Aston University

Some pages of this thesis may have been removed for copyright restrictions.

If you have discovered material in Aston Research Explorer which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our Takedown policy and contact the service immediately (openaccess@aston.ac.uk)

PULSE TIME-CODE MODULATION SYSTEM

Thesis for the degree of

Doctor of Philosophy

by

MAWLOOD MUSTAFA AL-HAMAD

University of Aston in Birmingham,
July 1981

PULSE TIME-CODE MODULATION SYSTEM

by

Mawlood Mustafa Al-Hamad

A thesis submitted to the
University of Aston in Birmingham
for the degree of
Doctor of Philosophy 1981

SUMMARY

This research is concerned with the development of a system designed to economise on the bandwidth required for the digital transmission of speech signals. The sampling rate in the PCM system is constant and related to the maximum frequency of the sampled signal.then, at times when lower frequencies only are being transmitted, the sampling rate is unnecessarily high, and there would thus appear to be the possibility of reducing the bandwidth. This led to the thought of a new method of sampling. The new method has been called pulse time-code modulation (PTCM). The sampling in PTCM is achieved by dividing the full range of the signal amplitude into quantizing levels. When the signal crosses one of these levels the time measurement starts. It continues until another crossing of the adjacent level occurs, then a digital word is completed and a new measurement of time starts for a new digital word. Every digital word consists of one bit indicating the direction of crossing up or down and the rest of the word represents the measurement of time. The transmitting rate should be the average of the sampling rate. This can be done by storing the samples and averaging their rate. At the receiving end the process is reversed.

It was found that it is possible to achieve appreciable saving in transmission rate. A large store is necessary for the purpose of averaging. The presence of such a store will necessarily cause a delay in the transmitted signal, which is a disadvantage when the PTCM system is used for duplex transmission.

During the course of the research into the above system, it became apparent that the same principles of averaging could be applied to the conventional PCM system. This was investigated, and it was found that an appreciable saving in the bandwidth can be achieved. We have called this system AVERAGE RATE PULSE CODE MODULATION (ARPCM).

Key Words : Coding, Modulation, Communications.

CONTENTS

Page No.

Page No.

Page No.

ACKNOWLEDGEMENTS

The author is highly indebed to Dr. P.N. Nield for his supervision, his guidance and encouragement throughout the work.

The author would like to express his gratitude to the Iraqi Government, for their financial support during this research.

HISTORICAL BACKGROUND

The history of this system goes back to 1972 when M.A. Binal(1) started it under the heading of, Variable Time-Scale Information Processing(1). Binal made some tests and investigations to find out the benefit and practicality of this system. From his tests and investigations he concluded that appreciable benefit over conventional PCM systems can be achieved by the application of this system.

After Binal, M.G. Mountis(2) came, he built the first prototype of the system. Mountis called the system Pulse Time Code Modulation and abbreviated it to PTCM. He depended mainly on Binal's results. He also carried out some further tests which seemed to him to re-inforce the results obtained by Binal. Mountis built a model of the system but he was unable to get conclusive results. It was concluded that the work done by Mountis needed extending.

The first part of the present research was devoted to measurements on the speech signal properties. Owing to shortage of time by previous researchers, they had to assume a sinusoidal signal. For instance, the maximum slope of the speech voltage, which dominates the design of the system, had previously been evaluated by calculation, assuming the speech voltage to be sinusoidal signal with amplitude and frequency having their maximum values. The measurement of these parameters in the actual speech required a design and construction of special test equipment, and occupied a substantial part of the research time.

CHAPTER 1

INTRODUCTION

One of the most important fields of pulse techniques is digital communications. Digital communications meet extensive applications at the present time. Although the bandwidth occupied when applying digital communications is wider than the bandwidth occupied by the analogue signal to be transmitted, yet the advantages of digital communications make it very attractive to use[3]. The most important advantage is that the digital communications such as PCM are more resistant to noise than the analogue ones. Another advantage of digital communications is the simple detection of the signal at the receiving end. This comes from the fact that in a digital signal it is only needed to determine whether the signal exists or not. Hence when the signal exceeds a certain threshold level it is present and when it does not exceed that threshold it is not present. The situation in analogue communications is different, because an infinite number of levels need to be detected faithfully[3]. It is clear that the only disadvantage of digital communications is the wider bandwidth occupied.

In this research an attempt was made to reduce the bandwidth occupied by the channel in digital communications. It is obvious that the bandwidth is an essential factor in the capacity of the line. Hence reducing the occupied bandwidth will increase the line capacity.

In the system developed a new way of sampling was used. This new method of sampling does not depend on the maximum frequency of the signal directly. It was described in detail by M.A. Binal(1) and M.G. Mountis(2). Instead of sampling the signal at a fixed rate as it is done in PCM and delta modulation, the sample in PTCM represents the time measurement between two successive level crossings of predetermined levels. Since the coding is done in binary counting basis, the resulting code will be in logarthmic form.

This thesis is divided into two parts. The first part covers the investigation and tests concerning the Pulse Time Code Modulation (PTCM) System. It is an extension of the work done by previous researchers. The second part deals with the investigation of the Average Rate Pulse Code Modulation (ARPCM) System.

Since the PTCM system was explained and investigated by previous researchers, Binal(1) and Mountis(2), the investigation in this research was to extend the work done by those researchers. This research was concentrating on the problems which were not deeply investigated before.

The ARPCM system is a purely new system, suggested by the author, and came as a direct result of the research done on the PTCM system.

CHAPTER 2

PULSE TIME CODE MODULATION SYSTEM (PTCM)

2.1 DESCRIPTION OF PTCM SYSTEM

PTCM system is similar to other digital communication systems in the representation of the data to be conveyed in digital form. However, PTCM system differs completely in the way of coding the data and the information represented by this data. In conventional PCM system for instance, the digital word represents an amplitude, whereas in PTCM the digital word represents time-interval. In PTCM the full range of the signal is divided into a number of quantization levels. The time-interval spent by the signal between two adjacent quantization levels is coded. The coded time-interval represents the data word in PTCM system. The addition of another bit to indicate whether the signal is rising or falling is necessary. When the signal rises or falls sharply, the number of levels crossed by the signal will be high and the time-intervals to be coded will be short. On the other hand, when the signal rises or falls slowly the number of crossings will be low and the coded time intervals will be long. By this fact, one can notice that the signal slope plays an important role in PTCM system. If the transmission of data is to be done directly as the words are generated, the PTCM system will not be useful in reducing the band-width occupied by the message in spite of the low rate of generation of words during silence periods and low slopes signal. This results from the fact that in direct transmission, the system should be capable of transmitting the fastest rate of word generation. In order to make use of the slow rate of word generation during pauses and slowly varying signal, the samples which are coded time-intervals need to be stored. The transmission rate will be the average of the words generation rate.

At the receiver's side the data should be stored in a similar store of the transmitter, so that it will be possible to decode the samples. The input to the receiver's store will come at a constant rate. The rate at which the samples are discharged out of the store of the receiver will be variable according to the time-interval the word represents. Each time a word is taken out of the receiver's store, time measurement starts. When this time measurement becomes equal to the time-interval represented by the word, a step up or a step down occurs, depending on the sign bit in the word itself. The stepping up or down can be done at the beginning of the time measurement, or at the end of it. PTCM system operation principles are illustrated in Figure 1, and will be discussed in the following items.

2.2 OPERATION OF THE PTCM SYSTEM

2.2.1 The Transmitter

2.2.1.1 The Sampler

As the proposed system differs from other pulse

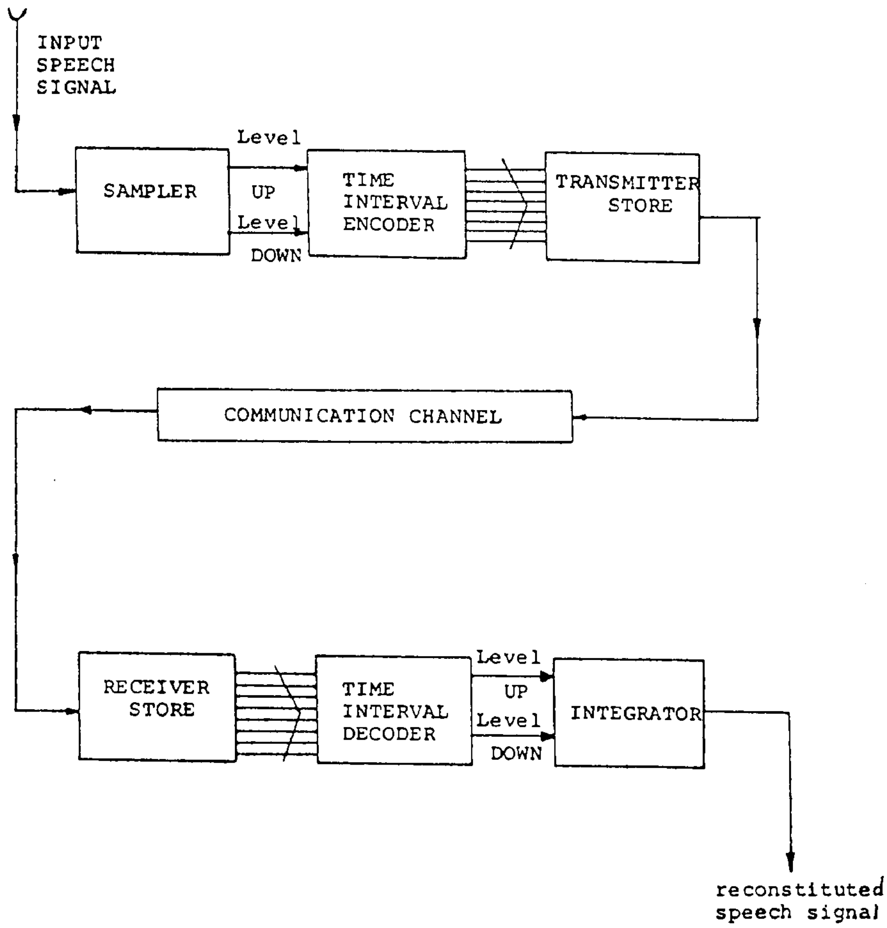


FIG. 1 THE PROPOSED PTCM SYSTEM[2]

code modulation systems mainly in the sampling technique, it becomes necessary to build a special sampler for it. The function of this sampler is to generate a pulse whenever the signal voltage crosses one of the levels at which the full range of the signal is divided. The generated pulse indicates either "Level Up" or "Level Down". Hence two output outlets are required. A positive pulse and a negative pulse are to be generated if one output outlet is provided. The sampler should be able to follow the fastest rate of rise or fall, i.e. the highest slope of speech signal if the system is used for speech transmission. If the sampler slew rate was less than the maximum slope of speech signal voltage, then it will not be possible for this sampler to follow the signal properly, and its performance will not be satisfactory. The sampler was designed to accept speech signals of one volt peak amplitude.[2,5]

2.2.1.2 The Time-interval Encoder

When the signal indicating a level change, either "Level Up" or "Level Down" is received from the sampler, the time interval encoder starts the time measurement. Time measurement is done by pulse counting of a fast crystal oscillator in binary form. The encoder continues counting until another level change pulse comes out of the sampler. At the moment when this second pulse of a level change is received, the coded time-interval is fed into the store. The coded time-interval and the bit indicating the direction of crossing "Level Up" or "Level Down" form the complete digital word which is kept in the store in an addressed location.

The encoder should be able to encode the minimum time-interval to be coded Δt_{min}, which is determined by the maximum slope of the signal. The degree of accuracy of encoding Δt_{min} should be reasonable. The maximum time-interval needs to be coded Δt_{max} is mainly limited by the minimum slope of the signal in this system. Previously the determination of Δt_{min} was calculated by assuming that the signal is a sinusoidal waveform, having frequency and amplitude at their maximum values of speech signal voltage. Δt_{max} was also calculated by previous researchers by assuming that the sinusoidal waveform has an amplitude of voltage which varies between two successive levels of the quantizations levels and a frequency equal to the minimum frequency of speech signal voltage.

In this research practical measurements for the determination of Δt_{min} and Δt_{max} have been done. These measurements will be discussed later.

The encoding of pauses or silence gaps is restricted by the length of digital word to be used. The digital words representing silence gaps or pauses, should be well distinguished from the longest digital word representing active speech.

2.2.1.3 Transmitter Store

Temporary storage of data is essential in PTCM systems, in order to have the variable rate of digital word generated averaged. The store can be in different forms. First-in-First out 'FIFO' memory can be used. Also it is possible to use random-access memory, 'RAM'. The important thing to be observed in the construction of the store is to keep the data in their correct order, so that the word entered into the store first, should be taken out of it first, and the last word to enter the store should leave it last.

First-in-First-out memory was used by Mountis[2]. It would be better to use random-access memory for the following reasons:

(a) No delay will be introduced more than the access time of the memory, while in the case of the FIFO memory, the shifting of the words from the first location to the last one will cause some additional delay over the access time.

(b) In the case of RAM, there is no need to have a shifting clock, while in the case of the FIFO it is necessary.

(c) Less changes of states inside the RAM device, compared to what happens in the FIFO device, means less noise, and less decoupling for the power supply is necessary.

(d) The availability of a single chip RAM of high capacity is more than that of the FIFO. That means less devices of RAMs will be required to build the store.

(e) The RAM devices are in general, cheaper than the FIFO devices, if they were of the same capacity.

The FIFO devices offer one advantage over the RAM devices, which is the elimination of the address circuitry. This results from the fact that the data in the FIFO are fed from one end and taken out from the other end by fast shifting. In the case of the RAM, it is necessary to have the locations addressed, and when the data of certain location is called out, the proper address should be connected to the address lines.

The rate at which the data will flow out of the store to the transmission line should be equal to the average rate of data entering the store. It is very difficult to find exactly the average rate of data. The difficulty arises from the fact that the speakers differ widely and even one speaker may speak in different ways from time to time. It is possible to determine the mean average rate which can be considered suitable for most speakers. In addition, it could be useful to use some sort of feedback to control the transmission rate by the contents of the store. If this feedback was not used and the transmission was made at constant rate, then there would be two possibilities. The first possibility happens when the transmission rate is less than the average rate of the incoming data. The result is an accumulation of data gradually in the store, until a case of what is called 'overflow' happens. In this case some of the data will be lost because there will be no room to accommodate them in the store. The result will be distortion in the signal after reconstitution. The second case happens when the transmission rate is higher than the average rate of the incoming data. The result in this case will be frequent 'underflow' which means that the transmission frequency is more than required. That means the bandwidth used is greater than required, while the aim of this research is to reduce the bandwidth as much as possible.

If it is required to make the transmission frequency constant, the store size should be large. The larger the store size is, the less the possibility of overflow and underflow will be. On the other hand, the large store size will cause longer delay to the data which is undesirable. The cost of the store also will be increased by making it larger, and the physical size and the power consumption will be increased. When smaller store size is used the possibility of overflow and underflow will be increased. The effect of overflow, underflow and delay will be discussed later.

If feedback is used to control the transmission frequency the bandwidth occupied by the transmission will be limited by the maximum transmission frequency. The result of that efficiency of transmission line will be less than the efficiency in the case of the constant transmission frequency. It is possible to reduce the store size by using a variable transmission frequency, while keeping the possibility of overflow and underflow unaffected. In order to make use of variable word length technique, that is to get rid of any non-used bits, the use of a marking signal is essential. The marking signal is used to indicate an end of one word and the beginning of another. It can be called Start-stop signal. Start-stop signal is a pulse width modulation which will be discussed later in another chapter. The use of this start-stop technique adds another difficulty to the use of variable frequency transmission. The variation of transmission frequency should be kept within certain limits, so that it will not affect the discrimination of start-stop signal.

The store can be any memory device provided it is possible to arrange it in first-in-first-out basis. It can be a magnetic disc, addressable random-access-memory, whether it is semiconductor or magnetic, or any other kind of device. The important feature of the store used, is its speed. It should be able to accept the maximum rate of incoming data. The number of peripheral devices used in the store should be kept to its minimum. The optimum store size should be determined by taking into account all these factors, in addition to the physical size, the cost and power consumption.

2.2.2 The Receiver

2.2.2.1 The Receiver Store

The receiver store should be similar to that of the transmitter. It accepts the digital words coming from the transmitter via the transmission line. The incoming words contain information about the input signal. If the rate of transmission was made variable, the store should be fast enough to follow the highest transmission frequency f_T.

Each digital word will be stored in the receiver store in a defined location. The locations occupied by the digital words should be addressed in such a way that the first word fed into the store should be taken out of the store first and the second word fed in should be taken secondly, and so on. In other words the store should be arranged in first-in-first-out basis. Each word will suffer a short delay when entering the store and getting out of it. This delay is the store access time and it will be the same for all the locations. The access time of the store is dominated by the fastest bit rate required, which is similar to that of the transmitter store. The data are entering the store of the receiver at the transmission rate and taken out of it at a rate which depends on the decoding. Since the digital words represent coded time-interval, the time required to decode these digital words will be proportional to the encoded time-interval. This means that the rate at which the words are taken out of the store depends on how fast the decoder decodes each particular word. If it happened that the data input rate was high, while the decoder was decoding long time-interval words, the possibility of overflow would be high. When the incoming rate of data becomes low and the decoder decodes short time-interval words the possibility of underflow will be higher.

Underflow and overflow phenomena are similar for both transmitter and receiver store. Overflow can be reduced by selecting the proper store size. Underflow and overflow can be minimized by controlling the decoding process by the store occupancy, i.e. when the store occupancy decreases, the decoding process is slowed down and vice-versa.

2.2.2.2 Time-interval Decoder

The function of time-interval decoder is to decode the digital words which represents time-intervals. It accepts the digital words from the receiver store and generates a level change signal 'Level Up' or 'Level Down, according to the signal in the digital word, then the time described by the digital word is decoded. The timing signals for the decoder should be the same as those for the encoder so that the encoder time-interval is correctly decoded to its original form.

2.2.2.3 Integrator

The decoder sends pulses which generate an accurately defined quanta of charge of the appropriate polarity. These pulses are delivered to the integrator. Each quantum of charge causes the integrator to change by a unit amplitude. With this method it will be possible to reconstitute the original waveform, see Figures 2 and 3.

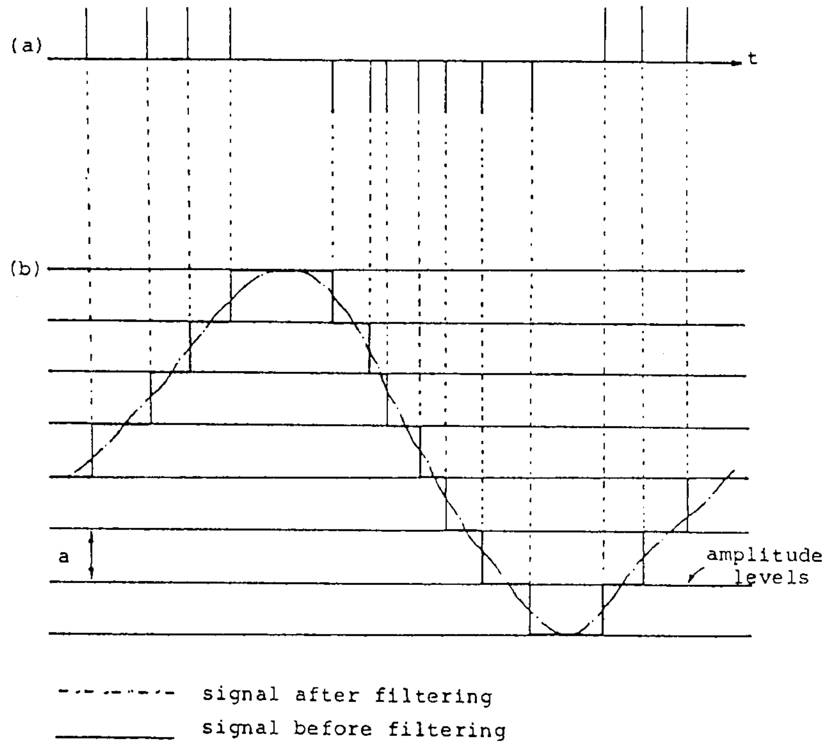


FIG. 2 RECONSTITUTION OF THE SIGNAL IN PTCM SYSTEM[2]

(a) Quanta of charge delivered to the integrator

(b) Reconstituted signal.

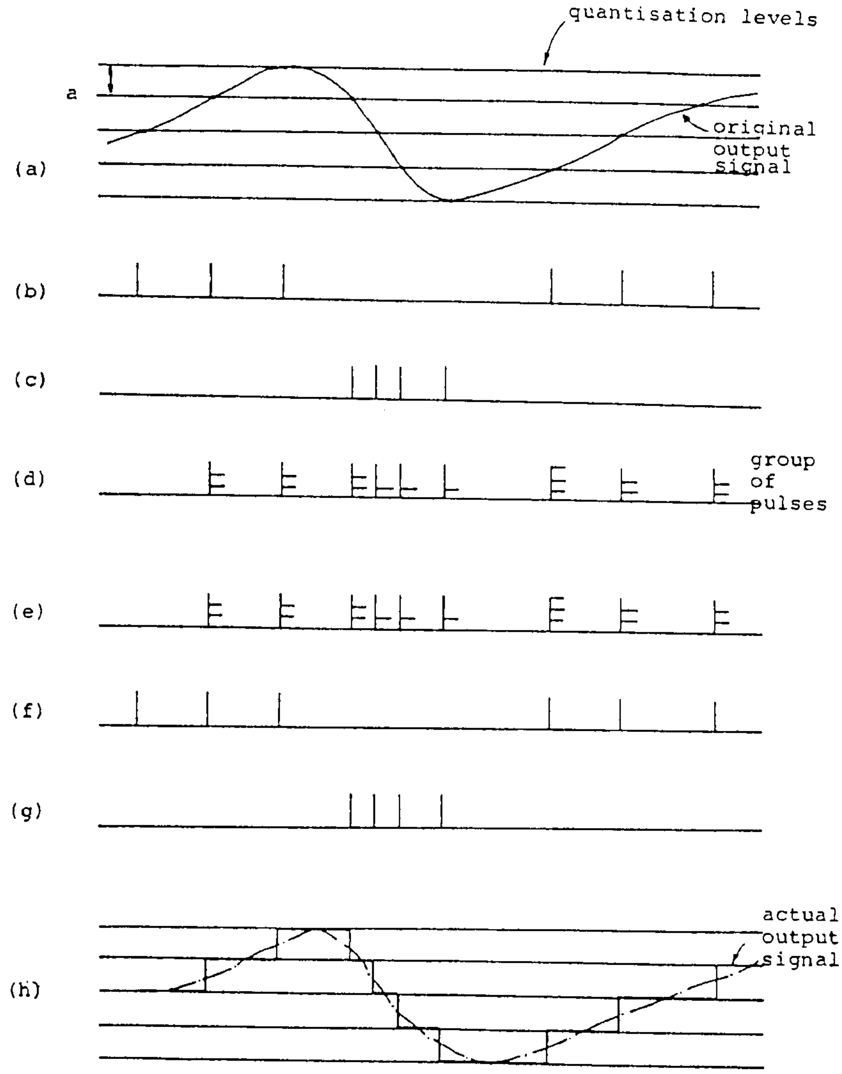


FIG. 3 WAVEFORMS OF THE PROPOSED PTCM SYSTEM (2)

(a) Input signal to the sampler
(b),(c) Output of sampler (Level UP, Level DOWN)
(d),(e) Output of encoder, input of decoder
(f),(g) Output of decoder. (Level UP, Level DOWN)
(h) Output of integrator (Reconstituted signal)

2.3 BANDWIDTH IN PTCM SYSTEM

The economy in bandwidth by using PTCM system is the important feature of this system. In order to compare the bandwidth occupied by PTCM system with that occupied by the conventional PCM system, the signal transmitted by both systems should be the same.

Let the merit factor be R:

$$R = \frac{\text{Bandwidth occupied by one channel PCM system}}{\text{Bandwidth occupied by one channel PTCM system}}$$

The higher the factor R, the greater is the economy of the PTCM system.

Suppose the number of quantization levels for both systems were the same and equal to q. The peak value of the signal is V_p.

Then PCM maximum bits in the sample P will be

$$P = \log_2 q$$

If the sampling frequency was f_s, then the number of bits per second transmitted will be:

Bit rate in PCM system = $P.f_s$ bits/sec.

If the signal is bandwidth limited to a maximum frequency f_{max}, then

Bit rate in PCM system $\geqslant 2f_{max}.P$ bits/sec

where $f_s \geqslant 2f_{max}$ according to the sampling theorem.

On the other hand, suppose that the PTCM system uses the same number of quantization levels and the same number of bits per word then

Bit rate in the PTCM system = $f_T.P$ bits/sec

where f_T is the transmission frequency of the PTCM system. Therefore

$$R = \frac{f_s \cdot P}{f_T \cdot P} = \frac{f_s}{f_T}$$

To achieve any economy in the PTCM system over the PCM system R should be > 1. If $f_s = 8$ kHz, then $f_T < 8$ kHz.

Since the speech activity in telephony was determined experimentally and was found to be between 0.4 to 0.45[2] and in the PTCM system the generation of data during pauses is very little, considerable improvement in bandwidth occupancy over the PCM system can therefore be expected.

CHAPTER 3

SPEECH SIGNAL SPECIFICATIONS

3.1 GENERAL

Although the speech signal can vary widely from one speaker to another, there are still common specifications which can be used to describe it.

The most important and probably the earliest property of speech signal ever known was its bandwidth. It was found that the bandwidth of a speech signal lies between 400 Hz and 3400 Hz. The maximum frequency of speech signal dominates the sampling frequency. The well-known sampling theorem states that the sampling frequency must be at least double the maximum frequency of the sampled signal. 8 kHz sampling rate was taken as a standard rate in PCM for sampling.

The specification of speech signal depends mainly on the way of looking at it. The stream of speech can be observed and analysed in many different ways according to the purpose of the analysis[4].

For the present system "PTCM", Binal[1] stated that the most important factor which will dominate the system design is the maximum slope of the speech signal[1]. Since the speech signal is considered as a random signal, the statistical way of determination of signal slope distribution is the most convenient way of dealing with such a signal.

Another very important property of speech signal for this project and has not been determined by either Binal [1] or Mountis [2], is the distribution of the coded time-intervals of the signal when sampled in the PTCM sampling method.

These two properties, the slope distribution of the signal and the coded time-intervals of the PTCM sampling method for speech signal are the main investigations done to determine the speech signal specifications from the PTCM viewpoint.

The store size for PTCM was also determined by measurement, which was found to be the most important factor in the PTCM system, because it determines the practicality of this system. Although store size was found to be essential, yet it was not given the required attention before.

3.2 STANDARDIZATION OF SPEECH SIGNAL

Because the speech signal voltage is non-uniform signal and random, the standardization of it is somewhat relative. The method used in the PCM system can also be applied to the PTCM system. It is desirable to use the same method for both systems in order to compare easily between these two systems.

The level of speech signal voltage is essential for the PTCM system. In conventional PCM systems the signal level only affects the quantized levels, while in the PTCM system, in addition to its effect on the quantized levels, the signal level affects the maximum rate of change of the signal, which in turn affects the number of times of level crossings. That means the rate at which the samples are generated will be affected by the speech voltage level.

The method used in PCM to standardize the speech signal voltage is to clip the signal level that exceeds 12 to 14 dB over the long term RMS value of the loudest talker[2,3]. The signal level depends on the dynamic range of the encoder or the quqntizer.

This same method of standardization was applied to the PTCM system. The level of the signal was kept between + 1V and -1V range.

Before having the speech signal fed to the standardization circuit, the signal was passed through a band-limiting filter. This filter was used to prevent any hum or picked-up signal which might come with the original signal below 400 Hz or over 3400 Hx. The signal itself comes from a high quality tape recorder. The speech represents standard telephone calls recorded on two-track tape by the Post Office Research department. Male and female voices were the two speakers in these calls. The period of each call was approximately 5 minutes.

More details about standardization of speech signals can be found in references (2) and (3).

3.3 DISTRIBUTION OF SPEECH SIGNAL SLOPES

The importance of the determination of speech signal slopes distribution can be summarised in the following points:

(a) To determine the maximum time-interval which can be considered as speech and when exceeding that interval it will be considered as non-speech, i.e. silence.

(b) To find the minimum time-interval to be coded so that negligible distortion in the speech signal will be produced by neglecting the shorter intervals.

(c) It can be useful to determine the distribution of the bit groups of the coded time-intervals.

The distribution of slopes of the speech signal voltage was measured using the system shown in block diagram of Figure 4. The test procedure is explained below.

The standardized speech signal voltage comes to a differentiator. The output of the differentiator is proportional to the slope of the signal (v). Then $\frac{dv}{dt}$ (the signal slope) is clipped from both peaks (upper and lower) symmetrically. The reference voltage of the clipper can be varied continuously. The comparator uses the output of the clipper to open a gate. This gate enables or inhibits the pulses generated by the crystal oscillator from getting to the accumulative counter. The frequency of this crystal oscillator is 1 MHz. The counter will show the real time when the gate opens. By inversion, it can be used to count the time when the gate is closed. The time measured will be in microseconds. The clipped slope of the signal are then integrated to eliminate the effect of the differentiator

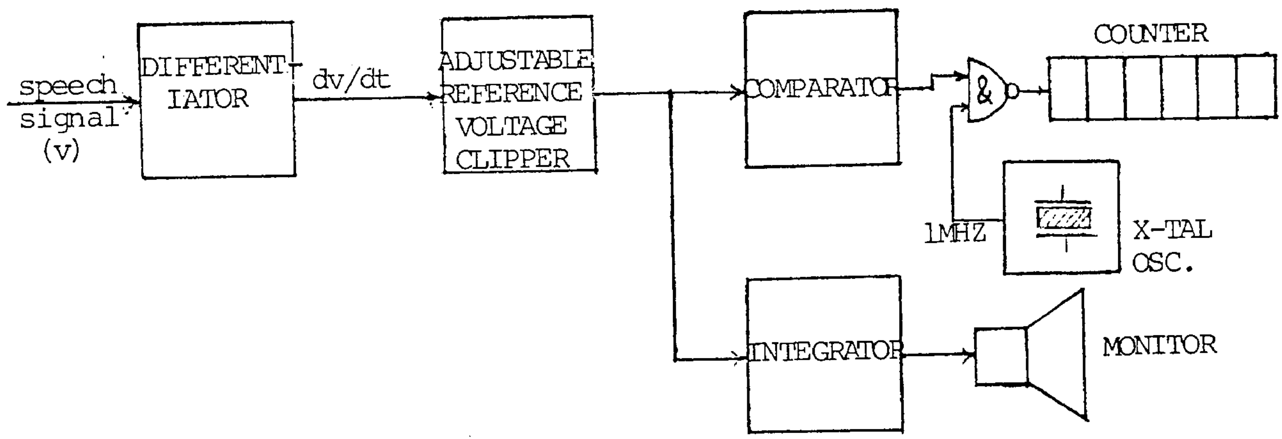


FIG. 4 SPEECH SIGNAL SLOPE DISTRIBUTION MEASURING SYSTEM

and to bring back the original signal. The integrated signal is monitored to notice the effect of clipping the slope on the signal. By this method it is possible to define the slope at which negligible distortion occurs when clipping the slopes over that defined slope.

Figure 5 shows the curve obtained by this measurement for the probability of the slopes of speech signal, (only positive slopes are shown, the negative slopes are similar to the positive side).

The following results can be derived from the graph of Figure 5.

(a) The minimum slope cannot be determined easily of this curve. The reason is that values of slopes below 0.2V/ms become very hard to measure. The result of that is the difficulty of determination of maximum time interval needed to be coded. Hence this method of maximum time-interval measurement is not accurate enough. Another method was used, and will be explained later. The other method is more accurate, and was useful for other measurements also. It was a digital way of measurement.

The importance of determination of maximum time-interval to be coded, comes from the fact that it determines whether the code represents speech signal or silence gap. Also it is important to determine the maximum word length required in coding, which is very important in store construction.

(b) The maximum slope at which the clipping will not affect the speech quality to a noticeable degree was found to be

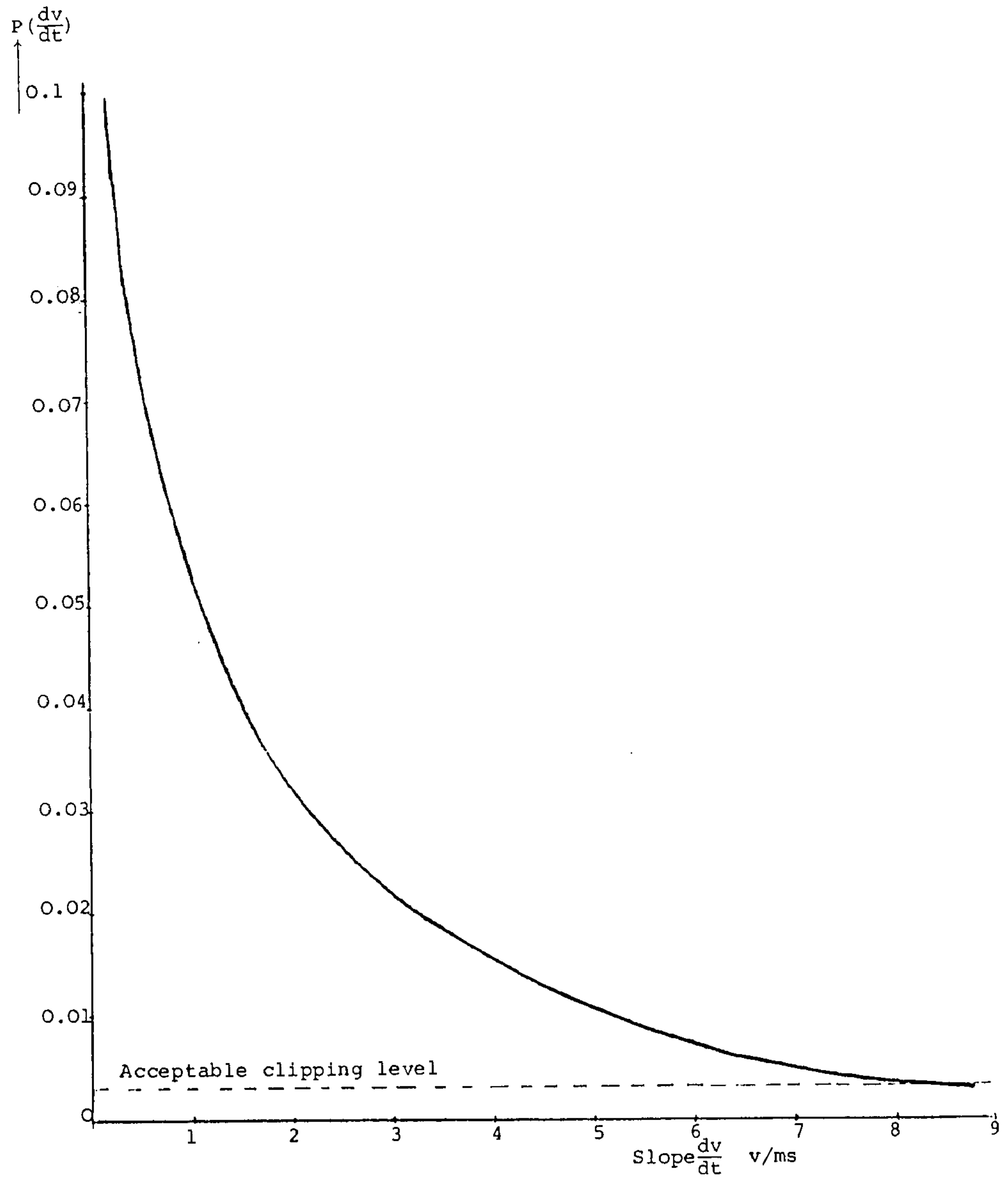


FIG.5 PROBABILITY DENSITY FUNCTION OF SPEECH SIGNAL SLOPES (MALE VOICE)

(For negative slopes a similar curve obtained)

8 volts/ms. From this value of slope it is possible to calculate the minimum time-interval which has to be coded. By the help of Figure 6 the minimum time-interval (Δt_{min}) can be calculated as below:

$$\Delta t_{min} = \frac{h}{(\tan\alpha)_{max}}$$

$$= \frac{1/128 \text{ v}}{8 \frac{\text{v}}{\text{ms}}} \times 10^3 = 0.98 \ \mu\text{s}$$

$$\cong 1 \ \text{s}$$

where Δt_{min} : the minimum time-interval to be coded

$(\tan\alpha)_{max}$ = maximum slope at which clipping to slope will not distort the speech signal to a noticeable degree

h = the voltage between the successive levels.

In order to compare this result with that calculated by Binal[1], it will be useful to repeat the calculation done by Binal[1].

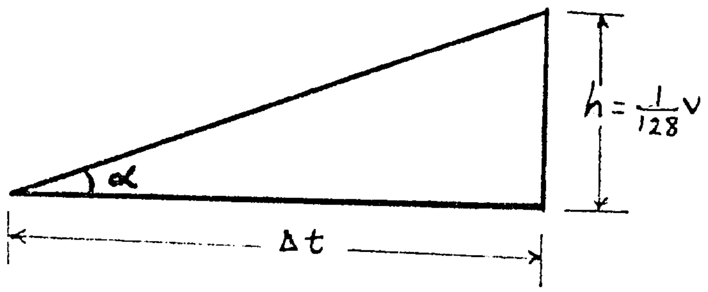
If a sinusoidal waveform of maximum frequency (3400 Hz) and a peak voltage of 1 volt then by the help of Figure 7;

$$v = 1. \sin\omega t$$

$$\tan\alpha \Big|_{max} = \frac{dv}{dt}\Big|_{max} = w$$

$$= 3400.\ 2\pi = 21363 \text{ volt/sec.}$$

$$= 0.021363 \text{ v}/\mu\text{s}$$



For:

$$\Delta t_{min} \tan \alpha = 8v/ms$$

$$\Delta t_{max} \tan \alpha = 0.01v/ms$$

FIG. 6

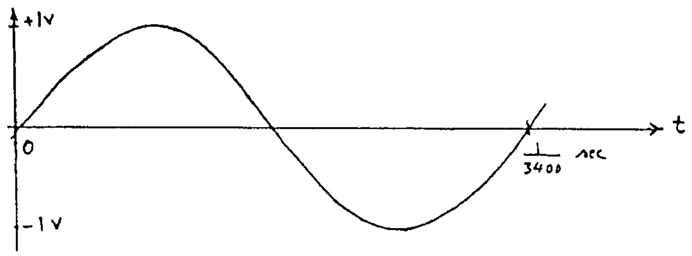


FIG. 7 WORST CASE IN SINE WAVE SIGNAL FOR Δt_{min} CALCULATION

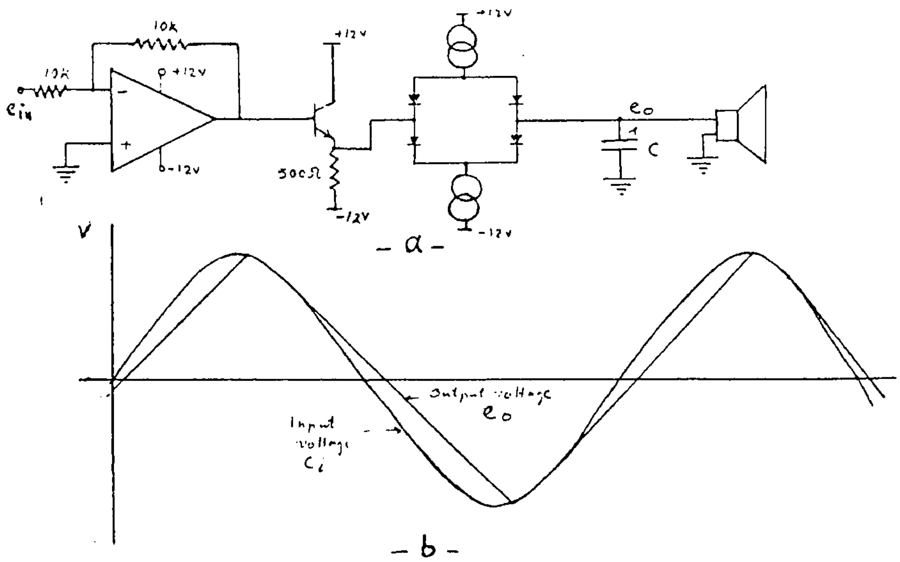


FIG. 8 MAXIMUM SLOP DETERMINATION BY SLEW RATE CONTROL METHOD

(a) System used (b) Typical input and output waveforms.

$$\Delta t_{min} = \frac{h}{(\tan\alpha)_{max}}$$

$$= \frac{(1/128)v}{0.021363}$$

$$= 0.37 \ \mu s$$

where w = angular frequency = $2\pi f$

f = signal frequency.

The measured value for speech signal voltage for Δt_{min} was found to be 1.0 μs. That means the assumption of sinusoidal waveform gives a result for Δt_{min} nearly one third of the truly measured value for the speech signal. This result of measurement is encouraging from the viewpoint of Δt_{min}.

In order to verify the measured result of the maximum slope and hence Δt_{min} for the speech signal, another method was used. This method is a slope clipping without differentiating the signal. It is in fact an operational amplifier with controlled slew rate. The system used in this method is shown in Figure 8.

The operation of this system can be explained in the following way. The capacitor C is varied until no noticeable distortion can be heard in the output when applying the standardized speech signal to the input. Then the slope can be determined for that value of capacitor C by applying a sinewave to the input, keeping the value of C constant, the limiting slope can be noticed by the oscilloscope. The result obtained by this method was nearly one third of the maximum slope of 1 volt peak sinusoidal waveform of 3400 Hz frequency. This result reinforces the results obtained by the first method.

By measuring the value of the maximum slope of speech signal voltage, the minimum time-interval Δt_{min} to be coded was defined. This was one of the aims of determination of slope distribution of the speech signal.

(c) Only a rough idea can be obtained from the graph of Figure 5 about the distribution of bit groups. It can be seen from the graph that the short intervals of time to be coded are not coming very frequently, while the long intervals which are related to low slope values are coming more frequently.

The rough idea that can be obtained from the curve is that the long digital words are more frequent than the short digital words. This indicates that the bit rate is high.

A digital method was invented to measure the distribution of words or bit groups. This method was also useful to determine the maximum and minimim time-intervals of speech signals to be coded Δt_{max} and Δt_{min}. This method will be explained in detail in the next chapter.

CHAPTER 4

DISTRIBUTION OF DIGITAL WORDS "BIT GROUPS"

4.1 GENERAL

It was mentioned before and also by Binal[1] and Mountis[2] that the PTCM system is based on coding the time-intervals between successive crossings of predetermined levels through the whole range of speech signal voltage. The factors which need to be determined in order to be able to design a system capable of performing the requirements of the PTCM system are:

1. The average rate of crossing.
2. The average bit rate.

The sampler used in the test was the same sampler used by Mountis[2]. It was the Mendenhall[5] instrument.

4.1.1 The Average Number of Crossings

This measurement was carried out by using Mendenhall[5] box. The number of crossings during a period of 300 seconds (5 minutes) was counted. An external counter was used in addition to the internal counter of the instrument. The external counter function was to make sure that the internal counter of the sampler was not faulty, by comparing the readings of both counters. It was found that the two readings of the counters were the same.

The filtered and standardized speech signal was fed to the sampler. The number of crossings measured in 5 minutes was found to be 13×10^6 times.

This number means that the average crossing per second is

$$= \frac{13 \times 10^6}{300}$$

$$= 43 \times 10^3$$

This result is not encouraging because it also represents the average word rate. Even if the average word length is supposed to be 2 bits, which is unlikely, the average bit rate will still be higher than that of the conventional PCM bit rate, while at least an extra 1½ bits need to be added to indicate the end of one word and the start of another.

In order to reduce this figure, Mountis used a clever way. He used an integrator between the clipping amplifier and the sampler(2). It was a carefully designed integrator which by using it, it was possible to reduce the above figure by a factor of 10.

Hence the average word rate (level crossings) after using the integrator was 4.3 kword/sec.

The explanation of the effect of the integrator on the signal is that the amplitude of the high frequency components of the signal will be attenuated. Because these high frequency components are the main source of frequent level crossings, then attenuation of the amplitude of these components will lead to a reduction in the average level crossings. The original signal can be brought back to its original form by passing it through an equivalent differentiator at the receiving end. The differentiator must have exactly the inverse characteristics of the integrator. This method is similar to the pre-emphasis and the de-emphasis used in radio communication.

Since the average rate of crossings can be reduced to 4.3 kword/sec. the possibility of using the PTCM system seems practical. A saving of 3.7 kword/sec. over the conventional PCM system rate made it attractive to build a PTCM system. It is not the time yet to decide whether the PTCM system is a truly practical system or not. More investigations are required to find out that the PTCM system is practical from all sides. The following step is to determine the average bit rate.

4.1.2 Average bit rate

The most important factor in the PTCM system is the average bit rate, since it will determine the feasibility of the system. It can be determined if the average word length is determined. Since the average rate of word was found as mentioned in 4.1.1, the average rate is then the produce of average bit in the word (average word length) by the average word rate. By measurement it was found that the average word length lies in the range of 5 to 6 bits per word.

If it is required to make use of the variable word length facility, it is necessary to have a start or stop signal of a length which can be easily distinguished from other bits. Hence if 1.5 bit stop signal is used, the total length of the word on average will become:

$$5.5 + 1.5 = 7 \text{ bit/word.}$$

Then

$$\text{the average bit rate} = 7 \text{ bit/word} \times 4.3 \text{ kword/sec.}$$
$$= 30.1 \text{ kbit/sec.}$$

These measurements were done on male voices. The average bit rate for the female voice was slightly higher.

This result is encouraging. It is less than 50% of the bit rate in the PCM system. Up to now it was found that the investigations on the PTCM system are promising. Other major investigations need to be done to decide finally whether to build the PTCM system or not.

4.2 SPECTRUM OF WORD DISTRIBUTION

It is necessary to determine how the words in the PTCM system are distributed.

Figure 9 shows the system used to find the distribution of bit groups (words).

A long synchronous counter (12-stage) was used to count the pulses of the crystal oscillator. The crystal oscillator frequency is 1.0 MHz. The sampler generates pulses of 200 ns duration at any level crossing[(5)]. This sampler pulse indicates the end of a digital word and the beginning of another digital word. The 200 ns pulse width is reduced to about 60 ns by the monostable multivibrator MMV1. The reason for this reduction of width is

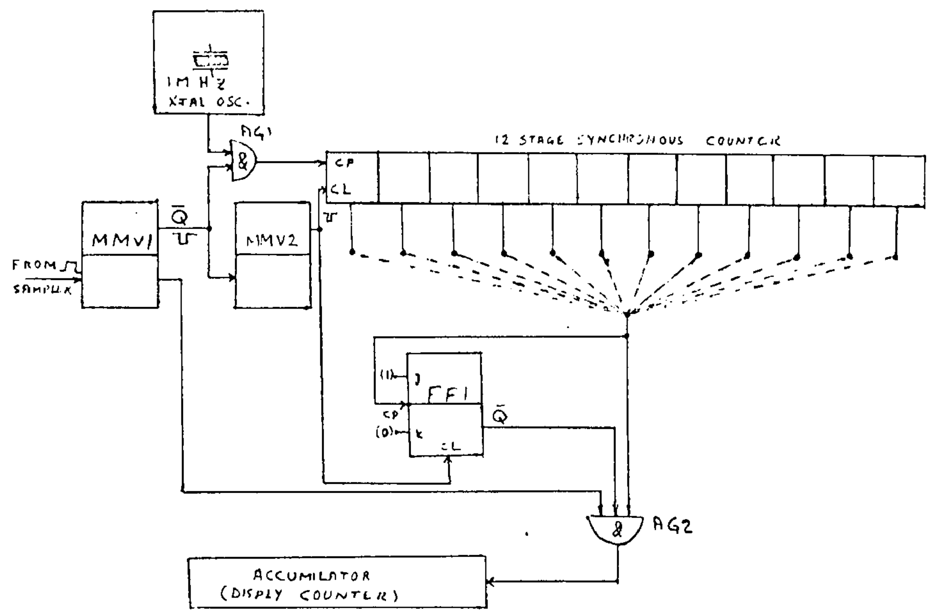


FIG. 9 <u>SYSTEM USED TO MEASURE THE DISTRIBUTION OF BIT GROUPS (WORDS</u>

to eliminate any possibility of race hazard since the same pulse is used to do other functions.

The $\overline{Q}$ of the MMV1 is used to stop the 1 MHz pulses from getting to the input of the counter. Also it is used to trigger another monostable multivibrator MMV2. $\overline{Q}$ of MMV2 which is of the same width as the first monostable multivibrator output pulse, is used to clear the counter and the flip-flop FF1. The clock input of FF1 is connected to one of the counter outputs at a time. This connection is moved to all the outputs of the counter, one after the other. The function of FF1 is to keep the gate AG2 open until the connected output of the counter goes from 1 to 0, then the gate AG2 is closed because the flip-flop FF1 changes state. When the counter output changes from 1 to 0 that means the word is longer than the one required in that test. The gate AG2 will stay closed even if the output connected to FF1 changes from 0 to 1 again. FF1 needs to be re-set or cleared to return to its original state. If the word is shorter than the one already under test, the direct connection from the output of the counter to the gate AG2 will keep this gate closed. To avoid any race hazard the gate AG2 is opened only for a short period when a new word is senses to be started. Namely, a word has been finished and a new one is coming. This action is done by using the Q output of the monostable multivibrator MMV1.

The accumulator counts the number of times the selected word occurs during the period of measurement. The procedure was repeated for every word length from 1-bit word length to

12-bit word length. The time for each run was 300 sec. This test was carried out without using the integrator between the clipping amplifier and the sampler and was repeated after having the integrator introduced before the sampler. The tests results are shown in the graphs of Figure 10 and Figure 11.

The average bit rate can be calculated from these graphs. Since Figure 10 obviously gives higher bit rate than the conventional PCM rate, it is not necessary to use it. The average bit rate obtained from the graph of Figure 11, i.e. by using the integrator was:

Male average bit rate = 32.6 kbit/sec.

Female average bit rate = 46.4 kbit/sec.

∴ The average bit rate of male and female = 39.5 kbit/sec.

This value of average bit rate includes 1.5 bit as a sign of separation between words.

This figure is optimistic. It shows that the bit rate can be reduced to less than 2/3 of the original PCM rate which is 64 kbit/sec. It was found by deep investigation that the average bit rate is not the most important factor in the system. It will be shown later that the most important factor in fact is the store size.

Comparing the results got by both methods of measurement; average word length method and word distribution method, it can be seen that they are enhancing each other. The first method gave a result for average bit rate of 30.1 kbit/sec for male voice, while the second one gave an average bit rate

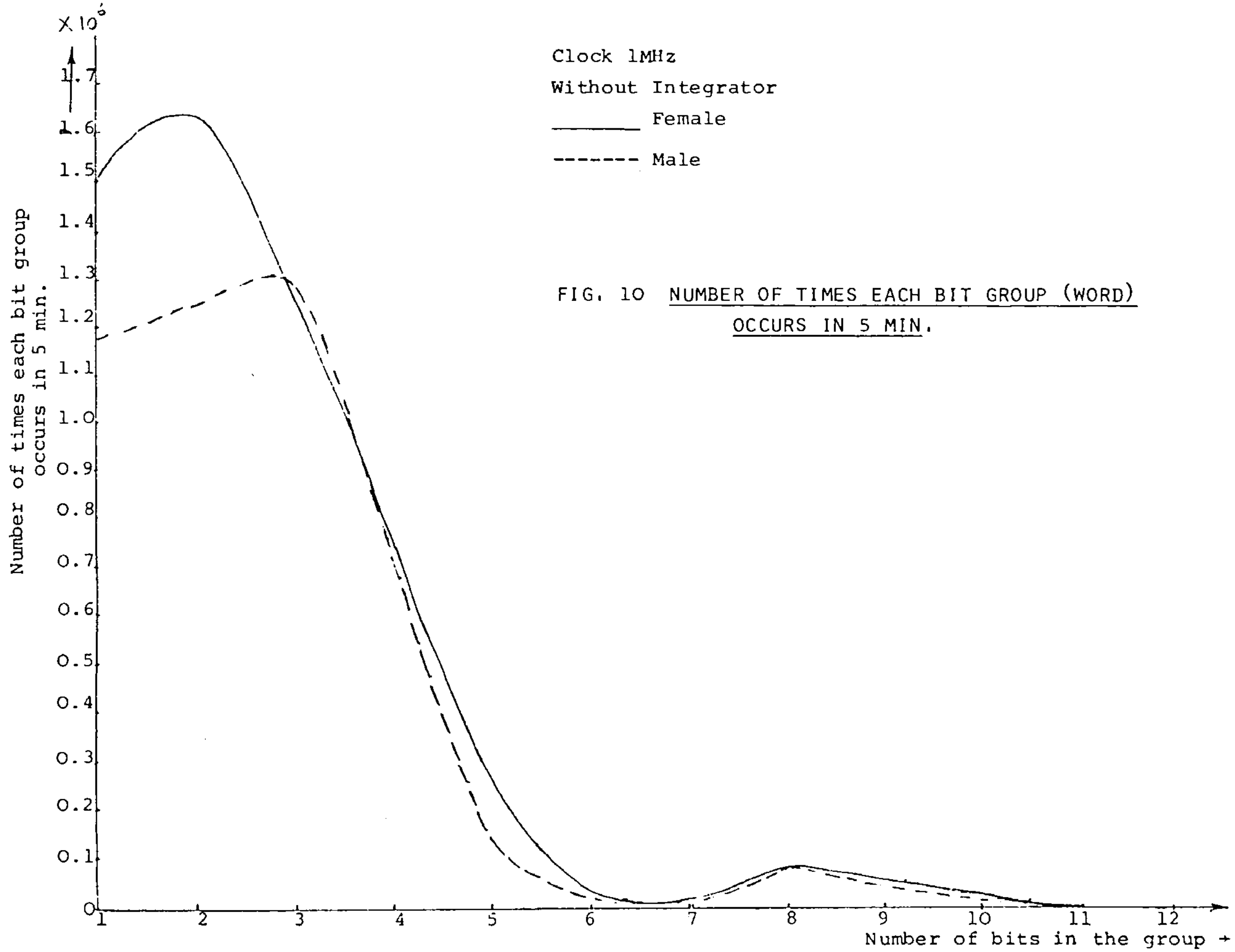


FIG. 10 NUMBER OF TIMES EACH BIT GROUP (WORD) OCCURS IN 5 MIN.

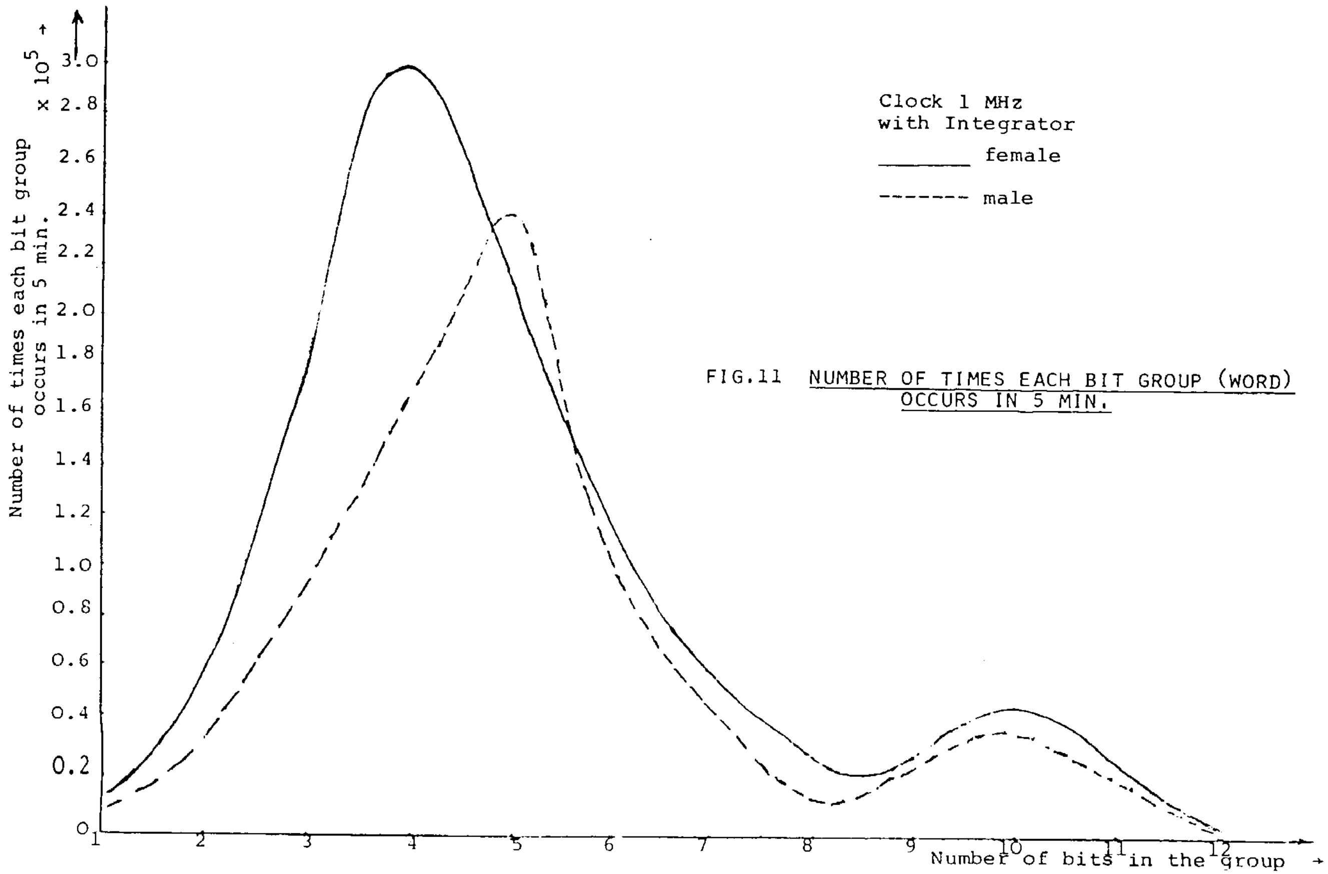


FIG.11 NUMBER OF TIMES EACH BIT GROUP (WORD) OCCURS IN 5 MIN.

of 32.6 kbit/sec. for male voice also. The two results are not widely different.

There was no satisfactory explanation for the second peak shown in the graphs of Figure 10 and Figure 11. Trials were done to get rid of this peak without success. A good filter was used to suppress the noise when there was no actual signal from the recorder. All these trials did not eliminate this peak. It could be produced by 50 Hz hum picked by the leads and the test circuit, or probably cuased by some sort of noise within the speech signal bandwidth, caused by the tape recorder or in the tape itself, when it was originally recorded.

In the case of measurement without using the integrator it was necessary to stop the noise coming when there was no speech. This was done by analogue switch controlled by the speech signal voltage itself. When the signal exists the analogue switch comes on and lets the signal pass through. When the signal is absent the switch turns off and stops any noise from getting through. This circuit of the analogue switch was not necessary when using the integrator, because the noise level was so low and hardly effective.

From the graphs in Figure 10 and Figure 11, the following results can be obtained in addition to average bit rates.

(a) The minimum group of bits is 1. This means that the minimum time-interval that must be measured and coded 'Δt_{min}' is 1 μs, since the measuring clock period is 1 μs. Again this result is the same result found by the method of Chapter 3. 'Δt_{min}' was found in paragraph 3.3b to be 1 μs sec.

(b) The maximum group of bits is 11 for Figure 10 and 12 for Figure 11. This means that the maximum time-interval to be coded and considered as speech signal 'Δt_{max}' = 2^{12} μs ≃ 4 ms. Any coded interval over 4 ms will be considered as silence.

(c) The most frequent groups are 2-3 bits for Figure 10 - without the integrator - and 4 - 5 bits for Figure 11 - with the integrator.

CHAPTER 5

STORE SIZE

5.1 STORE SIZE MEASUREMENT

Having the slope distribution of the speech signal voltage and the word distribution of the sampled speech signal voltage determined, the other important factor left is the store size. Previously it was thought that once the maximum slope of speech signal voltage was defined, it would be possible to calculate the required size of the store. Later it was found that it was not so simple to determine the store size from the maximum slope of the speech signal voltage. Actual measurement is essential to find the required store capacity. A practical method of measuring the store size was developed. The simple system shown in Figure 12 was used for this purpose.

The operation of this circuit can be explained as follows. The capacitor C represents the store. It is a good quality capacitor. The leakage current should not be high so that the results obtained will be not far from reality. The voltage developed across the capacitor C represents the data in the store at any instant. The pulses generated by the sampler triggers a monostable multivibrator. The output of the monostable multivibrator is a constant duration pulse and is generated after each triggering. The constant duration pulse generated by the monostable multivibrator charges the capacitor C with constant quantum

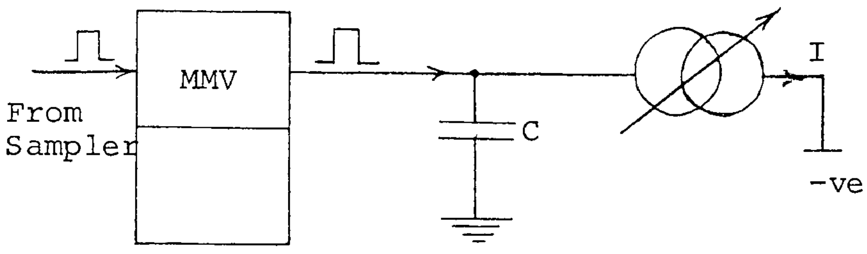


FIG. 12 SYSTEM USED TO DETERMINE STORE CAPACITY FOR PTCM SYSTEM

of charge. The constant current source discharges the capacitor C at constant current. This discharge represents the transmitted data coming out of the store. The discharge current can be adjusted until it becomes equal to the long term average of the charging current. The test using this system was repeated for both male voice and female voice. The discharging current was adjusted for each case to be equivalent to the average of the charging current. The output is simulated by the voltage developed across the capacitor C. This output voltage was plotted by means of slow moving X-Y plotter. The procedure was done for a period of about 150 sec. for each male and female voice. The plots are shown in Figure 13.

The results which can be obtained from Figure 13 are: for male voice a store size is required of 76 kword. For the female voice, the store size required is 87 kword. The word length in both cases is 12 bit/word.

Hence the average store size = 81.5 kword

Word length = 12 bits/word

This result for store size does not take into account that there is an acceptable degree of overflow. If the overflow of one per cent was considered to be acceptable, only a slight reduction in the store size will result.

The store size can be reduced to about 90 per cent of its actual size by using some sort of feedback to control the transmission rate by the contents of the store. This method will reduce the expected economy of the system since

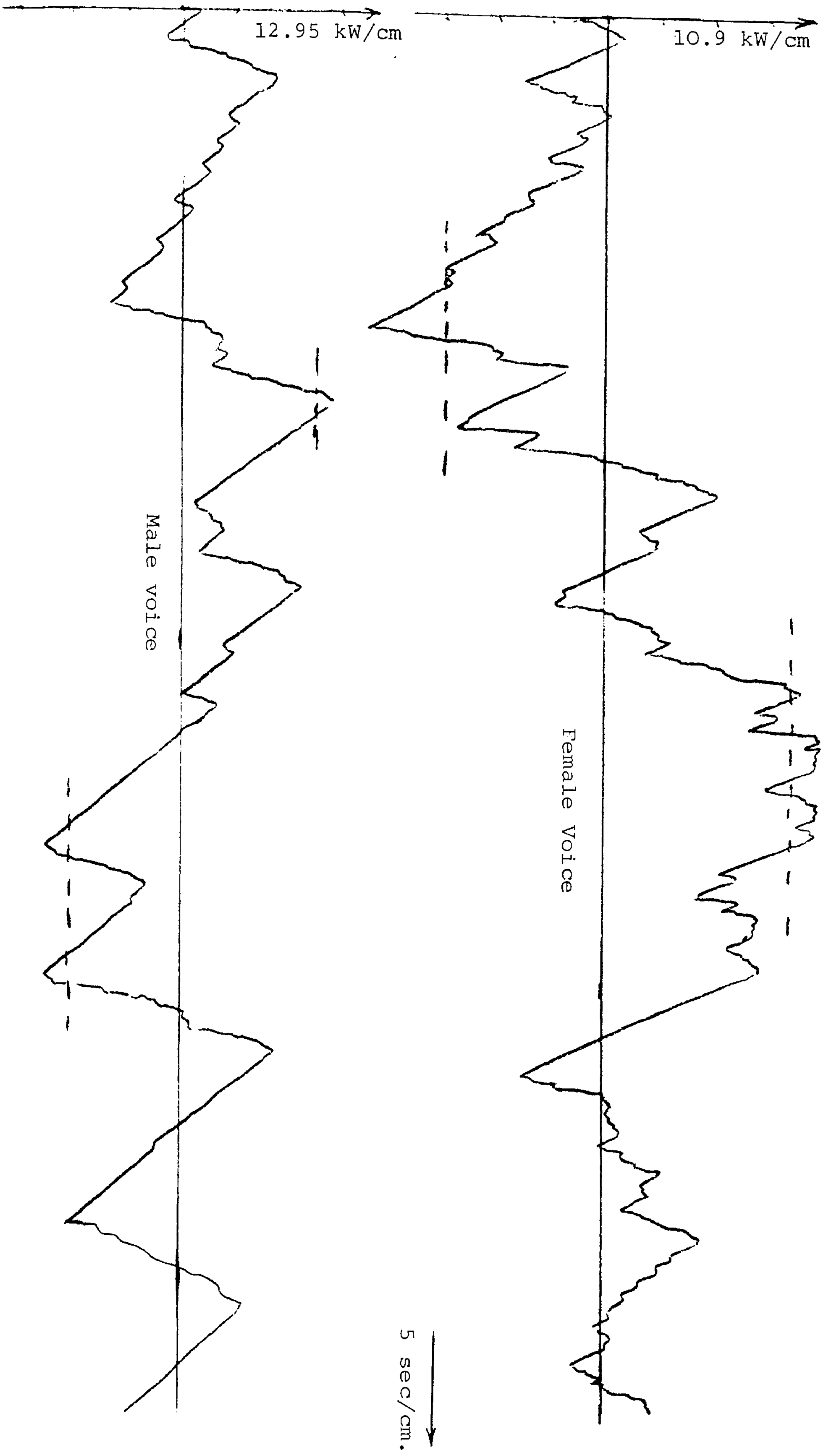


FIG. 13 X-Y PLOTS SHOWING THE INSTANTANEOUS STORE OCCUPANCY IN PTCM SYSTEM

the bit rate will vary to higher than the average rate. If the overflow beyond the dotted lines in Figure 13 is taken as acceptable to happen, the store size will become:

Store size for male voice = 58.3 kword.

Store size for female voice = 70.85 k word

Therefore the average store size required = 64.6 kword

5.2 DELAY CAUSED BY THE STORE

The delay in the store of the PTCM system was discussed by Binal(1) and by Mountis(2). It would be useful to repeat the formula they derived for the maximum delay in the store of this system:

$$\Delta t_{store(PTCM)} = \frac{S_{max}}{f_{wo}}$$

where

$\Delta t_{store(PCM)}$: The maximum delay in the store of PTCM system in seconds.

S_{max} : The maximum store capacity

f_{wo} : Word transmission rate.

f_{wo} is found in paragraph 4.1.1 to be 4.3 kw/sec.

$$\text{Then } \Delta t_{store(PTCM)} = \frac{64.6 \text{ kw}}{4.3 \ \frac{\text{kw}}{\text{sec}}} = 15 \text{ sec.}$$

This means that the signal may suffer from a delay of up to 15 sec. in the store only. The effect will be worse if the full store size was taken into account without allowance for any overflow. The delay in this case will become:

$$\Delta t_{max} = \frac{81.5}{4.3} \simeq 19 \text{ sec.}$$

The effect of delay on the PTCM system will be discussed in the following item.

5.3 EFFECT OF DELAY IN THE PTCM SYSTEM

In telephony there is a specified allowable delay in each direction. This allowable delay was found experimentally[8]. The speaker in one side must get a reply from the other side within a certain time or he may think that the other side did not get his message. The acceptable delay for each direction in telephony was found to be 0.5 sec. or even less[4][8]. This value of acceptable delay when compared to the delay which is going to take place in the store of the PTCM system, 15 sec. or 19 sec., is widely different. The effect of this delay on the application of the PTCM system will be discussed later in the following chapter. The delay of 15 sec. is only in one side of the system. Certainly an equivalent delay in the other side will take place, because there is a store in the receiver also like the transmitter store. This means that the total delay over the system will exceed 30 sec.

CHAPTER 6

DISCUSSIONS AND CONCLUSIONS OF PTCM SYSTEM

6.1 SAVING IN BIT RATE BY USING PTCM SYSTEM

As mentioned previously, the aim of this system is to find a way of reducing the bit rate compared to conventional PCM. Theoretically this system was discussed in detail by M.E. Binal[1] 1974 and completed by M.G. Mountis[2] 1978. Also a lot of practical tests were done by both of them in order to investigate the feasibility of this system. Their investigations led Binal[1] and Mountis[2] to the conclusion that this system is practical and probably a breakthrough in its subject. It was found from the investigations they did that the PTCM system will produce great saving in the bit rate over the conventional PCM. Their results showed that at least 30% saving will be possible. In spite of the complexity in the suggested PTCM system the hope of saving around 30% in bit rate over the PCM system makes it an attractive one.

The results obtained in this research provided a solid ground on which one can rely. It is possible now after having the required data to give a final decision concerning the feasibility and practicality of the PTCM system. Without these valuable data being obtained for this research the decision was not final.

6.2 MAXIMUM SLOPE OF SPEECH SIGNAL VOLTAGE

In the early stages of this research it was thought that once the maximum slop of standardized speech signal voltage was determined it would be possible to determine the most important factors of the system. It was though that the bit rate and the store size would be easily determined from the maximum slope, in addition to the minimum time interval to be coded Δt_{min}. After having this maximum slope determined it was found that it was useful only in the determination of Δt_{min}. Although the average bit rate and the store size are in direct relation with maximum slope of speech signal voltage, it is not easy to derive a formula to calculate them with the help of maximum slope of the signal. It was found that the speech activity is of more importance to the average bit rate and the store size than the maximum slope of the standardized speech signal voltage. Referring to Figure 13, it is easy to see how the data are building up in the store in the presence of the speech signal, irrespective of the slope of this signal. This building up varies according to the slope of the signal but in general it fills up the store. The discharge of the store happens only in the case of pure silence. Previously it was believed that the data will fluctuate above and below the middle of the store, according to the slope of the signal. The reason for this unexpected result will be discussed later. The conclusion concerning the maximum slope of the speech signal voltage is that it is not so important in the PTCM system, as it was assumed to be previously. This assumption will be correct only in the case of taking the average number of crossings of continuous speech without including the pauses. It is obvious in this case that the word rate is at least double the average rate of single side of telephone conversation. This problem will be discussed in the next paragraph.

6.3 LONG TERM AVERAGE RATE

All the tests carried out to find the average rate of the number of crossings (word rate) were done for a period of 5 minutes (300 sec.) or there about. This obviously included the silence gaps of one side on the telephone line while the other side was speaking. The measurements using this technique gave optimistic results. It was found that a saving of 30% to 40% over the PCM system is possible.

Looking deeply into Figure 13 again, the method of long terms averaging needs to build a store capable of doing this averaging. It is possible to see that the data takes a long time to build up in the store (10 seconds or more). A pause of the same length of time or slightly more often takes place after each talkspurt. Henc the time for averaging one talkspurt is around 20 seconds.

Since the average word rate was found experimentally to be around 4.3 kword/sec., as mentioned in paragraph 4.1.1, taking into account that during the pause there is no data entering the store, then the actual rate at which the data are coming into the store during the presence of speech is double the above figure, i.e. 8.6 kword/sec.

The average length of talkspurt was found to be 4.14 seconds[2]. This figure cannot be taken as it is, to calculate the required store size for the following reasons:

(a) The store size must be large enough to accept the longest talkspurt and not the average one. To explain this, the following example can be used;

Suppose a store of sufficient capacity to handle data for 4.5 sec., which is slightly above the average talkspurt length. If a talkspurt of 10 seconds occurs, then the store will hold 4.5 sec. of it provided that it is completely empty. The 5.5 sec. of this talkspurt will be definitely lost, because there is not enough room in the store to hold it. That means more than one half of the talkspurt is already missed. This great loss of data will damage the whole talkspurt. Unless the store has enough capacity to hold the longest talkspurt, then there will be a great loss of the message and the speech will be badly affected. There is a certain limit of overflow which must not be exceeded. A loss of coded interval of time that can be equivalent for the time consumed in saying one word of the speech will not be acceptable. For example, if the shortest word in the English language 'no' was lost because of overflow, then the total meaning of the sentence will be reversed. To say the word 'no' the required time is less than one second. Hence if the maximum talkspurt is 10 seconds, it is not possible to reduce the store to hold 9 seconds of information. In fact, this is a great difficulty, even if the store will be large enough to hold the longest talkspurt. The reasons for this is the difficulty in predicting the longest talkspurt. It may happen that a speaker will keep speaking continuously for 15 sec. or 20 sec. or more, that means the store size will be very hard to limit. Then the store size required will not be easy to determine.

(b) The calculation to find the store capacity cannot be done by multiplying the average digital word rate by the average talkspurt. The reason for this is that the average digital word rate was determined by long-term averaging of the generated digital words. Hence the pauses had been included in finding the digital word rate. From (a) it was shown that the maximum talkspurt must be taken into account to find the required store size. In addition to the maximum talkspurt period, an equivalent pause period at least should be added to it, and then multiplied by the average digital word rate, or in mathematical expression:

Store size required = Average word rate x (Maximum talkspurt period) x 2).

If the talkspurt taken to be 10 sec.

then the required store size = 4.3 kw/s x 10 x 2
= 86 kword.

This figure is not far from that counted before in paragraph 5.1, which was 81.5 kword.

6.4 SILENCE GAPS

The PTCM system was based on the fact that there are long silence gaps during telephone conversations. Since the time intervals are coded in binary form, the silence gaps will be a great source of coding time interval, with little increase in the number of bits of the code. This is because an increase of one bit in the word means doubling the time-interval coded. Because of this feature it was believed that the silence gaps would be the main source of reduction in the transmission bit rate. Unfortunately, it was found that these silence gaps are of very little use in the PTCM system. To explain this conclusion, take the following example.

Starting with an empty store, say a talkspurt of 10 seconds was loaded into the store. The transmission starts as soon as the data is available in the store. Apparently the first group of data will be transmitted without delay. The flow of data into the store is quicker than the flow of data out of the store during the presence of speech, since the transmission of the digital words of this talks-spurt is at the average rate. Suppose that the telephone call consisted of 50% speech and 50% silence, the time required to get the talkspurt completely transmitted from the store will be twice the time of the talkspurt itself. The delay to the last word of the talkspurt will be nearly 10 seconds. The effect of the delay will be discussed in the following part.

6.5 SYSTEM DELAY

6.5.1 System Delay Effect on Duplex Applications

As mentioned before, PTCM system introduces a delay to the signal. This is a result of using a store in this system. Telephony specifications set certain limits for the allowable delay. The allowable delay of the message from the transmitter to the receiver in telephony, is nearly 0.5 seconds or even less[4,8]. If the delay in the PTCM system is equal on both sides, then the allowable delay for each side will be 0.25 sec..

If the data are flowing out of the store at a rate of 40 kbit/sec., the average length of each digital word was found to be 4.5 bit/word, hence

$$\text{The average digital word rate} = \frac{40}{4.5} = 8.9 \text{ kword/sec.}$$

Now it is possible to calculate the allowable store size of the PTCM system:

$$8.9 \, \frac{\text{kw}}{\text{Sec}} \times 0.25 \text{ sec.} = 2.25 \text{ kword}$$

This is the store allowable to be able to operate the PTCM system within the specified limit of delay in telephony. Unfortunately it was found that the store size required to get satisfactory operation of the PTCM system in Chapter 5, was 81.5 kword, which is too much greater than 2.25 kword. It is clear by comparing these two figures that the application of the PTCM system to duplex communication is limited.

6.5.2 System Delay Effect in Simplex Applications

Simplex transmission such as telegraphy, radio, television, etc., has no restriction concerning the delay of the signal.

Binal[1] and Mountis[2] were worried about the delay in the PTCM system, although they did not investigate it deeply. They suggested that in the case of failure to make use of the PTCM system in duplex system, there would be no problem in applying this system in simplex transmission systems[1,2]. This suggestion was not investigated. It is possible now to examine the possibility of using the PTCM system in simplex transmission.

If in a telephone call the speech activity is 50% or less, in simplex transmission the speech activity is much higher. Tests were done by applying the speech of the telephone call of both sides simultaneously to the sampler. It was found that the average bit rate goes up to 80 kbit/sec instead of 40kbit/sec. This is obviously higher than the conventional PCM system bit rate which is 64 kbit/sec. Then the trial to use the PTCM system in simplex transmission was not successful.

The conclusion must be that the PTCM system shows little advantage over the conventional PCM system, either for simplex or for duplex transmission of speech signals.

6.6 OVERFLOW IN THE PTCM SYSTEM

The digital word in the PTCM system represents the coded time-interval. This time-interval can vary from 1.0 μs to around 4 ms for active speech. When a pause is coded, it will be more than 4 ms time-interval. These time-intervals are represented by one sample. The overflow will cause a loss of one sample or more. If the sample represents short time-interval the effect on the reconstituted signal will not be great, but when the sample represents a time-interval of a few milli-seconds, then the distortion on the signal will be effective. This is the case of the loss of one sample, so the effect of the loss of several samples will cause great damage to the orignal signal, especially when the samples represent long time-intervals. For example, a loss of 10 samples of 3.5 ms interval means a loss of a period of 35 ms.

6.7 THE RESULTS OF THE PTCM SYSTEM RESEARCH

It is disappointing to find that, after this long period of research, the PTCM system offers so little real advantage over the conventional PCM system. However, in the course of this research a great deal of new information and deeper understanding of the properties of speech signals have been acquired, which may well be of use in a related field of research.

In the course of the research previously described, it becomes apparent that some of the principles developed might be applied with advantage to conventional PCM system. An investigation was therefore carried out to determine the possibilities of a new type of PCM system, utilizing these principles, and this work is described in the following chapters. The new system is called Averaged Rate Pulse Code Modulation (ARPCM).

CHAPTER 7

DESCRIPTION OF THE AVERAGED RATE PULSE CODE MODULATION SYSTEM (ARPCM)

7.1 VARIABLE WORD LENGTH TRANSMISSION

The word length in the PCM system is fixed. Eight bits per word becoming standard word length in the PCM system, although in special cases other word lengths are used. In all cases the word of the PCM system can be divided into two parts. The first part is the code for real data. This part carries the whole information and it varies from one bit to the full length of the word, depending upon the amplitude of the sample taken. The second part is the non-used bits of the word. These bits do not carry any useful information. They are used to complete the word to its full length. If these non-used bits were removed from the words, then there would be a possibility of reducing the transmission bit rate. In order to make use of the improvement in the average rate of the used bits, it is necessary to average them for a certain time. It is necessary to store the samples and then transmit them at the average rate of the used bits of the words. Then the word transmission rate will be variable while the bit transmission rate will be constant and equal to the average bit rate of the used bits of the words over the period of storage. The longer the word is, the longer the time is required to transmit it, and vice-versa. Since the words now have variable length it is necessary to indicate the beginning and the end of each word. This can be done in the same way as used in the teleprinter. A stop signal of 1.5 bit length space, to indicate the end of the word, followed by a start signal of one bit long mark to indicate the beginning of a new word[6]. This method decreases the advantage gained by the removal of the non-used bits. This decrease is regrettable but there is no way to avoid it.

7.2 AVERAGING THE BIT RATE OF THE PCM SYSTEM BY STORAGE

The PCM samples come continuously at constant rate of 8 kword/sec. Normally the word length in the PCM system is 8 bit/word[3,7]. Apparently the eight bits are not always used to convey useful data. Moreover the silence periods are sampled at the same rate. These silence periods samples are transmitted as an eight bits word, while there is very little information in these words. Hence the word length which is carrying real information varies from zero to eight bits.

Variable word length transmission technique mentioned in 7.1 will help to remove the non-used bits from the digital words of the samples. If the removal of these bits is done instantaneously, then the word to be transmitted will be variable in length, and according to that the bit rate will be variable. The word rate will stay at the sampling rate and the bit rate will vary from 8 kb/sec. to 64 kb/sec. according to the word length. So no benefit will be gained by this method because the maximum bit rate will be as before and then the line capacity will not be increased. Hence the samples should be stored for a sufficient period to get the average of the used bits. The storage can be done in a digital or analogue way. Since there is no analogue store similar to the random access memory, the digital store should be used. The samples are fed to the store at the sampling rate. The transmission will be at a constant bit rate but the word rate will vary as mentioned in 7.1. The ideal store should not give any chance for overflow so that no loss of any part of the data will take place. It is hard to judge to what limit it is possible to accept overflow. This case of overflow is similar to that in the PTCM system. It was discussed in the PTCM system paragraph 6.3.9, that a one second loss of data is considered too long, since this time might be enough to cause one word of speech to be lost. According to this fact it is better to ignore the possibility of overflow. The design of the store should be done in such a way that there is no overflow allowed in the system. Overflow problem can be controlled by some sort of feedback to adjust the transmission rate. Underflow is also not allowed in this system, because it will introduce samples into the receiver's store which are not originally in existence. The underflow effect can be stopped by proper design in the transmitter.

The transmission rate is the average rate of the actual bits entering the store, plus the signal which indicates the end of each word which is 1.5 bit length.

7.3 STORE IN THE ARPCM SYSTEM

As in the PTCM system the store in the ARPCM system is the important element in this system. The function of the store has been mentioned in the previous paragraph 7.2, i.e. to average the bit rate of the samples over a period enough for this purpose. The number of bits in the store per location is equal to the number of bits in the word of each sample. In the PCM system the number of bits in the sample is generally 8. Namely, the store width required is also 8 bits. It is very hard to find a theoretical means of calculating the required store size. A practical method is the easiest way to find the suitable store capacity. This method of measurement will be explained later.

7.4 COMPARISON BETWEEN THE TRANSMISSION IN THE PTCM AND THE ARPCM SYSTEMS

In the PTCM system, the sampler generates a pulse each time the signal crosses one of the predetermined threshold levels as mentioned before, and also described by Binal[1] and Mountis[2]. This generated pulse indicates an end of a digital word and a beginning of a new one. When the rate of change of speech signal voltage is high, the digital words will come fast, while when the rate of change of speech signal voltage is low, the digital words will come slowly. When there is no signal, the words will be generated much more slowly, depending on how long it is possible to make the word. These words will be stored temporarily in a store large enough to give long term averaging of the rate of the used bits. The information stored in the store are the coded time intervals from the beginning of a word to the beginning of another. The words will be transmitted at constant bit rate after the addition of a marking signal. This marking signal is a space of 1.5 bit long, always followed by a mark which is the MSB of the following word. As mentioned before, this marking signal is essential to distinguish between a word and the following one. This technique of marking signal is important only when serial transmission is to be used. If parallel transmission is to be employed, the marking pulse will be unnecessary, but many unnecessary bits will be transmitted. The average bit rate will rise by the presence of these non-used bits and could spoil any possibility of reduction in the bit rate.

The transmission rate of words will be inversely proportional to the number of bits in the word. The words are coming into the store irregularly and coming out of the store also irregularly. In the ARPCM system, the encoder is generating a digital word representing one sample at a rate of 8 kword/sec. When the sample comes at a peak of the signal, it will occupy the whole bits of the word. In all other cases when the sample does not come on the peak, some of the bits will not be utilized. Then the words will be stored in a store in the same way mentioned in the PTCM system. The transmission is similar to that of the PTCM system. Each word is taken out of the store, the unused bits will be dropped. It is also necessary to use serial transmission. The bit transmission rate will be the average rate of the useful bits. The same way mentioned in the PTCM system for the separation between words has to be used in the ARPCM system. Also the word transmission rate is inversely proportional to the word length, i.e. the used bits of the word.

The difference between the PTCM system and the ARPCM system is that the flow of data words into the store of the transmitter in the first system comes randomly, while in the second system these data words are coming at a fixed rate. In the receiver side the data words are coming out of the store randomly in the case of the PTCM system and they are coming out at a constant rate in the ARPCM system. In both systems the transmission of data is done by the constant bit rate and variable word rate.

7.5 TRANSMISSION RATE AND STORE RELATIONSHIP

Although the ARPCM system can be used for transmission of any kind of signal, this research was concentrated on the transmission of speech signal. The properties of speech signal has been studied carefully during the investigation of the PTCM system and ARPCM system. Since the speech signal is non-uniform, it is useful to assume it as a stationary random signal. An exact defined mathematical formula for store parameters cannot be derived since the number of bits at any sample is random.

The transmitter store is receiving data at a regular rate which is the sampling rate, normally 8 kHz. The data are transmitted out of the store via the transmission channel at variable and random rate, as mentioned previously. The average rate of transmission is equal to the input rate. The rate of change of data in the store depends on the difference between the rate of the data coming into the store and the rate of data leaving the store at any instant.

Let us consider two successive samples, i, j. The time interval between these two samples equals $\frac{1}{f_s}$ sec., where f_s is the sampling frequency which is equal to the word input rate into the store f_{wi}. Then the instantaneous store occupany at the sample j will be incremented by one word and decreased by the rate of transmission multiplied by the time interval from the occupancy at sample i, or in mathematical form

$$S_j = (S_i + 1) - f_{wo} \cdot \frac{1}{f_s}$$

where S_i and S_j are the store occupancy at samples i and j respectively.

Since $f_s = f_{wi}$ the above equation becomes

$$S_j = S_i + 1 - \frac{f_{wo}}{f_{wi}} \tag{7.1}$$

f_{wi} is constant, then f_{wo} is the factor which effects the store occupancy.

$$f_{wo} = \frac{f_{bt}}{n}$$

where f_{bt} is the bit rate of transmission before adding the signalling pulse.

n is the number of bits in digital word.

Theoretically f_{bt} is required to be constant but in practice it should be necessary to have it slightly varying with the contents of the store, in order to avoid the problem of building up or drying out in the store, caused by the difference between the actual average rate and the transmission rate. The number of bits in the word varies from one sample to another. Equation 1 can be rewritten as:

$$S_j = S_i + 1 - \frac{f_{bt}}{nf_{wi}} \tag{7.2}$$

If the term $\frac{f_{bt}}{nf_{wi}} = 1$

$$S_j = S_i$$

This is the equilibrium state condition where the store occupancy does not change and this state is the most desirable one. In practice it is very seldom that this state is obtained for long periods of time. The larger the store is, the smaller the deviation of store occupancy from this state will be.

The number of bits in the word varies between one to 8 bits. When n is high the occupancy of the store increases and vice versa. This is obvious, since the word with a high number of bits requires a longer time to be transmitted and the word with a small number of bits requires a shorter time.

Overflow condition occurs when $S_i = S_{max}$ where S_{max} is the maximum store capacity and

$$\frac{f_{bt}}{nf_{wi}} < 1$$

The store occupancy is affected by the transmission rate f_{bt}. When f_{bt} is high there will be no overflow, but rather underflow will be more frequent, while when f_{bt} is low overflow will be more frequent. If the store capacity is small, overflow and underflow may happen frequently.

To get satisfactory results, it is necessary to pay great attention to the design of transmission rate and store size required in this system.

7.6 PROBABILITY DENSITY FUNCTION OF DATA IN THE STORE

From the previous discussion, it was shown that the store occupancy at sample j which comes directly after sample i is

$$S_j = (S_i + 1) - \frac{f_{bt}}{njf_{wi}}$$

The state of store occupancy is a first order Markovian relationship because it depends only on the previous state[10,11].

Let $S_i > 0$

Then the next store occupancy S_j depends on the net amount of data during the period from event i to event j, which is one sample period. By averaging the store occupancy over a long period of time it is possible to find the steady-state probability function of the store occupancy.

From the definition of the probability density function it is possible to determine the probability density function of the store occupancy in the following way:

By definition[10,11] $$P(a \leqslant X \leqslant b) = \sum_{X=a}^{b} f(x) \qquad (7.3)$$

In the present case the variable X is represented by the instantaneous store occupancy S. Then for a period of one sample:

$$P(S_i < S < S_j) = \sum_{S=S_i}^{S_j} f(S) \qquad (7.4)$$

The store occupancy S is a dependant variable, it depends on the number of bits in each sample n and the bit transmission rate f_{bt}. If f_{bt} is kept constant, the only random variable affecting the store occupancy is the number of bits in each digital word, n.

The probability function of n cannot be derived straightforwardly, because n depends on the amplitude of the signal at the moment of sampling. Although the

LIBRARY

distribution of the signal amplitude, follows gamma distribution[4], as will be shown later, yet the companding for this amplitude complicates the matter. It is easier to measure the mean value of n.

Since the store occupancy depends directly on the number of bits n, it follows that the probability function of store occupancy s will have a probability function similar to that of the number of bits n.

The total probability of store occupancy throughout the store is the sum of the probabilities of the samples[10] assuming that:

(a) Each $P_i(s)$ is non-negative, $P_i(s) \geq 0$

(b) The sum of $P_i(s)$ is unity, $P_1(s) + P_2(s) + \cdots + P_n(s) = 1$

Therefore if $P_i(s)$ is the probability of store occupancy for one sample, then the store occupancy for the whole store is:

$$P(s) = \sum_{S=S_{min}}^{S_{max}} P_i(s) \qquad (7.5)$$

According to the condition (b) above this summation should equal to unity. $S_{min} = 0$ and S_{max} is the total store capacity. If there is underflow and/or overflow in the system, condition (b) cannot be correct.

7.7 UNDERFLOW AND OVERFLOW CONDITIONS

The underflow condition means that the store becomes empty, i.e. transmission rate becomes greater than the rate of incoming data into the store. The store will stay empty for the period of underflow. The probability function of underflow or zero store occupancy can be written as

$$P(S < S_{min}) = \sum_{S=-\infty}^{S_{min}} P(S)$$

where P(S) is the store occupancy density function.

Underflow could produce false samples at the receiving side if the transmission continues as usual. It should be necessary to stop the transmission when this condition occurs.

The overflow condition is the opposite of the underflow condition. It occurs when the store becomes full and the transmission rate is lower than the rate of data flowing into the store. The probability of store occupancy being larger than S_{max} can be written as:

$$P(S > S_{max}) = \sum_{S=S_{max}}^{+\infty} P(S)$$

where S_{max} is the maximum store size and P(S) is the store occupancy density function. The overflow condition is not desirable. It could affect the reconstituted signal badly. If the lost samples by overflow were samples of real signal their effect will be worse than thoseof silence. The silence samples when lost will bring the successive talkspurts closer, while the samples of active signal will distort the signal itself. When the loss of samples exceeds the time required to say the shortest word in speech, it becomes unacceptable. The overflow condition can be reduced greatly by means of feedback to control the transmission frequency by the store occupancy.

The previous discussion was all about the transmitter store. The receiver store does not differ from the transmitter store, so all the above discussion is also valid for the receiver store.

CHAPTER 8

INVESTIGATION OF THE ARPCM SYSTEM

8.1 SPEECH SIGNAL VOLTAGE DISTRIBUTION

In the ARPCM system there are two properties of speech signal voltage which dominate the design of the system. The first property is the maximum frequency, and the second one is the distribution of speech signal voltage. Since the speech signal is random, it is necessary to put some rules for the standardization of it. It was mentioned in the PTCM system, paragraph 3.2, the method of standardization normally followed in the PCM system, which was applied to the PTCM system.

Speech signal voltage distribution can be found in many references. For example, in reference 3 and reference 4, the matter has been discussed and explained. The following articles are copied from reference 4, page 64, 65 and 66:

" When the physical conditions are constant, vocal levels are distributed approximately according to the Gaussian probability distribution function with a standard deviation of about 2 - 8 dB. When handset telephones are used, the distribution of speech levels observed on telephone lines is influenced also by variations in the manner of holding the handset, and by the characteristics of the microphone, so that the standard deviation becomes about 4 - 6 dB for a fixed commercial telephone connection.

Under field conditions, measurements at local exchange or group switching centres are affected also by variation in the transmission loss from the talker's mouth to the point of measurement and then the standard deviation may be as great as 5 - 6 dB. Measurements on telephone connections containing carbon microphones are usually more dispersed than those on connections that contain linear microphones".[4]

"The amplitude distribution of speech can be expressed in various ways; using different integrating periods or none.

Different types of speech can also be considered. Fig. 14 shows probability density function for instantaneous amplitude of speech from a given talker with the microphone at a fixed position in front of the lips. For comparison certain theoretical distributions are shown on the same scale.

The exponential distribution has been used for certain purposes as a theoretical model to represent the distribution of instantaneous speech amplitudes. As can be seen in Fig. 14, the approximation is poor at very low levels and so it has been suggested that a more suitable model is to accept the exponential form for the vowel sounds which are taken to represent 40% of the total speech duration and to add to this a Gaussian distribution about 20 dB lower in power level to represent the consonants for the remaining 60%. Alternatively two Gaussian signal 23 dB different in level and equally probable (representing vowels and consonants) has been used.

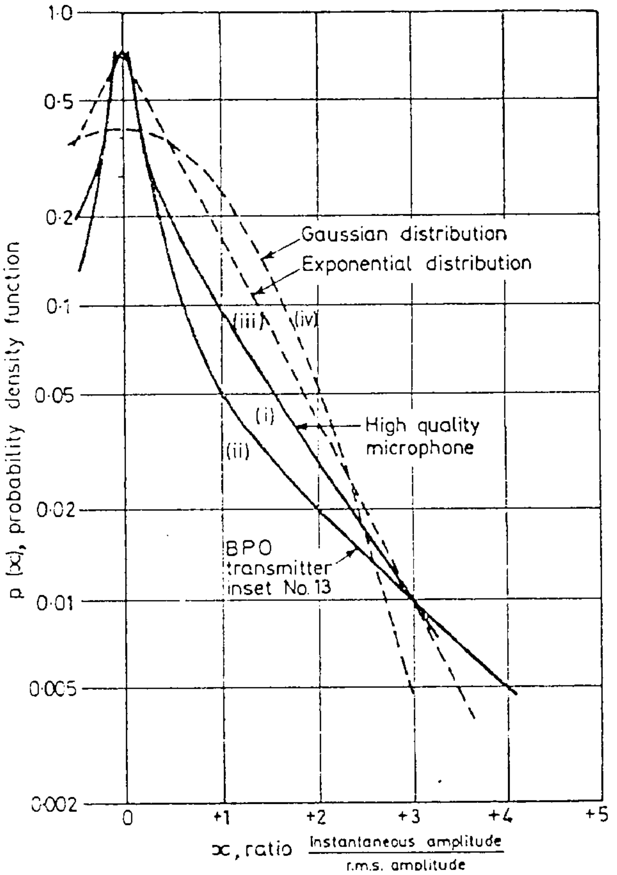


Figure 14. Probability density functions for instantaneous amplitudes of speech. (4)
Note. Negative values as for x positive

A much better fit to the distribution shown in Fig. 14 can be obtained by use of the gamma distribution with a suitable choice of value for the parameter. The gamma distribution can be represented by the expression

$$P(x) = \frac{k}{2\Gamma(\ell)} \cdot (kx)^{\ell-1} . e^{-kx} \qquad (8.1)$$

where ℓ is a parameter and

$$k = |\ell(\ell-1|^{\frac{1}{2}}$$

This expression gives the probability density for positive values of x. If $x = v/v_o$, where v is the amplitude and v_o is the r.m.s. value of v, the r.m.s. value of x is unity and the integral of P(x) from zero to infinity is 0.5. The negative values of amplitude are represented by an identical expression.

The waveformsof speech are often found to be slightly unsymmetrical even when derived from high quality linear microphones, the extent of asymmetry being more markered for some talkers than others. The characteristics of non-linear microphones, like carbon types, may increase this asymmetry to some extent.*

* This article is the footnote on page 66.

When $\ell = 1$, the exponential distribution is obtained but a value of $\ell = 0.5$ gives a good fit for high quality speech waveform and $\ell = 0.2$ can be used to represent the waveform of speech from one, rather old, type of carbon microphone. Fig. 14 illusrates these distribution functions and includes the Gaussian distribution for comparison.

Modern types of telephone microphone may be expected to conform to gamma distributions having the parameter ℓ lying in the range 0.2 - 0.5."[4]

(Whenever Fig. 14 is found in this paragraph copied from reference 4 it is Figure 2.6 in that reference).

The voltage distribution of the present telephone calls under investigation needs to be determined practically. The test to find this distribution was done by using the system shown in Fig. 15. The operation of this test system can be explained as follows.

The standardized speech signal is fed to a comparator of adjustable reference voltage. The output of the comparator is used to open a gate which enables the output of 1 MHz crystal oscillator. The output of the gate is fed to an accumulating counter. The gate is also controlled by a start-stop signal. For any setting of the reference voltage the test was taken for a period of 5 minutes. It was repeated for another setting of reference voltage of the comparator and so one, until a smooth graph was obtained. This test was done for both male and female voices. Analogue switching circuit was used before the comparator to stop the noise from getting to the comparator input when there is no speech. This analogue switching circuit was the same one used during the PTCM system tests, see Appendix B.

The results of these tests are shown in the graph of Figure 16 (a and b). These graphs in general are similar

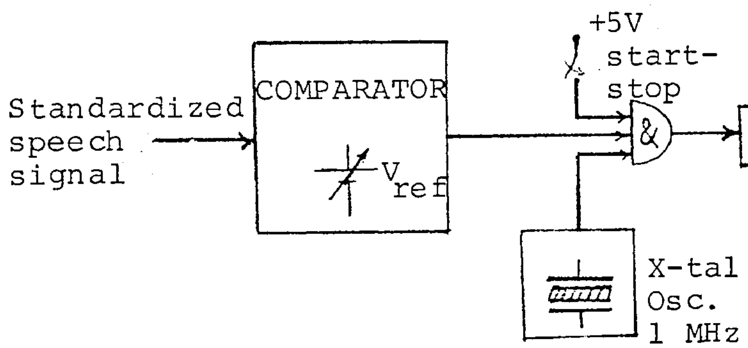


FIG. 15 <u>SYSTEM USED TO MEASURE SPEECH VOLTAGE DISTRIBUTION</u>

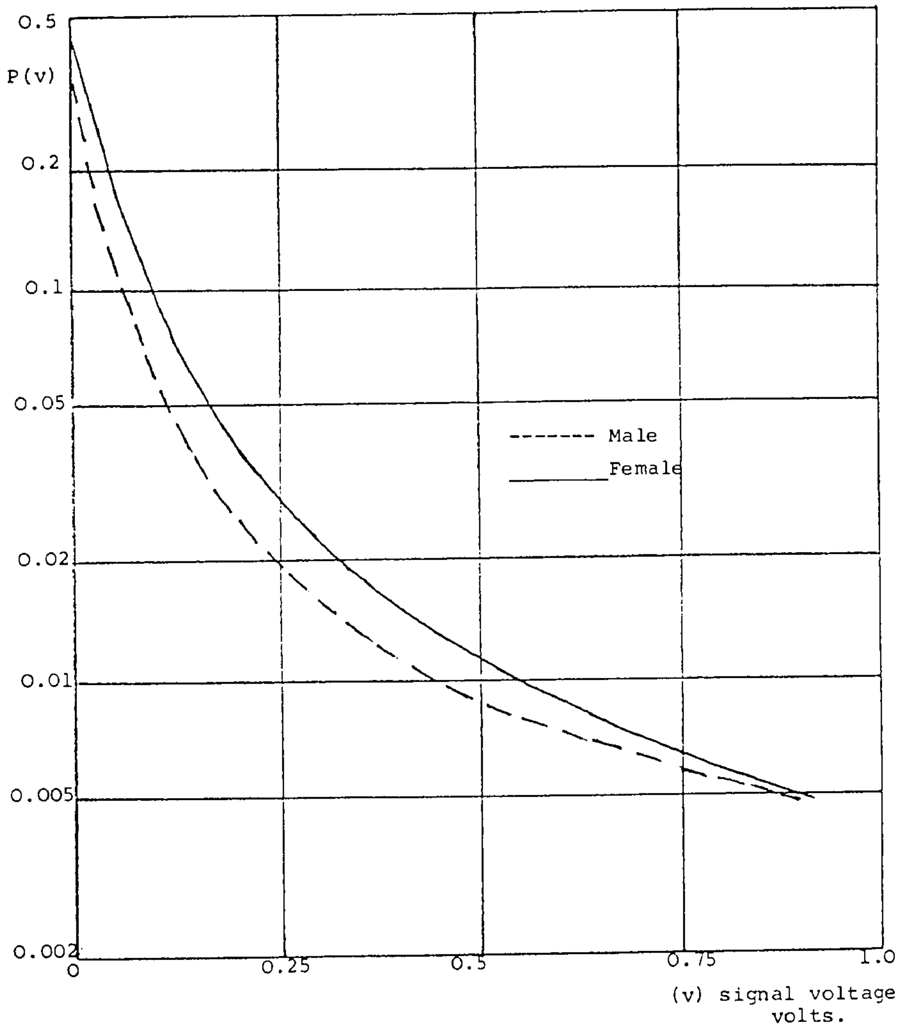


FIG. 16 PROBABILITY DISTRIBUTION OF SPEECH SIGNAL VOLTAGE

to those obtained by other authors (see references 3 and 4).

It can be seen from these graphs that the low levels of the signal are more frequent than the high levels. This means that the short digital words are more frequent than the long ones. From this one can predict that the reduction in bit rate will be considerable. Although the companding technique used in the PCM system will eliminate part of the benefit obtained, there will still be a considerable saving in bit rate. Companding is amplification to the low levels of the signal and attenuation to the high levels. This will cause the average bit rate to be increased. The addition of stop signal of 1.5 bit length will again increase the average bit rate to be used for transmission. All these factors will be discussed and taken into consideration in the following discussion.

8.2 AVERAGE BIT RATE

8.2.1 Average Bit Rate Calculation

The distribution of speech signal voltage makes it possible to calculate the average bit rate in the PCM system. The average bit rate in this case represents the average bit rate of the used bits after the removal of the unused bits. The voltage values taken from Figure 16 to represent the sample voltage must be changed to new values according to the law of companing used in the system. The commonly used laws of companding are the A-Law and the μ-Law. In the present case the A-Law will be considered as it is the standard law used in European systems.

The graph of Figure 16 was used to calculate the average bit rate for both male and female voices. The voltage values were treated according to the A-Law companding technique. Addition of 1.5 bit was important to have the true average bit rate to be used for transmission. The results of these calculations were:

$$\text{Average bit rate} \begin{cases} \text{Male} = 28 \text{ kb/sec.} \\ \text{Female} = 32 \text{ kb/sec.} \end{cases}$$

From these results it is easy to notice that the overall average bit rate is 30.0 kbit/sec. This value is less than half the bit rate in the conventional PCM system. That means great saving in transmission rate can be achieved by the application of the ARPCM technique. In order to reinforce these results and to make sure that they are correct, a practical method of measurement was used. This method will be discussed below.

8.2.2 Average Bit Rate Measurement

In order to measure the average bit rate for the ARPCM system it was necessary to build special test equipment. The system used for this purpose is shown in Figure 17. The data coming from the PCM encoder DF341 are fed into 8 bit shift register. The shift register is used to convert the data from serial form to parallel form. When eight bits enter the shift register these bits are transferred to the latch for holding them temporarily. From the latch

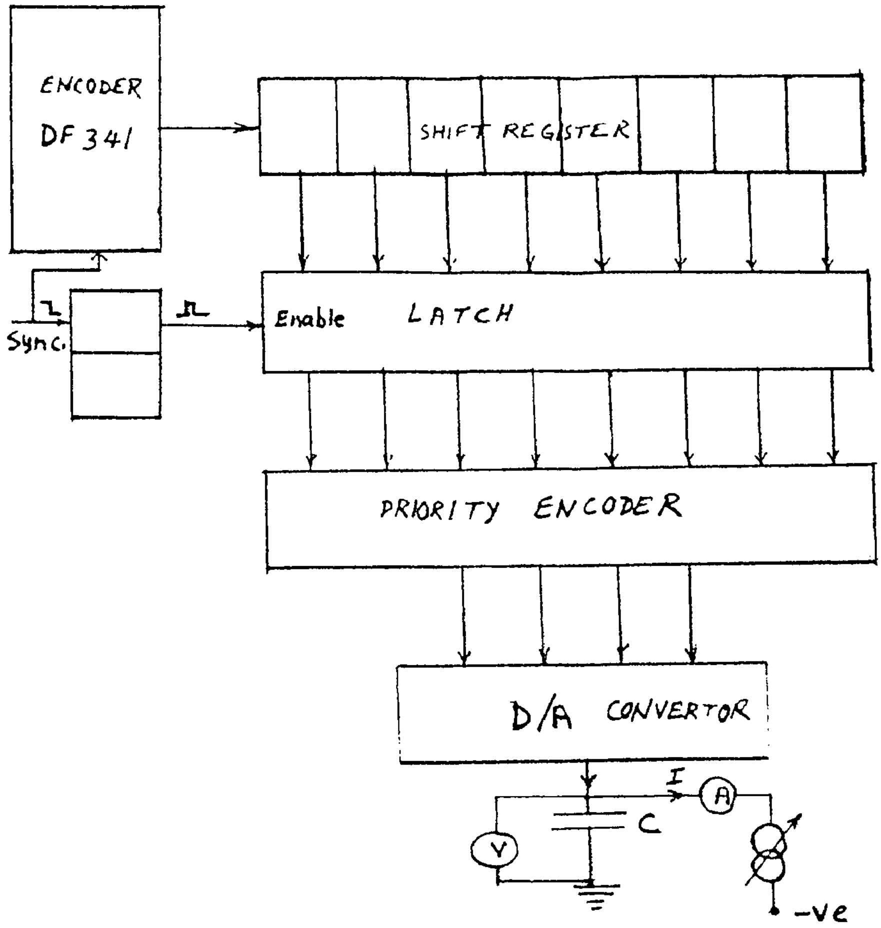


FIG. 17 BLOCK DIAGRAM OF THE TEST CIRCUIT USED TO FIND THE AVERAGE BIT RATE OF ARPCM SYSTEM

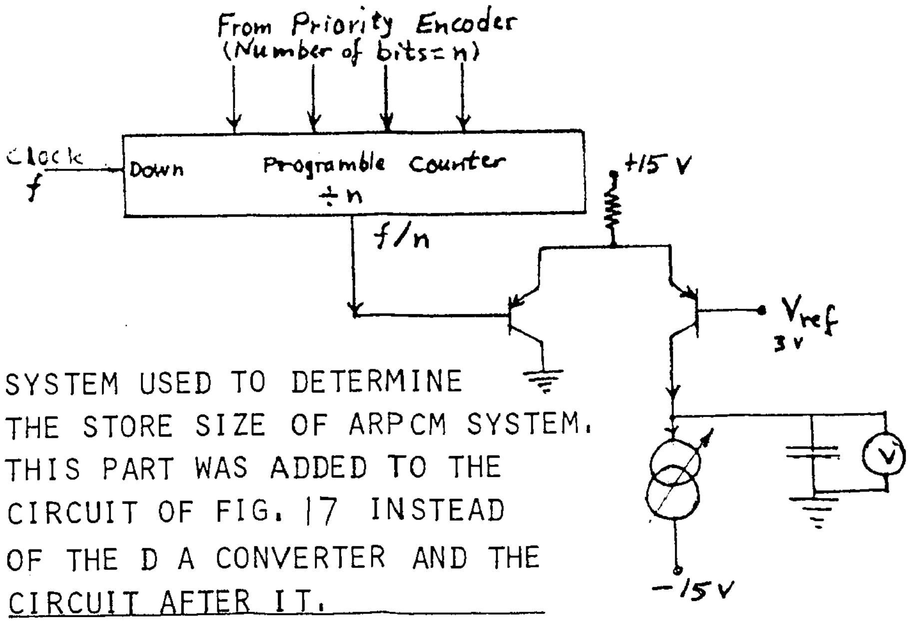


FIG. 18 SYSTEM USED TO DETERMINE THE STORE SIZE OF ARPCM SYSTEM. THIS PART WAS ADDED TO THE CIRCUIT OF FIG. 17 INSTEAD OF THE D A CONVERTER AND THE CIRCUIT AFTER IT.

the data is fed to the priority encoder. The priority encoder will look to the MSB of this group of bits and encode it in binary code. The D/A converter changes this MSB code to analogue voltage. The rest of the circuit is similar to Figure 12 mentioned before in the PTCM system, Chapter 5, section 5.1. The capacitor should be large enough to do the required averaging over the determined period. The current I was varied until the voltage across the capacitor C started to vary above and below zero equally, and not building up to any side. This setting of current I represents the average bit rate. Bearing in mind that the value of current I stays constant, the average bit rate can be found by applying high logic level to different inputs of the priority encoder, until the voltage across capacitor C stays constant. The results obtained by this way of measurement were not far from those calculated before from the curve of speech signal level distribution. The results were:

$$\text{Average bit rate} \begin{cases} \text{Male} = 28 \text{ kb/sec.} \\ \\ \text{Female} = 36 \text{ kb/sec.} \end{cases}$$

The overall average bit rate = 32 kb/sec.

The operation of the PCM encoder DF341 can be seen in the appendix A in the end of this thesis.

The above result is encouraging. A saving of 50% or thereabouts should be possible. However, we must examine the store size required and ensure that the signal delay

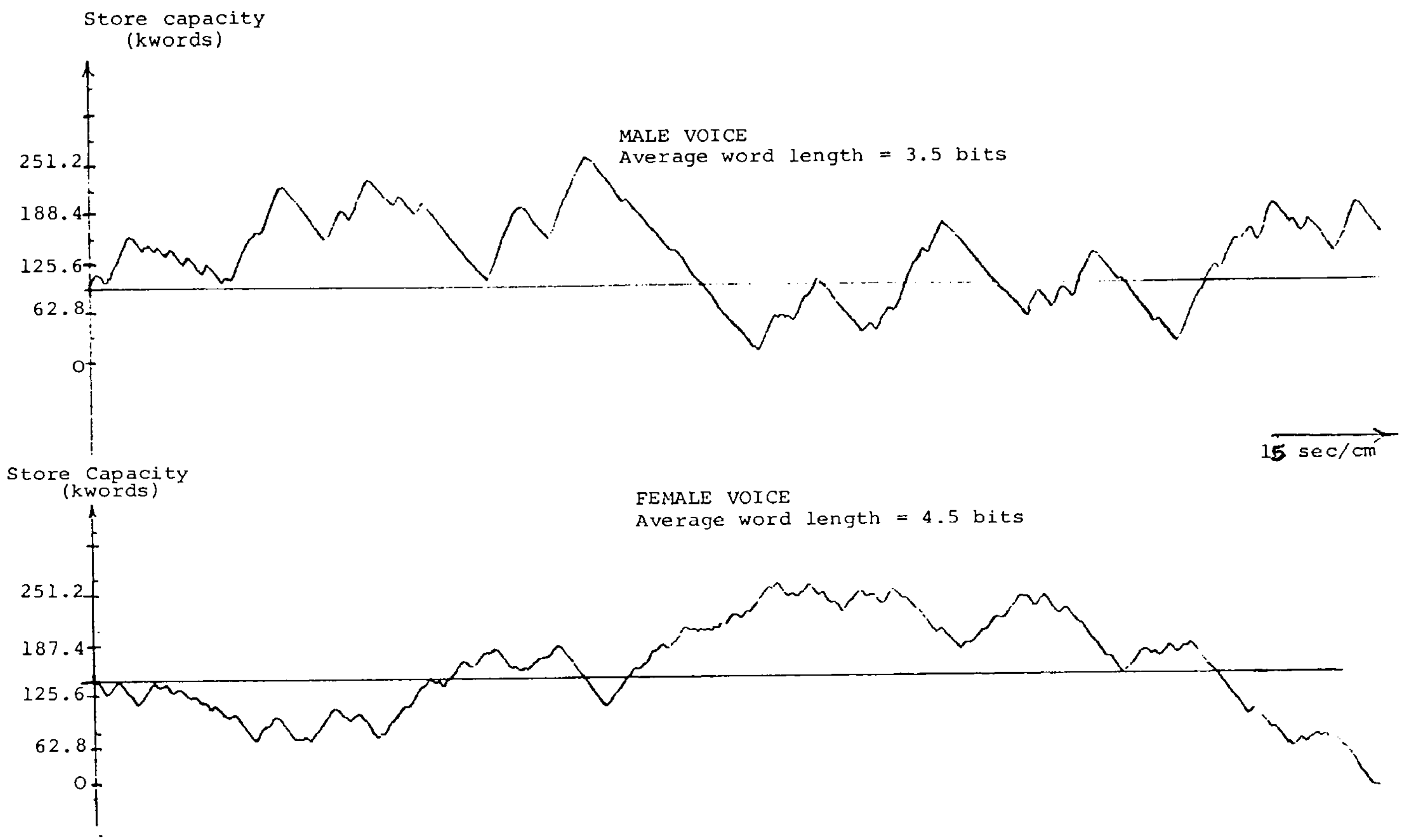


FIG. 19 ARPCM SYSTEM
STORE CAPACITY MEASUREMENTS "TELEPHONE CALLS"

is acceptable.

8.3 STORE SIZE MEASUREMENT IN ARPCM SYSTEM

The importance of the store in the ARPCM system is similar to that in the PTCM system. It is necessary for the purpose of averaging the bit rate. Its size must be enough to do this action of averaging without overflow.

In the case of ARPCM system, the samples are coming continuously at a fixed rate, 8k samples per second. Some samples may carry no real information. Then the samples will be taken out of the store at a constant bit rate but variable word rate, as mentioned before. The word rate is inversely proportional to the number of the used bits in each word. To measure the store capacity it is important to take this fact into account. A measuring system was developed to perform the required measurement. Figure 18 shows this system. The system of Figure 18 is to be added to that of Figure 17, instead of the D/A converter and the elements after it. The programmable counter is used to divide the frequency of the clock f by the number of bits n. The output of the counter will be proportional to the word rate required. The rest of the circuit is similar to the circuit used to measure the store size in the PTCM system.

The voltage across the capacitor C was plotted by using a slow X-Y plotter. The chart obtained is shown in Figure 19. The store capacity required was found to be 251.2 kword, of 8 bit per word. This value of store capacity is very large and the ARPCM system shows no improvement in signal delay over the PTCM system; the application of ARPCM will therefore be mainly in simplex operation, as described in the following section.

8.4 APPLICATIONS OF THE ARPCM SYSTEM

In the case of simplex transmission the speech activity is double that of duplex transmission. It was found that the PTCM system is not useful in simplex transmission applications because the average bit rate becomes higher than that of the conventional PCM system.

Investigation of the possibility of using the ARPCM system in simplex transmission required the determination of the average bit rate and the suitable store size for this purpose. The same measuring system of Figure 17 and Figure 18 was used to determine these factor. In order to obtain a standardized continuous speech signal, both channels of the tape recorder were applied simultaneously. The two outputs of the channels were added together. Then the telephone call and the reply from the other side form a continuous speech signal. The measurements were carried out for a period of 5 minutes to find the average bit rate. This test was repeated several times, the results were averaged. It was found that:

The average bit rate of continuous speech = 40 kbit/sec.
The average word length of continuous speech = 5 bit/word

Although the average word rate stays at 8 kw/sec, the

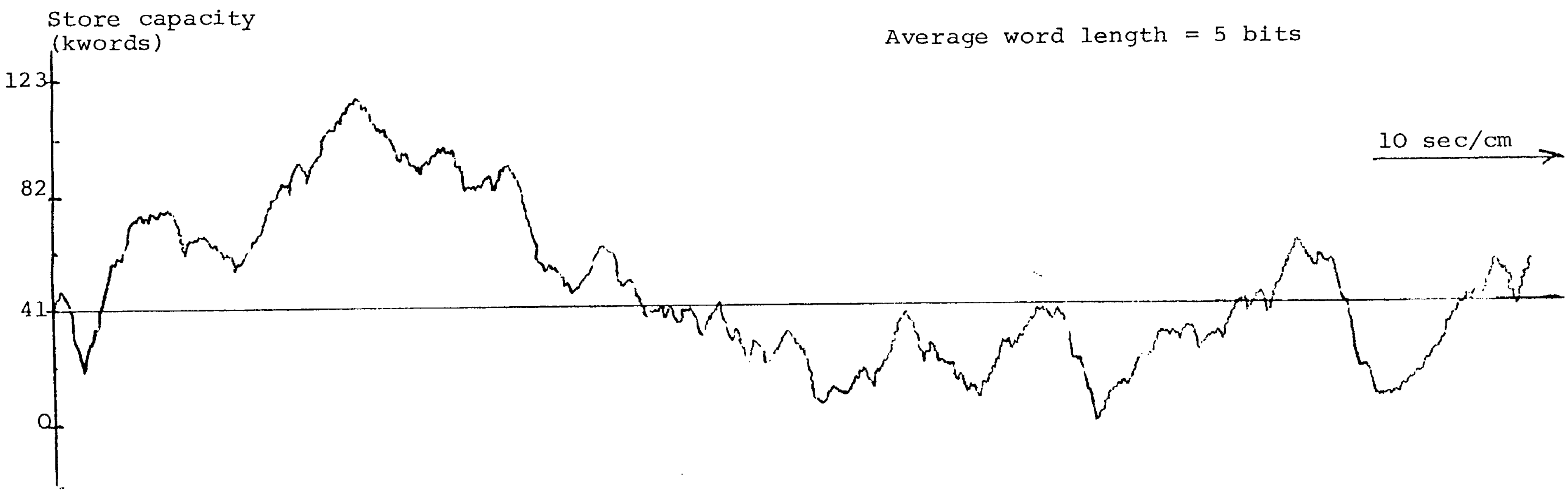


FIG. 20 ARPCM SYSTEM - STORE CAPACITY MEASUREMENT "CONTINUOUS SPEECH"

average word length is less than 8 bit/word and hence there will be saving in the bit rate. This result is different from that of the PTCM system for continuous speech showing that there is a possibility of using the ARPCM system in simplex transmission from this point of view.

The other element which is essential to know is the store size. This was determined in the same way as that used for the PTCM system. The plot obtained is shown in Figure 20. The result obtained by measurement is:

The store size required for simplex transmission in the ARPCM system = 123 kword.

The width of the store = 8bit/word.

This length of the store in the ARPCM system is more than that of the PTCM system, but the width is less. Since the delay in simplex transmission is not a problem, then the ARPCM system offers a considerable saving in bit rate over the conventional PCM system. The only difficulty in building this system in the university is the high cost of the store.

The reason why the store size in continuous speech is less than that for telephone call speech can be explained now. In telephony the silence gap lasts for an appreciable length of time. During this silence gap the transmission is continuous at its constant rate. In order to get proper averaging over a long period of time the transmission rate was adjusted to prevent underflow in the store. The transmission rate will then be low but the store capacity required will be high. In the case of continuous speech it is necessary to increase the transmission rate. The silence gap will be short so the possibilities of underflow are less. This results in a smaller store but with a higher bit rate.

CHAPTER 9

DESIGN AND OPERATION OF THE ARPCM SYSTEM

9.1 OPERATION OF THE ARPCM SYSTEM

Generally the principles of operation of the ARPCM system are similar to those of the PTCM system. The main difference between them is the encoding method used in each system. In the PTCM system the time interval between successive crossings is coded while in the ARPCM normal PCM coding is used. The block diagram shown in Figure 21 is used to illustrate the principle of operation of the ARPCM system.

The standardized speech signal is fed to the encoder. The output of the encoder is a conventional PCM code in serial form. It is required to have this coded data changed from serial form into parallel form. Then the parallel PCM words are fed into the store. The data flows into the store at a fixed rate of 8,000 words per second. The words are taken out of the store and transferred from the parallel form back to its original serial form again. Then these words are transmitted after the removal of all the unused bits in these coded words, i.e. all zeros after the MSB are ignored. After the used bits of the word are all transmitted a stop signal is introduced to the line. This is in the form of space of length 1.5 bits, which will be followed by a mark representing the MSB of the next word.

At the receiver side, the incoming data passes through

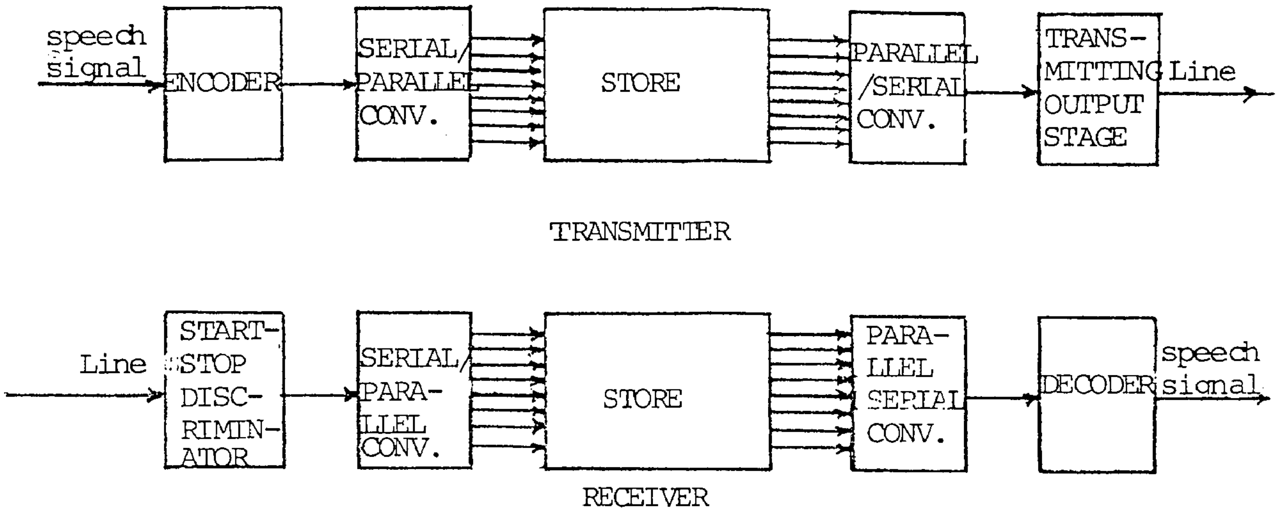


FIG. 21 PRINCIPAL OF OPERATION OF THE ARPCM SYSTEM

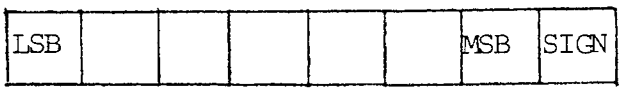


FIG. 23a ONE WORD DATA IN THE SHIFT REGISTER SR1

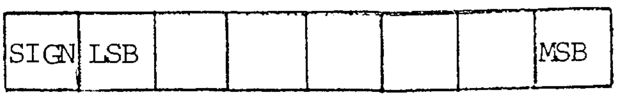


FIG. 23b THE REQUIRED ARRANGEMENT OF THE WORD BITS

a circuit which can discriminate the stop signal. Then this stop signal is removed from the coded word and the rest of it is converted from the serial form to a parallel form. These parallel formed words are stored in the store in the same way at the transmitter. From the store the words are taken at the normal fixed rate of the PCM system. Then these words are transferred back to serial form. The serial form words are fed to the decoder which brings them back to their original analogue form. The reduction in bit rate is a result of the removal of the unused bits of the words. The words transmitted through the line will be variable in length. In the conventional PCM system the word length, the bit rate and the word rate are all fixed. In the present system the bit rate is fixed but the length and the word rate are variable.

9.2 CIRCUIT DESIGN AND OPERATION FIGURE 22

9.2.1 The Transmitter Fig. 22a

9.2.1.1 The Encoder

The encoder DF341 receives the standardized speech signal. The maximum amplitude of the signal fed to the analogue input of the encoder (Pin4) is limited by the reference voltages $+V_r$ and $-V_r$. According to the characteristics of the encoder these reference voltages should be in the range of $\pm$ 2.5V to $\pm$ 3.5V. It is wise to use $\pm$ 3.0V for reference voltages, so that the dynamic range of the encoder will be restored and at the same time

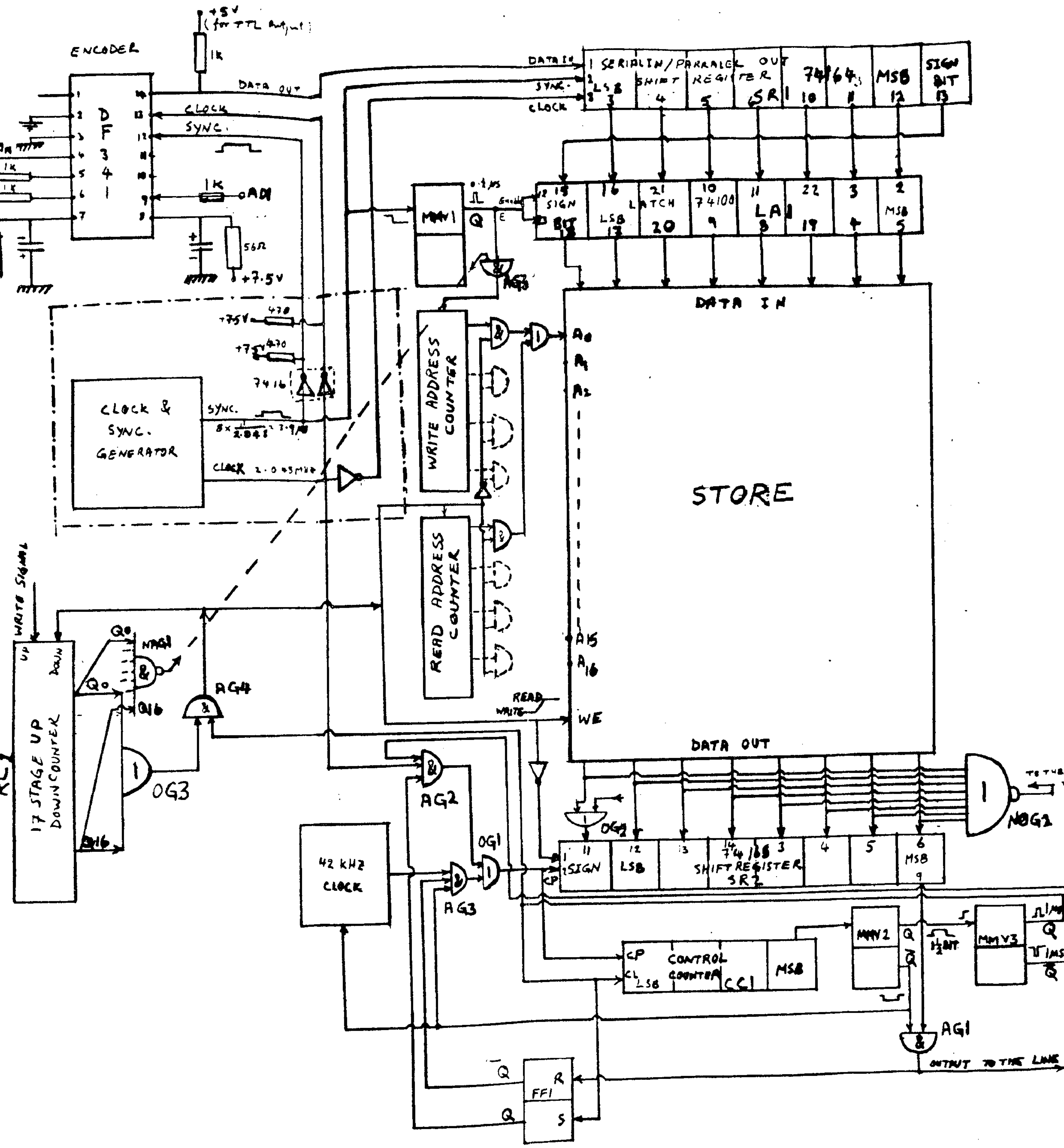


FIG.22a ARPCM SYSTEM TRANSMITTER

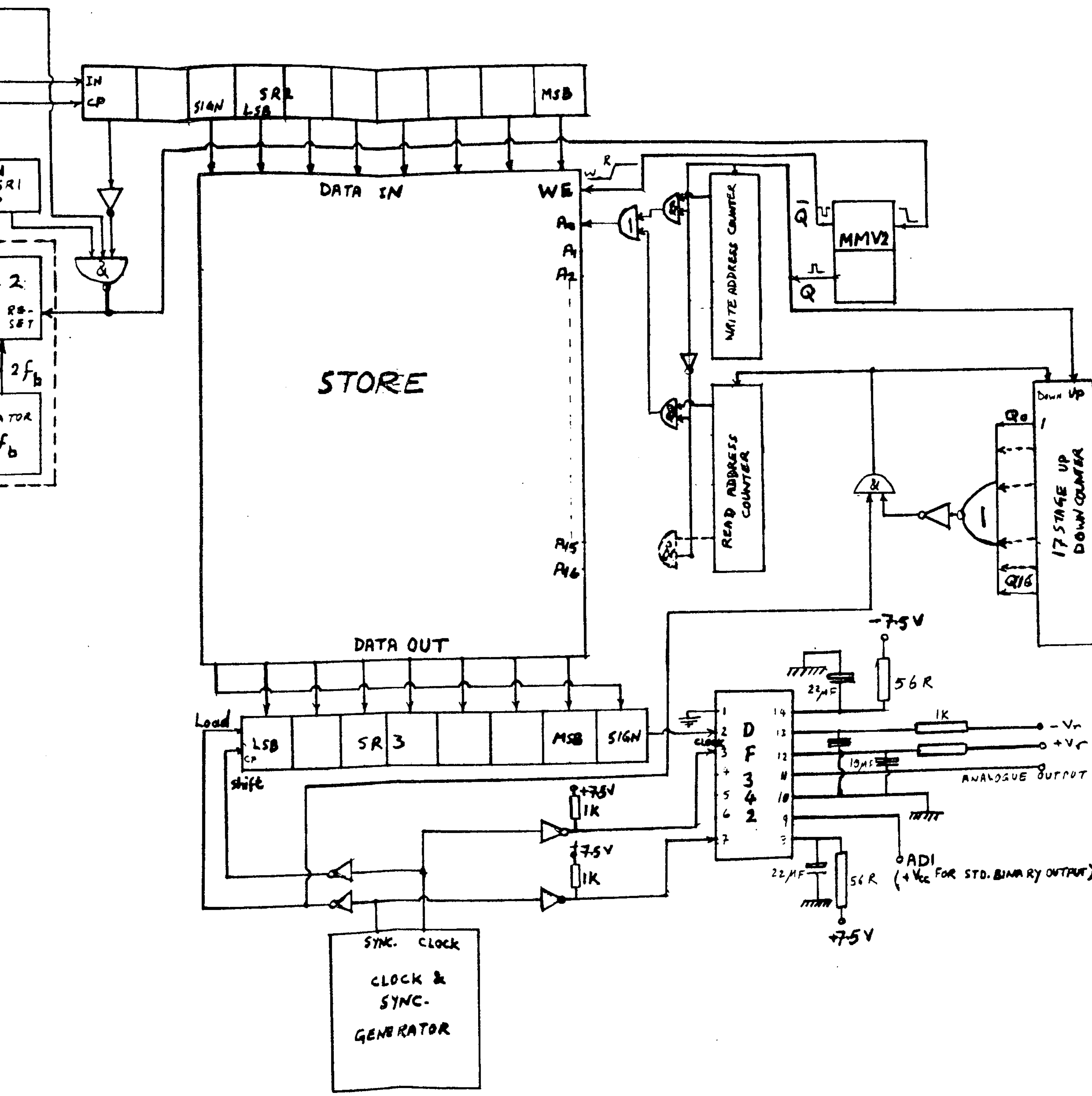


FIG. 22b ARPCM SYSTEM RECEIVER

excessive harmonic distortion will be avoided. More details can be found in Appendix A and the manufacturer's specifications of these coder-decoder devices.

The alternate digit inversion ADI pin number 9 of the encoder should be connected to the positive supply via 1.0k resistance. This will make the output date of the encoder in standard binary form.

In order to make the output data level compatible with the TTL inputs, the data output pin number 14 is connected to +5.0V via 1.0K resistor.

The clock and the sync signals can be generated as advised by the manufacturer. Circuit diagram and wave-forms can be found in Appendix A. The crystal oscillator can be used also and the sync signal will be generated in the same way.

9.2.1.2 Serial to Parallel Converter

The shift register SR1 (74164) is used to do the serial to parallel conversion. The data are fed into the serial input of this shift register SR1 pin number 1. The level of the data should be compatible to the TTL input, i.e. 5 volts. The sync signal is fed to the other serial input to stop any random noise from entering the shift register while the sync signal is low. The same clock of the encoder is used as a shifting clock.

The data in the shift register SR1 is arranged as shown in Figure 23a. The sign bit probability is shared between zero and one equally if the speech signal is considered to be symmetrical. This will make it impossible to remove the unused bits in the word for half the range when the sign bit is one. To overcome this difficulty the sign bit should be transferred from its position in MSB side to the other side of LSB. Figure 23b shows the required arrangement of the bits. This required arrangement of locations of the word bits is done when the data are loaded into the latch LA1 (74100), Figure 22a. The sign bit is transposed from its position in the shift register SR1 to the position of the LSB and the whole bits of the word are shifted to the right by one location, as in Figure 23b.

The data are loaded from the shift register SR1 into the latch LA1 after the completion of one word data shifting in SR1 by a short period. The shifting in SR1 stops when the sync signal changes from high to low. The negative going edge of the sync signal is used to trigger the monostable multivibrator MMV1. This monostable generates a positive pulse of suitable duration. The duration of this pulse is not critical. It can be as short as 20 ns or as long as 1 μs. This pulse is used as a load command to the latch LA1. It has also got another function which is the incrementation of the write address counter by one. The write address counter is incremented by one with the lagging edge of this pulse. Hence the data during the positive period of this pulse will get enough time to move from the shift register TR1 to the latch LA1. There will be no problem of race hazard in this case.

9.2.1.3 The Store

It was found experimentally that the required store size is 123 kword, each word of 8 bit length. If a Random Excess Memory (RAM) is to be used, the nearest size to this figure can be calculated in this way:

Let S represent the store size

A represent the number of address line

$$S = 2^A \tag{8.1}$$

$$\therefore \quad A = \frac{\ell n S}{\ell n 2} \tag{8.2}$$

$$A = \frac{\ell n 123 x 10^3}{\ell n 2}$$

$$= 16.9$$

Henc:

The number of address lines must be 17.0 lines.

That means the RAM size $= 2^{17}$

$= 131072$ words by 8 bits.

The address lines are arranged as shown in the diagram, Figure 22. The read and write addresses are similar. They are 17 stage counters. They are arranged such that only one address is available to the address input of the store at a time. The write enable WE of the store is supposed to be in read mode when it is high in the present case. Write mode is when the WE is low. This can be reversed depending on the RAM device used. When the write enable input WE is in its read mode the read address counter outputs are connected to the address inputs of the store while the write address counter outputs are inhibited from reaching the address inputs of the store and vice-versa. For simplicity it was supposed that the data inputs and data outputs are separate in this RAM. If this were not the case, then it will be necessary to do similar arrangement to the address lines for the data lines.

9.2.1.4 Parallel to Serial Conversion and Transmitting Output Stage

The shift register SR2 (74165) is used as a parallel to serial converter. When a read signal appears on the WE input of the store, it is inverted and used as a load signal to let the data transferred from the store to the shift register SR2. The inverted read signal is fed to the shift/load input pin number 1 of shift register SR2. When the read signal is high the input shift/load of SR2 will go low and the data from the output bus of the store will be loaded into the shift register SR2. When the read signal goes low the shift/load input of SR2 will be high and the shift register SR2 will be in its shift mode. When a word is loaded from the store into the shift register SR2, no shifting of data will take place in this shift register. The shifting clock is inhibited momentarily. After the loading takes place a fast clock, the sampler clock 2.048 MHz can be used to shift the data from LSB side, extreme left of the shift register SR2, to the right which is the MSB side. When one appears in the extreme right position together with the high level output of the monostable multivibrator MMV2 output, to ensure that it is not the write modethen, the one on the MSB position and the high level output of MMV2 opens the AND gate AG1. The output of the gate AG1 is fed to the line to indicate the start of transmission of one word. The output of AG1 is also used to reset the flip flop FF1. The Q output of FF1 will be low after reset and $\overline{Q}$ will be high. Any change on the reset input of FF1, high or low will not affect its state until it is set again. Q output of FF1 is used to inhibit the fast clock by means of the AND gate AG2 so that the clock will not get to the clock input of the shift register SR2. The slow clock 42 kHz is enabled by the $\overline{Q}$ output of the flip flop FF1 to get to the clock input in the shift register SR2. At this moment the actual transmission of data starts. Any clock pulse gets to the input of the shift register SR2, fast or slow, is counted by the control counter CC1. The data will continue flowing out at the same rate of the slow clock until the control counter CC1 counts eight. When CC1 counts eight that means a complete word has been transmitted. It is now necessary to generate the stop pulse before starting the transmission of a new word. When the fourth bit of the control counter CC1 changes state from zero to one, that means the control counter has counted eight counts, and this change of state is used to trigger the monostable multivibrator MMV2. The monostable multivibrator MMV2 generates a pulse of duration equal to 1.5 bit long. If the transmitting clock frequency is 42 kHz, then the duration of one bit is 23.8 μs, and the stop pulse duration should be

$$23.8 \times 1.5 = 35.4 \ \mu s.$$

So the stop signal is a space of 35.4 μs followed always by a mark which represents the MSB of the next word.

The Q output of the monostable multivibrator MMV2 is used to trigger another monostable MMV3, which generates a short duration pulse, for instance 1.0 μs or even less. The positive outputs pulse of MMV3 is used to do the following:

(a) Increments the read address counter by one.

(b) Puts the write enable input of the store WE in read mode.

(c) Sets the flip flop FF1.

(d) Clears the control counter CC1.

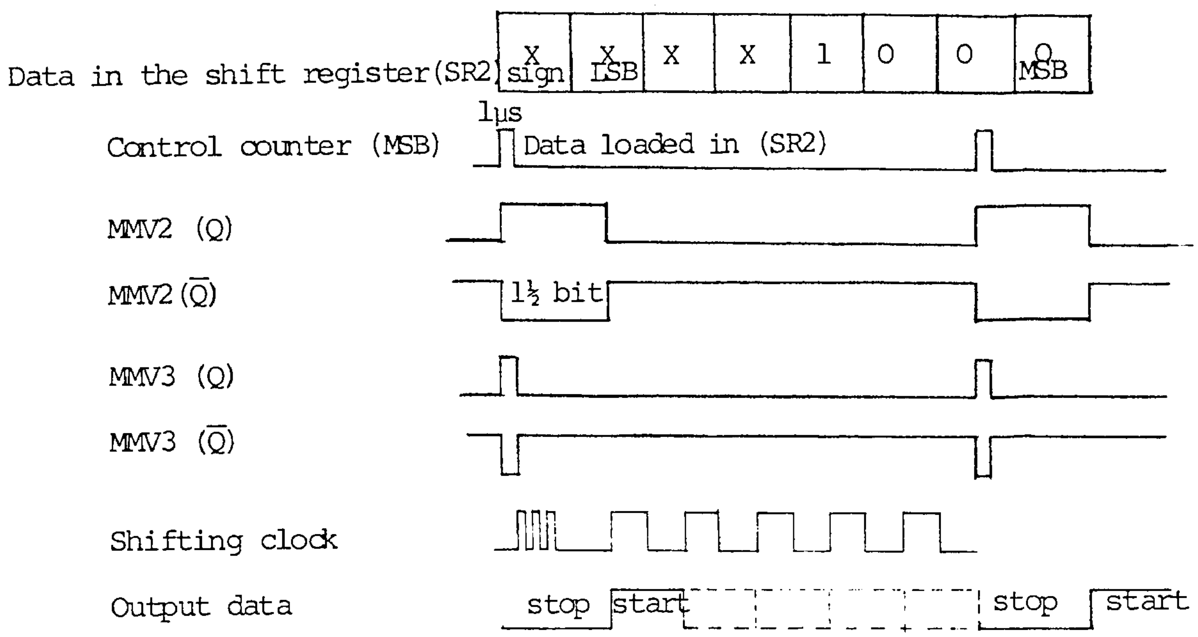
The negative going pulse of MMV3 outputs is used to inhibit the fast shifting clock from getting to the shift register SR2 clock input. Namely, no shifting is going to take place during the loading time. The AND gate AG2 is used to stop the fast shifting clock mentioned above. The write command signal is inverted and fed to the shift/load input of the shift register SR2. This means that during the read mode in the store, the shift register SR2 will be in load mode.

It could happen that the fast shifting of data in the shift register has finished while the stop pulse of the previous word is still not finished. In this case the slow clock for instant the 42 kHz is not enabled to the clock input of the shift register SR2 until the stop pulse is completed. This is done by means of the AND gate AG3. Having the stop pulse completed the AND gate AG3 is opened and the 42 kHz clock resets to ensure that it is starting from the beginning of the cycle. After that the transmission is carried out normally until the word is finished completely by counting eight in the control counter. A new stop pulse is generated and the events will be repeated again and so on. Figure 24a shows the timing of transmission events.

9.2.1.5 All-zero Code Condition

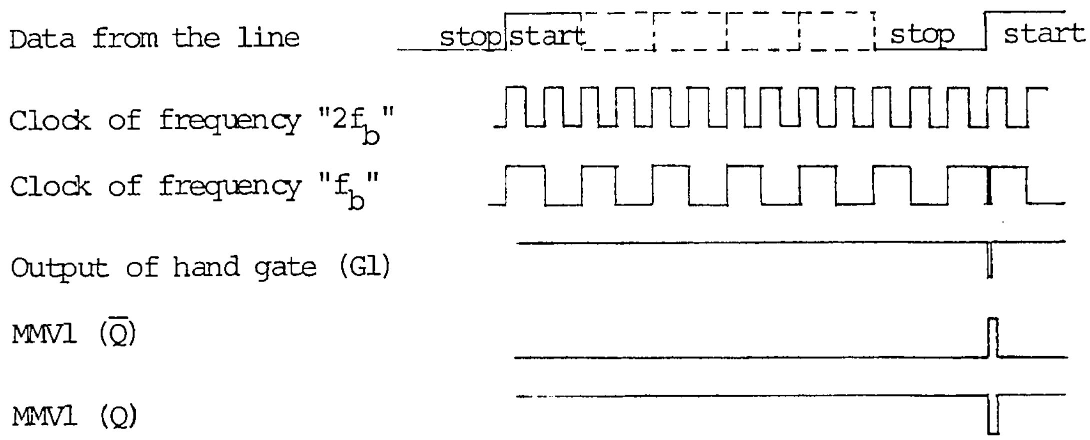
In the PCM system the arrangement of alternate digit inversion is used to prevent the all zero code condition. It is possible in the PCM system to use this method because the words transmitted are of constant length. In the case of the ARPCM system the situation is different because the alternate digit inversion will cancel the benefit gained by the removal of the unused bits. The all-zero code condition is a problem, because it represents an actual sample, while it is not transmitted. This will alter the signal when it is reconstituted at the receiver because of the loss of the samples of all zero codes. Hence it is necessary to find a way to indicate the case of the all-zero code.

From that mentioned previously in section 4, it can be seen that the indication to start transmission is the appearance of a digit one in the last stage of the shift register SR2. In the case of all-sero code the appearance of the digit one in the last stage of SR2 will not happen during the fast shifting of the word. That means no command to begin the transmission is going to be generated. After eight counts by the control counter CC1 a new stop pulse will be generated, while the stop pulse of the previous word is not yet finished. The



(a)

TRANSMITTER EVENTS TIMING



(b)

RECEIVER EVENTS TIMING

FIG. 24

f_b = bit rate

result of this situation is no transmission is going to take place. The last word is a true sample separated by 125 μs from the next sample. When this word is lost, all the positions of the samples will be affected. If this is repeated many times, the result will be a great distortion in the signal. In order to avoid the loss of these samples, the NOR gate (NOG2) senses the data lines. When the code coming out of the store is all-zero code, the NOR gate (NOG2) generates a logic one. In all other cases the output of NOG2 is logic zero. This one is fed to the shift register SR2 by means of the OR gate OG2. The position of it will be in the sign bit position. This will change the code of the all-zero code to one-bit code. Hence the transmission of this case will be in the form of a stop signal followed by a start signal and followed by another stop signal which is necessarily followed by a start signal. The result later at the receiver will be similar to the all-zero case because even the sign bit had been changed but the amplitude is still all-zero code.

9.2.1.6 Underflow Condition

If the word transmission rate from the store goes faster than the word flow rate into the store, the store will be emptied. If the store becomes completely empty the case of underflow happens.

If underflow condition occurs, the transmission should be stopped immediately. If the transmission does not stop the read address counter will pass the write address counter and a complete store content must be transmitted while it is not actually in existence. This case will introduce long silence gaps which are not in the original speech, and will also affect the store of the receiver by storing these data unnecessarily. In order to prevent this problem occurring, it is necessary to stop the transmission when underflow condition happens. The reversible counter RC1, sometimes called up-down counter, is provided to control the transmission when underflow condition occurs. The number of stages of the reversible counter RC1, depends on the number of address lines of the store. In the present case it is a 17-stage counter. When the reversible counter RC1 outputs all zeros, it means that the store becomes empty. The OR gate OG3 senses the outputs of the control counter RC1. When the outputs of RC1 become all zero, the output of OG3 will become zero. The output of OG3 is used to stope the read pulse from getting to the read address counter and the control counter. Hence the transmission will be stopped until new data comes into the store.

9.2.1.7 Overflow Condition

When the store becomes full, the reversible counter RC1 will be in its maximum count, i.e. all the outputs will be ones. This condition means that no more space is available in the store for new data to enter. Once this condition happens, there will necessarily be a loss of data. It is necessary to keep the control counter and the write address counter stationary while the store is full. If these two counters left to count while the store is full, the address and the data in the store will not be correct. To keep these two counters idel, the NAND gate NAG1 is used. The inputs of NAG1 are connected to the outputs of the control counter RC1. When the outputs of RC1 become all ones, the output of NAG1 will be zero. This output will be used to stop the write signal by means of the AND gate AG3 from getting to the inputs of the write address counter and of the control counter.

9.2.2 The Receiver (Figure 22b)

The function of the receiver is to admit the coded words and bring them back to their original analogue form. The receiver consists of the following parts:

1. Stop-Start pulse discriminator and serial to parallel converter.
2. Store
3. Parallel to serial converter
4. Decoder.

9.2.2.1 Stop-Start Discriminator and Serial to Parallel Converter

It was mentioned before in the transmitter discussion that the way to distinguish between two successive words is a space of 1.5 bit length, was called the stop pulse to indicate the end of one word, is always followed by a mark of one bit length, called the start pulse which indicates the MSB of the following word. The discriminating circuit consists of two shift registers SR1 and SR2. SR1 is a single stage shift register, its clock frequency is equal to double the bit rate at which data are coming, i.e. $2f_b$ where f_b is the bit rate. SR2 is a ten stage shift register. Its clock frequency is the same as the bit rate, f_b. SR2 is performing two functions, the first function is to be a part of the stop-start discriminator and the second function is to be a serial to parallel converter. The frequency f_b is derived by dividing the oscillator frequency $2f_b$ by two. The oscillator which generates the frequency $2f_b$ is triggered by the incoming data, so that when a change from low to high occurs in the data line, it will cause the oscillator to start at high level. This will keep the oscillator synchronized with the frequency at which the data are coming.

The discriminator of the stop pulse is performed in the following way. When the data are flowing normally without stop pulse, the two shift registers SR1 and SR2 will have similar data in their first stages. When the stop signal finishes and it has lasted for 1.5 of the normal time of one bit, then the start pulse starts. The shift register SR1 will be affected by this change because its clock frequency is $2f_b$, while the shift register SR2 will not be affected by this change, because its clock frequency is f_b. In this case the first stage of SR1 will go high while the first stage of SR2 will stay low. This case happens only after the finishing of a stop signal and the beginning of a start signal. Hence by inverting the output of the first stage of the shift register SR2 and feeding it with the output of the shift register SR1 to the inputs of the NAND gate NAG1, the required signal to indicate an end of one word and the start of another word will be generated.

It is necessary to eliminate any possibility of race hazard due to the different delays in the elements. The monostable multivibrator MMV1 is used for this purpose. MMV1 when triggered by rising edge generates a negative going pulse. The rising edge, when the stop pulse is finished and the start pulse started, will trigger the monostable multivibrator MMV1. The output of MMV1 is a negative pulse of duration longer than the time at which there is a possibility of race hazard, and at the same time this negative pulse should not be long to such a degree that it could effect the operation of stop-start discrimination. A pulse of duration around 50 ns is suitable. This pulse is fed to one of the inputs of the NAND gate NAG1. The output of NAG1 is always high and it will go low only when a change from stop to start happens and after about 50 ns of that change.

The output of the NAND gate NAG1 is used to reset the divide by 2 bistable in order to keep the synchronism of the frequency f_b and $2f_b$ with the coming data. The output of NAG1 is also used as a write signal. The monostable multivibrator MMV2 is triggered by this output and generates the required write signal. The duration of MMV2 output pulse depends on the access time of the store elements. It is not critical, it can be made as long as 5 μs.

Figure 24b shows the timing of the receiver events.

9.2.2.2 <u>The Store</u>

The store in the receiver is similar to the one used in the transmitter. It is a random access memory, RAM. Its capacity is also equal to that of the transmitter, i.e. 131072 words of 8 bit per word. The 17 line address input are multiplexed between the read address counter and the write address counter outputs, in the same way as used in the transmitter. If the data in and the data out are connected to a single bus it will be necessary to do the same multiplexing used in the address lines for the data lines.

Underflow and overflow conditions are also undesirable in the same way as disucced in the transmitter. Hence the same way explained before in the transmitter can be used to prevent the effect of these two conditions. The underflow effect eliminator is shown in Fig. 22b. The overflow effect eliminator is not shown, but it is similar to that in the transmitter, Figure 22a.

9.2.2.3 <u>Parallel to Serial Converter</u>

The parallel to serial converter is an eight stage shift register, SR3. This shift register acts as a latch and parallel to serial converter. The clock for SR3 is the same clock of the decoder. The decoder clock is similar to that of the encoder. The load input of the shift register SR3 is connected to the same signal. The sync signal also used to increment the read address counter by one count each time one word is started. During the sync signal being low the loading of one word from the store into the shift register SR3 will take place. When the sync signal goes high, the shift mode will start. Before feeding the data into the decoder the sign bit should be brough back to its original place as in Figure 23a. This is done by proper wiring of the data bus from the store to the shift register SR3, see Figure 22b.

9.2.2.4 The Decoder

The decoder is DF342 siliconix, which is the complement of the encoder DF41, if the A-law companing is used. If μ-law is to be used the encoder DF331 can be used and the decoder in this case is DF332.

With the decoder DF342 all the signal levels, clock, sync alternate digit inversion connection AD1 and reference voltages must be similar to those used with the encoder DF341. The output of the decoder is the original speech signal fed to the encoder.

CHAPTER 10

DISCUSSIONS AND CONCLUSIONS

10.1 DISCUSSIONS AND CONCLUSIONS OF THE ARPCM SYSTEM

It has been found that a considerable saving in bit rate can be obtained by the application of the ARPCM techniques. In duplex transmission such as telephony up to 50% economy in bit rate can be achieved. However, because of the excessive signal delay the system is of limited application. For those application using continuous speech, the economy is substantial, and may well justify the use of this system.

10.2 GENERAL CONCLUSIONS FOR BOTH THE PTCM AND THE ARPCM SYSTEMS

Two systems for the reduction of the bandwidth required to transmit a digital speech signal have been investigated. The first system, which has been called PULSE TIME CODE MODULATION (PTCM) is a development of the work done by previous workers at this university(1),(2). The principles of the second system, called AVERAGED RATE PULSE CODE MODULATION (ARPCM) were discovered while research on the PTCM was in progress.

The advantages of the PTCM system, when compared with conventional PCM, proved to be less than had been expected. For duplex transmission a substantial economy in bandwidth is achievable, but the delay in transmission necessarily caused by the buffer store, makes the use of this system inconvenient. For simplex transmission, the greater occupancy of the speech channel so increases the mean sampling rate that the PTCM system shows little superiority in bandwidth required over conventional PCM system.

The ARPCM system is much more useful however, for simplex operation when transmission delay is unimportant, the bandwidth reduction can be as high as 30%, and for this type of work the use of this system should be well justified. However, for duplex operation, the system suffers the same disadvantages as the PTCM system, namely, the time delay in transmission, and ARPCM system offers little advantage over PCM for this type of operation.

The ARPCM system is believed to be of great potential for simplex transmission. All the essential investigation had been completed during this research period. Also a complete design of the system has been included in this thesis. The reason why this system was not built in the university is the financial problem. The store cost is beyond the financial possibility of the university at the present time. The author feels very sorry that he was unable to build this project. If financial resources become available or the cost of the store is reduced by the advanced technology, there will be no problem in building this project.

APPENDICES

APPENDIX A

DESCRIPTION OF ENCODER-DECODER

SILICONIX DF341-DF342

From the manufacturer's leaflets

1. INTRODUCTION

The Encoder-Decoder DF341-DF342 for A-law and its equivalent DF331-DF332 for μ-Law are siliconix designed chips. These per channel CODEC using a pair of CMOS chips operating on a capacitive charge redistribution technique. Designed specifically for D4 channel banks, but ideal also for various PCM telephone systems, the compressing A-to-D converter (coder) and expanding D-to-A converter (decoder) produce the standard digital approximation to both the North American μ 255 characteristics and the European A-law characteristic. Coders and decoders for both μ-law and A-law specifications are supplied under type numbers DF331/332 and DF341/342 respectively.

2. THE A -TO-D and D-TO-A CONVERSION PROCESS

The coding and decoding process which conforms to the μ-law and A-law (figure 25) characteristics is a combination of a compressing A-to-D coder and an expanding D-to-A. The 8 bit output from the A-to-D coder, by virtue of the conversion technique used is a compressed digital representation of the analogue. The compression technique results in a nearly constant signal-to-noise ratio over a wide range of analog voice amplitudes. This, in telephone applications makes possible a high degree of intelligibility when someone is speaking softly or loudly into the telephone. A straightforward linear A-to-D conversion, without compression, would of course result in a loss of voice quality at the lower anlog signal levels.

3. PRINCIPLE OF OPERATION

The analog is converted into an 8 bit digital code. The particular digital code for example is 11001101. The first digit indicates the polarity of the input analog. The next 3 digits known as segmentation digits, indicate which segment the analog falls into. The last of the 4 digits known as 'bit inside segment' digits, indicate which of the 16 levels inside a given segment correspond to the exact analog value.

The A/D and D/A codec requires the application of an externally applied clock and sync. The nominal operating clock frequency required for telephone link is 1.544 MHz and 2.048 MHz for the Mu-law and A-law respectively. The sync in both cases would be nominally run at 8 kHz (slightly more than twice the maximum analog frequency for voice signals).

Under practical conditions, the coders (sending end) and the decoders (receiving end) will be an appreciable distance apart, e.g. the distance between two public telephone exchanges. The coder and decoder will use separate clock and sync generators and consequently synchronisation between the two could seemingly be difficult to implement. The normal method therefore is to use the PCM digits, arriving at the decoder input, to generate the clock and sync pulses at the decoder end. Obviously, a source of problems could exist during light traffic periods, when a number of sequential coders could be generating no output at all. In this case, the clock at the decoder end could stop operating. It is necessary to implement a system of PCM coding other than the standard binary output, so that when the output is zero, we get at least a logic '1' appearing in a train of 8 digits appearing at the coder output. It is therefore necessary that the PCM code has logic transitions (0 to 1 or 1 to 0) at reasonably regular intervals. These transitions in the PCM code are used to regenerate and regulate a clock which is kept in step with the incoming pulse train.

In Mu-law device, an all zero code suppression is implemented and in the A-law a somewhat different facility known as Alternative Digit Inversion is implemented (Figure 26). Alternatively, a standard binary output can be obtained on the A-law device by taking Pin 9 to +V.

4. DESIGN FEATURES

The separation of the encoding and decoding functions onto two chips has many advantages. Such a configuration is low cost, guarantees high isolation between the transmitting and receiving operations, and allows the codec user considerable freedom of design. Each chip of such a pair is smaller than a one-chip codec, so that reliability and yields are higher and cost lower. Each fits in a 14 pin package, so that the board layout is compact. Being isolated from one another, coding and decoding can be carried out asynchronously, as needed, and crosstalk between the two directions of voice travel is completely eliminated. Moreover, using minimum support circuitry, the designer can layout his system in a variety of ways - on single channel codec cards, for example, or on multiple channel receive-only and transmit cards that minimise the amount of digital busing that will be needed on mother boards.

The $\pm$ 7.5 ($\pm$ 10%) power supplied required for each chip keep power dissipation low, to an average of only 40 to 45 milliwatts per device. The supply voltages are easily obtained from a discrete regular supply circuit or from standard three-terminal regulator chips.

These reasonable power dissipation power levels do not impair the codec's analog system performance. Both the Bell System's D-3 specification and the CCITT recommendations are readily met. What is more, any encoder is guaranteed to function properly with any decoder since the analog performance is specified on a per-part basis.

The reference voltages are important to the performance of a codec's analog system. They serve both the A-D and D-A conversion and are crucial to achieving proper gain levels throughout the system. For the siliconix design, external voltage references were chosen for their cost-effectiveness, accuracy over the long term, and independance of the codec chips. They minimise the effect of codec manufacturing tolerances on system gain and temperature stability since the responsibility for gain and long-term drift is shifted to an external and therefore easily controlled, maintained, and designed source.

The approach yields an absolute gain accuracy of typically $\pm$ 0.2 dB from device to device, making the codecs interchangeable without additional channel gain adjustment. The $\pm$ 3.0V external voltage reference also sets the maximum allowable input signal level. Given a system dynamic range of 72 dB, this also means the minimum signal-to-noise ratio will be relatively large. Figure 27A shows the circuit connection of the codec chips. Figure 27B shows the clock and sync generator and the respective waveforms.

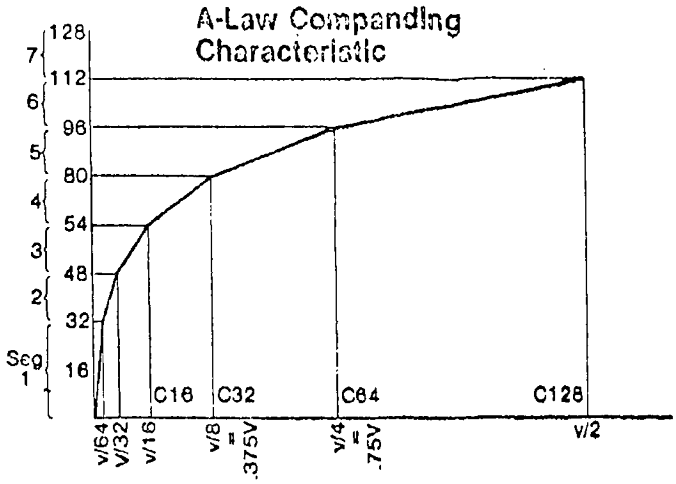


Figure 25

All Zeros Code Suppression Mu-Law

0 0 0 0 0 1 1 1
0 0 0 0 0 1 1 0
0 0 0 0 0 1 0 1
0 0 0 0 0 1 0 0
0 0 0 0 0 0 1 1
0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 1
0 0 0 0 0 0 1 0

Alternate Digit Inversion A-Law

Coder Binary Output at Pin 14 when Pin 9 is programmed to logic 1	Coder Output at Pin 14 when Pin 9 is programmed to logic 0
0 0 0 0 0 1 1 1	0 0 1 0 1 1 0 1
0 0 0 0 0 1 1 0	0 0 1 0 1 1 0 0
0 0 0 0 0 1 0 1	0 0 1 0 1 1 1 1
0 0 0 0 0 1 0 0	0 0 1 0 1 1 1 0
0 0 0 0 0 0 1 1	0 0 1 0 1 0 0 1
0 0 0 0 0 0 1 0	0 0 1 0 1 0 0 0
0 0 0 0 0 0 0 1	0 0 1 0 1 0 1 1
0 0 0 0 0 0 0 0	0 0 1 0 1 0 1 0

All Zeros Code Suppression and Alternate Digit Inversion

Figure 26

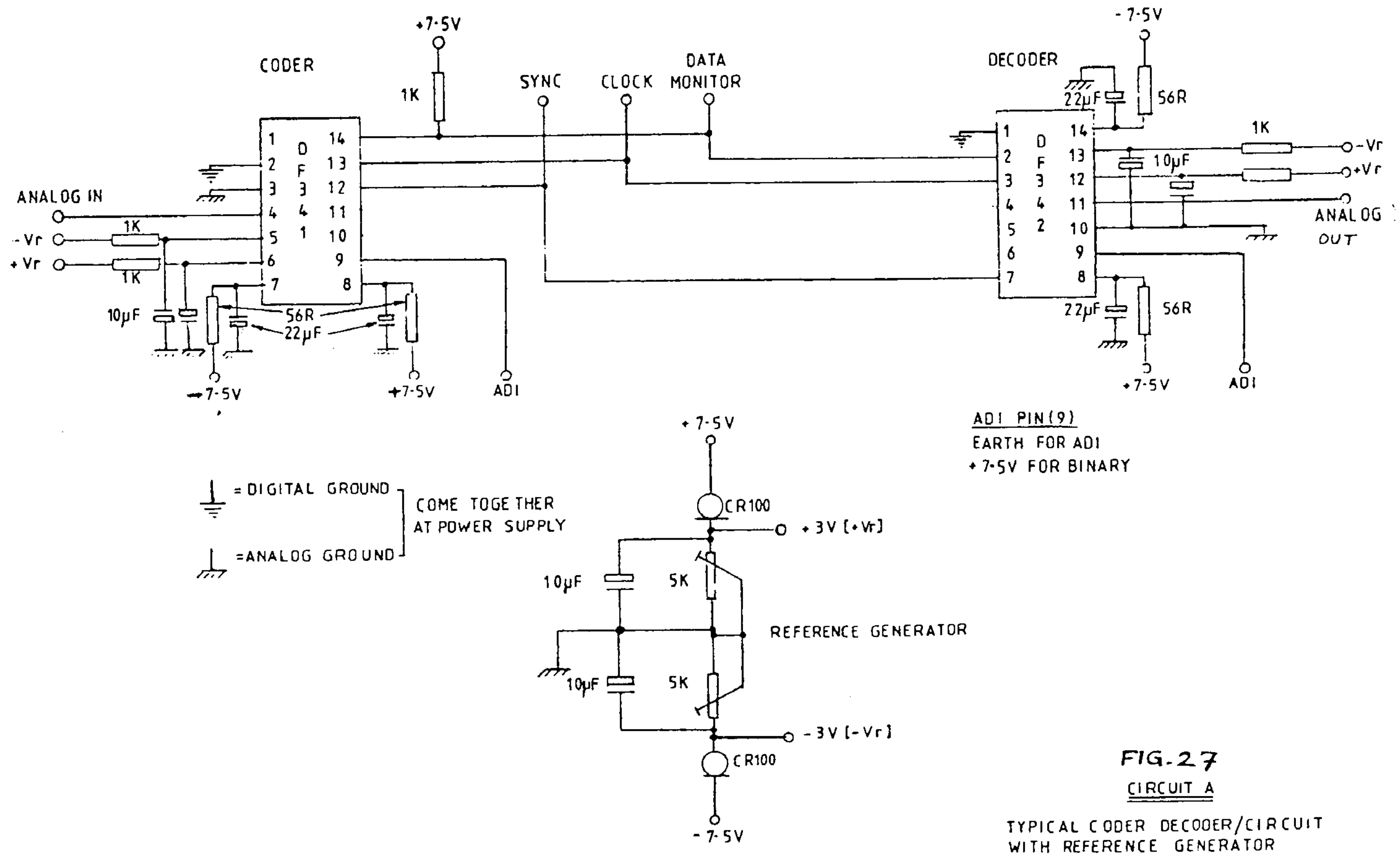
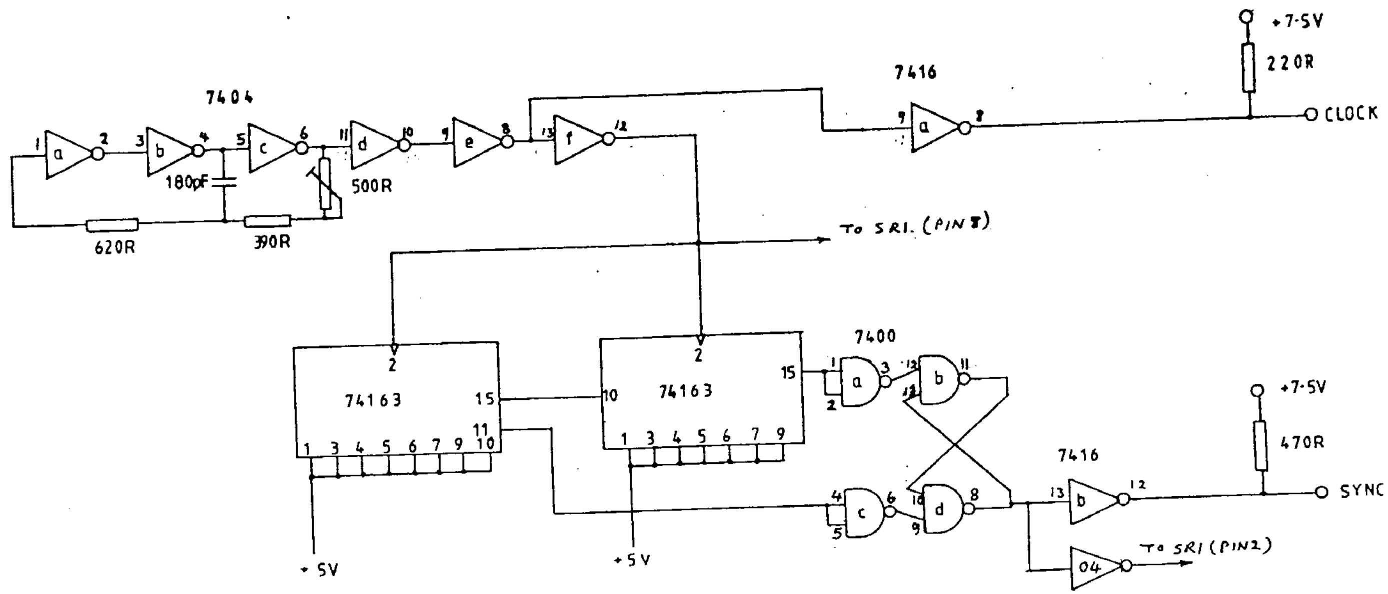


FIG. 27

CIRCUIT A

TYPICAL CODER DECODER/CIRCUIT
WITH REFERENCE GENERATOR



CIRCUIT B

CLOCK AND SYNC GENERATOR

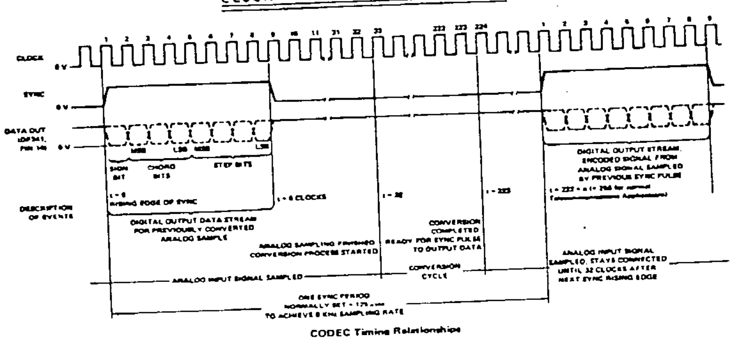


FIG. 27

IC'S REQUIRED

7400
7404
7416
74163 X 2

APPENDIX B

SIGNAL CONTROLLED ANALOGUE SWITCH

When there is no signal it was found that a lot of background noise passes through the filter to the sampler. In order to eliminate the effect of this sort of noise an analogue switch circuit which is controlled by the signal was used. This system is effective in stopping the noise only when there is no signal. When there is a signal the noise passes through because the switch is on. The circuit diagram of this system is shown in Figure 28.

Operation of this circuit is quite simple. The input signal is divided into two parts. One part of the input signal is fed to the input of the bilateral switch and the other part is fed to a high gain common emitter amplifier via a gain control VR1 and a d.c. blocking capacitor C2. The transistor Tr1 is the amplifier and it has collector load resistor R1 and base resister R2.

The output of the amplifier is fed to a rectifier and smoothing network, using C3, D1, D2, C4 and R3. The output of this network is used as a gating signal to the switch.

The variable resistor VR1 is adjusted so that the control voltage fed to the bilateral switch is insufficient to turn the switch on with only a background signal being fed to the input, but it is adjusted so that this voltage is only just about to turn the switch on. Therefore, when a proper input signal is present, the trigger voltage will be exceeded and the switch will be enabled.

The transition from off-state to on-state or vice versa should be complete, i.e. fully off or fully on, otherwise the switch will act as an attenuator to the signal. Fortunately, however, this range of transition is very restricted, and the 4016 or 4066 chip has a fairly well-defined changeover voltage. This circuit therefore works satisfactorily in practice without the need to incorporate any triggering.

The circuit has a fast attack and slow decay, hysteresis, which is necessary to ensure that the unit does not cut off during the brief pauses which occur during normal speech. The decay time can be altered to suit invididual requirements by modifying the value of C4. The decay time is proportional to the value of this component, and it is something less than one second, with the value shown in Figure 28.

Satisfactory operation of the device is possible with a nominal signal input level of less than 100 mV r.m.s. to a little more than 1 volt r.m.s. sinewave. The input impedance is about 10K. Tr2 and associated components form an emitter follower output stage which provides the unit with low output impedance.

C3 must be a high quality component if it is electrolytic type, and it would probably be best to use a tantalum bead or plastic foil component here. Some electrolytic types have relatively high leakage currents which would result in the delay time of the circuit being greatly prolonged.

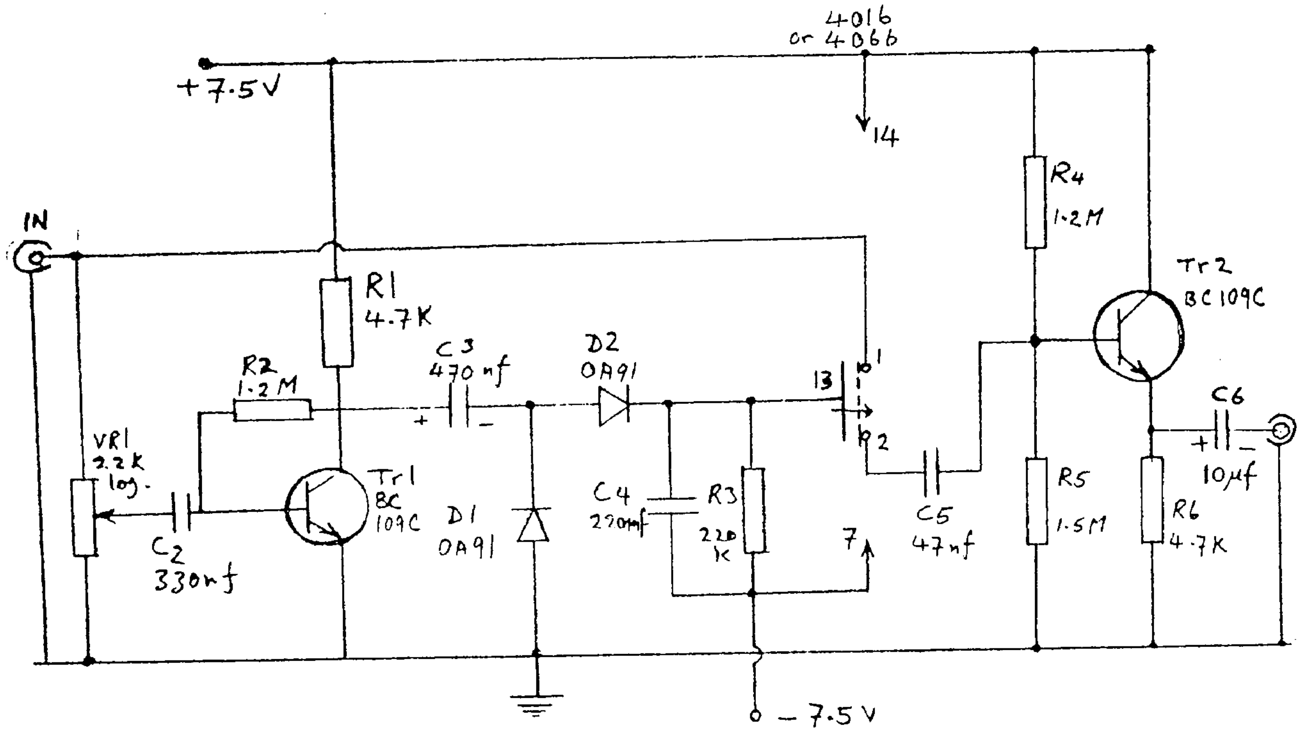


FIG. 28 SIGNAL CONTROLLED ANALOG SWITCH

LIST OF SYMBOLS

Δt_{min}	minimum time interval to be coded in PTCM system
Δt_{max}	maximum time interval to be coded in PTCM system
$\tan \alpha$	signal slope at any moment
h	voltage between two quantization levels
v	signal voltage
ω	angular frequency = $2\pi f$
f	signal frequency
$\Delta t_{store\ PTCM}$	delay in the store in PTCM system
S	store capacity
f_{wi}	word input rate to the store
f_{wo}	word output rate of the store
$P(x)$	probability density function of variable x
A	number of address lines
f_b	bit rate
f_t	transmission frequency
R	merit factor
q	number of quantization levels
f_s	sampling frequency
S_i, S_j	store occupancy
f_{bt}	bit transmission frequency
n	number of bits in a digital word.

REFERENCES

1. BINAL, M.E. 'Variable time-scale information processing', Ph.D. thesis, Department of Electrical Engineering, University of Aston, 1974.

2. MOUNTIS, M.G. 'Pulse time-code modulation system', Ph.D. thesis, Department of Electrical Engineering, University of Aston, 1978.

3. BETTS, J.A. 'Signal Processing, Modulation and Noise', Hodder and Stroughton Educational, Kent.

4. RICHARDS, D.L. 'Telecommunication by Speech', Butterworth and Co. (Publisher) Ltd., London.

5. MENDENHALL, P.C. 'Speech samples and analyser', M.Sc. Thesis, Dept. of Electrical Engineering, University of Aston, 1974.

6. GOACHER, D.J. and DENNY, J.G. 'The Teleprinter Handbook', First edition, Radio Society of Great Britain.

7. CATTERMOLE, K.G. 'Principles of Pulse Code Modulation', Iliffe books, 1969.

8. STEPHENS, N.W.F. 'Character recognition system using incoherent Fourier transform', Aston University, 1971.

9. SILICONIX, LTD. 'Leaflet describing the encoder-decoder devices DF331-DF332 and DF341-DF342.

10. LIPSCHUTZ, S. 'Probability' SI (metric) edition, Shaum's outline series, Mc Graw-Hill Book Co.

11. PFEIFFER, P.E. and SHUM, D.A. 'Introduction to applied probability'.

12. TEXAS INSTRUMENTS, The TTL Book for Design Engineers.

13. TEXAS INSTRUMENTS, The Linear and Interface Circuits Data Book for Design Engineers.

14. MALMSTADT, H.V., ENEK, Ch.G. and CROUCK, S.R. Electronic Measurements for Scientists, 4 volumes, W.A. Benjamin, Inc.

15. KERSHAW, J.D., Digital Electronics: Logic and Systems. Duxbury Press, North Scituate, Massachusetts.

16. HAMILTON, D.S., Handbook of Linear Integrated Electronics for Research. Mc Graw-Hill Book Company, (UK), Ltd.

17. MILLMAN, J. and TAUR, H. Pulse Digital and Switching Waveforms, Mc Graw-Hill, Kogakusha.

18. MEADOWS, R.G. Electrical Communications, The Macmillan Press, Ltd.

19. SCHWARTZ, M. and SHAW, L. Signal Processing, Mc Graw-Hill, Kogakusha, Ltd.