programme material the high-frequency components are of smaller amplitude than those of low frequency. Tests have shown that to avoid overloading the system the general level of the programme must be reduced so that the overall improvement is slight.

3.2 Companding

3.2.1 The Companding Principle

The noise level in a sound transmission system is usually specified with reference to the maximum signal level that can be transmitted without distortion. However, the permissible level of noise is determined by its subjective effect when the signal level is low. It is therefore possible to increase the effective signal-to-noise ratio of a system by using automatic gain controls with complementary characteristics at the sending and receiving ends of the system. The automatic gain control at the sending end reduces the amplitude range of signals applied to it and is known as a compressor; the complementary device at the receiving end of the system restores the signal to its original dynamic range and is known as an expander. The two devices are jointly known as a compandor.

The gain-variation characteristic of the compressor and expander must be accurately matched to avoid distortion of the signal.

3.2.2 Instantaneous Companding

In p.c.m. telephony systems a companding process is applied to the instantaneous value of the signal and the quantising levels are spaced at unequal intervals so that the smaller amplitude signals cover more quantising steps than they would if linear spacing were used.

If the quantising steps are spaced in a logarithmic manner, it is possible to obtain an almost constant signal-to-noise ratio for a wide range of signal amplitude; moreover, this can be done using fewer quantising steps than would be needed to obtain the same minimum signal-to-noise ratio with linear quantising.

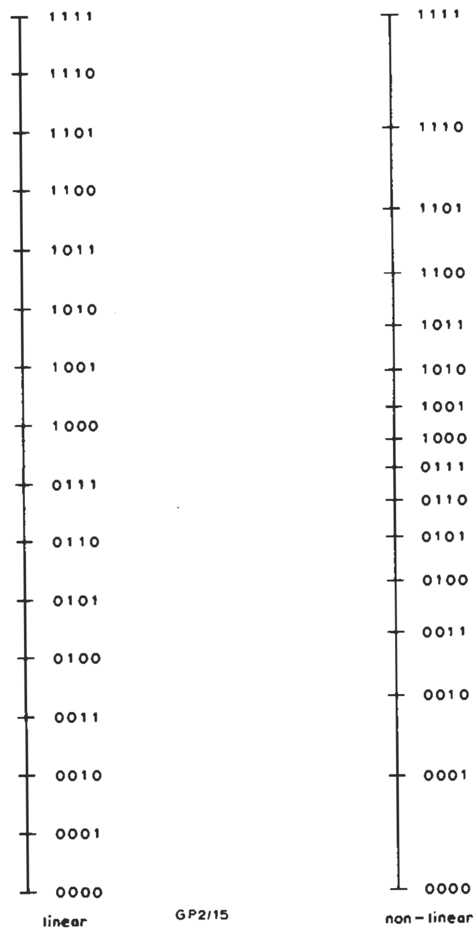

Fig. 3.1. Comparison of Linear and Non-linear Quantising Steps for a 4-bit System

A truly logarithmic quantising characteristic would require an infinite number of steps and, in practice, so-called logarithmic quantising is logarithmic for large signal amplitudes and linear for very small signal amplitudes. A comparison of linear and non-linear quantising steps for a four-bit p.c.m. system is given in Fig. 3.1. (see also Fig. 2.5).

Another method of obtaining non-linear quantising is to use a system in which equally spaced quantising levels are fed with a signal which has been passed through a circuit with a non-linear input/output

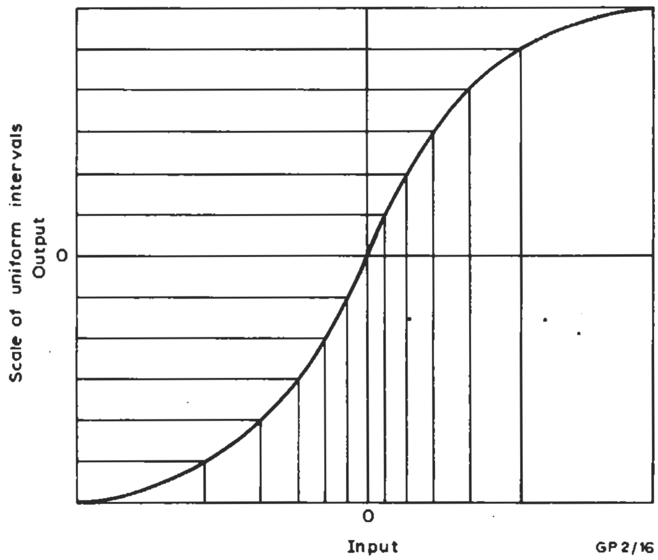

Fig. 3.2. Effect of Passing a Signal through a Circuit with a Non-linear Input/Output Characteristic

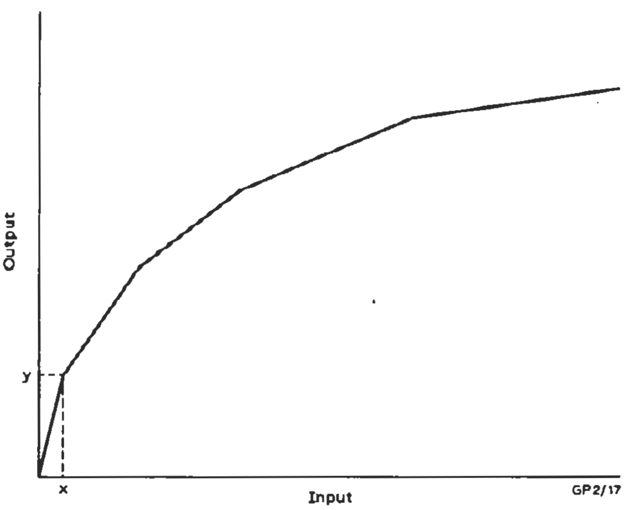

Fig. 3.3. Non-linear Input/Output Characteristic formed from a Number of Linear Segments

characteristic, as shown in Fig. 3.2. Such a characteristic can be obtained from an amplifier with a non-linear feedback network, or by making use of the non-linear characteristics of diodes. A variation of this system of companding is used by the British Post Office in which the input/output characteristic consists of a number of linear segments and the upper half of such a characteristic is shown in Fig. 3.3. The initial slope of the characteristic (x/y in Fig. 3.3) indicates the improvement obtained on the smallest amplitude signals and is known as the companding advantage. For the Post Office system x/y = 16 which gives a companding advantage of 20 log 16 = 24 dB.

For distortionless performance, the decoder must contain a non-linear circuit which complements that contained in the coder. This requirement is difficult to achieve in practice and though instantaneous companding gives adequate results for telephony it is not suitable for high-quality sound programmes, because of the difficulty of matching the non-linear elements in the coder and decoder.

3.2.3 Syllabic Companding

With this method of companding a limiter, which reduces the gain of high-level signals and so acts as a compressor, is introduced before the coder. An expander with a complementary performance to that of the compressor is used after the decoder to restore the signal to its original level. Because the compressor and expander are relatively slow in action compared with the instantaneous type they are jointly referred to as a syllabic compander.

With syllabic compandors the gain does not vary appreciably during one cycle of the signal waveform. However, any noise originating in the system between the compressor and the expander (which in a p.c.m. system is mainly quantising noise) becomes modulated with the programme level. This fluctuating programme-modulated noise, commonly referred to as hush-hush noise, is aesthetically objectionable but can be reduced by the use of suitable techniques.

In telephony p.c.m. systems programme-modulated quantising noise is to a large extent masked by the signal itself and, even when audible, can be tolerated so long as intelligibility is not impared. However, in wide-band sound systems the predominant signal and noise components often lie too far apart in the a.f. spectrum for masking to be effective.

The subjective effect of programme-modulated noise can be mitigated by dividing the signal into several frequency bands by means of filters and providing each band with a separate compressor and expander. The outputs from the individual compressors are combined at the sending end of the system. Any noise components which appear in a frequency band containing little or no signal energy are thus subject to the maximum attenuation in the expander.

Fortunately the noise in a p.c.m. system is predominantly high-pitched and so a simplified arrangement can be adopted in which only the higher frequency signal components are subject to compression and expansion. This can be done by introducing a large amount of pre-emphasis before the compressor and a corresponding amount of de-emphasis after the expander, (for details see Instruction P.15, Sound-in-syncs).

3.3 Adding a Dither Waveform and White Noise

Granular distortion (see sub-section 2.2.2) occurs when the signal amplitude is low and only a few of the available quantising levels are used. This form of distortion can be reduced by adding a *dither* waveform to the analogue signal. The dither waveform is a square wave at half sampling frequency and with an amplitude which covers more than one step in the quantising scale.

Subjective tests have shown that the granular distortion of low-level signals can be further reduced by adding to the dither waveform a certain amount of white noise. The white noise content must be low, otherwise the overall signal-to-noise ratio of the system is reduced.

The dither waveform method of noise reduction depends on the use of quantising levels and so applies only to p.c.m. systems. The three previously mentioned methods of noise reduction apply equally to analogue systems.

SECTION 4

MISCELLANEOUS

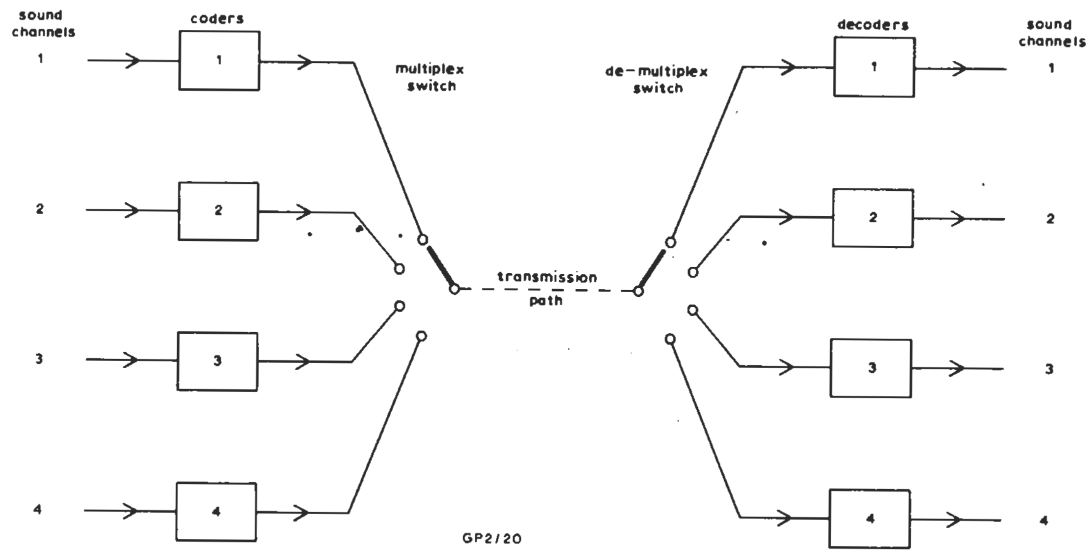

Fig. 4.1. Simplified Block Diagram of a P.C.M. Multiplexing System

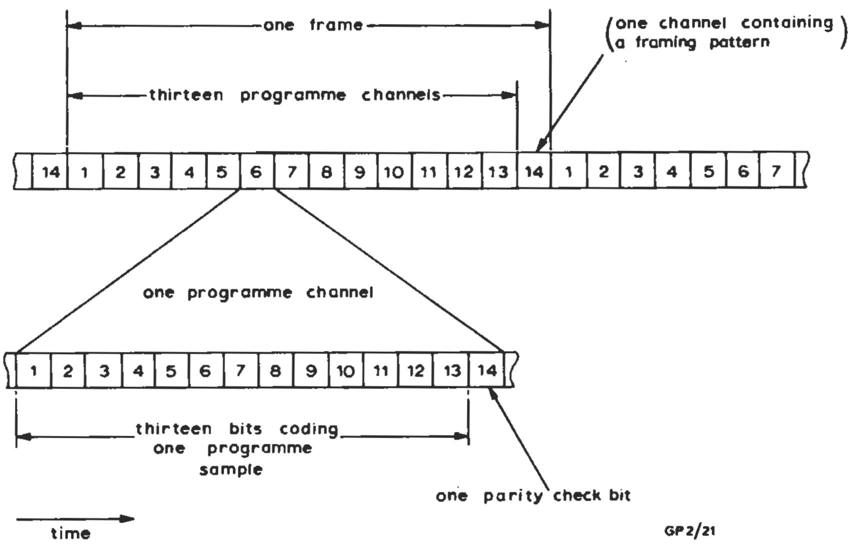

Fig. 4.2. Format of a 14-channel Multiplex Signal

4.1 Multiplexing

Because p.c.m. signals are in pulse form, time-division multiplex techniques can be used to share the same transmission path between several p.c.m. channels. A simplified block diagram of a p.c.m. multiplexing system is shown in Fig. 4.1. The multiplex and de-multiplex operations are performed electronically although, for the sake of simplicity, these operations are denoted by switches in the diagram •

During each complete cycle of the multiplex operation all the contributing channels are scanned once; the information transmitted by one complete cycle of the multiplexer is called a frame. Each frame contains a word (i.e. a binary-coded sample) from each channel, together with some extra bits. These extra bits are used:

(a) To synchronise the multiplex and de-multiplex operations; this action is known as framing or frame synchronisation.

(b) To check errors in the transmission: this is done by a process known as parity checking.

To synchronise the de-multiplexer to the multiplexer a regularly recurring pattern of 1's and 0's is transmitted during each framing interval, and this pattern is recognised by the de-multiplexer. False synchronising information could be transmitted if the framing pattern occurred accidentally in the channel information, but special measures are taken to prevent the de-multiplexer responding to such false information. Fig. 4.2 shows the format of a multiplex signal containing 14 channels per frame (13 programme channels and one framing channel) and illustrates the way in which framing and parity-checking bits can be provided.

4.2 Error Detection

The presence of errors in p.c.m. signals can be detected by transmitting extra bits (called parity bits) for checking purposes.

A parity bit is added to a given code to express the parity of that code. The parity may be either odd or even according to the system chosen; for an even parity the polarity of the parity bit is arranged so that an even number of 1's occurs in each complete word and for an odd parity the polarity of the parity bit is arranged so that an odd number of 1's appear in each complete word. A partial truth table which gives the 8 least significant numbers for a 4-bit system to which an even parity bit has been added is given in Table 4.1.

TABLE 4.1

Parity Bit	*Information Bits*				*Decimal Number*	
1	0	0	0	1	=	1
1	0	0	1	0	=	2
0	0	0	1	1	=	3
1	0	1	0	0	=	4
0	0	1	0	1	=	5
0	0	1	1	0	=	6
1	0	1	1	1	=	7
1	1	0	0	0	=	8

4.3 References for Further Reading

1. Pulse Code Modulation for High Quality Sound-signal Distribution. BBC Engineering Division Monograph, Number 75 (December 1968)

2. Principles of Pulse Code Modulation. K.W. Cattermole. Published by Iliffe Books Ltd. (1969)

3 Techniques of Pulse Code Modulation I.E.E. monograph, Series 1. Published by Cambridge University Press (1967).

APPENDIX A

DELTA MODULATION

Delta modulation (also known as one-digit differential p.c.m.) is a variation of pulse code modulation. In delta modulation systems a single digit is generated during each sampling period to indicate the direction in which the signal amplitude has changed since the last sampling period. The transmitted pulses thus indicate only whether the signal amplitude has increased or decreased during the period between samples. At the receiving terminal the pulses are integrated and filtered to obtain the original signal.

The comparator compares the input signal e_0 with a feedback signal e_1, obtained by integrating the output pulses, and provides a difference signal which is fed to the modulator. The modulator operates on the sampling pulses in such a way that positive-going pulses are generated if the difference signal is positive and negative-going pulses are generated if the difference signal is negative. The transmitter pulse chain therefore consists of positive or negative-going pulses at sampling pulse rate.

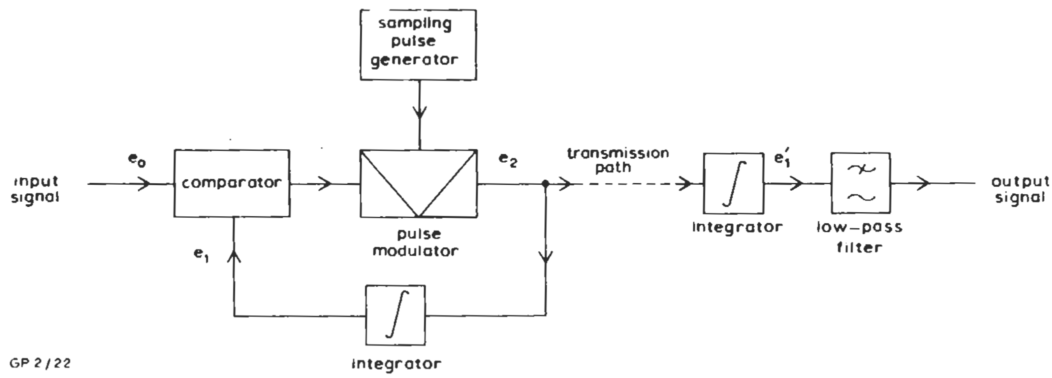

Fig. A.1. Basic Circuit for Delta Modulation

Delta modulation has the advantage over conventional p.c.m. of cheapness, simpler coding and decoding, and the ability to use low-bandwidth circuits; moreover it does not require synchronisation between coder and decoder. However, for equivalent noise performance in high-quality applications, delta modulation requires a higher transmission bit rate than p.c.m.

Basic System

A block diagram of a basic delta modulation system is given in Fig. A.1.

In the decoder the received pulses e_2 are integrated in a network similar to that in the coder to give a signal e'_1 which is a close approximation of the input signal. Sampling components are removed by means of a low-pass filter. The difference between the original signal and the reconstructed signal can be regarded as noise and can be reduced by increasing the sampling frequency.

For delta modulation the transmitted bit rate is equal to the sampling rate; hence the sampling rate must be several times greater than the highest frequency contained in the input signal if a satisfactory signal-to-noise ratio is to be obtained.

APPENDIX B

GLOSSARY OF DIGITAL TERMS

Bit: A binary digit; representing logic 0 or logic 1.

Bit Rate: Rate at which bits are produced or transmitted. e.g. 50 bits per second. 64 kilo-bits per second, 120 Meg-bits per second.

Bit Stream: A stream of bits, usually on one wire and of defined bit rate.

Channel: Usually refers to the time slots available for one particular signal in a multiplex system.

Character: See *Word.*

Clock Pulse: A pulse, usually one of a continuous stream of pulses, which is used to define the time at which a digital operation takes place.

Code: Patterns or sequences of bits with a defined meaning; e.g. letters of the alphabet, numbers, instructions.
To convert signals or data into a digital code.

Coder: Device for coding information; e.g. turns audio or video signals into a digital form. Often used as an alternative to Analogue-to-digital converter (A.D.C.)

Codec: A coder and a decoder

Complement Logic 0 in place of logic 1 and logic 1 in place of logic 0.

Decoder: Device which turns the digital signal back into the original information. Often used as an alternative to Digital-to-analogue converter (D.A.C.).

Eye Height: The range of signal level within which a pulse height discriminator can correctly sort out logic 1's and logic 0's in a received signal; the smallest difference between any voltage which means logic 0 in a received signal.

Frame: A defined number of words forming a block of coded data. There may be, for example, 24 words of 8 bits in each frame. Each word position can be used to carry one information channel.

Framing Pattern: A special sequence of bits used to identify the end or the beginning of a frame and used to synchronise the multiplexing and de-multiplexing process.

Dither: A waveform which can be added to the analogue signal prior to coding; its purpose is to reduce the subjective impairment caused by the quantising process.

Multiplexing: The transmission of several independent signals over a single path. In digital systems this is achieved by interleaving the pulses from the different signals, a process known as time-division multiplexing.

Parity Check: A method of detecting an error by checking the state of a *parity* bit added to a binary word. Parity bits added to a word either make the total number of 1's in the word even (even parity) or odd (odd parity).

Pulse Code (Modulation) A method of turning analogue signals into digital form. The analogue signal is first sampled at regular intervals in time; digital words are then derived which indicate the amplitude of each sample as a binary number using a fixed number of bits per sample

Quantising: Expressing the instantaneous amplitude of an analogue signal as one of a set of pre-determined amplitude levels.

Quantising Noise: This is a general phase referring to the errors introduced into a signal by the quantising process.

Sampling: The process of measuring the amplitude of an analogue signal at regular time intervals.

Word: A group of bits, fixed in number, describing one element in coded form. A word (of eight bits, for example) could represent the coded form of one sample of an analogue signal.