A DIGITAL SYSTEM FOR MULTIPLEX SPEECH TRANSMISSION

Thesis for the degree of Doctor of Philosophy

by

Keith Geoffrey Whitehouse

University of Aston April 1974

SUMMARY

This thesis describes the investigation of a digital transmission system intended for multichannel telephony. The object of the system is to achieve a bandwidth economy compared with conventional PCM (Pulse Code Modulation) systems, while maintaining good voice quality. This is accomplished by using simple predictive techniques to remove redundant samples from the digitized speech signal. A computer aided study of conversational speech revealed that such a strategy would eliminate a substantial fraction of samples.

The transmitted digit-rate is reduced, compared with PCM, by having fewer timeslots in the frame of transmitted pulses than channels. Mutual interference between channels occurs when the number of samples to be transmitted exceeds the number of available timeslots. This is termed frame overflow and places an upper limit on the number of channels for a given number of timeslots. Some methods of minimising the effects of frame overflow are described. The performance of these methods are compared by deriving signal-to-overflow noise ratios. It is concluded that the number of channels can be doubled, compared with conventional PCM, for the same bandwidth.

A simulation model of the proposed transmission system was constructed. Signal-to-noise ratio measurements and subjective tests on the model confirmed the findings of the theoretical analysis.

ACKNOWLEDGEMENTS

The work described in this thesis was supervised by Professor J. E. Flood, whose advice and encouragement is gratefully acknowledged. Thanks are also due to Mr. R. W. Whorwood for his invaluable guidance.

CONTENTS

Chapter

Chapter

Chapter

C H A P T E R 1

INTRODUCTION

INTRODUCTION

Digital modulation systems, and Pulse Code Modulation (PCM) in particular, are finding increasing application in the public telephone network. The one important disadvantage of digital systems is the large transmission bandwidth required compared with the original baseband signal. This is a result of the high digit rate required to achieve an acceptable voice fidelity. Techniques which reduce the digit rate while maintaining voice quality are therefore desirable.

One such technique is to remove redundant information from the digit stream. This process is reversible; information removed at the transmitter can be replaced at the receiver. Hence there is no loss of voice quality. An early bandwidth economy system employing redundancy reduction is Time Assignment Speech Interpolation (TASI)[1,2,3]. In this system, a subscriber is assigned a transmission channel only when he is speaking. Since, in normal telephone conversation, speech is present for less than half the time, TASI enables the number of channels provided to be at least doubled for the same bandwidth. Recent advances in digital integrated circuits and techniques have enabled the TASI system to be simplified and provided more economically[4,5]. Although TASI eliminates a lot of redundancy it is possible to remove even more using sample elimination techniques. A sample which is defined as being redundant is simply removed, thus creating a vacant timeslot. Applying this idea to a multiplex system enables the vacant timeslots so created to be used for non-redundant samples from other channels.

Sciulli[6] has investigated the statistical distributions of samples from a PCM encoder using both linear and non-linear encoding.

From the results obtained he concluded that simple predictive and differential techniques offer the best bandwidth compression potential. Dorgello and van der Veer [7] have investigated a multiplex system which achieves a bandwidth economy by eliminating zero-valued samples. Their work revealed that up to 60 channels could be accommodated in the bandwidth required for a standard 30 channel PCM system.

New forms of multiplex systems are possible if the baseband signal of each channel is stored for a short period before transmission [8,9]. Moreover, it should be possible to analyse the stored information to eliminate redundancy prior to transmission [10]. This principle forms the basis of the proposed digital multiplex telephone transmission system investigated in this thesis. In this system a sample is defined as redundant when it is identical to the immediately previous sample. To test for redundancy a one-sample store per channel is required at the sending terminal. A similar store is needed at the receiving terminal to replace redundant samples which have been removed at the sending terminal.

CHAPTER 2

THE PROPOSED SYSTEM

THE PROPOSED SYSTEM

2.1 General description

A block diagram of the system[11] is given in FIG. 2.1. At the sending terminal, the baseband signals of the channels are sampled sequentially (at Nyquist rate) and the samples are encoded in the analogue-to-digital (A/D) converter, as in PCM. Each encoded sample is sent to an appropriate address in the store; there is one location in the store for each channel. A sample is only sent to line if it differs from the previous sample stored at the address for that channel. The frame of transmitted pulses can contain fewer timeslots than channels, if there is a high probability of the encoded signals changing by less than one quantum step. This condition of "inactivity" is met both by silent channels and channels with a relatively small instantaneous rate of change.

At the receiving terminal there is a similar store; the digits received replace those already stored at the address corresponding to the channel. If no sample is received, the store contents remain unchanged for that channel. When all transmitted samples have been stored, the stored sample values are read out sequentially via the digital-to-analogue (D/A) converter to the baseband channels.

Since channel samples do not occupy fixed timeslots in the frame, it is necessary to transmit additional information to cause each coded sample to be written into the correct address in the receiving store. Two methods of "sample addressing" were considered, the first being:

a) Full Addressing

The format of a frame of transmitted pulses using full addressing is given in FIG. 2.2. The pulses from one channel consist of two binary words; the address word and the sample word.

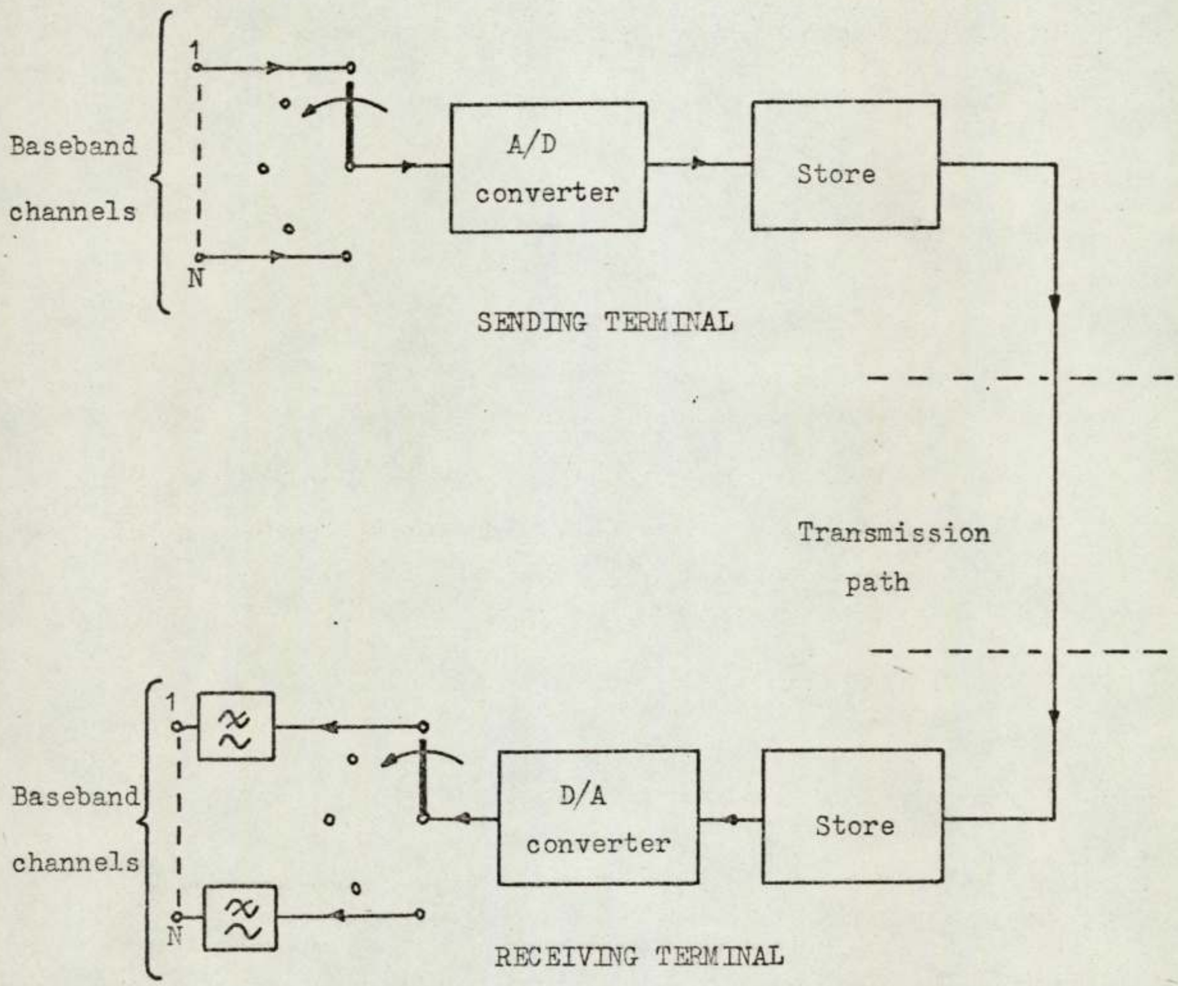


FIG. 2.1 The proposed multiplex transmission system

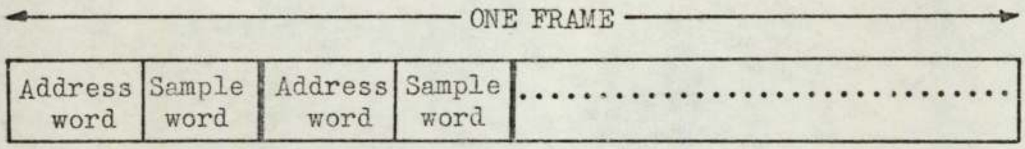


FIG. 2.2 Full addressing

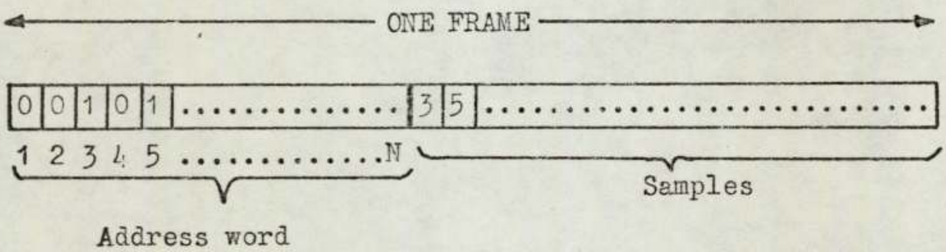


FIG. 2.3 Pre-addressing

The address word represents the channel number to which the sample is to be routed at the receiving terminal.

An alternative form of sample addressing is :

b) Pre-addressing

The format of a typical frame of transmitted pulses is given in FIG. 2.3. The channel samples are preceded by an N-bit address word (assuming an N-channel system). Each bit of the address word is associated with one channel; the first bit with channel one,etc. If the address bit is a binary '1' then the channel associated with that bit is "active" and a sample is transmitted. If the address bit is a '0' then no sample is sent. In the example illustrated in FIG. 2.3, channels 1 and 2 are inactive,but channels 3 and 5 are each sending a sample.

If the system has N channels and K timeslots per frame, method (a) requires $K\log_2 N$ bits, whereas method (b) only requires N bits. It will be shown that in a practical system K is approximately equal to N/2.5. Hence, for $N > 4$, pre-addressing uses less transmission-path capacity than full-addressing. However, pre-addressing is more vulnerable to bit errors on the transmission line. An error in one bit of the sample address word causes all subsequent channels in that frame to receive incorrect samples. However, an error only persists until the next sample is sent. Errors can be detected by appending a parity-check bit to the address word. If the parity-check fails at the receiving terminal, the incoming samples are discarded and the existing stored values are read out to the channels. This procedure produces less disturbance than if incorrect samples were sent to the receiving store.

Sciulli and Campanella[12] have made a theoretical investigation of the noise produced by errors in the sample address word protected by a single parity-check bit. Their calculations show that, for a 64-channel system, the noise produced is negligible compared with quantizing noise, provided the bit error rate is less than 1 in 10^5.

When using the full-addressing method, a bit error in one address word causes only two channels to receive an incorrect sample. However, provided the bit error rate is small, the pre-addressing method is to be preferred because of its economy in frame pulses.

2.2 Bandwidth economy

The bandwidth economy of the proposed system can be calculated quite easily, as follows :

$$\text{Let } R = \frac{\text{Bandwidth required for N-channel PCM system}}{\text{Bandwidth required for N-channel proposed system}}$$

If digits for synchronization are ignored, this can be written as :

$$R = \frac{\text{Number of bits per frame for PCM}}{\text{Number of bits per frame for proposed system}}$$

If pre-addressing is used then :

$$R = \frac{Na}{Ka+N}$$

Where :

N = Number of channels

K = Number of timeslots per frame

a = Number of bits per sample

R can be re-written as :

$$R = \frac{1}{\frac{K}{N} + \frac{1}{a}} \qquad (2.1)$$

To obtain a large bandwidth advantage, the ratio of $\frac{K}{N}$ should be small. However if the number of 'active' channels in any frame exceeds K, one or more channels lose a sample. The smaller $\frac{K}{N}$, the more samples are lost, because the number of active channels exceeds K more often. The probability of losing samples which can be tolerated,thus,places a lower limit on the value of $\frac{K}{N}$ that can be used.

2.3 Problems requiring investigation

The initial stage of the project was to determine the feasibility of the proposed system. Clearly, if the system is to be efficient (in bandwidth economy) then there needs to be a high probability of zero change between successive encoded samples from each channel. This probability was measured, for conversational speech, with the aid of a digital computer and analogue-to-digital converter. The probability of zero change is a function of the instantaneous rate of change (i.e. slope) of the speech signal. Thus, to gain further insight into the problem the statistical distribution of speech slopes was investigated. Again using the digital computer and A/D converter the probability density function (pdf) of speech slopes was obtained.

In common with existing PCM systems, the proposed system requires some form of companding to accommodate variations in talker volumes. The compressor, at the sending terminal, changes the relative amplitudes of large and small components of the speech waveform; consequently it also changes the probability of zero change between successive encoded samples. Thus, companding directly affects the bandwidth saving potential of the proposed system. The companding in practical PCM systems is of the instantaneous type; however it was considered possible that companding at syllabic rate might give a higher probability of zero change. To test this hypothesis, both types of companding were simulated on a digital computer in conjunction with an A/D converter.

The bandwidth economy of the proposed system is achieved by having fewer timeslots per frame than input channels. An important feature of the system is the probability that the number of "active" channels in any frame exceeds the number of available timeslots. When this occurs, the frame is said to have "overflowed"; this causes one or more channels to lose a sample. Let the number of timeslots be K and the number of channels be N. If timeslots are assigned to channels sequentially on a fixed priority basis then the first K channels will never experience frame overflow. However, channels (K + 1) to N will experience increasing sample losses with increasing channel number. This situation is obviously undesirable, and some method of distributing frame overflow throughout all channels is required. Various methods were considered and their merits investigated. Clearly, the amount of frame overflow that can be tolerated, places an upper limit on N for a fixed number of timeslots.

The effect of frame overflow can be considered as a degradation of signal-to-noise ratio. The amount of degradation can be calculated for each of the overflow error-distribution systems considered. However, the ultimate performance measure for evaluating voice communication systems is the subjective quality of the received speech.

To compare the subjective performance of the various overflow error-distribution systems, a N-channel model (with N variable) of the proposed system was built. The model consists of one speech channel with simulation of the interference caused (due to frame overflow) by the remaining (N-1) channels. The subjective performance of the error-distribution systems was considered in terms of listeners' preference. One error-distribution system was used as a reference, and the others compared with it, using isopreference tests.

The proposed system uses zero-order prediction (ZOP) to eliminate redundant information from the encoded samples. The next sample is predicted to be the same as the last sample. An alternative is to predict that the next sample is zero-valued; this is termed zero-amplitude prediction (ZAP). This alternative does not require a store at the sending terminal. However, the receiving store is retained in case of errors in the sample address word. To compare the subjective performance of this alternative scheme with ZOP, it was also built into the model.

This investigation started in October 1970 and a preliminary description of the system was published in January 1972 (J.E. Flood and K. G. Whitehouse : 'Digital Multiplex Telephone Transmission System Using Statistical Properties of Speech', Electronics Letters, vol. 8, no. 4, p.96, 24th February, 1972). During the preparation of this thesis, it was discovered that other workers had been investigating a similar system. Sciulli and Campanella described in the July 1973 issue of the I.E.E.E. Transactions on Communications (Vol. COM-21, no. 7) a system called SPEC (Speech Predictive Encoding Communication System). This is a digital multiplex telephone system basically the same as that shown in FIG.2.1. Their system uses ZOP to eliminate redundancy and the pre-addressing method of sample addressing is used. They mention only one type of overflow error-distribution system.

CHAPTER 3

SPEECH WAVEFORM ANALYSIS

SPEECH WAVEFORM ANALYSIS

3.1 Introduction

The performance of the proposed communication system (in terms of bandwidth economy) is a function of certain statistical properties of the baseband signal. The statistics of interest concern the rate of change, or slope, of the speech waveform. Of particular interest is the probability of the speech waveform changing by less than one quantum between sampling instants. This probability must be relatively large for the proposed communication system to be viable.

Statistics on the instantaneous amplitude of speech waveforms have been obtained by many workers including Davenport [13] and Richards [14]. However, little attention has been paid to speech slopes. Sciulli [6] has investigated the statistics of digitally encoded speech samples for a 15 second speech passage, and gives a histogram of first difference probabilities. Although this information was useful, a more detailed study of speech slopes was required on longer speech passages. This chapter describes the investigation that was carried out to obtain the required data.

The investigation can be divided into two parts:

a) Measurement of the probability of zero change ($P(d_o)$) between successive quantized samples of the speech waveform.

b) Determination of the probability density function (p.d.f.) of the instantaneous slope of the speech waveform.

The second part of the investigation, although not essential, gives further insight into the results obtained in part one.

Extensive use was made of a digital computer for processing data and presentation of results.

3.2 Measurement of the probability of zero difference P (d_o)

To avoid the time and expense of constructing equipment, the PDP-9 digital computer in the Electrical Engineering Department was used to make the measurement. An analogue-to-digital (A/D) converter was used in conjunction with the computer.

A block diagram of the experimental apparatus used to measure P (d_o) is given in FIG. 3.1. The speech passage to be analysed is recorded on standard $\frac{1}{4}$" magnetic tape. The tape is played back on the tape recorder, and the tape recorder output signal is amplified to a level suitable for the A/D converter. The calibrated attenuator is included to set the mean level of the speech signal accurately. The signal is then passed through a 300 H_z to 3400 H_z band-pass filter to simulate a practical telephone channel. The bandlimited signal then passes to the A/D converter, where it is sampled at the standard 8 $_kH_z$ rate, and quantized uniformly in a 128-level range.

Encoded speech samples from the A/D converter are processed by the computer and the results printed out on a teletype.

In effect the apparatus up to, and including, the A/D converter simulates the transmit terminal of a PCM system with 7-bit linear encoding.

3.2.1 Computer programmes and flowcharts

The computer programs comprise two parts:

a) Main program (written in FORTRAN 4), which executes most of the arithmetic and controls the printout of results.

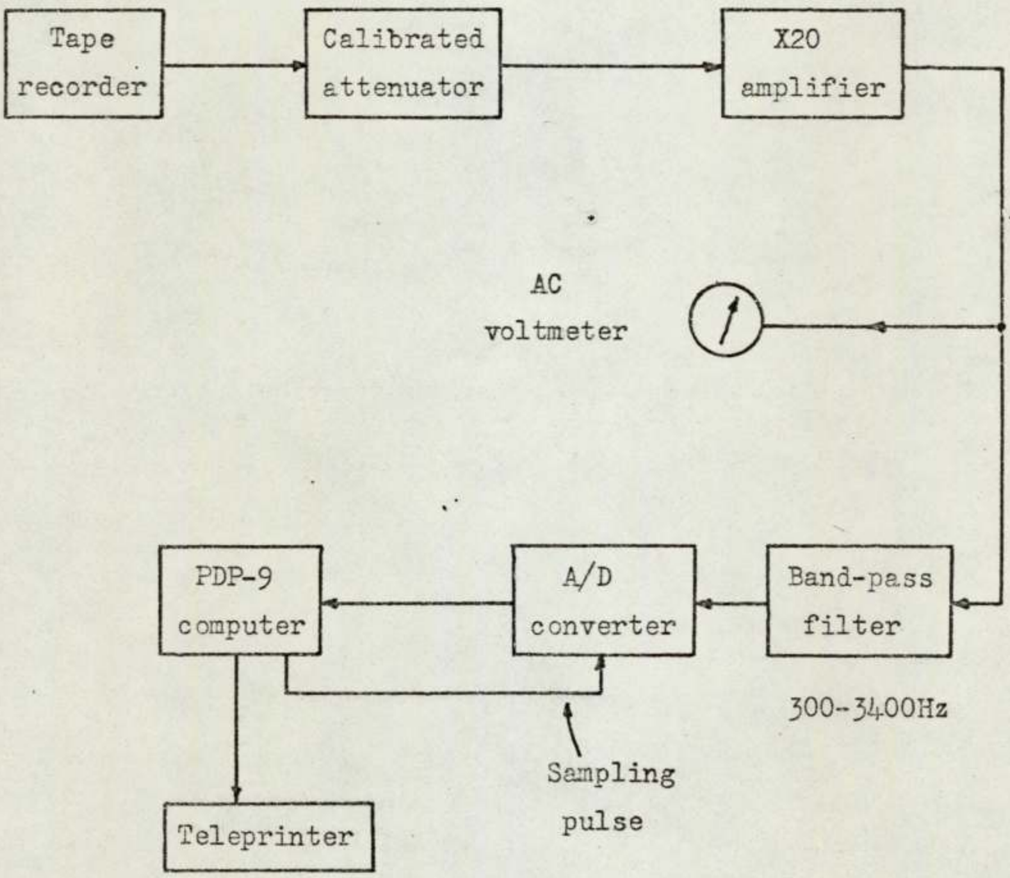


FIG. 3.1 Experimental apparatus used for measuring $P(d_o)$

b) A subroutine (written in MACRO) within the main program. This services the A/D converter and controls the sampling interval.

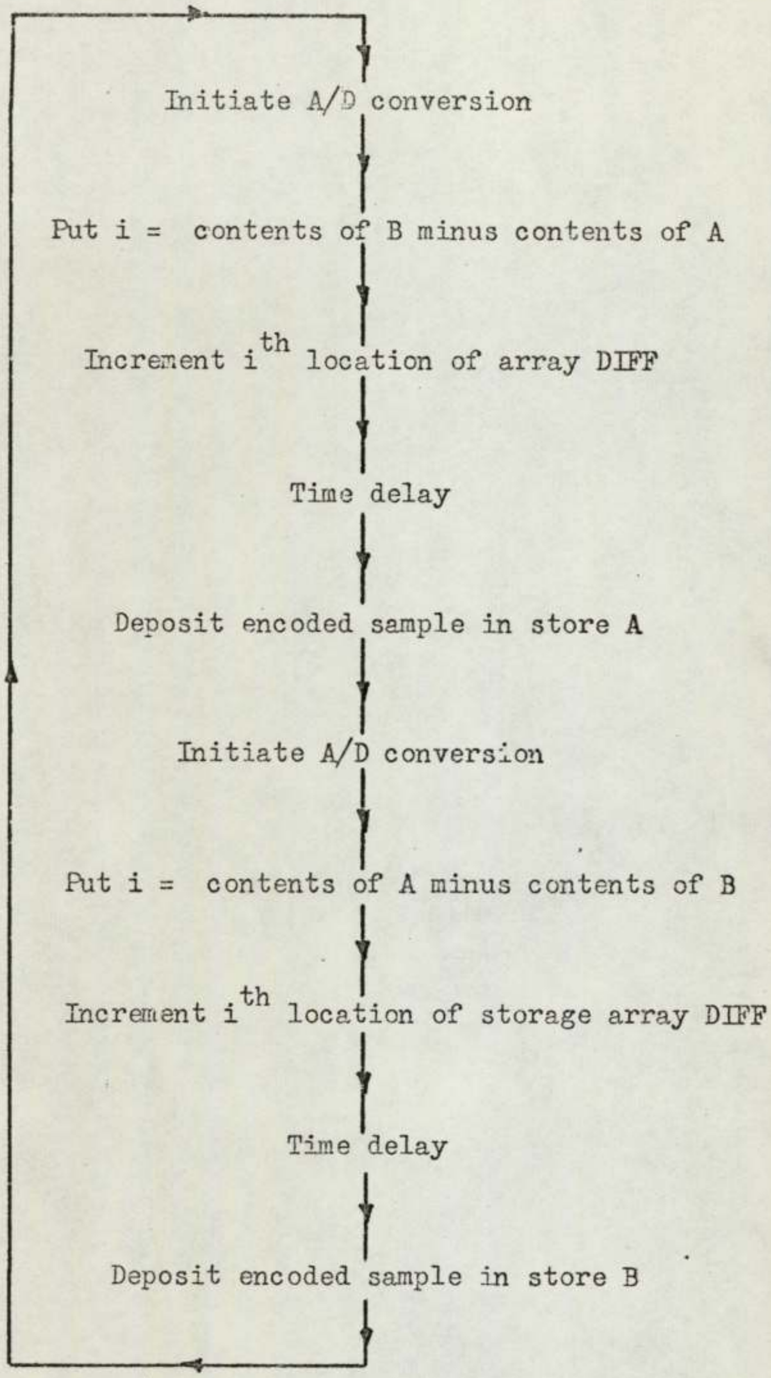
A copy of both programs is given in Appendix A1.

A simplified flowchart indicating the operation of the MACRO subroutine is given in FIG. 3.2. The time between sampling instants of the A/D converter is set by this subroutine. It has been arranged that the time between successive instructions 'initiate A/D conversion' is exactly 125μS. This is achieved by using the fact that each program instruction takes a multiple of 1μS to execute. By adding up the total time for all instructions and adding an equalising time delay the required 125μS can be obtained.

The storage array DIFF has 128 locations and is associated with the difference statistics of the speech waveform. Each location is associated with one particular value of difference. To see how it is used, let the difference between successive samples at some sampling interval be i quanta, where i varies from zero to 128. Then, location i of DIFF is incremented. This procedure is repeated at every sampling instant.

At the end of a test run, location i in DIFF contains the total number of times that a difference of i quanta occured. Then the probability ($P(d_i)$) of difference i occurring is calculated as:

$$P(d_i) = \frac{\text{contents of location i in DIFF}}{\text{Total samples taken}} \qquad (3.1)$$



<u>KEY</u>

A&B : Temporary storage for encoded samples.

DIFF : One-dimensional storage array with 128 locations.

<u>FIG. 3.2</u> <u>Flowchart of MACRO subroutine</u>

3.2.2 Speech passages

The speech passages used in the measurement were recorded on standard $\frac{1}{4}$" magnetic tape. This allows the two procedures of obtaining the passages and analysing them to be conveniently separated.

Since the purpose of this investigation is to provide data to check the viability of a telephone communication system, telephone conversations were chosen for analysis. The author is indebted to Mr. D. L. Richards of the Post Office Research Station who made several hours of recorded conversations available. The conversations were not recorded in the public telephone network, but were produced in a laboratory by the Post Office to simulate conversation typical of the network.

The two participants in each conversation were given identical sets of numbered pictures prior to the conversation. They each place the pictures in their order of preference. Then they discuss their choice over a telephone link, eventually achieving a mutual order of preference. The conversations are typically between three and five minutes in length.

Each participant is recorded on a separate tape track with no crosstalk between tracks. At the beginning of each track there is a 1 kHz calibration tone one minute in length. A chart, provided with the tape, indicates the mean speech level of each talker relative to the calibration tone. Thus, on playback, the calibration tone can be set to some convenient level with the volume control on the tape recorder. Then, by reference to the chart, the mean speech level of any talker can be adjusted, with the calibrated attenuator, to the desired level.

The choice of mean speech level is influenced by two factors:

a) Maximising $P(d_o)$

b) Achieving an acceptable signal-to-quantizing-noise ratio. The mean speech level (S) needs to be as low as possible to maximise $P(d_o)$. However a lower limit to S is imposed by the required signal-to-noise ratio.

It can be shown that the quantizing noise N_q with uniform quanta size δV is given by [15]:

$$N_q = \frac{1}{12} (\delta V)^2 \tag{3.2}$$

then

$$\frac{S}{N_q} = \frac{12\ S}{(\delta V)^2} \tag{3.3}$$

Purton [15] argues that 28db is an acceptable value for S/N_q; substituting this value into equation (3.3) gives S a value of 19db below the peak quantizer level for 128-level quantizing. This value of S also gives a more than adequate clipping margin.

3.2.3 Speech detector

The values of $P(d_o)$ obtained from initial tests were disappointingly low. The cause was traced to noise during pauses in speech. The noise occurred in two forms :

a) Thermal noise, the peaks of which were greater than the first quantum level.

b) Room noise caused by the listener.

Both types of noise, of course, also occur in a practical telephone network.

To eradicate the effects of this noise, a simple speech detector was fitted in the signal path, between the calibrated attenuator and the X20 amplifier. This device allows speech to pass but rejects low-level noise. A block diagram of the speech detector is given in FIG. 3.3. A detailed circuit diagram is given in Appendix A2. The incoming signal takes two paths. One is to a transmission gate which is opened and closed by a control signal from a voltage comparator. The input signal is also rectified by diode D, and the resulting d.c. voltage is averaged by a single-section RC filter of 10mS time constant. The output voltage Vs from the filter is compared with a fixed threshold voltage V_t ; when V_s exceeds V_t, the comparator output signal opens the transmission gate.

The setting of V_t is a compromise between speech quality and efficient noise rejection. The value of V_t was increased while listening to the speech at the detector output. The final setting of V_t was at the point that the effects of speech detection on speech quality just became evident.

After fitting the speech detector, the values of P (d_o) obtained substantially increased.

3.2.4 Speech activity

The output from a PCM encoder during silent pauses in speech will consist of long "runs" of zero-valued samples. Obviously, there will also be equally long runs of zero-differences between these encoded samples. Clearly, silent pauses will provide a substantial fraction of zero differences that occur.

P (d_o) can be considered as the sum of two fractions:

a) The fraction of time occupied by silent pauses.

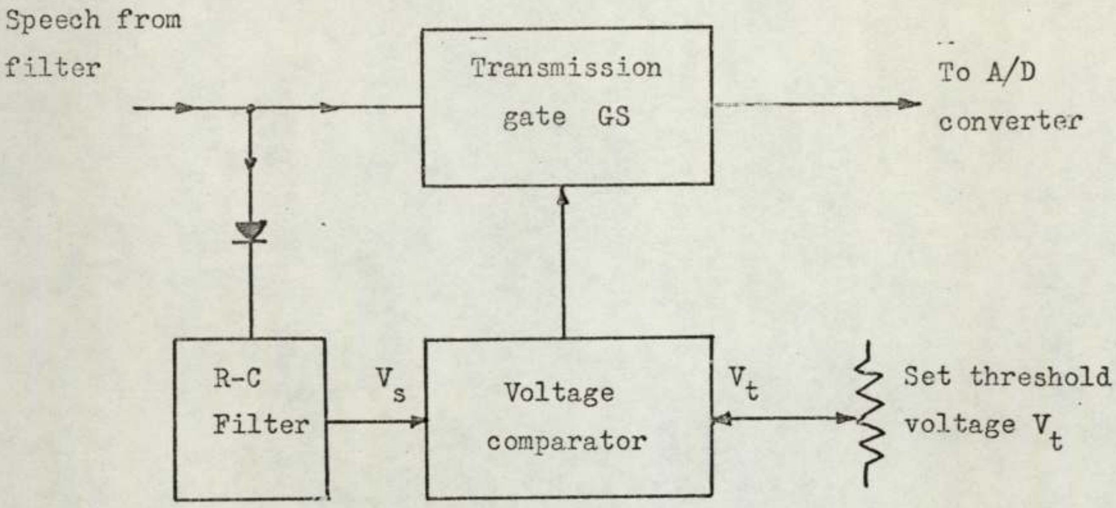


FIG. 3.3 Speech detector

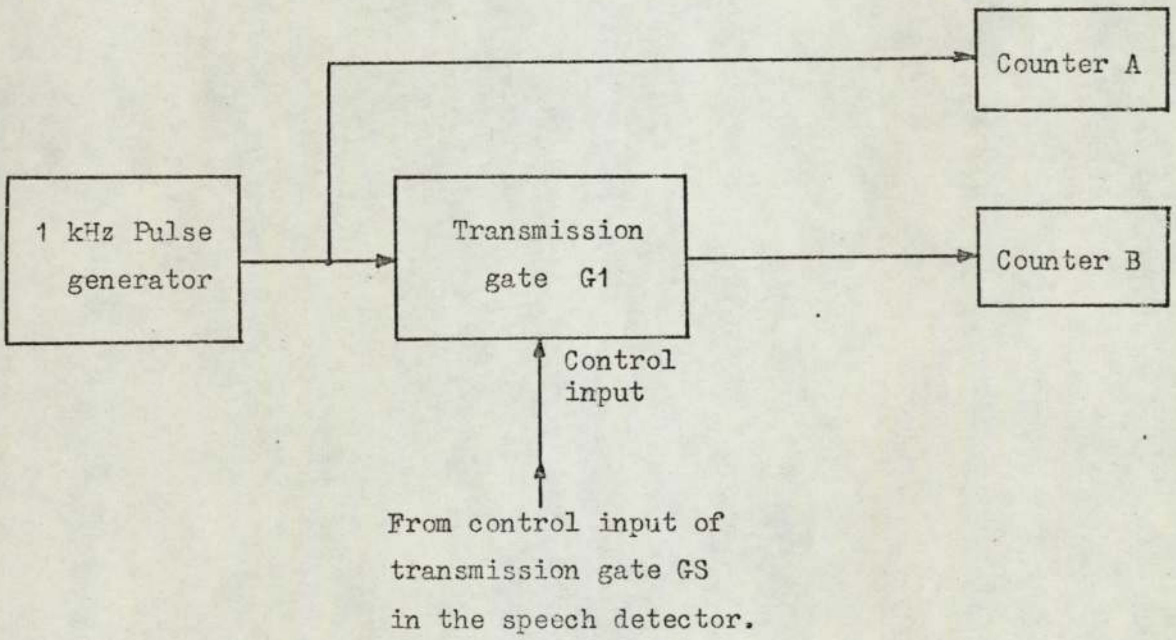


FIG. 3.4 Measurement of τ_A

b) The fraction of time during "active speech" that zero differences occur.

By "active speech" is meant those signals that operate the speech detector.

Let the fraction of time that the speech detector is operated be τ_A ; then P (d_o) can be written as:

$$P(d_o) = (1 - \tau_A) + C\tau_A \tag{3.4}$$

where C is a constant and :

$$0 < C < 1 \tag{3.5}$$

Since P (d_o) and τ_A can be measured , the value of C can be calculated by re-arranging equation (4) i.e.

$$C = \frac{P(d_o) - (1 - \tau_A)}{\tau_A} \tag{3.6}$$

C is, thus, the probability of zero difference occurring in active speech.

3.2.5 Measurement of τ_A

The fraction of time occupied by active speech, τ_A, was measured by the method shown in FIG.3.4. The output from the pulse generator goes directly to counter A and via a transmission gate G1 to counter B. The transmission gate G1 is controlled by the same signal that controls the similar gate GS in the speech detector of FIG.3.3. Hence G1 opens and closes in unison with GS. Initially, both counters are reset to zero and held reset until the conversation begins. The counter A continually accumulates the number of pulses from the pulse generator, but counter B only accumulates pulses when gate G1 is open. At the end of the conversation, the contents of the two counters enable the value of τ_A to be calculated.

$$\tau_A = \frac{\text{Contents of counter B}}{\text{Contents of counter A}} \tag{3.7}$$

3.2.6 Results

The results of analysing five conversations are presented in TABLE 3.1. The values of C were calculated using equation (3.6).

TABLE 3.1

Conv. No.	TALKER A			TALKER B			Length of conv. (mins.)
	$P(d_o)$	$1-\tau_A$	C	$P(d_o)$	$1-\tau_A$	C	
1	0.812	0.745	0.263	0.619	0.485	0.260	5.5
2	0.675	0.557	0.266	0.568	0.395	0.286	5.0
3	0.452	0.241	0.278	0.739	0.617	0.318	4.5
4	0.641	0.530	0.236	0.868	0.797	0.350	4.5
5	0.660	0.502	0.317	0.691	0.563	0.293	4.5

Results of speech analysis

Overall averages of the data listed in TABLE 3.1 are given below:

$$P(d_o) = 0.673$$

$$(1-\tau_A) = 0.545$$

$$C = 0.285$$

The results in TABLE 3.1 clearly show the dependence of $P(d_o)$ on speech inactivity $(1-\tau_A)$. Also, the values of C are quite consistent, there being only a few values differing by more than 10% from the average.

The computer program also provides statistics on non-zero differences. FIG. 3.5 is a histogram of the probabilities of the first ten differences. Each point is an average of the data from all five conversations listed in TABLE 3.1. The same data are presented as a cumulative distribution in FIG. 3.6.

FIG.3.5 also gives the approximate shape of the probability density function of the instantaneous slope of the speech waveform. The large peak at zero difference and long "tail" are also typical of the p.d.f. of instantaneous amplitude [13].

3.3 Prediction of P (d_0) for a practical telephone network

The results given in section 3.2.5 enable a value of P (d_0) to be predicted for a practical telephone network. The only other information required is the average speech activity (let this be $\bar{\tau}_A$) in the network. Then from equation (3.4) :

$$\text{Predicted } P(d_0) = (1 - \bar{\tau}_A) + C\,\bar{\tau}_A \qquad (3.8)$$

Where C is the probability of zero-differences occurring in active speech as defined in section 3.2.4.

In a typical conversation, each talker speaks for about 50% of the time. However, even during speech there are short pauses and therefore τ_A will have a value somewhat less than 0.5. This point was confirmed by measurements made on the recorded conversations investigated in this chapter. They yielded an average speech activity of 0.454. In a practical telephone network there are other factors which reduce τ_A even below this figure. These include:

a) A subscriber waiting for a number to ring out.

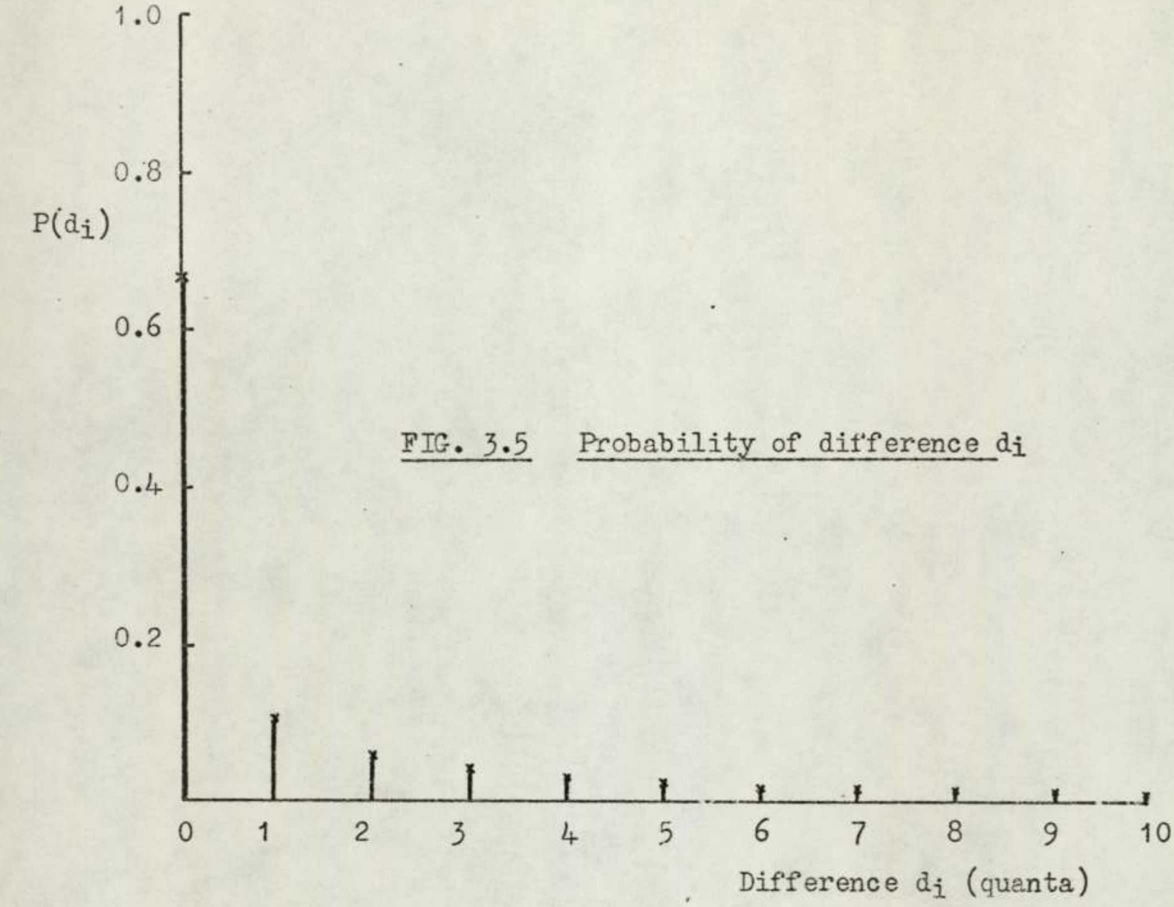


FIG. 3.5 Probability of difference d_i

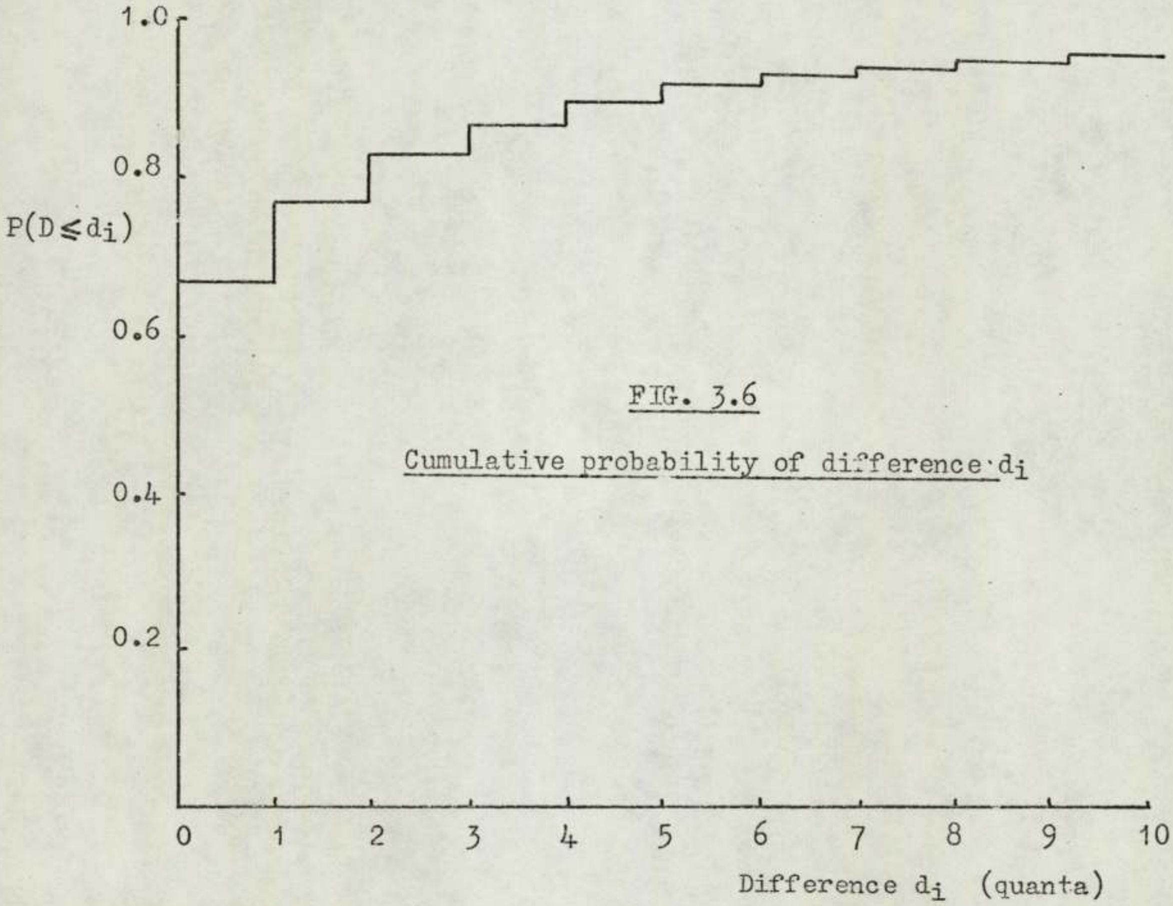


FIG. 3.6

Cumulative probability of difference d_i

b) A subscriber waiting for someone to be brought to the telephone at the other end of the line.

On the other hand, bursts of noise and echoes will tend to increase τ_A.

Measurement of speech activity on transatlantic telephone circuits have been published by Bullington [1] and Berry [16]. These measurements used speech detection devices similar to that described in section 3.2.3. Under busy-hour conditions, average values of approximately 0.4 were obtained.

Assuming $\bar{\tau}_A$ to be 0.4, the predicted value of $P(d_o)$ from equation (3.8) is:

$$P(d_o) = 0.6 + 0.4C \qquad (3.9)$$

For 7-bit linear encoding $C = 0.285$

Then, $P(d_o) = 0.6 + 0.285 \times 0.4 = 0.714$

3.4 Probability density function of the slope of speech waveforms

3.4.1 Theory of measurement

The probability density function (p.d.f.) of a random variable x is defined by the equation:

$$p(x) = \frac{P(x_1 < x < x_1 + d_{x_1})}{d_{x_1}}$$

Where $P(x_1 < x < x_1 + dx_1)$ is the probability that x lies in the range x_1 to $(x_1 + dx_1)$.

In this case, the random variable is the slope of the speech waveform.

$P(x_1 < x < x_1 + dx_1)$ is measured by approximating it to the relative frequency of occurrence $f_r(x_1)$. If x lies in the range x_1 to $(x_1 + dx_1)$ t times in a total batch of T samples then:

$$f_r(x_1) = \frac{t}{T} \simeq P(x_1 < x < x_1 + dx_1)$$

Since $f_r(x_1)$ is formed by counting samples of a random process that fall in a fixed range, then $f_r(x_1)$ is also a random variable. The variance of $f_r(x_1)$ will limit the accuracy with which it can approximate to $P(x_1 < x < x_1 + dx_1)$. Acceptable accuracy can be obtained by taking a large number of samples.

The technique used to measure slope probabilities is very similar to that used for measuring $P(d_o)$. The signal is band-limited to the range 300-3400 Hz and sampled at a rate of 8KH$_z$. The slope of the waveform is measured as shown in FIG.3.7. The times t_A and t_B indicate two instants at which samples were taken. The sample values at t_A and t_B are V_A and V_B respectively.

The slope at a point midway between t_A and t_B is measured as:

$$\text{Slope } x' = \frac{V_B - V_A}{t_B - t_A} = \frac{D}{\Delta t}$$

Δt is a constant 125 μS.

The quantity D is a discrete number (because of quantization) and the probability of it taking on any particular value can be measured exactly as for $P(d_i)$.

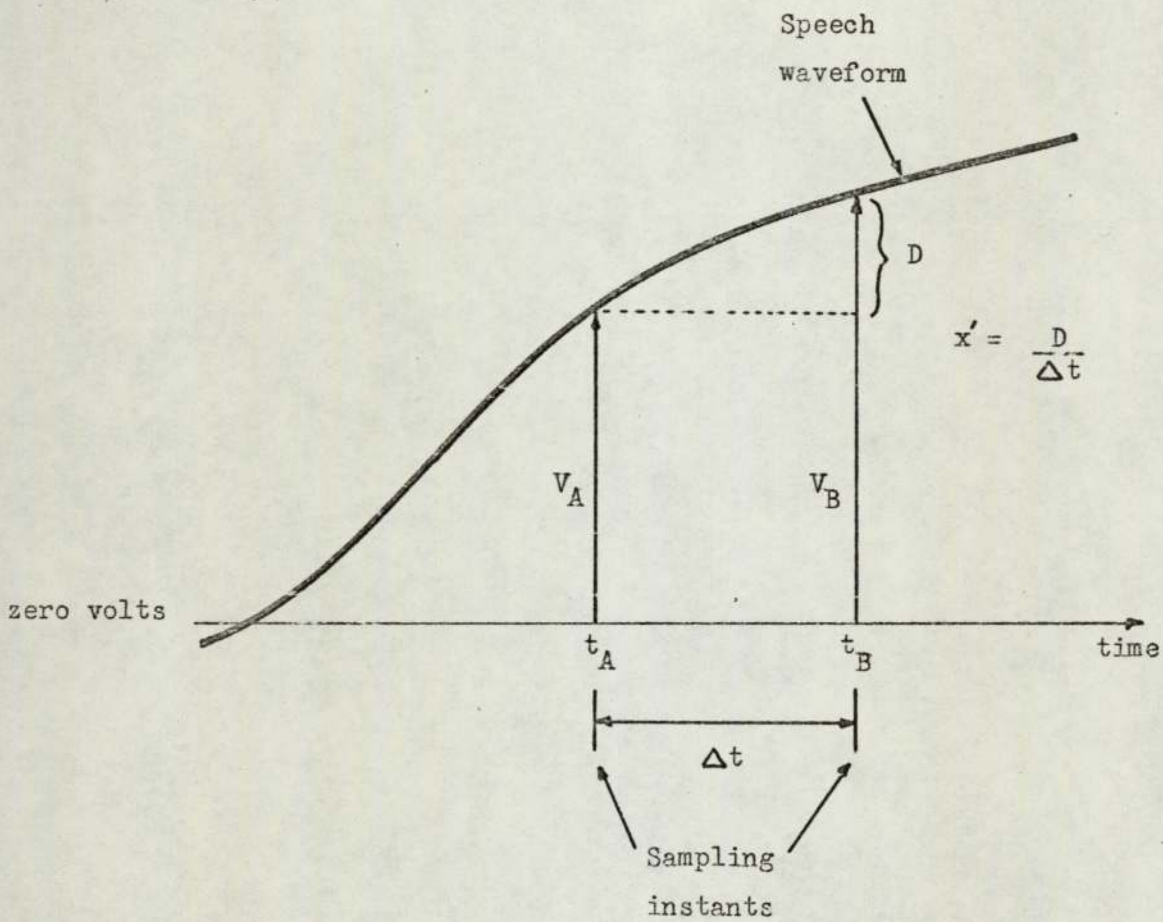


FIG. 3.7 Measurement of slope

3.4.2 Experimental procedure

The experimental equipment is identical to that for measuring $P(d_o)$ (see FIG. 3.1). The computer programs used are also similar, the main differences being:

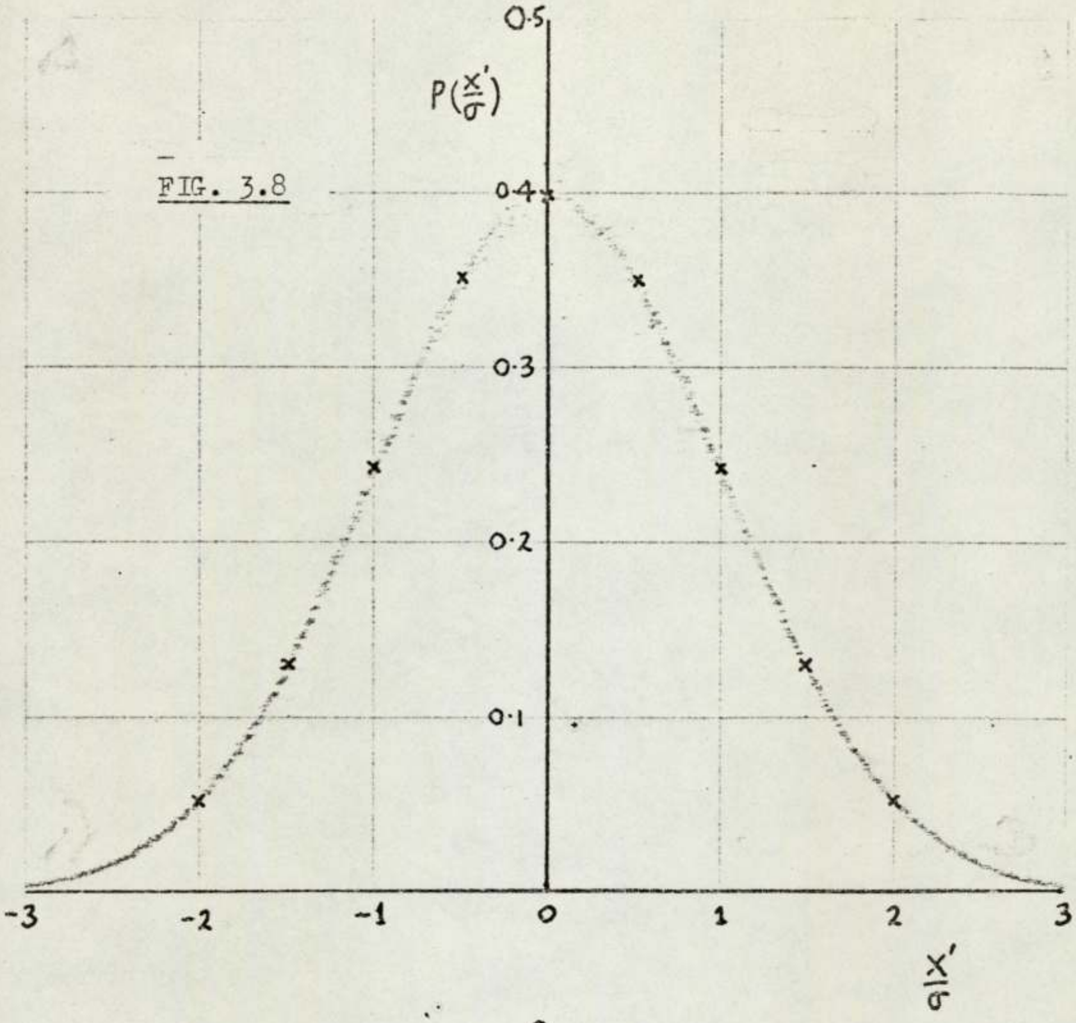
a) 4096 - level quantization instead of 128 - levels. This enables the slope to be measured more accurately.

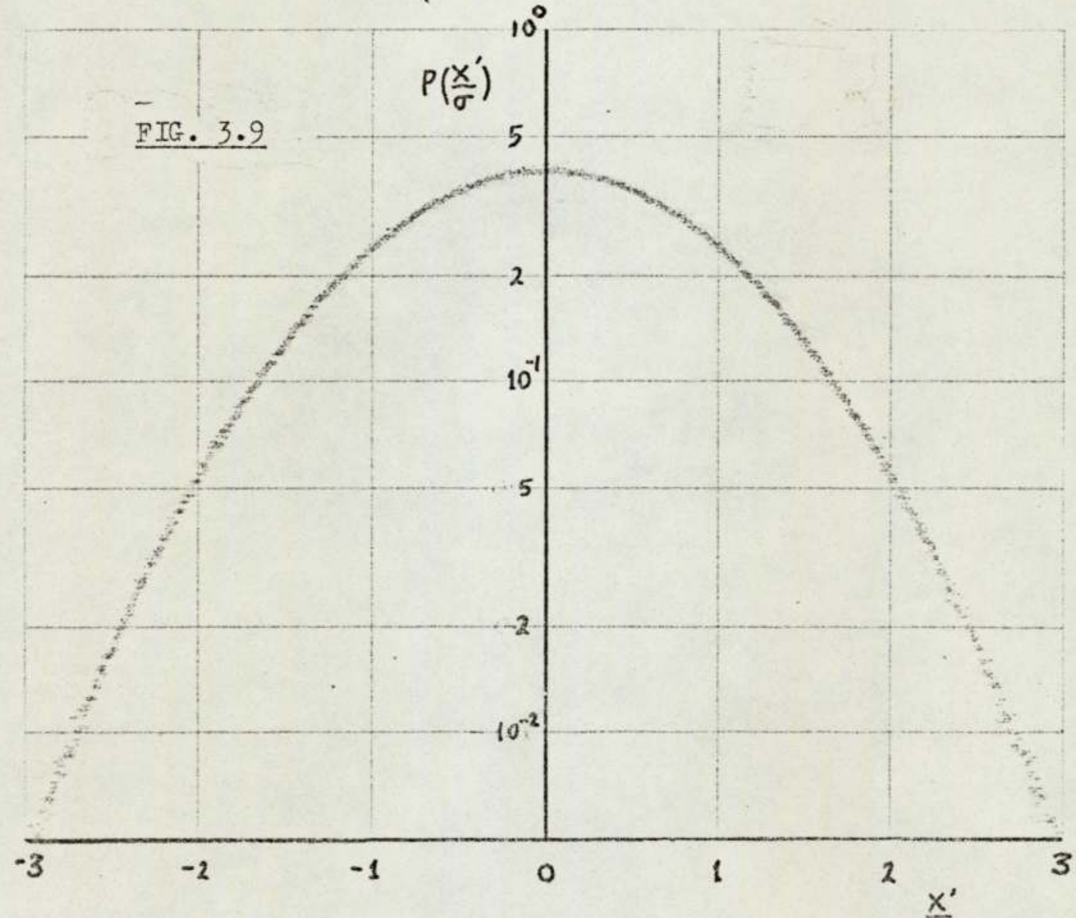
b) The results were plotted directly on graph paper, using a computer-controlled x-y plotter.

As an overall check on the measuring system, a waveform was analysed for which the p.d.f., was known. A convenient random process is a "white" noise voltage waveform.

The tape recorder in FIG. 3.1., was replaced by a General Radio type 1383 white noise generator. A five-minute segment of the voltage waveform was analysed. The resultant p.d.f., is illustrated to a linear scale in FIG. 3.8 and to a semi-log scale in FIG.3.9.

The autocorrelation function of white noise ($R(\tau)$) bandlimited to f_m is zero at $\tau = \frac{1}{2f_m}$. Hence if the noise is sampled at a rate of $2f_m$ the samples are uncorrelated. Furthermore, since the noise is a normally-distributed process the samples are independent. The slope of the waveform, as measured here, is proportional to the difference between two successive samples. Hence the slope is proportional to the difference between two independent normal processes, which is another normally-distributed process. Thus, the p.d.f. illustrated in FIG. 3.8 and FIG. 3.9 should be that of a normal process. The larger crosses in FIG.3.8 are data taken from a published table of values for a normal p.d.f. It can be seen that there is good agreement between the measured and published data.





Pdf of the instantaneous slope of band-limited white noise

It will be noted that the measured data in FIG.3.8 and FIG. 3.9 form "bands" rather than a single line. This is due to the variance in $f_r(x_1)$, as mentioned in section 3.4.1. The "spread" in data is quite small because of the large number of samples that are taken. All the speech passages analysed were at least five minutes in length, so approximately 2.4×10^6 samples were taken.

The computer program also calculates the r.m.s. slope (σ) over the entire speech passage. The instantaneous slope is then normalised by dividing by σ.

3.4.3 Results

Having established the accuracy of the measuring system, several speech passages were analysed. These included:

a) Two of the telephone conversations used in the first part of this investigation.

b) Continuous speech read from text into a high quality ribbon microphone (male and female talkers)

The results obtained are plotted in FIGS. 3.10 to 3.17. The results from each speech passage are plotted on a linear scale and again on a semi-log scale.

The striking features of the p.d.f.s illustrated in FIGS. 3.10 to 3.17 are:

a) The large central peak and

b) A long "tail" as $\frac{x'}{\sigma}$ increases.

These features are also typical of the amplitude p.d.f. for speech[13].

Male voice (KGW)

High quality ribbon microphone

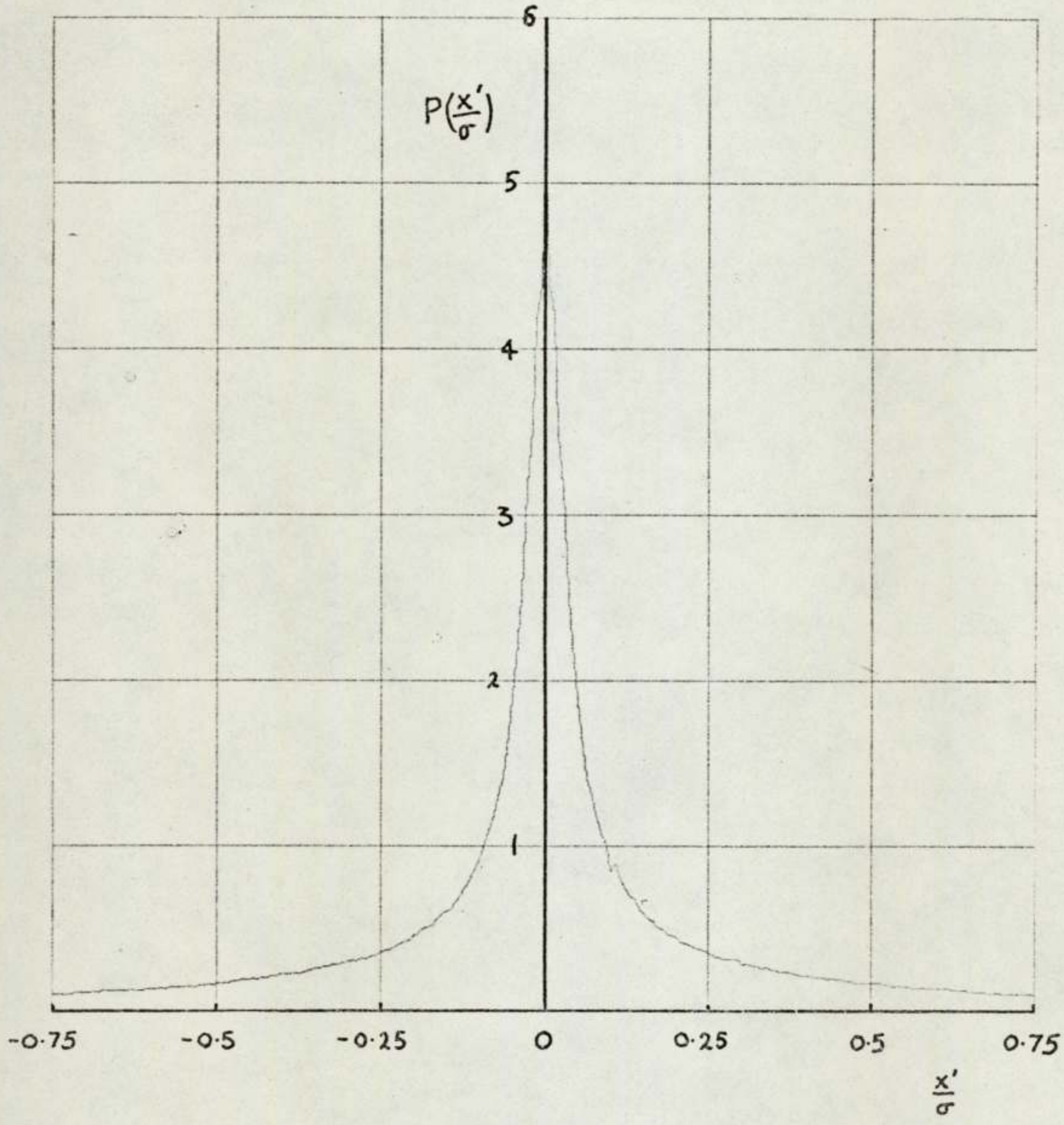


FIG. 3.10 Probabilty density function of the instantaneous slope of speech

Male voice (KGW)

High quality ribbon microphone

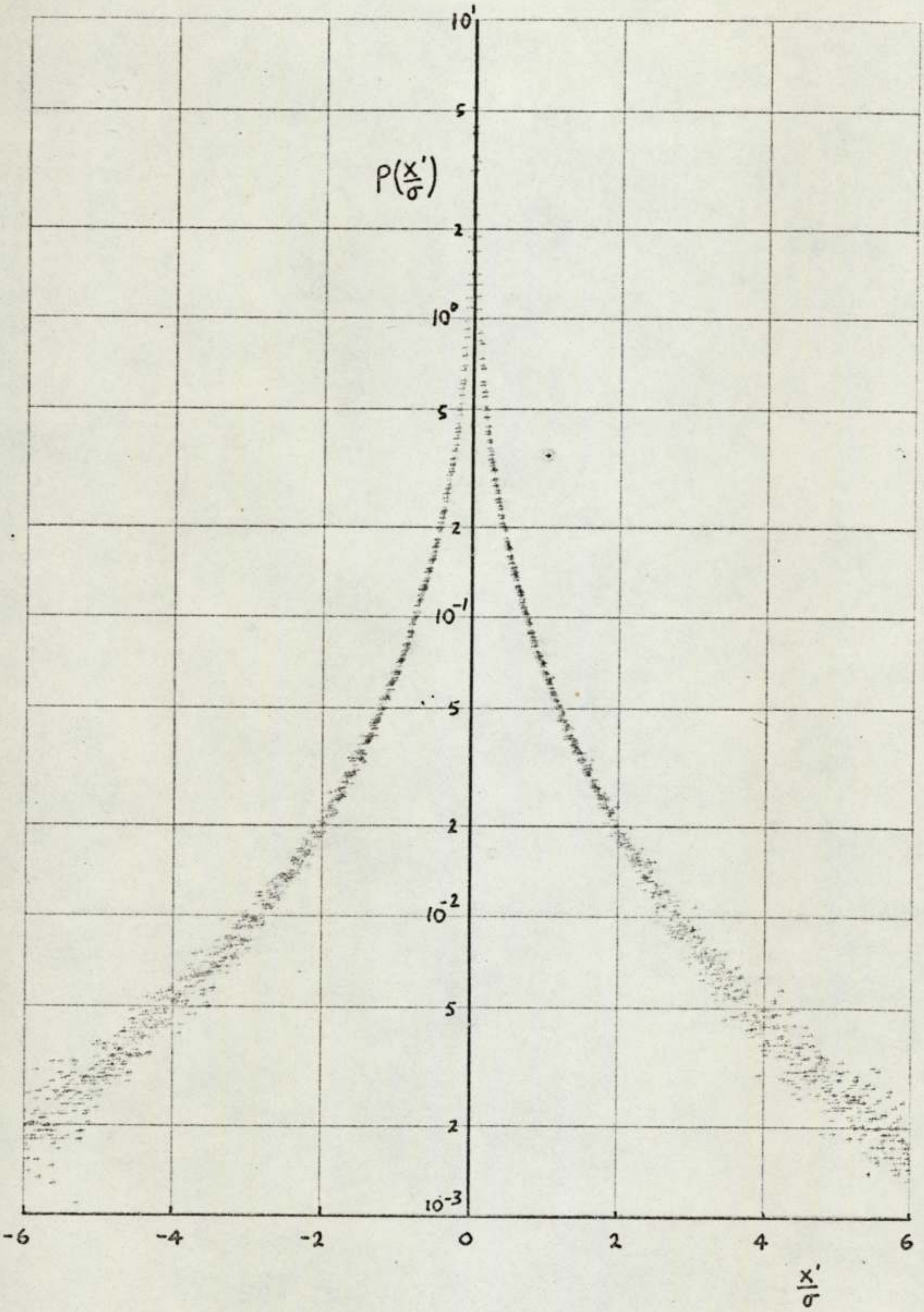


FIG. 3.11 Probability density function of the instantaneous slope of speech

Female voice (CMW)

High quality ribbon microphone

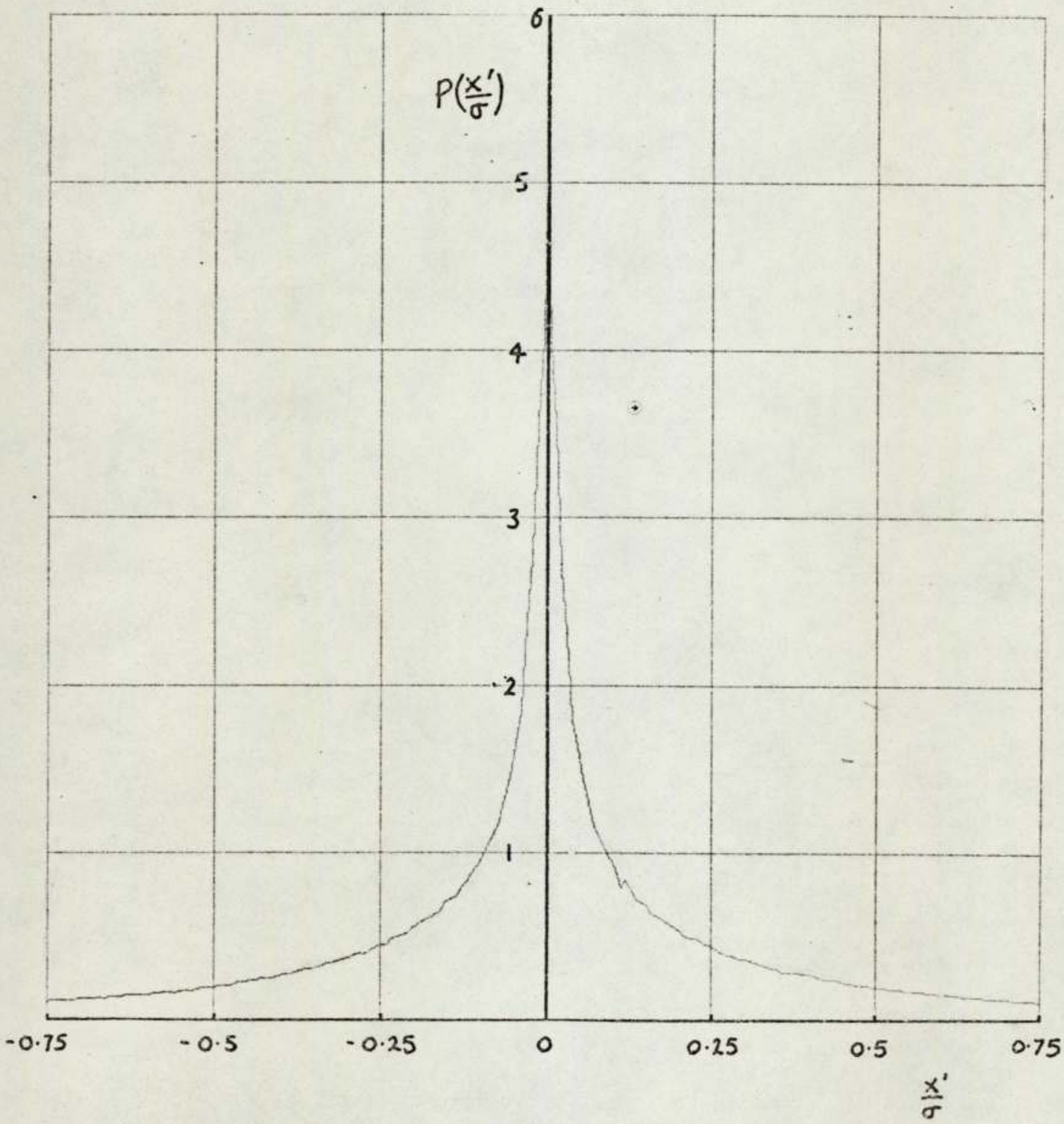


FIG. 3.12 Probability density function of the instantaneous slope of speech

Female voice (CMW)

High quality ribbon microphone

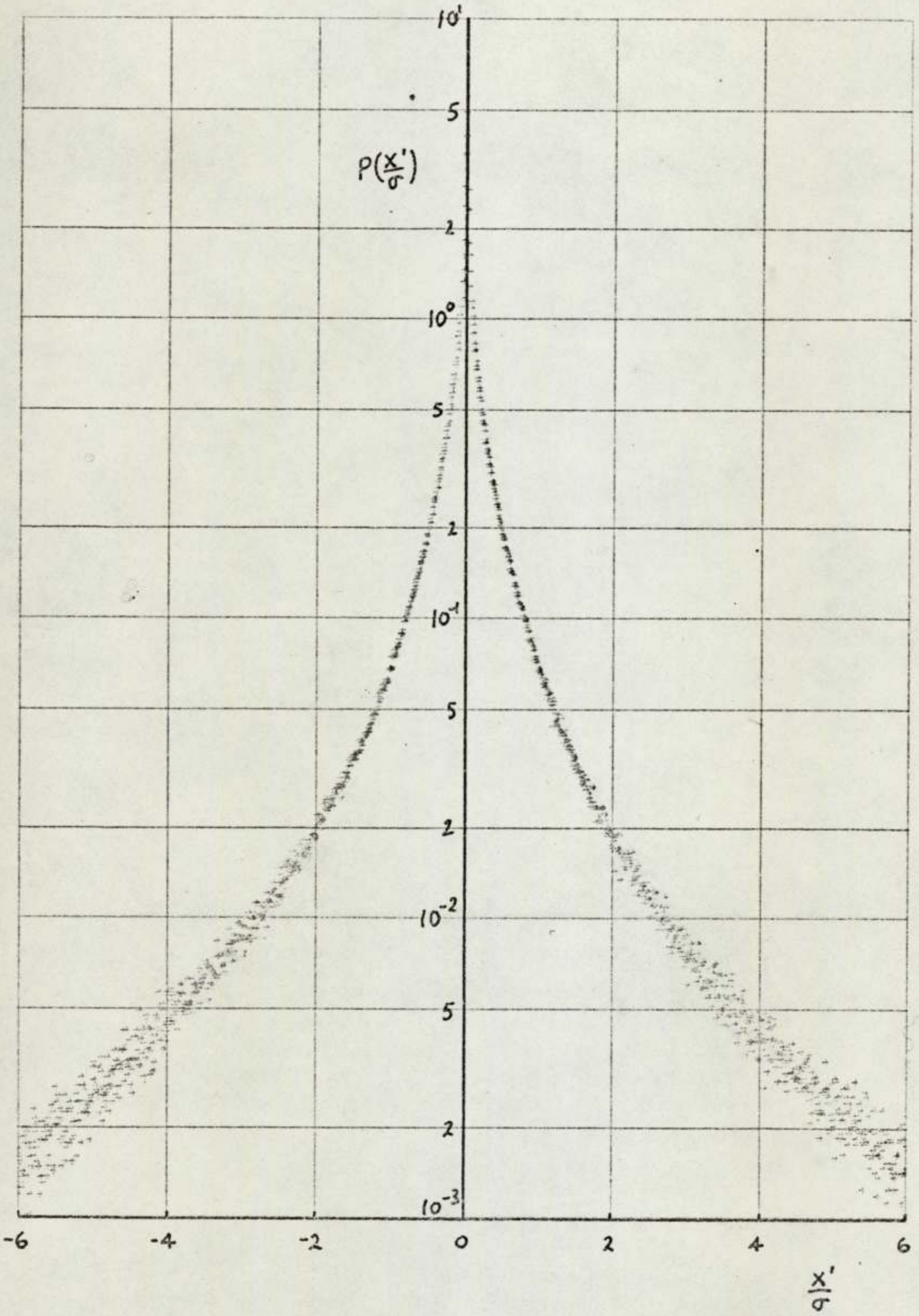


FIG. 3.13 Probability density function of the instantaneous slope of speech

Male voice

Telephone conversation

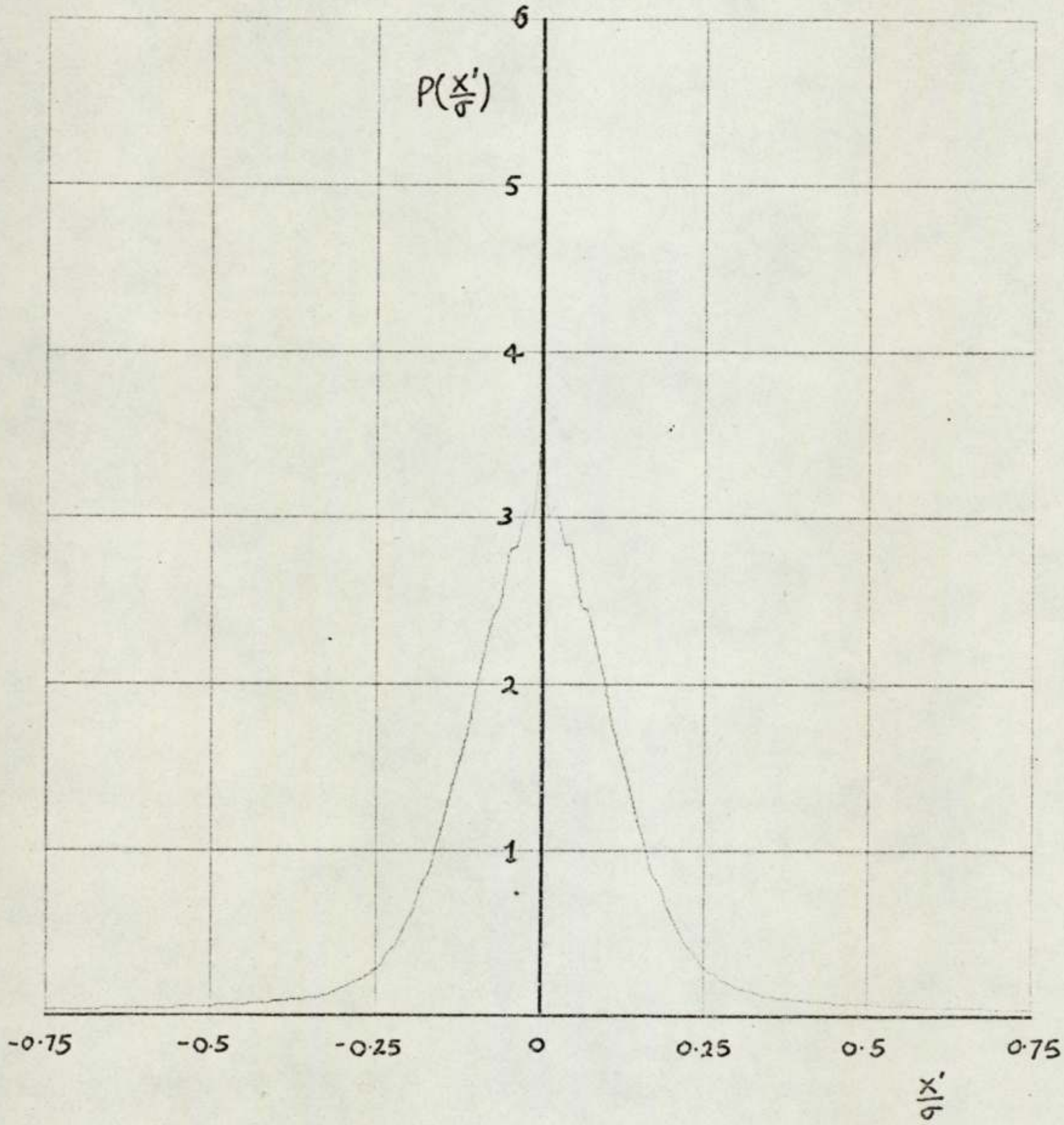


FIG. 3.14 Probability density function of the instantaneous slope of speech

Male voice

Telephone conversation

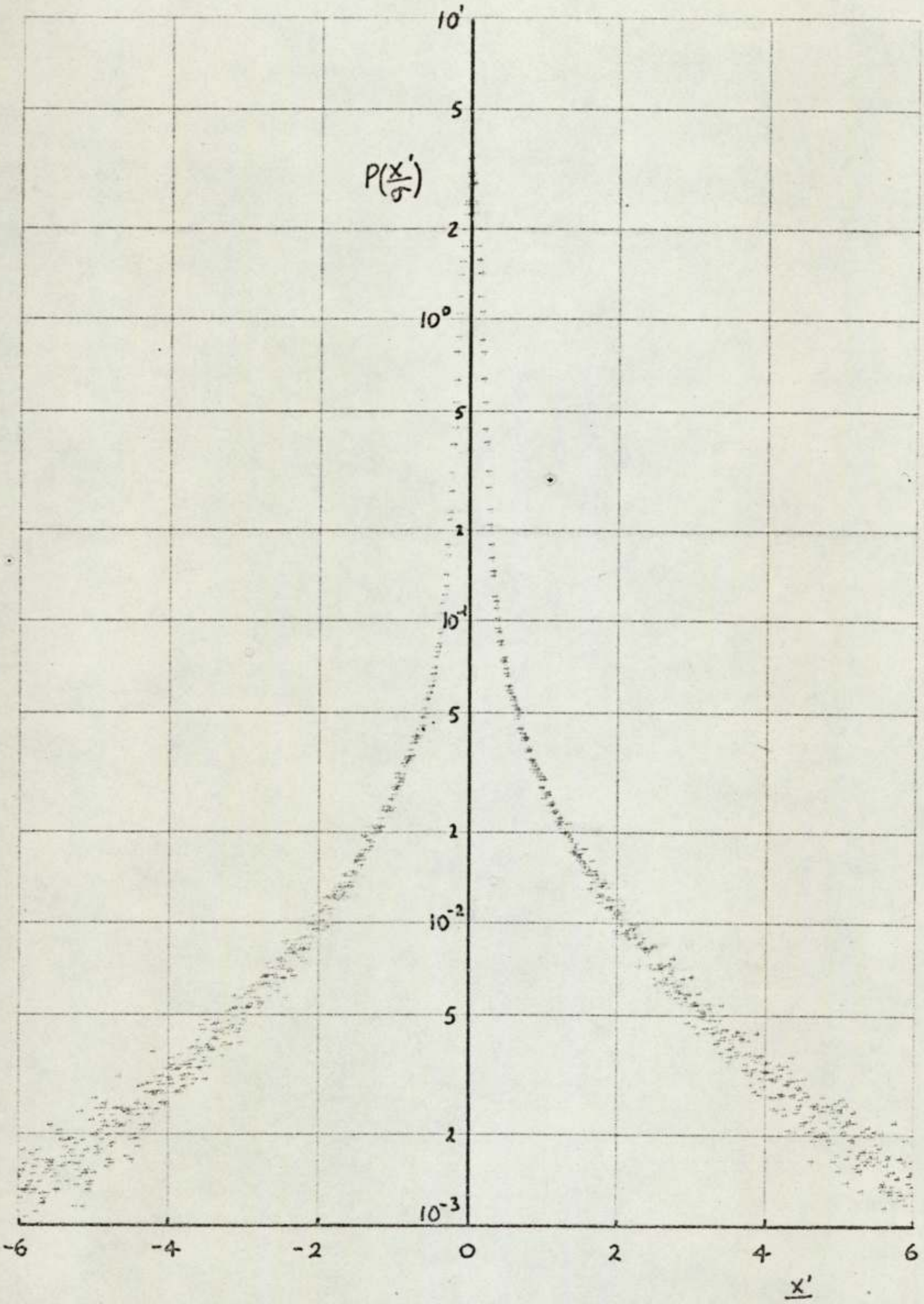


FIG. 3.15 Probability density function of the instantaneous slope of speech

Female voice

Telephone conversation

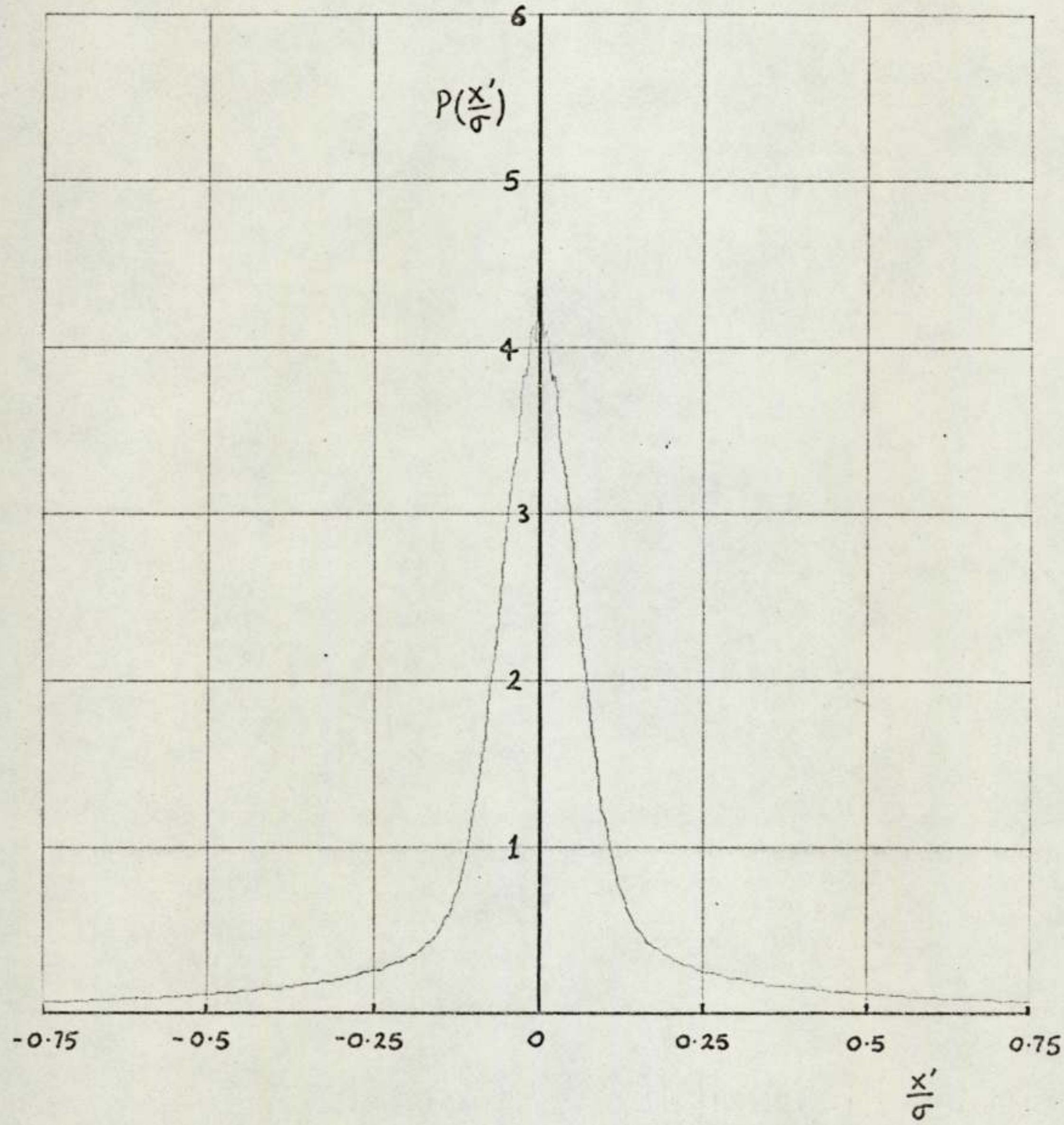


FIG. 3.16 Probability density function of the instantaneous slope of speech

Female voice

Telephone conversation

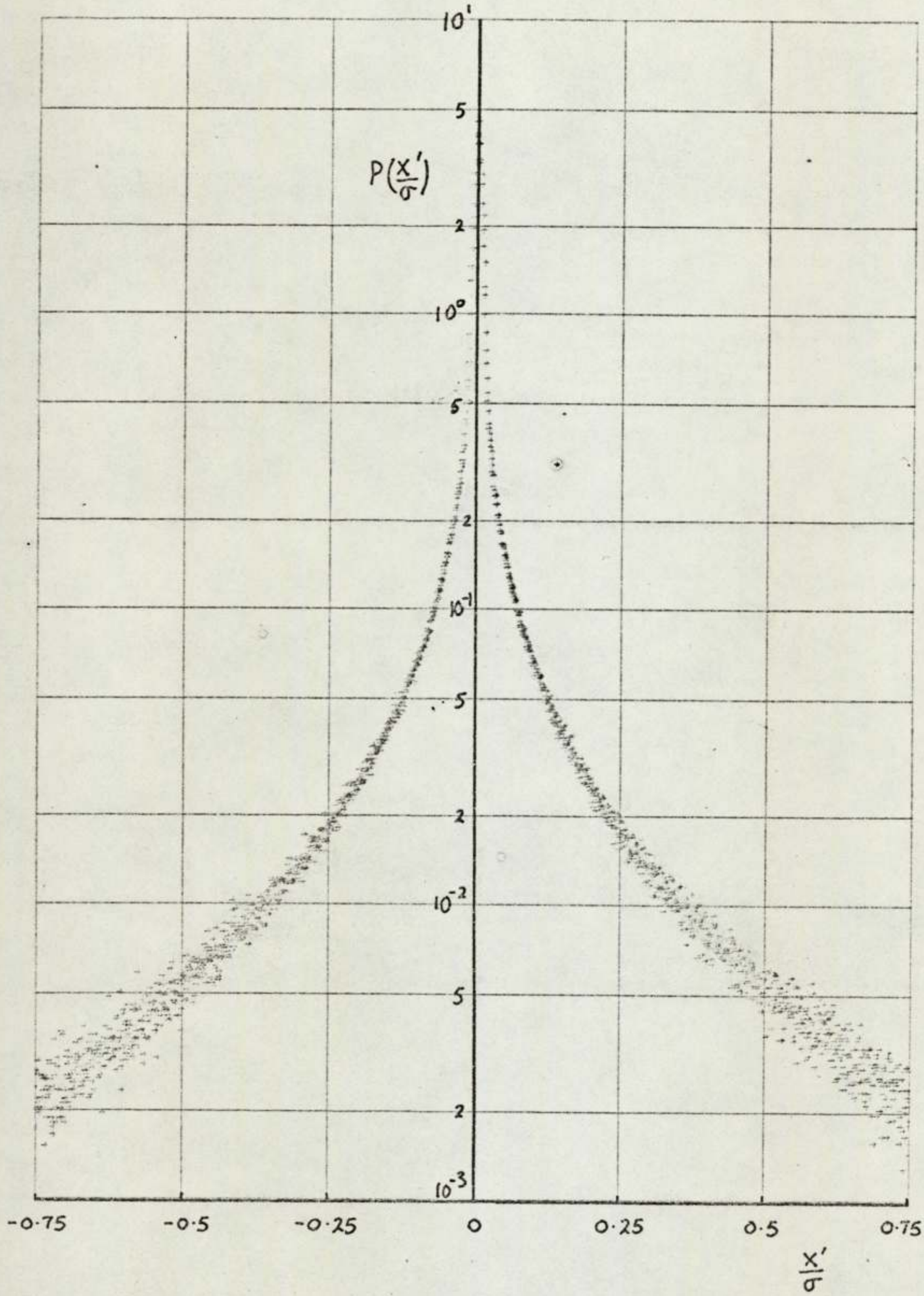


FIG. 3.17 Probability density function of the instantaneous slope of speech

In the case of the telephone conversations (FIGS. 3.14 to 3.17) the central peak is shorter and wider than for the other speech passages. This is due to noise during silent pauses, which also explains why the peak is approximately normally distributed.

The p.d.fs from the high-quality microphone are almost identical, despite the fact that the talkers are of different sex.

The single central peak, which is typical of all of the speech slope p.d.fs provides some useful information. It indicates that zero or small slopes are far more probable than large slopes. Thus, the most probable difference between successive quantized samples of a speech waveform is zero. This will be true for uniform quantization regardless of quantum size.

C H A P T E R 4

COMPANDING SYSTEMS

COMPANDING SYSTEMS

4.1 Introduction

In Chapter 3 the probability of zero difference P (d_o) was measured for constant-volume speech. Unfortunately, in practical telephone networks, speech volume varies from talker to talker. This volume variation is caused by variations in transmission line loss and in the loudness with which subscribers speak. The range of volumes is quite large; Purton [15] states that a junction PCM system should be capable of handling a volume range of 30dB.

Volume variations can be catered for in two ways :

a) Provide very fine-grain uniform quantization. The quantization must be fine enough to provide an acceptable signal-to-quantization-noise ratio for the weakest talker.

b) Compress the volume range at the sending terminal and re-expand it at the receiving terminal. (i.e. companding).

The first solution is unacceptable since an 11 or 12 bit encoding is required.

It is now standard practice in PCM terminals to use instantaneous companding. Speech samples are compressed and expanded on a sample to sample basis.

An alternative method is to compand at syllabic rate, although this is a more complex method. It was considered possible that syllabic companding might give a substantially larger P (d_o) compared with instantaneous companding. To test this hypothesis both forms of companding were investigated.

To avoid constructing equipment, the digital computer in the Electrical Engineering Department was again extensively used. Using an analogue-to-digital (A/D) converter in conjunction with the computer the following systems were simulated :

a) Sending terminal of a PCM system using 7-bit non-linear encoding (Instantaneous companding).

b) Sending terminal of a PCM system with linear encoding and syllabic companding.

The instantaneous-amplitude and first-difference probabilities of encoded speech samples from both simulations were measured.

4.2 Syllabic companding

A generalised single-channel syllabic companding system is illustrated in FIG.4.1. In the compressor, the speech-level analyser measures the mean level of the speech at the output of the variable attenuator VA1. This is usually accomplished by rectifying the signal and smoothing the resulting d.c. voltage with a short time constant RC filter. The amount of smoothing introduced is such that the control signal follows the variations of level at syllabic rate, but does not follow the actual speech waveform. The control signal at the output of the speech-level analyser controls the loss introduced by the variable attenuator VA1; the larger the control signal the larger the loss.

In the expander, the variable attenuator VA2 has the inverse property that loss decreases with increasing control signal. The speech-level analysers in both compressor and expander are identical. Also, since the input to the analysers is from the same information stream, it is evident that the control signals developed at both compressor and

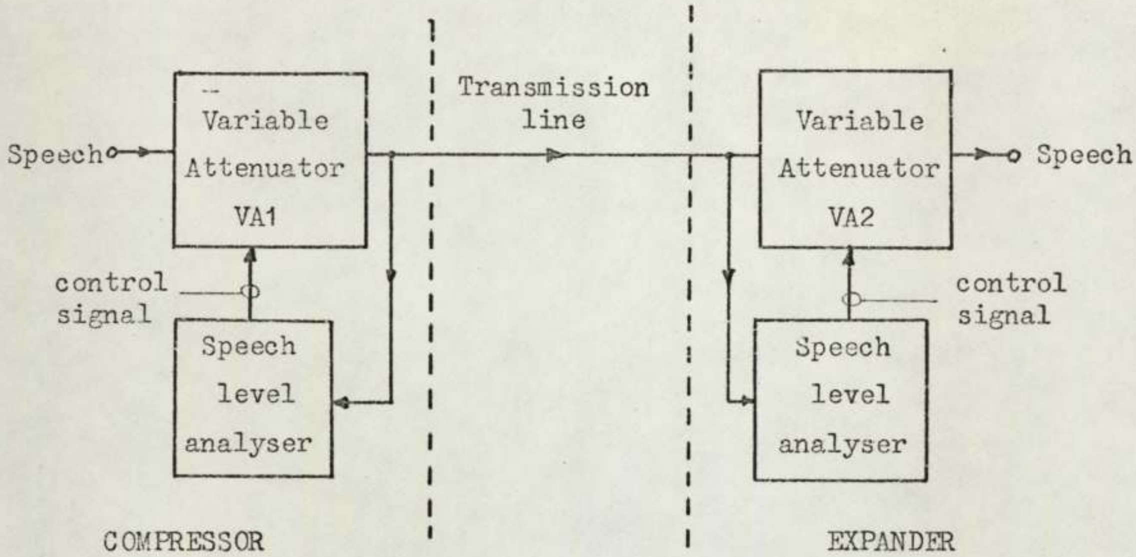


FIG. 4.1 A generalised syllabic companding system

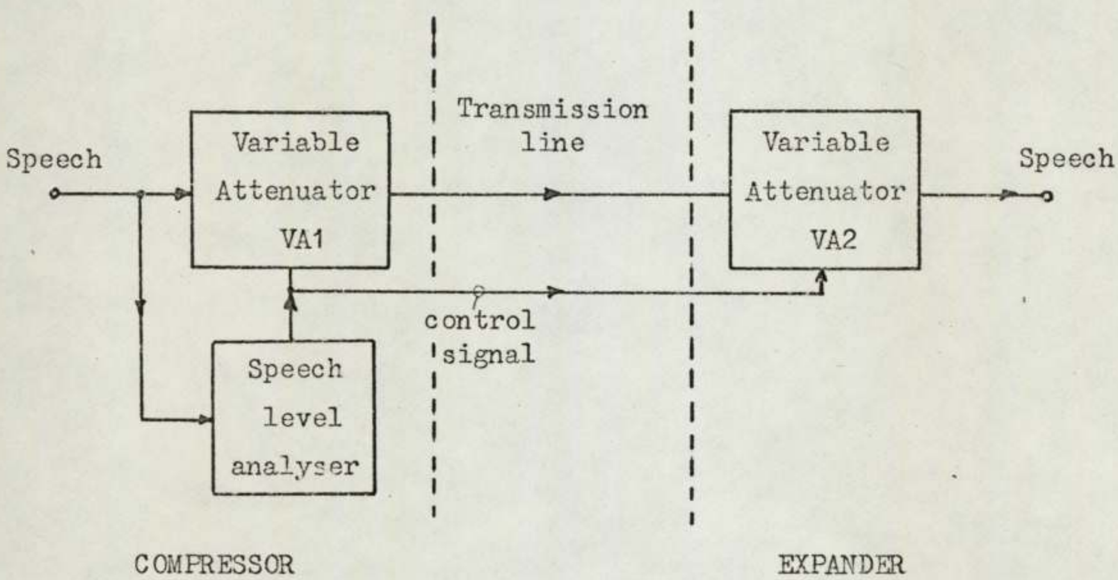


FIG. 4.2 An alternative syllabic companding system

expander are identical. Thus, provided the two variable attenuators have complementary characteristics, the signals at the input to the compressor and the output of the expander are identical.

The speech on the transmission-line is not constant volume but the volume range is smaller than at the input to the compressor. The amount of volume compression achieved depends upon the characteristics of the variable attenuators.

An alternative method of syllabic companding is illustrated in FIG.4.2. The variable attenuators and the speech level analyser are the same as in FIG. 4.1.

The basic difference is that the control signal is derived from the signal at the input to the compressor. This is in contrast to the system of FIG.4.1 where the control signal is derived from the output of the compressor.

The advantage of "forward feed" companding shown in FIG. 4.2 is that the speech at the output of the compressor is constant volume. However, this system has a drawback compared with the system of FIG. 4.1; the control signal for the expander cannot be derived from the received speech. Thus, it is necessary to send to the receiver additional information as to how much expansion is required. This type of companding is used in the Lincompex system [17].

The bandwidth required for the expansion information is not large, since the syllabic envelope does not vary rapidly. In a digital system, the control signal at the compressor could be sampled and encoded. The required digit rate would be small and each binary word could be transmitted over several frames by appending one bit to each signal sample.

The latter type of syllabic companding was chosen for investigation, despite its extra complexity. The constant-volume nature of the signal at the compressor output is ideal for linear encoding.

4.3 Simulation of a syllabic companding system

The syllabic companding system chosen for simulation is that illustrated in FIG. 4.2. The sending end of such a system was simulated with the aid of an A/D converter and digital computer.

The model on which the computer program was based is illustrated in FIG. 4.3. Comparing FIG.4.3 with FIG.4.2 the only additional items are the speech detector and the gate G1. The speech detector is included to reject noise during silent pauses in speech. When speech is detected, the gate G1 is opened and samples from the attenuator are allowed to pass. During silent pauses, the speech detector closes G1 and zero-valued samples are fed out.

The model operates as follows. The incoming bandlimited speech is sampled at 8KHz rate and encoded linearly in the 11-bit A/D converter. The digital speech samples are then attenuated, using digital techniques, in the digital attenuator before passing to G1. Volume variations in the incoming speech are detected by the mean-level measurer, the output of which controls the loss introduced by the digital attenuator.

The operation of the individual parts of the model are now described in more detail.

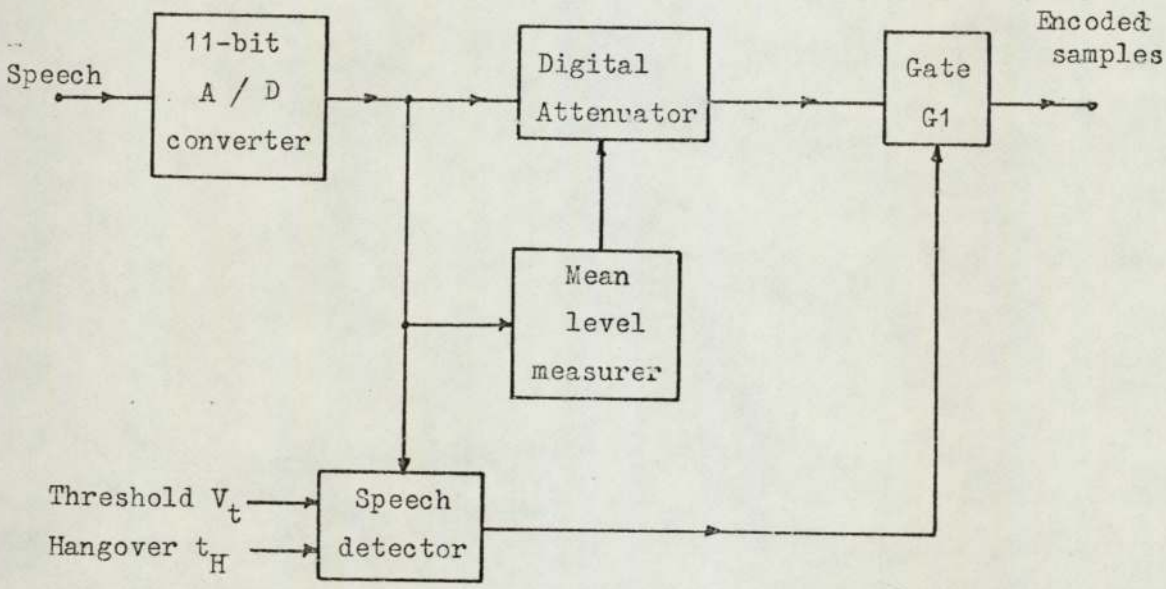


FIG. 4.3 Simulation of a syllabically companded PCM terminal

4.3.1 Mean-level measurer

A more detailed block diagram of the mean-level measurer is given in FIG. 4.4. The measurement of mean speech level is carried out on a sample-to-sample basis. During each sampling interval, the latest sample from the A/D converter is loaded, in parallel, into register A. Then the contents (X_C) of counter C are compared with the contents (X_A) of register A by the digital comparator B. The output of B is taken to the 'up/down' control of counter C. If $|X_A| > X_C$, the counter C is set to count up and the frame clock increments X_C. If $|X_A| \leqslant X_C$, then C is set to count down and X_C is decremented. Thus, the circuit action continuously attempts to force X_C to follow the speech signal. However, X_C cannot change rapidly since counter C receives only one count pulse per frame and X_C tends to follow the low-frequency syllabic envelope.

Although X_C is a measure of the volume of a speech signal X (t), it corresponds to the mean value of its magnitude $\overline{|X|}$, whereas the object of the compressor is to achieve constant mean power $\overline{X^2}$. The mean-square value can be related to the mean magnitude by considering the instantaneous amplitude statistics of the speech waveform. The probability density function of the speech waveform can be approximately represented by an exponential distribution[13].

i.e. : $p(x) = a e^{-ax}$

Where : $x \geqslant o$

For convenience, the distribution is assumed symmetrical about the axis x = o and only positive x are considered. The area under the curve is doubled to take account of the negative half of the distribution.

The rms value of the speech waveform is given by :

$$X_{rms} = \sqrt{\int_o^\infty x^2 \, ae^{-ax} \, dx}$$

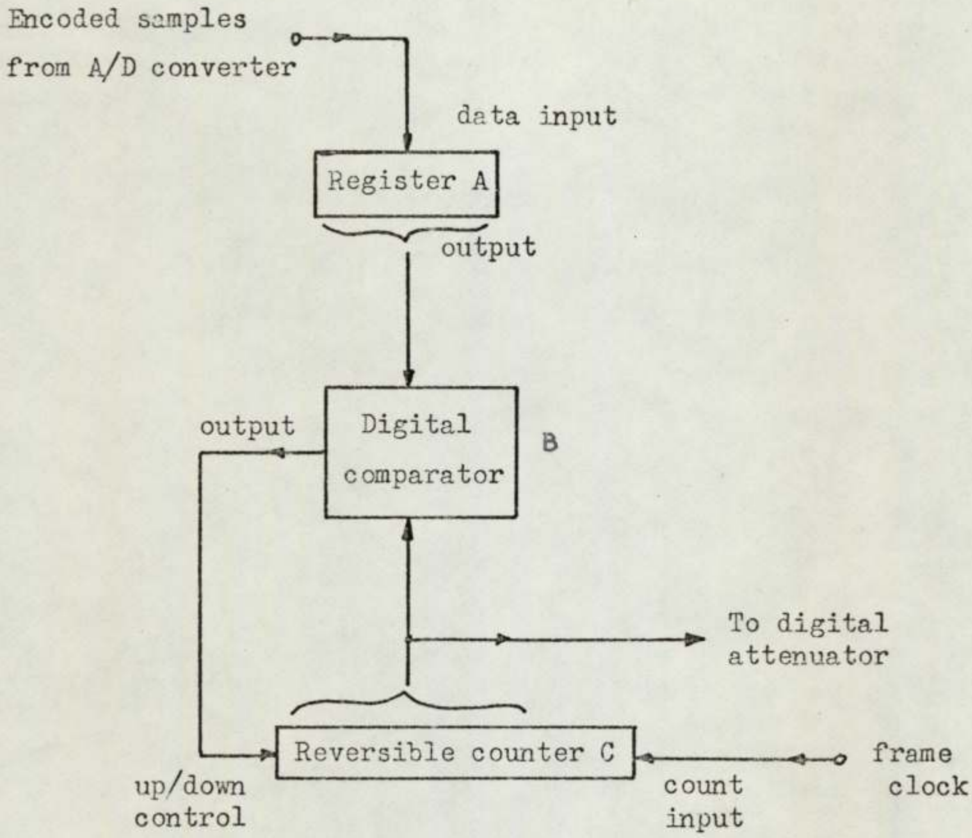


FIG. 4.4 Mean-level measurer

Then : $$X_{rms} = \frac{\sqrt{2}}{a} \qquad (4.1)$$

The number, X_C, in the mean-level measurer will tend to settle at a value which is exceeded by the signal for 50% of the time. In terms of the cumulative distribution of the instantaneous speech amplitude F (X), it is the level at which F (X) = 0.5.

Now F (X) is defined as :

$$F(X) = \int_o^X p(x)dx$$

Thus

$$F(X) = \int_o^X ae^{-ax}\,dx = 1 - e^{-aX}$$

Now $$F(X_C) = 0.5$$

$$\therefore \quad 0.5 = 1 - e^{-aX_C}$$

Then $$X_C = \frac{\log_e 2}{a} \qquad (4.2)$$

Thus, from equations (4.1) and (4.2) :

$$\frac{X_{rms}}{X_C} = \frac{\sqrt{2}}{a} \quad \frac{a}{\log_e 2} = 2.04$$

or $$X_{rms} = 2.04\, X_C \qquad (4.3)$$

4.3.2 Digital Attenuator

The digital attenuator shown in FIG.4.3 is illustrated in more detail in FIG.4.5. The digital attenuator uses a very simple principle. Consider each encoded signal sample as a binary number with its "binary point" to the right of the least significant bit. Attenuation is achieved by shifting the binary point an appropriate number of places to the left. Thus, for each left shift of the binary point the sample is attenuated by a factor of 2 (i.e. 6dB). If the point is shifted n places the attenuation factor is 2^n (6n dB).

Referring to FIG. 4.5, during each sampling interval the latest encoded sample from the A/D converter is loaded in the shift register SR. Then, pulses from the shift pulse generator shift the sample in SR an appropriate number of places to the right. This effectively shifts the binary point to the left. The number of pulses from the shift pulse generator, during each sampling interval, is governed by the number X_C in the mean-level measurer. The digital translator converts the number X_C into the required number of shift pulses.

The digital translator consists of an array of stored data. Each location in the array corresponds to one possible value of the number X_C in the level measurer. The contents of each location in the array is the required number of shift pulses to be applied to SR. The design of the data array was based upon the following factors :

a) Equation (4.3) i.e. $X_{rms} = 2.04\, X_C$

b) A signal-to-quantizing noise ratio of 28 dB.

i.e. $$\frac{X_{rms}^2}{N_q} = 28 \text{ dB}$$

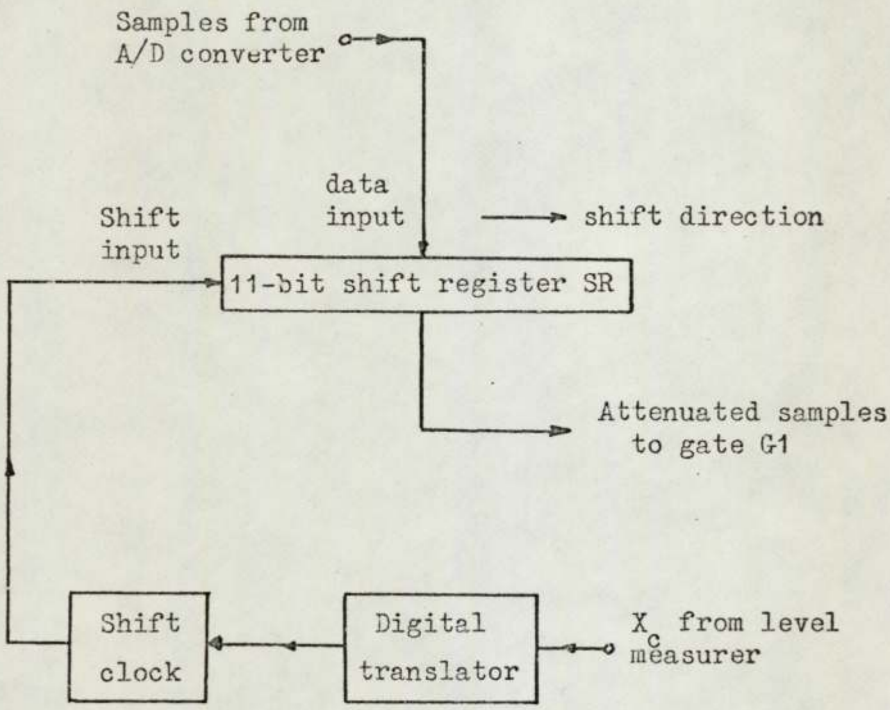


FIG. 4.5 Digital attenuator

Then, using equation (4.3) :

$$\frac{(2.04\ X_C)^2}{N_q} = 28 \text{ dB}$$

Thus : $$\frac{X_C{}^2}{N_q} \doteq 22 \text{ dB}$$

After the shift pulses have been applied, the number in SR is an attenuated version of the original sample.

4.3.3 Speech Detector

The speech detector shown in FIG. 4.3 is shown in more detail in FIG. 4.6. The modulus of each encoded sample from the A/D converter is compared with a threshold value V_t by the digital comparator. When a sample exceeds V_t, the output from the comparator causes the pulse generator to be triggered for a time t_H. When triggered, the pulse generator output causes gate G1 to open (see FIG. 4.3). If succeeding speech samples exceed V_t during the time t_H, the pulse generator is retriggered. Thus, the pulse generator will remain triggered continuously, provided samples greater than V_t occur at times less that t_H apart. This ensures that, during active speech, the speech detector does not release unnecessarily. The "hangover" time t_H was made 130mS[18] so that very short pauses in speech did not cause the speech detector to release.

The speech detector is very simple in design to enable it to be easily incorporated in the computer program.

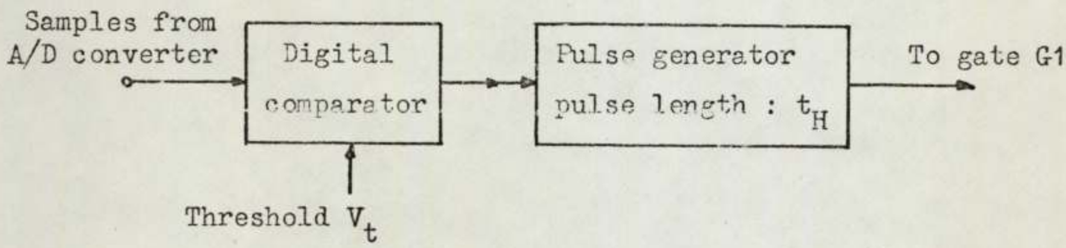


FIG. 4.6 Speech detector

4.4 Simulation of an instantaneous companding system

The system simulated was the sending terminal of a PCM system with 7-bit encoding and instantaneous companding. A block diagram of the model used in the simulation is given in FIG. 4.7. The speech is sampled at $8KH_z$ and linearly encoded in the 11-bit A/D converter. The 11-bit samples are then converted into 7-bit samples by the digital translator. The positive half of the non-linear transfer function of the digital translator is shown in FIG. 4.8. The characteristic is segmented; each segment forms a chord to a curve described by the following equations :

$$y = \frac{Ax}{1 + \log A} \qquad 0 \leqslant x \leqslant \frac{1}{A}$$

$$y = \frac{1 + \log Ax}{1 + \log A} \qquad \frac{1}{A} \leqslant x \leqslant 1$$

Where :

y is the normalised output of the translator

x is the normalised input to the translator

$A \doteq 87$

This is known as an A-law companding characteristic[19].

The translation of the 11-bit sample to the required 7-bit sample is executed by using an array of stored data. Each location of the array corresponds to one possible value of the 11-bit number. The contents of each location is the required 7-bit number, as defined by the characteristic given in FIG.4.8. For example, the content of location 512 is 56.

The speech detector shown in FIG.4.7 is identical to the one described in Section 4.3.3.

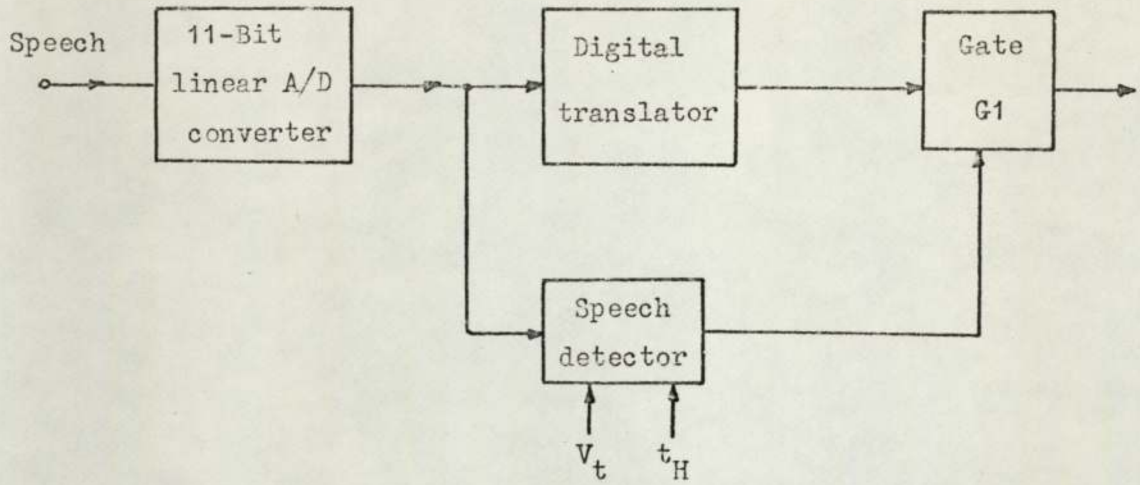


FIG. 4.7 Simulation of an instantaneous companding system

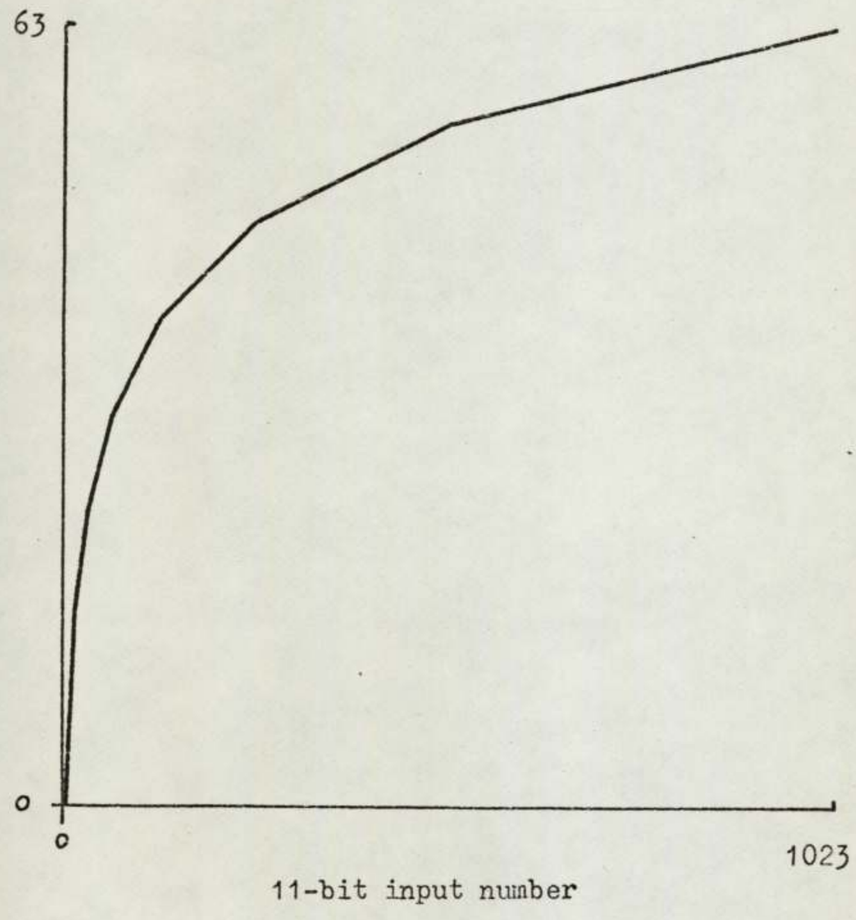


FIG. 4.8 Transfer function of the digital translator

4.5 Computer programs

The basic structure of the computer programs used in both simulations are similar. The main program controls input and output of data, generates arrays of stored data, and calculates probabilities from accumulated data. A subroutine within the main program controls the sampling interval and performs the following functions :

1) Functions common to both simulations :
   a) Accumulation of data for probability calculations
   b) Speech detection
   c) Digital translation

2) Functions for syllabic companding only :
   a) Digital attenuation
   b) Mean-level measurement

The technique used for measuring probabilities is identical to that described in Chapter 3. A copy of the programs used in both simulations is given in Appendix A3.

4.6 Experimental apparatus and procedure

The experimental equipment used in both experiments is shown in FIG. 4.9. The equipment is identical to that used in Chapter 3 and a full description of it is given in section 3.2. The speech passages analysed were the telephone conversations used in the investigation described in Chapter 3. The volume of each talker was set to 26dB below the peak quantizer level, using the calibrated attenuator and the calibration tone on the tape. This volume corresponds to that of a 'median' talker in a practical telephone network[15].

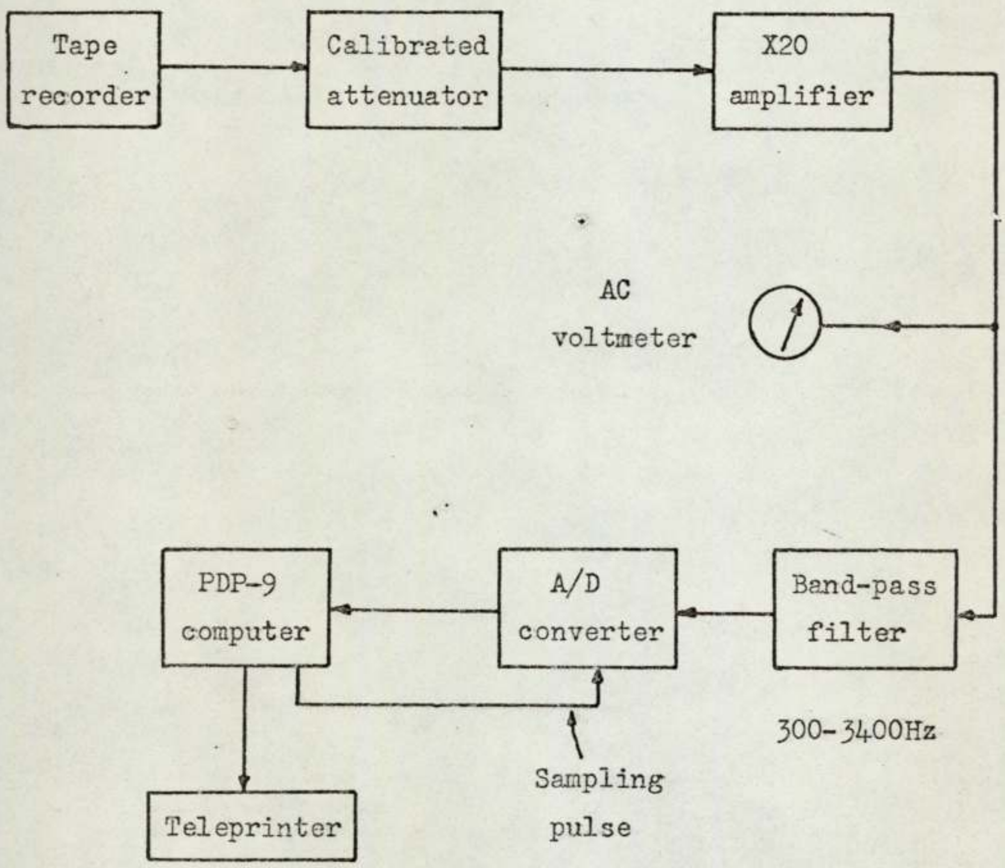


FIG. 4.9 Experimental apparatus

4.7 Results

The computer programs, for both simulations, provided the following output data :

a) Speech activity factor τ_A

b) Probability of a difference of i quanta between adjacent speech samples $P(d_i)$, where $0 \leq i \leq 128$

c) Probability of a sample assuming a value of j quanta $P(a_j)$, where $0 \leq j \leq 63$

The values obtained for τ_A, $P(d_o)$ and C are given in TABLE 4.1. The constant C is defined by the equation :

$$C = \frac{P(d_o) - (1 - \tau_A)}{\tau_A} \qquad \text{(see section 3.2.4)}$$

TABLE 4.1

SYSTEM	τ_A	$P(d_o)$	C
Instantaneous companding	0.481	0.554	0.073
Syllabic companding	0.488	0.553	0.084
Linear encoding	0.455	0.673	0.285

Effect of companding

Also included in TABLE 4.1 for comparison are the measured values of $P(d_o)$ and C, for linear encoding. From TABLE 4.1 it can be seen that companding causes a dramatic decrease in the factor C compared with linear encoding. This means that far fewer zero-differences occur in detected speech which has been compressed. This, of course, leads to a substantial reduction in $P(d_o)$ which is evident from TABLE 4.1.

The performance of the two companding systems, in terms of $P(d_o)$, are almost identical. Thus, instantaneous companding is to be preferred, because of its relative simplicity and cheapness.

The probability of zero-amplitude $P(a_o)$ is presented in TABLE 4.2, with $P(d_o)$ for comparison, for both companding systems.

TABLE 4.2

Companding system	$P(a_o)$	$P(d_o)$
Instantaneous	0.548	0.554
Syllabic	0.537	0.553

Comparison of $P(a_o)$ with $P(d_o)$

From TABLE 4.2 it is evident that $P(a_o)$ is only slightly smaller than $P(d_o)$ for both types of companding. Thus, zero-valued samples provide an alternative form of redundancy, whose elimination is almost as effective for reducing the digit-rate of the proposed transmission system.

C H A P T E R 5

FRAME OVERFLOW

FRAME OVERFLOW

5.1 Introduction

The proposed system achieves a bandwidth economy by having fewer timeslots per frame than channels. Thus, it is possible for the number of active channels (those channels requiring a sample to be sent) to exceed the available timeslots. This situation is known as frame overflow, and is undesirable since channel samples are lost and cannot be retrieved. The effects of frame overflow can be minimised to any required degree by limiting the number of channels. Thus, the upper limit on the number of channels is determined by the maximum error-rate that can be tolerated.

The relationship between frame-overflow error-rate and the number of channels (N), the number of timeslots (K) and the probability of zero difference ($P\ (d_o)$) is derived in this chapter.

If timeslots are assigned to channels sequentially on a fixed priority basis, then the first K channels will never experience frame overflow. However, channels (K+1) to N will experience increasing amounts of information loss with increasing channel number. This is undesirable, since satisfactory voice quality in channels towards the end of the frame would demand an uneconomically small number of channels. Clearly, overflow errors must be distributed so that all channels experience equal amounts of errors.

Various methods of distributing frame-overflow errors are described in the second part of this chapter, and their merits investigated.

5.2 Fraction of information lost due to frame overflow

The proposed system is designed to accommodate N channels in K timeslots, where each channel has a probability α (α = 1-P (d_o)) of needing a timeslot. This situation can be modelled on a binomial distribution. The probability of frame overflow as a function of N is given by :

$$P_{over}(N) = \sum_{i=K+1}^{N} \frac{N!}{i!(N-i)!} \alpha^{i} (1-\alpha)^{N-i} \qquad (5.1)$$

To calculate the fraction of samples lost, the size of the frame overflow must also be taken into account. Let the number of active channels in any frame be y. Then, when y exceeds K a frame overflow has occurred and the number of samples lost is y-K. The average number of samples lost per frame, L, is a weighted sum of all possible y-K, where y exceeds K.

$$\text{i.e.} \qquad L = \sum_{y=K+1}^{N} w_y(y-K)$$

Where the weight w_y is simply the probability of y channels being active. Hence :

$$L = \sum_{y=K+1}^{N} P_N(y).(y-K) \qquad (5.2)$$

where :

$$P_N(y) = \frac{N!}{y!(N-y)!} . \alpha^{y}.(1-\alpha)^{N-y} \qquad (5.3)$$

Now, L is the average of the total samples lost by all channels per frame. What is wanted is the fraction of samples lost by each channel (let this be Φ). Assuming overflow errors are distributed evenly, it is tempting to define Φ as simply $\frac{L}{N}$. This is correct in as much as it expresses the samples lost as a fraction of the total samples from that channel. However, this definition is misleading since it ignores the fact that samples are lost only during speech. Samples cannot be lost during silences, since no samples are transmitted. Thus, a more meaningful definition of Φ is :

$$\Phi = \frac{L}{N} \cdot \frac{1}{\tau_A} \qquad (5.4)$$

Scuilli and Campanella[12] make a similar calculation for the SPEC system, using a different approach. They calculate the probability of overflow in channels (K+1) to N and then sum these probabilities to get the average number of samples lost per frame. This derivation gives a formula which is equivalent to equation (5.2). Then, they define Φ as $\frac{L}{N}$ (rather than $\frac{L}{N\tau_A}$) which, as demonstrated above, gives a misleading and optimistic bias.

The loss of samples due to frame overflow is a random process, and Φ can be considered as the probability of frame overflow in one channel.

The proposed system requires a certain minimum number of channels to function efficiently. For example, if there were only two channels and these were multiplexed into one timeslot, the mutual interference would be unacceptable. FIG. 5.1 shows Φ plotted as a function of N with the ratio $\frac{K}{N}$ constant (from equation 5.4). The figure shows clearly the rapid increase in overflow error as N becomes small. Also, it can be seen that the overflow errors tend to flatten off for N greater than about 20. Hence, little advantage is to be gained by making N much larger than 20.

Consider now a practical situation. Suppose the proposed system is to be applied to a digital transmission system which has 25 timeslots available for channel samples. The number of channels that can be accommodated depends upon the overflow error-rate, Φ, which is tolerable. FIG. 5.2 shows Φ plotted as a function of N for K = 25 and α = 0.375. This is the predicted value of α for a practical telephone network.

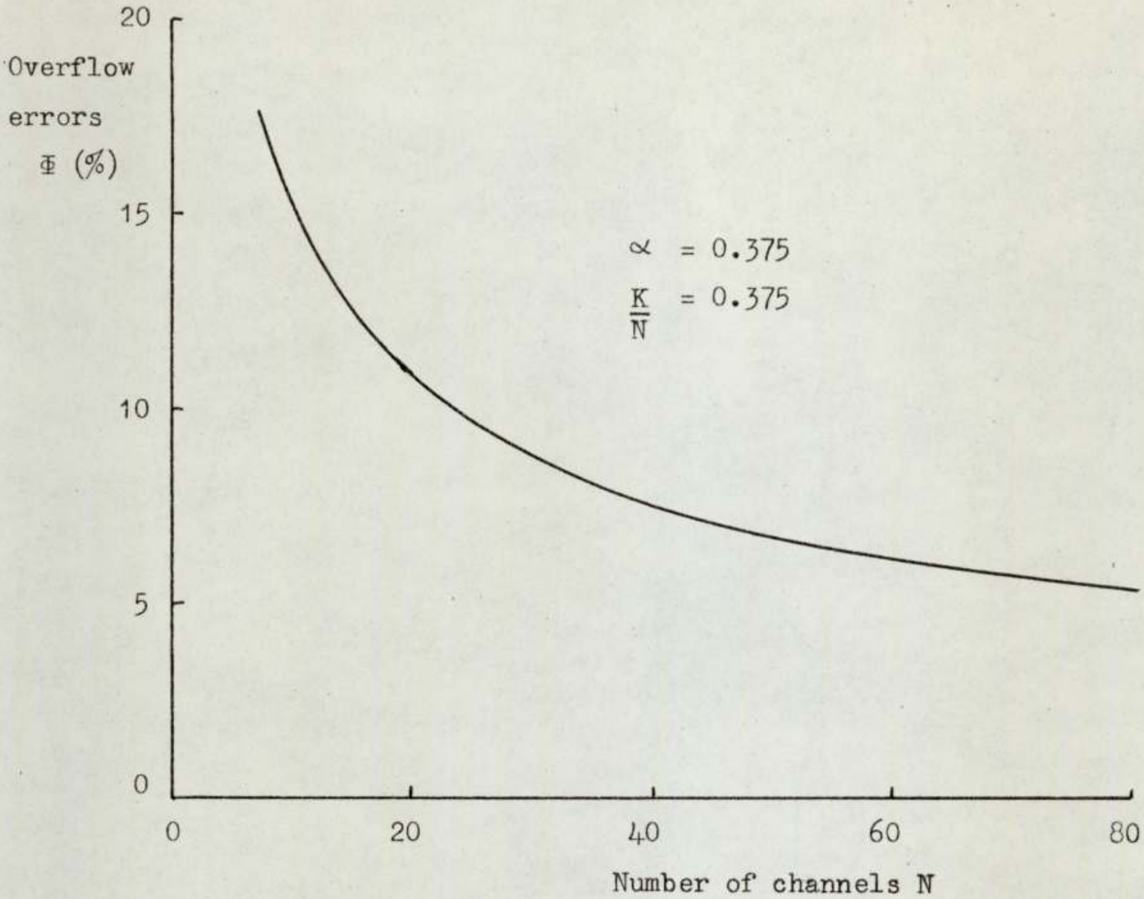


FIG. 5.1 Overflow errors with K/N constant

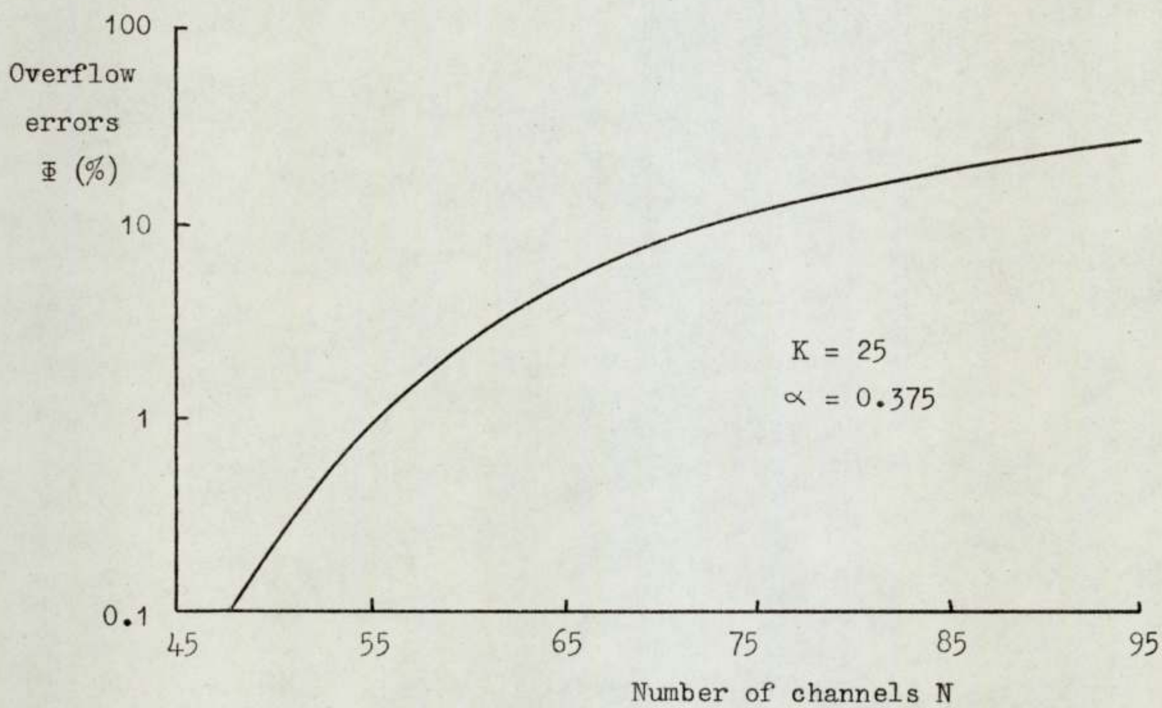


FIG. 5.2 Overflow errors with K constant

It was calculated from equation (3.9) (i.e. $P(d_o) = 0.6 + 0.4C$), using the value of C obtained in section 4.7 for instantaneous companding. A small safety margin was included in the calculation to allow for experimental error.

5.3 Distribution of frame overflow errors

The importance of distributing frame-overflow errors was mentioned in section 5.1. Three methods of implementing frame-overflow error distribution are described in this section.

5.3.1 Rota system

As the name suggests, this method causes overflow errors to be rotated throughout all channels. During each frame, channels are examined sequentially to detect whether they are active or not. Active channels are assigned a timeslot. The overflow errors are distributed, by starting the assignment sequence from a different channel number in each frame. There are two methods of implementing this idea. Sciulli and Campanella [12] have suggested that the starting channel number for the sequence of timeslot assignments is incremented frame by frame. This method is illustrated in FIG. 5.3. If channels are examined in the order 1,2,3, ... N in the i^{th} frame, then the order in the (i+1)th frame is 2,3,4, ... N, 1 and so on. The cycle is complete after N frames and is then continuously repeated. This method of timeslot assignment has a side effect. It tends to force overflow errors to appear at multiples of the assignment cycle period. This is because errors tend to occur as the channel nears the end of the assignment queue, where the probability of overflow is at its maximum. If the average error frequency is approximately the same as the assignment cycle

Frame i | 1 | 2 | 3 | .... | N |

Frame i+1 | 2 | 3 | 4 | .... N | 1 |

Frame i+2 | 3 | 4 | 5 | .... N | 1 | 2 |

⋮

Frame i+N-1 | N | 1 | 2 | .... | N-1 |

FIG. 5.3 Order of channel assignments to timeslots frequency the errors will even tend to occur periodically. This may make overflow errors more noticeable to the listener than if they occurred in a purely random fashion.

An alternative method is to start timeslot assignments at the channel after the last that obtained assignment in the previous frame. As an example, suppose channel C was the last channel to be assigned a timeslot in frame i. In frame (i+1), the timeslot assignment sequence would begin at channel (C+1). This method has the advantage that a channel which loses a sample in one frame is near the front of the assignment sequence in the next frame. Thus, the possibility of a channel losing two consecutive samples is extremely remote.

When a channel experiences frame overflow, no sample is transmitted. The receiving terminal assumes that, since no sample has been sent, a zero difference has occurred and it reads out the previous stored sample value. This is fortunate, since the most probable difference is zero. Hence, replacing the missing sample by the previous one is the best estimate that can be made with the available data.

It was mentioned in section 2.3 that as an alternative to ZOP, zero-amplitude prediction (ZAP) could be used to remove redundancy. If this idea is used in conjunction with the ROTA system, then frame overflow causes the affected samples to be replaced with zero-valued samples at the receiver.

The SPEC system, investigated by Sciulli and Campanella[12], uses the ROTA method of overflow-error distribution.

5.3.2 Error grouping system

When a frame overflow occurs in this system, the active channels are divided into two groups, as in a TASI system. These groups are :

a) Priority channels (which are guaranteed access to a timeslot)

b) Secondary channels (which queue for access to timeslots)

A channel which becomes speech active (speech-detector operated), initially joins the back of the secondary queue (if there is one). The secondary channel which has been speech active for the longest time is at the head of the queue. When the speech detector of a priority channel releases, this channel loses its status of priority. The first channel in the secondary queue then fills the vacant gap and becomes a priority channel.

A priority channel is guaranteed access to a timeslot if required. Clearly, if there are K timeslots there are a maximum of K priority channels. A priority channel does not always require a timeslot, since zero-differences do occur in active speech. These vacant timeslots are assigned to secondary channels in queue order.

Overflow errors occur only when a channel is in the secondary queue. Thus, all errors are grouped at the beginning of talkspurts.

This system operates in a similar way to TASI, except that the affected part of the talkspurt is not completely "frozen out" as happens in TASI.

5.3.3 Variable word length system (VWL) system

In the VWL system[7], frame overflow does not cause samples to be removed. Instead, the number of bits per sample is reduced until all samples fit into the frame of transmitted pulses. When the frame is not overflowing, the samples are encoded with the full number of bits (say 8), as in ordinary PCM.

It may appear at first sight that additional information must be transmitted indicating the sample word length b. This is not necessary, however, if pre-addressing is used. The receiving terminal only requires to know the number of active channels to compute b and this is simply the number of binary '1's in the sample-address word.

At the receiving terminal, the missing bits are replaced in a way that minimises the average error. The procedure is as follows :

a) The most significant missing bit is replaced with a '1'

b) All remaining missing bits are replaced with a '0'

If only one bit is missing a '1' is substituted. This method approximately halves the error compared with replacing all the missing bits with '0's. Alternatively, the most significant missing bit can be replaced with a '0' and all remaining missing bits with a '1'.

5.4 Comparison of overflow-error-distribution systems

The noise or distortion introduced by the various error-distribution systems produce different subjective effects. Overflow errors in the ZAP ROTA system causes complete samples to be removed, and this produces a "clicking" noise. The ZOP ROTA system replaces each missing sample with the previous sample. This also produces a clicking noise, but is not as noticeable as in the ZAP ROTA system for the same error-rate. The ERROR-GROUPING system causes the beginning of talkspurts to be distorted because of missing samples. This makes the speech sound a little abrupt, rather like TASI. Under overflow conditions, the VWL system causes the word length of samples to be reduced, which results in a increase in quantizing noise.

An important feature of overflow noise is its signal dependence. Overflow errors do not introduce idle-channel noise, since overflow errors do not occur under no-signal conditions. Also, when using the ROTA system, the amplitude of the error introduced by frame overflow is proportional to the amplitude of the signal.

The effects of frame overflow in the ROTA and VWL systems can be thought of as the addition of a continuous type of noise. Thus, a signal-to-noise ratio can be calculated for both of these systems. This idea cannot be applied, with any meaning, to the ERROR-GROUPING system since frame overflow does not produce a continuous noise.

5.4.1 Signal-to-noise ratios

a) ZAP ROTA system

Frame overflow in the ZAP ROTA system causes complete samples to be removed. This can be considered as the addition of noise impulses, equal in amplitude to the missing samples but opposite in sign. Thus, by calculating the power in this noise waveform it is possible to determine a signal-to-noise ratio.

The noise is a series of impulses, spaced randomly in time and with an amplitude distribution the same as the speech signal. Although the impulses are random, the time between impulses is some multiple of the sampling time τ_s. The probability of a noise impulse occurring at any sampling instant is Φ, the overflow-error rate.

The noise impulses can be considered as the product of the speech signal and a series of random impulses of unit area :

$$n(t) = X(t)I(t)$$

where -

$n(t)$ is the noise impulses due to frame overflow

$X(t)$ is the analogue voice signal

$I(t)$ is a random impulse waveform, each impulse having unit area

FIG. 5.4 illustrates typical $n(t)$, $X(t)$ and $I(t)$ waveforms.

The autocorrelation function of $I(t)$ is :

$$R_I(\tau) = \overline{I(t)I(t+\tau)}$$

Let each impulse be defined by the limiting process :

$$\delta(t) = A\,d\tau \Big|_{\substack{A \to \infty \\ A d\tau = 1}}$$

where A = the height of the impulse

$d\tau$ = the width of the impulse.

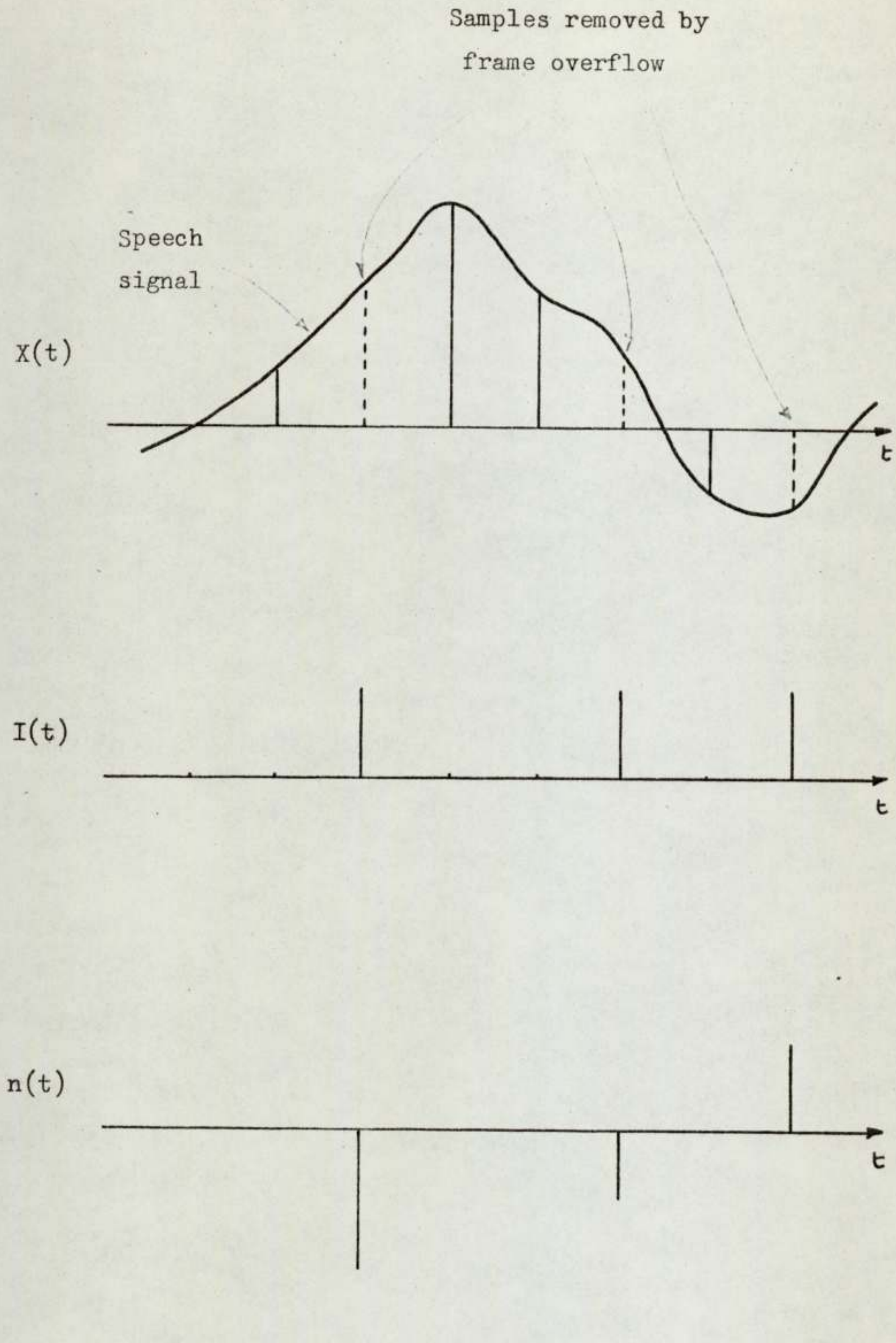


FIG. 5.4 Relationship between X(t), I(t) and n(t)

Now,

$$R_I(0) \;=\; \overline{I(t)^2} \;=\; \text{Mean square value of one impulse over } \tau_s\text{, multiplied by the probability of finding an impulse in any sampling interval.}$$

$$= \Phi \; \frac{A^2\, d\tau}{\tau_s} \Bigg|_{\substack{A \to \infty \\ Ad\tau = 1}} \;=\; \frac{\Phi}{\tau_s} \; A \Bigg|_{A \to \infty}$$

$$\therefore \quad R_I(0) \;=\; \frac{\Phi}{\tau_s}\, \delta(\tau)$$

Also, $R_I(n\tau_s) \;=\; \overline{I(t)I(t+n\tau_s)} \;=\; \Phi^2 \; \dfrac{A^2 d\tau}{\tau_s} \Bigg|_{\substack{A \to \infty \\ Ad\tau = 1}}$

$$R_I(n\tau_s) \;=\; \frac{\Phi^2}{\tau_s}\; \delta(\tau - n\tau_s) \qquad n \;=\; \pm 1,\ \pm 2, \ldots$$

Summarising the above data :

$$\left.\begin{aligned} R_I(\tau) \;&=\; \frac{\Phi}{\tau_s}\, \delta(\tau) \qquad \tau = 0 \\ &=\; \frac{\Phi^2}{\tau_s}\, \delta(\tau - n\tau_s) \quad n = \pm 1,\ \pm 2, \ldots. \\ &=\; 0 \quad \text{for all other } \tau \end{aligned}\right\} \qquad (5.5)$$

$R_I(\tau)$ is illustrated in FIG. 5.5

$R_I(\tau)$ can be divided into two components which are statistically independent.

$$\text{i.e. } R_I(\tau) \;=\; R_P(\tau) \;+\; R_R(\tau)$$

The two components are shown in FIG. 5.6.

$R_R(\tau)$ is the autocorrelation function of a random component with a "white" spectral density. $R_P(\tau)$ is the auto correlation function of a

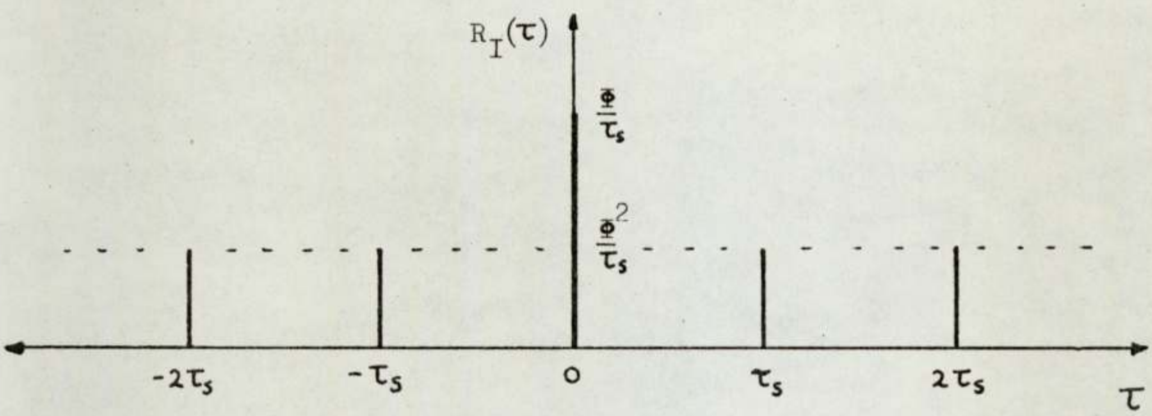


FIG. 5.5 Autocorrelation function of I(t)

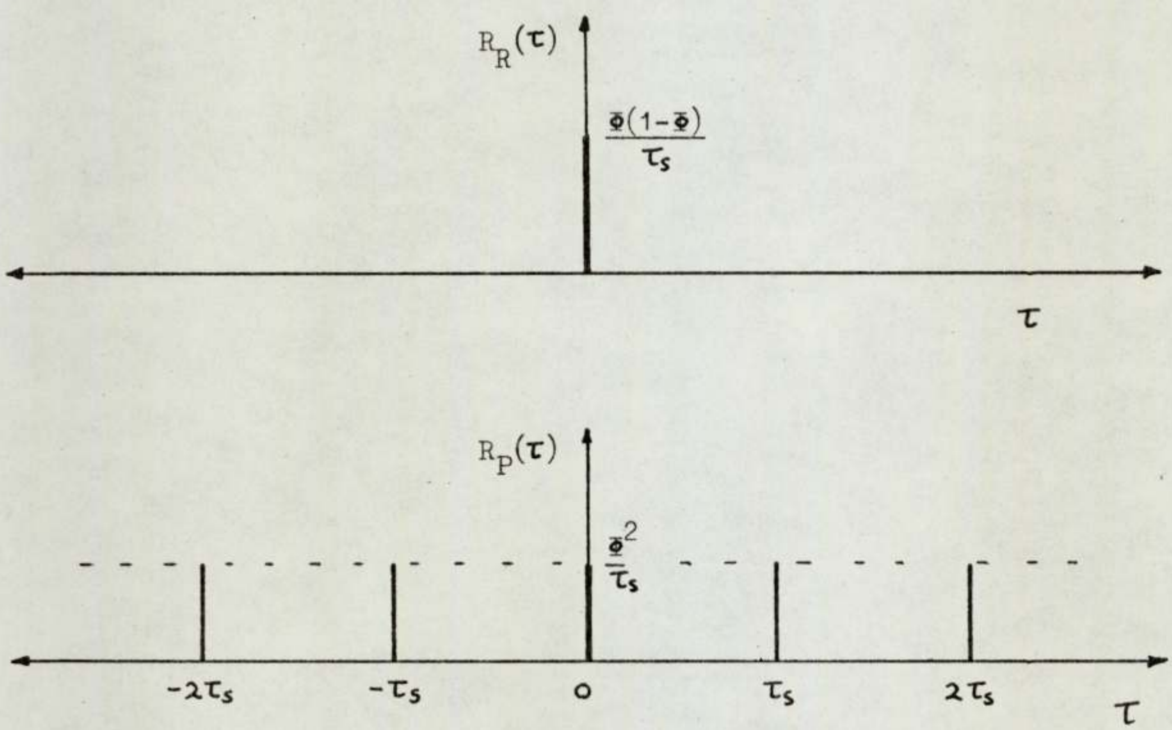


FIG. 5.6 The two components of $R_I(\tau)$

periodic impulse waveform. Thus, $I(t)$ is the sum of a periodic component and a random component. $I(t)$ can be written as :

$$I(t) = I_P(t) + I_R(t)$$

Now $$n(t) = X(t)I(t) = X(t)I_P(t) + X(t)I_R(t)$$

$$\therefore \quad n(t) = n_P(t) + n_R(t)$$

Where :

$$n_P(t) = X(t)I_P(t)$$

$$n_R(t) = X(t)I_R(t)$$

The autocorrelation function of $n_R(t)$ is :

$$R_{nR}(\tau) = \overline{n_R(t)\, n_R(t+\tau)} = \overline{X(t)I_R(t)\, X(t+\tau)I_R(t+\tau)}$$

Now, $X(t)$ and $I_R(t)$ are statistically independent, hence :

$$R_{nR}(\tau) = \overline{X(t)X(t+\tau)}\;\overline{I_R(t)I_R(t+\tau)} = R_X(\tau)R_R(\tau)$$

Where $R_X(\tau)$ is the auto correlation function of $X(t)$. From FIG.5.6 it can be seen that $R_R(\tau)$ is simply a Dirac impulse at $\tau = 0$, hence $R_{nR}(\tau)$ is of a similar form.

Now, $$R_R(\tau) = \frac{\Phi(1-\Phi)}{\tau_s}\,\delta(\tau)$$

Hence, $$R_{nR}(\tau) = \frac{\Phi(1-\Phi)}{\tau_s}\,R_X(0)\,\delta(\tau)$$

But $$R_X(0) = \overline{X(t)^2}$$

$$\therefore \quad R_{nR}(\tau) = \frac{\Phi(1-\Phi)}{\tau_s}\,\overline{X(t)^2}\,\delta(\tau) \qquad (5.6)$$

The power spectral density of $n_R(t)$ is given by :

$$G(f) = \int_{-\infty}^{\infty} R_{nR}(\tau)\, e^{-j\omega\tau} d\tau$$

$$= \int_{-\infty}^{\infty} \frac{\Phi(1-\Phi)}{\tau_s}\, \overline{X(t)^2}\, \delta(\tau) e^{-j\omega\tau} d\tau$$

$$\therefore \quad G(f) = \frac{\Phi(1-\Phi)\, \overline{X(t)^2}}{\tau_s} \qquad (5.7)$$

If the baseband filter has a bandwidth f_m, then the noise power due to $n_R(t)$ at the filter output is :

$$N_{OR} = \int_{-f_m}^{f_m} G(f) df = \int_{-f_m}^{f_m} \frac{\Phi(1-\Phi)}{\tau_s}\, \overline{X(t)^2}\, df$$

$$= \frac{2fm\, \Phi(1-\Phi)\, \overline{X(t)^2}}{\tau_s}$$

But $\quad fm \doteq \frac{1}{2\tau_s}$

$$\therefore \quad N_{OR} = \frac{\Phi(1-\Phi)\, \overline{X(t)^2}}{\tau_s^{\,2}} \qquad (5.8)$$

The sampled speech signal at the input to the baseband filter is given by :

$$X_{in}(t) = X(t) \sum_{k=-\infty}^{\infty} \delta(t-k\tau_s)$$

Where $\sum_{k=-\infty}^{\infty} \delta(t-k\tau_s)$ is the sampling function.

$X_{in}(t)$ can be re-written as a Fourier series :

$$X_{in}(t) = X(t) \frac{1}{\tau_s} \sum_{n=-\infty}^{\infty} e^{j2\pi n\tau_s}$$

The baseband filter removes all components of the series except the one at n = 0. Hence, the signal at the filter output is :

$$X_0(t) = \frac{X(t)}{\tau_s}$$

and the speech signal power is given by :

$$\overline{X_0(t)^2} = \frac{\overline{X(t)^2}}{\tau_s^2} \qquad (5.9)$$

The periodic component of the overflow noise $n_P(t)$ is in the form of a sampled speech signal. The noise power, at the output of the baseband filter, can be calculated using the same approach as for equation (5.9). The sampling function impulses in this case ($I_P(t)$) have an area Φ rather than unity (see FIG. 5.6).

$$n_P(t) = X(t)I_P(t) = X(t)\,\Phi \sum_{k=-\infty}^{\infty} \delta(t-k\tau_s)$$

Hence, using similar reasoning as for equation (5.9) above, the noise power at the baseband filter output due to $n_P(t)$ is :

$$N_{OP} = \frac{\overline{X(t)^2}\,\Phi^2}{\tau_s^2} \qquad (5.10)$$

The total noise power is :

$$N_O = N_{OR} + N_{OP} = \frac{\Phi(1-\Phi)\,\overline{X(t)^2}}{\tau_s^2} + \frac{\overline{X(t)^2}\,\Phi^2}{\tau_s^2}$$

$$\therefore \quad N_O = \frac{\Phi\,\overline{X(t)^2}}{\tau_s^2} \qquad (5.11)$$

The signal-to-overflow-noise ratio at the filter output is :

$$\frac{\overline{X_O(t)^2}}{N_O} = \frac{\overline{X(t)^2}}{\tau_s^2}\;\frac{\tau_s^2}{\Phi\,\overline{X(t)^2}} = \frac{1}{\Phi} \qquad (5.12)$$

b) <u>ZOP ROTA system</u>

Frame overflow in this system causes a sample to be removed and replaced with the immediately previous sample. This can be considered as adding a noise impulse equal in magnitude, but opposite in sign, to the <u>difference</u> between the sample removed and the previous sample.

The procedure for calculating the noise power is almost the same as for the ZAP ROTA system. The only difference is that $X(t)$ is replaced by the speech difference signal $X_D(t)$ where :

$$X_D(t) \quad = \quad X(t) - X(t-\tau_s)$$

Then:
$$N_0 \quad = \quad \frac{\Phi\, \overline{X_D(t)^2}}{\tau_s^2} \tag{5.13}$$

Hence, the signal-to-overflow-noise ratio is :

$$\frac{\overline{X_0(t)^2}}{N_0} \quad = \quad \frac{\overline{X(t)^2}}{\tau_s^2} \; \frac{\tau_s^2}{I\, \overline{X_D(t)^2}} \quad = \quad \frac{\overline{X(t)^2}}{\overline{X_D(t)^2}} \; \frac{1}{\Phi}$$

$$\frac{\overline{X_0(t)^2}}{N_0} \quad = \quad \frac{M}{\Phi} \tag{5.14}$$

Where :
$$M \quad = \quad \frac{\overline{X(t)^2}}{\overline{X_D(t)^2}}$$

The factor M was calculated using data obtained during the investigation described in Chapter 3. For the high-quality ribbon microphone, M was 1.9 for a male talker and 1.7 for a female talker.

The values obtained for the telephone conversations using the standard GPO carbon microphone were well scattered. M ranged in value from 1.0 to 2.0 with an average of 1.5. The smaller value of M for the carbon microphone is caused by its rising frequency response. This accentuates the higher frequencies and results in an increase in the difference power $\overline{X_D(t)^2}$.

The ZOP ROTA method of overflow-error distribution is used by Sciulli and Campanella in their SPEC system[12]. They have also derived a signal-to-overflow-noise ratio. Using a somewhat different approach to that outlined above, they have obtained the signal-to-noise ratio as

$$\frac{S}{N_0} = \frac{\overline{X(t)^2}}{\overline{X_D(t)^2}\ \Phi}$$

which is identical to equation (5.14).

They extend the calculation by deriving an expression for $\overline{X(t)^2}$ in terms of E_o , which is stated to be the value of one quantum step. This implies that all quanta are of equal size (i.e. linear encoding). This cannot be so if instantaneous companding is used, since this involves an effective non-linear encoding. Furthermore, they assume that companding "flattens" the probability density function (p.d.f) of speech, so that at the codec input, the distribution can be assumed to be uniform. This assumption does not seem valid since, ideally, the signals at the compressor input and expander output are identical. Hence, there is no distortion of either the speech or its p.d.f.

$\overline{X(t)^2}$ is then calculated from

$$\overline{X(t)^2} = \int x^2 p(x)dx$$

where $p(x)$ is the p.d.f. of the speech waveform and $p(x) = \frac{1}{AE_o}$

(i.e. assumed uniform over all quanta). A is the total number of quantization steps. Then,

$$\overline{X(t)^2} = \frac{A^2 E_o^2}{12}$$

Also, they express the power in the difference signal as

$$\overline{X_D(t)^2} = Q E_o^{\,2}$$

where Q is a constant with an experimentally determined value of 800.

The signal-to-overflow-noise ratio is then given by :

$$\frac{S}{N_0} = \frac{A^2 E_o^{\,2}}{12} \cdot \frac{1}{Q E_o^{\,2}} \cdot \frac{1}{\Phi} = \frac{A^2}{12 Q \Phi}$$

For 8-bit encoding $A = 2^8$ and

$$\frac{S}{N_0} = \frac{M}{\Phi} \qquad \text{where } M = 6.5$$

This value of M is over three times larger than the value obtained from the experimental work described in chapter 3.

Taking into account the difference in the definition of Φ (see section 5.2) and the difference in M, Sciulli and Campanellas analysis gives a signal-to-noise ratio approximately 10dB greater than the value given by equations 5.4 and 5.14.

c) Variable word length system

Frame overflow in the VWL system causes the word length of encoded speech samples to be reduced. This results in increased quantizing noise. Let there be B bits per frame available for signal samples and y be the number of active channels in any frame. Then the number of bits per sample, b, is adjusted to satisfy :

$$yb \leq B \tag{5.15}$$

When the frame has not overflowed, let the speech samples be encoded with r bits and let the quantizing noise power under these conditions be N_q. Then the total quantizing noise power is [7]:

$$N_{qT} = N_q\, P(b=r) + N_q 4\, P(b=r-1) + N_q 4^2\, P(b=r-2) + \ldots + N_q 4^{(r-1)} P(b=1)$$

Then,

$$N_{qT} = N_q \sum_{i=1}^{r} P(b=i) 4^{(r-i)} \tag{5.16}$$

Where $P(b=i)$ is the probability that a sample will be encoded with i bits.

The noise due to frame overflow is simply :

$$N_0 = N_{qT} - N_q$$

$$\therefore \quad N_0 = N_q\left(\sum_{i=1}^{r} P(b=i) 4^{r-i} - 1\right) \tag{5.17}$$

The probabilities $P(b=i)$ can be calculated using equation (5.15) and simple binomial probability theory. The probability $P(b=r)$ is simply the probability that the frame is not overflowing.

i.e. $P(b=r) = P(y \leq K)$

In the general case :

$$P(b=i) = P(y_{i-1} < y \leq y_i)$$

Where y_i is the highest integer value of y which satisfies equation (5.15) when samples are encoded with i bits (similarly for y_{i-1}). $P(y_{i-1} < y \leq y_i)$ is the probability that y lies in the range $(y_{i-1} + 1)$ to y_i and is given by :

$$P(y_{i-1} < y \leq y_i) = \sum_{j=y_{i-1}+1}^{y_i} \frac{N!}{j!(N-j)!} \alpha^{j}(1-\alpha)^{N-j} \qquad (5.18)$$

FIG. 5.7 shows signal-to-overflow-noise ratio $(\frac{S}{N_0})$ plotted against number of channels (N) for both the ROTA and VWL systems. This figure shows the considerable advantage that the VWL system has over the ROTA system, in terms of signal-to-noise performance.

In addition to overflow noise N_0, there will also be quantizing noise N_q. The total noise power is the sum of these, so the resulting signal-to-noise-ratio is $\frac{S}{(N_0+N_q)}$. FIG. 5.8 shows this signal-to-noise ratio plotted against number of channels (N) for both the ROTA and VWL systems.

The signal-to-noise ratios calculated above are for "worst-case" conditions since it has been assumed that all channels have a call on them. In a practical telephone system this will occur only for short periods during the day (i.e. the busy hour). For most of the time the call-load will be lighter and the signal-to-noise ratios will be greater.

The VWL system is superior to the ROTA system in signal-to-noise performance, because of the selective way it removes information under frame overflow conditions. Signal samples are reduced in length by removing bits of low-order significance, thus minimising the power in the error. The ROTA system has no such selective action.

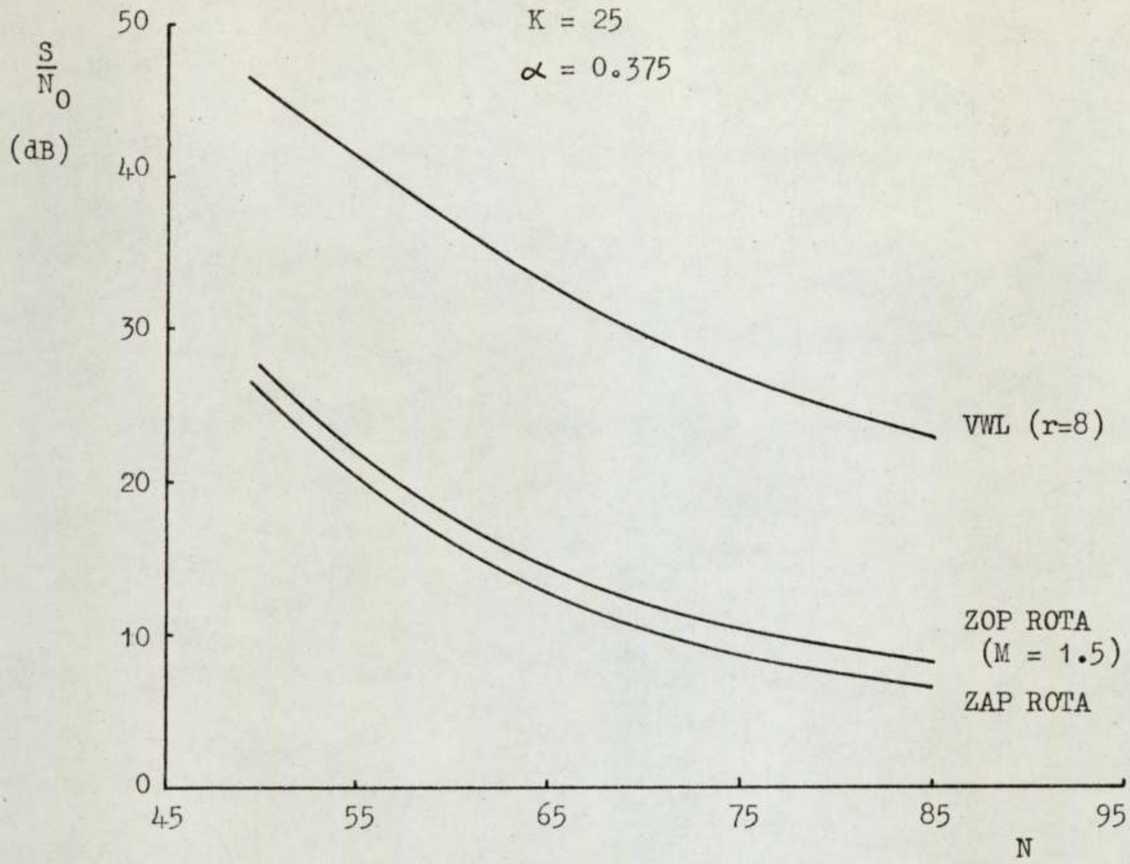


FIG. 5.7 Theoretical signal-to-overflow-noise ratio

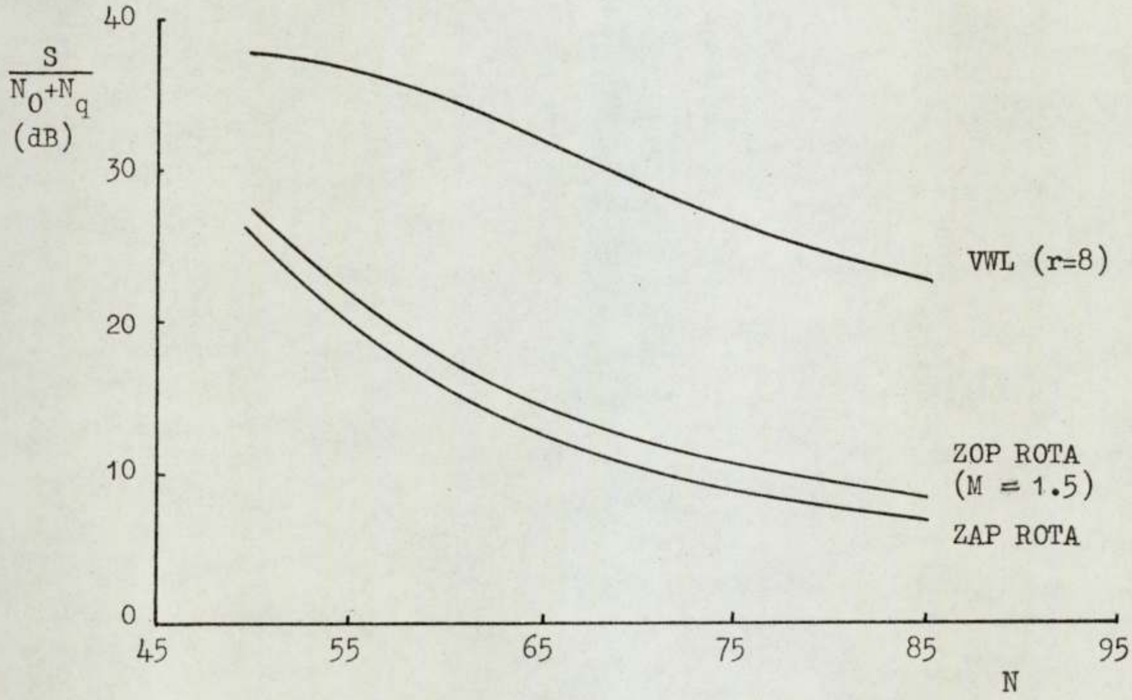


FIG. 5.8 Theoretical signal-to-(overflow + quantizing)-noise ratio

The VWL system does not, however, make maximum use of all frame bits when an overflow occurs. To maintain word synchronisation at the receiving terminal it is necessary that all sample words are of the same length. Thus, the number of signal bits per frame is a multiple of the number of active channels, and will rarely "fit" exactly into the available frame bits. This means that some frame bits are wasted in almost every frame, in addition to those lost because of overflow.

As an example, suppose there are 200 bits available for signal samples (i.e. 25 x 8-bit timeslots) and in one frame 26 channels are active. The frame has overflowed and the sample word length is reduced from 8 to 7 bits. The total number of signal bits is 7 x 26 = 182, and 18 bits (or over 2 timeslots) are wasted.

5.5 Frame filling

The previous sections have considered the effects of overflow when the number of active channels is greater than the number of timeslots. It is also necessary to consider the situation when the number of active channels is less than the number of timeslots. This condition will be called "underflow".

To maintain bit synchronisation throughout the frame, it is necessary to "fill in" any vacant timeslots when the frame has underflowed. This can be achieved by filling vacant timeslots with samples from non-active channels. When this technique is applied to the ROTA and ERROR-GROUPING systems, K samples are transmitted in every frame. This enables the parity-check digit, appended to the sample-address word, to be dispensed with, since parity is constant frame by frame. The parity-check digit cannot be removed when using the VWL system, however, since more than K samples can be transmitted.

CHAPTER 6

SIMULATION OF A PRACTICAL SYSTEM

SIMULATION OF A PRACTICAL SYSTEM

6.1 General considerations

Previous chapters have described the performance of the proposed multiplex transmission system in theoretical terms. Calculations have shown that a useful bandwith economy can be obtained with little degradation in signal-to-noise ratio. However, to make observations confirming theoretical findings and to make subjective tests, it was necessary to construct or simulate a practical system. With the limited resources available it was not possible to construct a practical version of the proposed system, nor was it necessary. A simpler solution is to consider an individual channel in the multiplex system and simulate the disturbing effects produced by the remaining channels.

This idea was adopted and a simplified block diagram of the model used in the simulation is given in FIG. 6.1. The model simulates an N-channel system with N variable between 33 and 95. The number of timeslots (K) is fixed at 25. The model consists of a single- channel transmission system with additional circuits to simulate the disturbing effects (frame overflow) caused by the remaining (N-1) channels.

The incoming speech is sampled, quantized and encoded in the sending terminal. The encoded samples pass to the receiving terminal, but may be disturbed en route by the output of the frame - overflow - error generator. In the receiving terminal the encoded samples are decoded and low-pass filtered to give the output speech.

As explained in chapter 5, all channels experience a disturbance caused by frame overflow. This is simulated by the frame-overflow error generator. The type of disturbance produced by frame overflow depends upon the method of overflow-error distribution. The ROTA and VWL methods described in section 5.3, together with the case where frame overflow

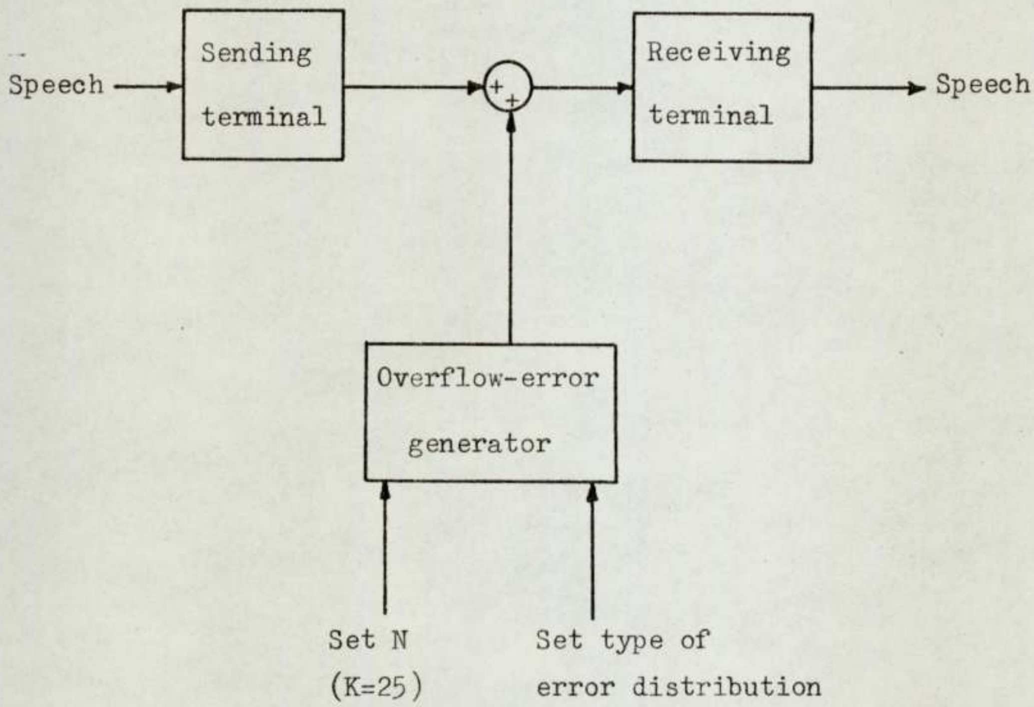


FIG. 6.1 Model used in the simulation of the proposed multiplex transmission system

errors are not distributed, were built into the simulation model. The ERROR GROUPING system was not built into the model, although the model was designed to allow it to be added later.

6.2 Active-channel generator

In a real system, frame overflow occurs when the number of active channels (y) in any particular frame exceeds the number of timeslots. The effect of the overflow depends upon how often it occurs and the size of the overflow. Thus, an essential feature of the overflow generator shown in FIG. 6.1 is a circuit which generates numbers representing y. This circuit is known as the active-channel generator.

The activity status of each channel (active or inactive) is a discrete (two-state) random process. Each channel has a probability α of being active. The number of active channels (y) is the sum of N such processes. Thus, y is also a discrete random process with N possible states. The individual processes which make up y are independent of each other. The process which generates y is analagous to tossing N coins and counting the number of heads. The probability of finding any particular number of heads (or y) follows a binomial distribution.

This intuitive argument is supported by the Central Limit Theorem. This theorem states that the sum of N independent terms asymptotically approaches a normal distribution if the terms are identically distributed and their variances exist. In this case, y is a discrete process, so it will tend towards a binomial rather than a normal distribution.

A convenient method of generating random numbers is by using feedback shift-registers. These generate pseudo-random binary signals in maximal-length sequences.[20,21] In an n-stage shift-register the sequence length is 2^n-1 digits. One method of generating random numbers using feedback shift-registers is to consider the contents of the register as a binary number. Unfortunately , the resulting random numbers have a uniform, rather than the required binomial, distribution. Alternatively, each stage of the shift-register can be considered as representing the activity status of one channel. For example, a binary '1' could represent activity and a '0' inactivity. Then the number of active channels is simply the number of '1's in the register. For an N-channel model the register would have N stages. The occurrence of a '1' in any stage is independent of any previous occurrence. Hence, the sum of '1's in the register is binomially distributed.

This method of generating the number of active channels has two drawbacks:

a) The probability of a '1' appearing in any stage of the shift-register is 0.5, whereas the probability of a channel being active is α.

b) The number of active channels (N) is required to be variable, which involves varying the number of stages in the register. This is not practicable, since the register has to have a specific number of stages to generate maximal-length sequences.[21]

These disadvantages can be overcome by using two shift-registers, as shown in FIG. 6.2.

The feedback shift-register SRA generates a pseudo-random binary sequence. The contents of several stages of SRA are examined by a combination of logic gates. By correct design of this logic (see section 6.4), the probability of a '1' at its output can be made equal to α . The output of the logic is taken to the input of shift-register SRB. Hence, the probability of a '1' appearing in any stage of SRB is α .

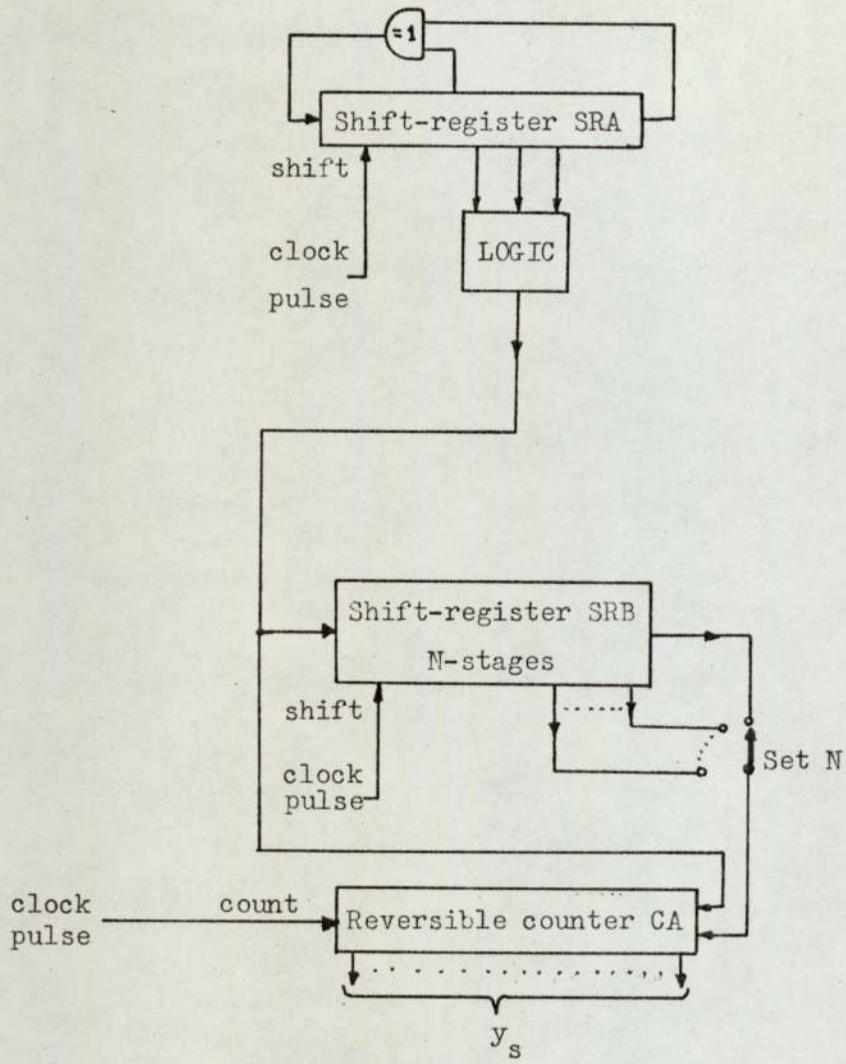


FIG. 6.2 Active-channel generator

As before, each stage of SRB represents one channel and SRB has a total of N stages. There is no restriction on the number of stages in SRB and N can be varied by a simple switch arrangement, as shown in FIG. 6.2.

The total number of '1's in SRB is accumulated in counter CA. Initially, CA and SRB are synchronised by clearing SRB and resetting CA to zero. At each clock pulse, CA examines what is about to enter and leave SRB. The contents of CA is incremented, decremented or left unchanged depending on whether the '1's population of SRB is increased, decreased or left unchanged. Thus, the contents of CA is the simulated number of active channels, y_s. The subscript s indicates that y_s is a simulation of y.

Another important feature of y is its dynamic behaviour. The activity status of an individual channel changes (i.e. active to non-active or vice versa) from time to time. Let the average rate of these transitions be λ. The transitions occur in "bursts," since they only appear during detected speech. However, since y is the sum of many such processes, the transitions in y will not have this burst characteristic. The arrival of a transition in y is independent of any previous arrivals; thus, the time between transitions has an exponential distribution. Also, the average transition arrival rate in y is simply $N\lambda$, for an N-channel system.

It is important that the dynamic behaviour of y_s is approximately the same as y. This is to ensure that the aural effect of frame overflow in the model is similar to a real system. The arrival of a transition in y_s is independent of any previous transitions, since the processes which make up y_s are independent. Thus, as for y, the transition arrival times have an exponential distribution.

The average transition arrival rate in y_s is a function of the clock pulse frequency (f_c) applied to SRB. The required value of f_c is calculated as follows. In shift-register SRB the contents will change after a clock pulse if a '1' enters and a '0' leaves, or if a '0' enters and a '1' leaves. The probability of a '1' entering or leaving is α and the probability of a '0' entering or leaving is $1-\alpha$. Hence, the probability of a transition in y_s at any clock pulse is :

$$P_t = \alpha(1-\alpha) + (1-\alpha)\alpha = 2\alpha(1-\alpha) \qquad (6.1)$$

If the frequency of the clock pulse applied to SRB is f_c, the average transition arrival rate of y_s is $f_c P_t$. Ideally, the average transition arrival rates of y and y_s should be identical. Hence,

$$N\lambda = f_c P_t = 2f_c\alpha(1-\alpha)$$

Then the required clock pulse frequency is :

$$f_c = \frac{N\lambda}{2\alpha(1-\alpha)} \qquad (6.2)$$

6.3 Simulation of the overflow-error distribution systems

This section is devoted to the modelling techniques used in the simulation of the various overflow-error distribution systems. For clarity, each of the systems is described and illustrated separately, although much of the equipment is common to all of them. Only the principles of the simulations are described here; the practical details may be somewhat different.

The model simulates systems using either zero-order prediction (ZOP) or zero-amplitude prediction (ZAP). However, for simplicity, only ZOP will be considered in the following descriptions. Only a small amount of additional circuitry is required to include ZAP.

In a practical transmission system using ZOP, the store in the sending terminal is used to detect and remove redundant samples (i.e. samples that do not differ from the immediately previous sample) that occur in active speech. This function is not required in the model and, for simplicity, the sending-terminal store is omitted. Thus, in the model, redundant samples that occur in active speech are sent to the receiving terminal. Under overflow conditions this causes the model to perform in a slightly different way from a practical system. The differences are :

a) The redundant samples that are transmitted may be affected by frame overflow. This would not occur in a practical system since redundant samples are not transmitted. This increases the amount of overflow noise in the model compared with a practical system.

b) In a practical system, a redundant sample is replaced with the previous sample at the receiving terminal. If the previous sample was affected by overflow noise then the latest sample will also be affected. This does not occur in the model, since redundant samples are transmitted. This reduces the amount of overflow noise in the model compared with the practical system.

Individually, these effects are relatively minor; also they tend to cancel each other out.

6.3.1 The SIMPLE system

In this system frame-overflow errors are not distributed (i.e channels are assigned timeslots on a fixed-priority basis). This system would not normally be used in practice and is included here for comparison purposes only. The number of overflow errors depends upon the position of the channel in the timeslot assignment sequence. The channel receiving the most interference is the last (N^{th}) in the sequence and is the one considered here.

A simplified block diagram of the simulation model is given in FIG.6.3 . The channel loses a sample when y_s exceeds the number of timeslots (K). To avoid having to subtract K repeatedly from y_s, the contents of CA is preset to minus K (instead of zero) when CA is synchronised with SRB. Thus, the contents of CA is the difference between y_s and K. An overflow is indicated when the contents of CA becomes positive; this condition is detected by examining the sign-bit of CA.

In a real system, a sample would be transmitted only when the speech detector was triggered and the frame had not overflowed. In the model, a sample from the A/D converter is loaded into the store only

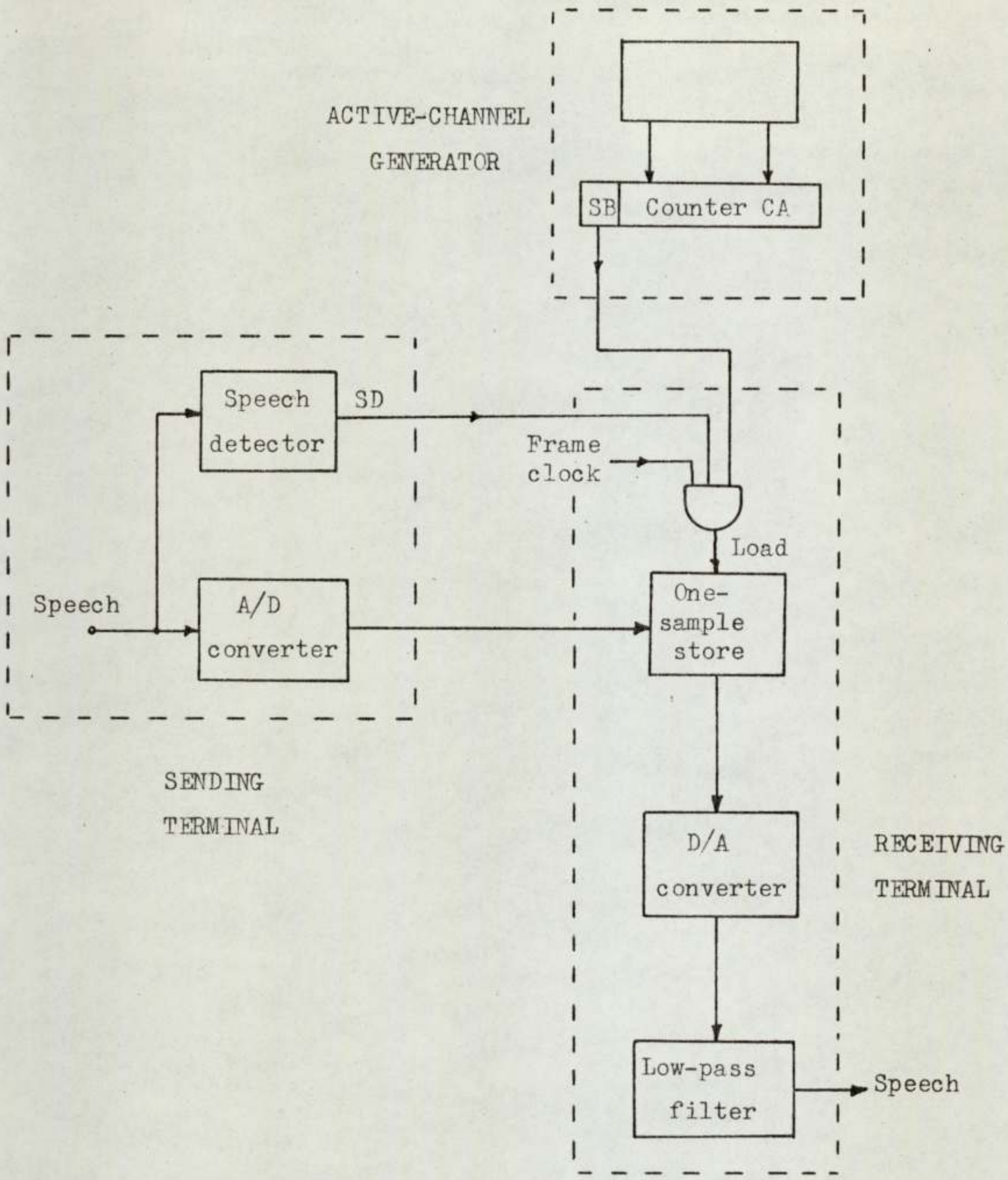


FIG. 6.3 Simulation of the SIMPLE system

when SD = '1' (speech detected) and SB = '1' (frame underflowed). When these conditions are detected by the 'AND' gate, the frame-frequency clock pulse causes the latest sample from the A/D converter to be loaded into the store.

6.3.2 The ROTA system

Only the ZOP version of the ROTA system is considered here; very little additional circuitry is required to simulate the ZAP ROTA system.

In the ZOP ROTA system, an overflow error causes a sample to be replaced by the immediately previous sample. The errors occur randomly in time with an average rate depending upon the values of N, K and α (see section 5.2). A possible method of generating these errors is given in FIG. 6.4. The logic detects patterns of '0's and '1's that occur in the feedback shift-register. When a particular pattern is detected, this represents the occurrence of an overflow error. A sample from the A/D converter is loaded into the store when the speech detector is triggered and there is no overflow (logic output = '0').

For a fixed K and α, the overflow-error rate varies with the number of input channels N. In the model, N is variable and hence the overflow-error rate also needs to be variable. The method used in generating errors is given in more detail in FIG. 6.5. The output of several stages of a feedback shift-register are taken, via switches, to an AND gate. As an example, suppose all of the switches are closed. The probability of a '1' occurring in any stage of the shift-register is 0.5. The probability of two '1's occurring simultaneously is $(0.5)^2$, and so on. In FIG. 6.5 the AND gate detects the occurrence of four '1's (all switches closed) which occurs with a probability of $(0.5)^4$. The probability of error can be varied in multiples of two by adjusting the number of closed switches.

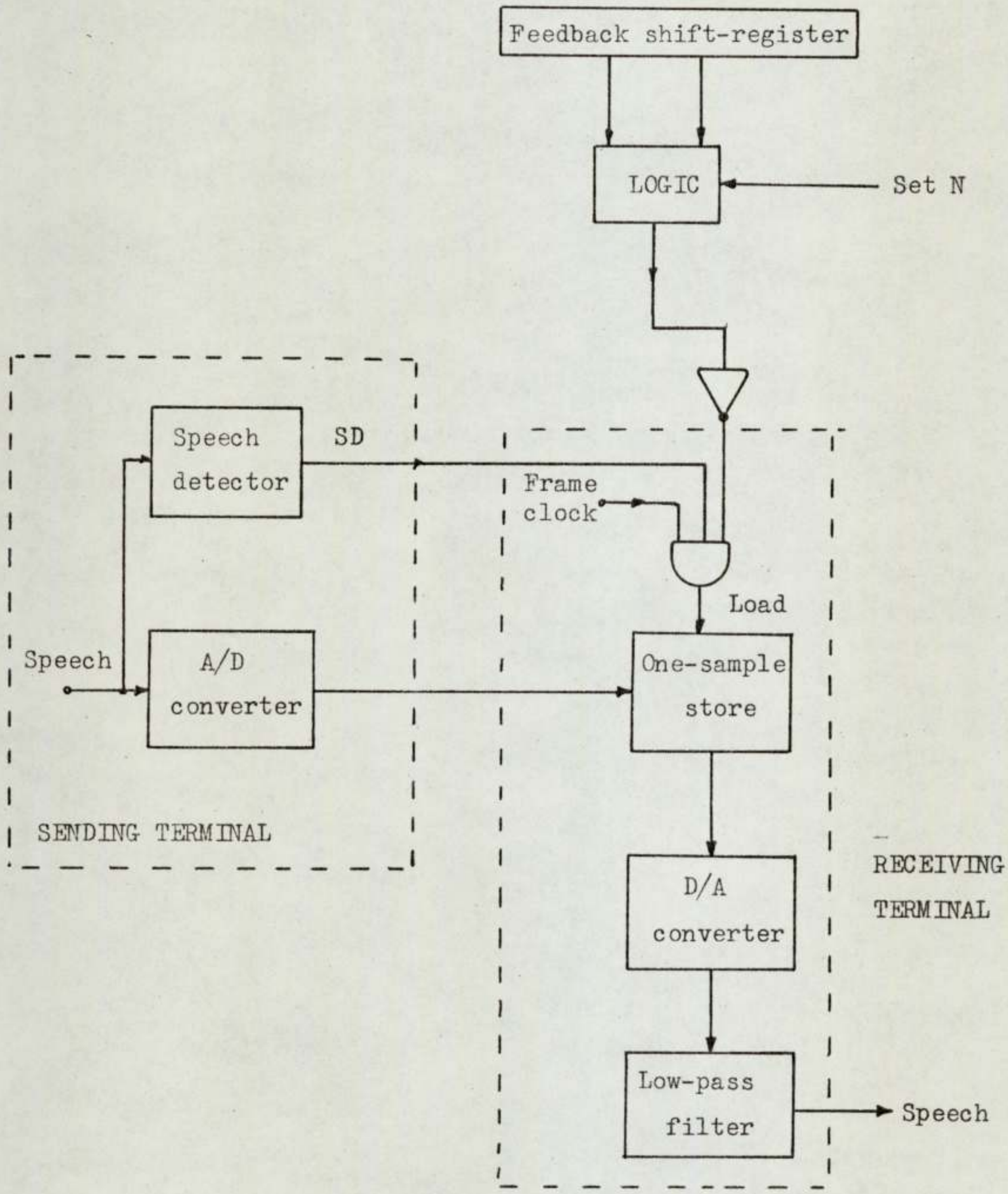


FIG. 6.4 Simulation of the ZOP ROTA system

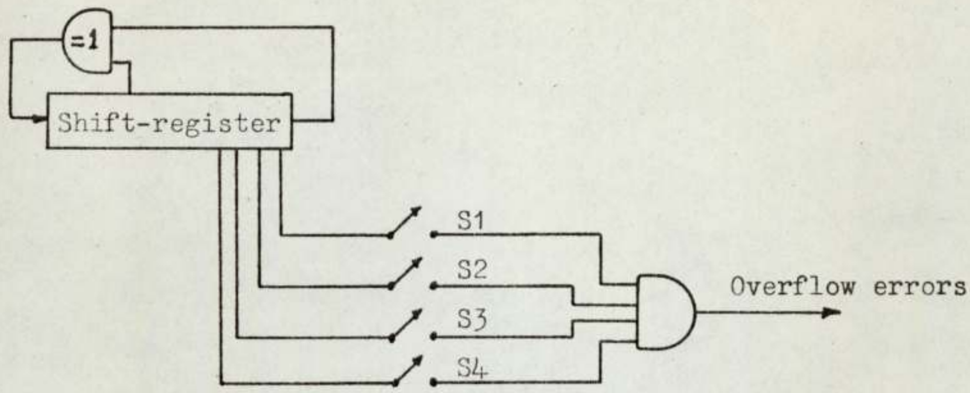


FIG. 6.5 Generation of ROTA system overflow errors

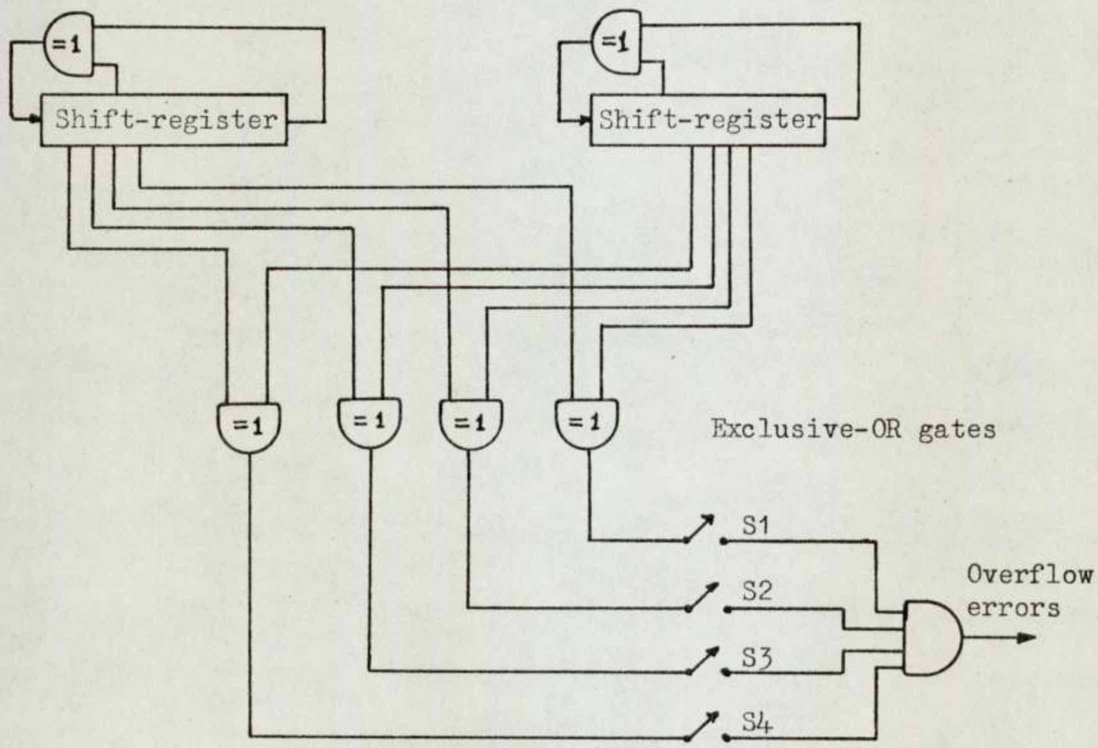


FIG. 6.6 An improved method of generating ROTA system overflow errors

This simple method of generating overflow errors has a disadvantage. The four outputs from the shift-register are related in time ; each one is a time-delayed (or advanced) version of the other outputs. This gives rise to a local non-randomness in the pattern of generated errors. In the example shown in FIG. 6.5, an error occurs when a '1' appears in four adjacent stages of the shift-register. The next clock pulse after the occurrence of an error causes the contents of the shift-register to be moved one stage to the right. Thus, three of the stages connected to the AND gate will contain a '1'. In this situation the probability of an error is the probability of a '1' entering the left-hand stage, which is 0.5 instead of the required $(0.5)^4$. This effect will tend to cause the generated pattern of overflow errors to occur in "bunches", which is obviously undesirable.

A solution to this problem is to use two independent feedback shift-registers as shown in FIG. 6.6. Each binary sequence to the AND gate is made up from the modulo-2 addition of one output from each of the shift-registers. Even using two shift-registers, the outputs have to be combined correctly to avoid one sequence being a time-delayed version of another sequence. Using this method, the sequences fed into the AND gate are independent [22] and error bunching does not occur.

It will be noted that this method of generating overflow errors does not require the active-channel generator.

6.3.3 The ERROR GROUPING system

Before describing the simulation model, a brief summary of the operation of this system will be given. Consider the sequence of events after the speech detector of channel i has been triggered.

a) The channel joins the rear of the secondary queue (if there is one). The length of the secondary queue in any frame is the difference between the number of speech-active channels and the number of timeslots K.

b) While in the secondary queue the channel has only occasional access to a timeslot. This occurs when the number of timeslots not required by priority channels exceeds the the number of channels ahead of channel i in the queue.

c) As priority channels release, secondary channels take their place and channel i approaches the front of the queue.

d) Eventually, channel i becomes a priority channel and then has access to a timeslot if required.

A simplified block diagram of a method of simulating the errors caused by this system is given in FIG. 6.7. The active-channel generator is set to generate the number of channels which are speech-active (i.e. speech detector triggered). Initially, CA is preset to minus K so that its contents represent the number of channels, in excess of K, that are speech active (i.e. the number of channels in the secondary queue).

During silent pauses in speech, the contents of CA is loaded into counter CB in each frame. When speech commences and the speech detector triggers, further loading into CB is inhibited. Hence, the contents of

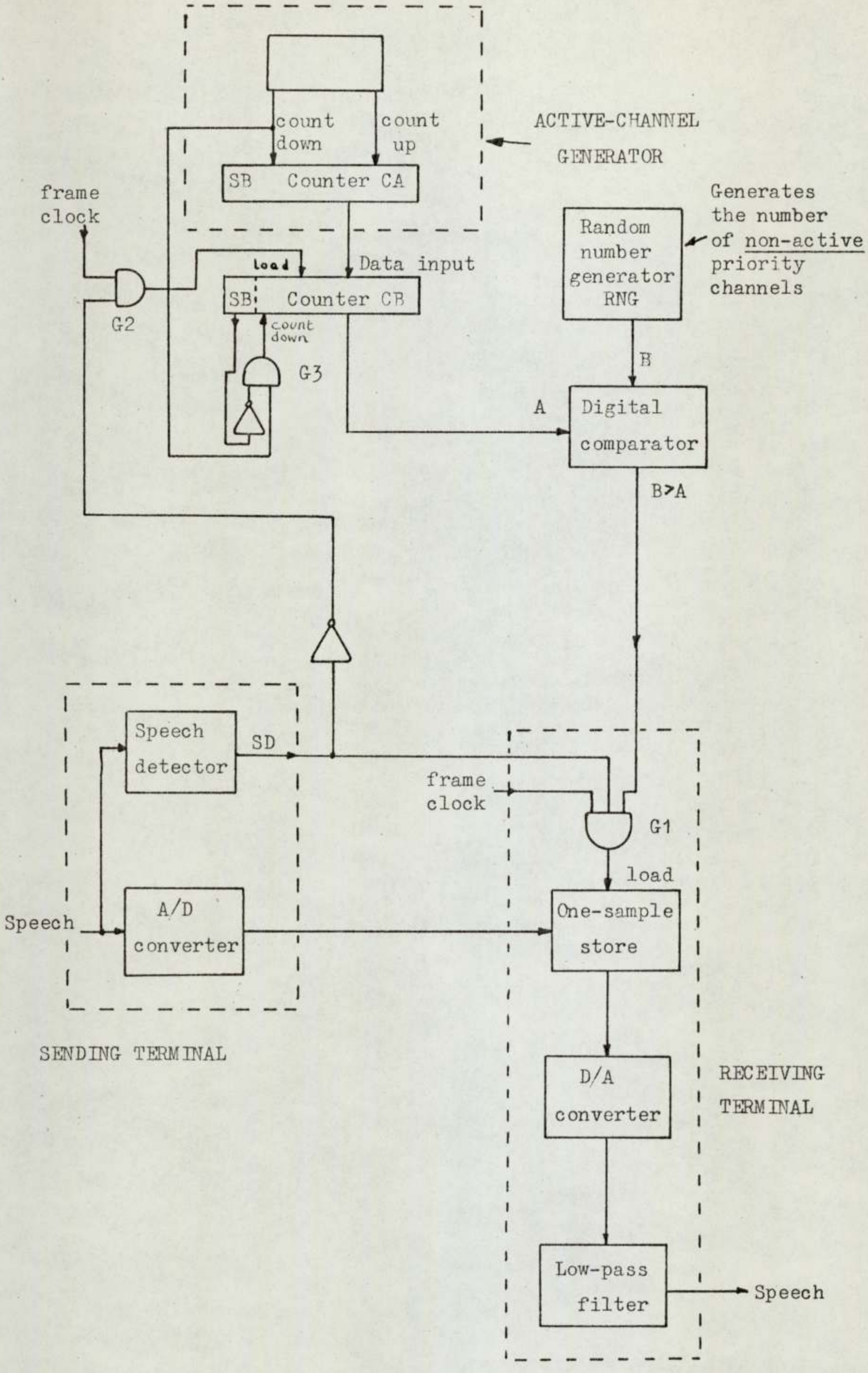


FIG. 6.7 Simulation of the ERROR-GROUPING system

CB (if positive) is the number of channels ahead in the secondary queue.

In subsequent frames, if the number of active channels decreases, it is assumed that a priority channel has become inactive (i.e. secondary channels never release). This is a reasonable assumption since:

a) Talkspurts are at least as long as the hangover time (t_H) of the speech detector.

b) The probability of a channel remaining in the secondary queue for a time longer than t_H is very small.

When a priority channel becomes inactive, a secondary channel replaces it as a priority channel; then, each secondary channel moves one place nearer the front of the queue. Thus, the contents of CB (i.e. the number of channels ahead in the secondary queue) is decremented each time the contents of CA decreases.

The number of spare timeslots (i.e. timeslots not required by priority channels) is generated by the random-number generator RNG. The output B of RNG is compared with the contents A of counter CB by the digital comparator. If B exceeds A, this indicates that there are sufficient spare timeslots for one to be allocated to this particular channel. The output of the comparator is a '1', which opens gate G1 and allows the frame clock to load a sample from the A/D converter into the store. If $B \leq A$ the output from the comparator is a '0'.

Eventually, the contents of CB becomes negative, indicating that the channel has attained priority status. Any further decrementing of CB is now inhibited by gate G3, since the sign-bit of CB has become a '1'. The overflow generating circuits have now reached a quiescent state, which they will remain in until the speech detector releases. The output of the comparator remains at a '1' (since A is negative) ; hence, G1 remains open until the speech detector releases.

The above description condensed into a flowchart is shown in FIG. 6.8. This diagram gives the cycle of operations that occur during each frame.

6.3.4 The VARIABLE WORD LENGTH system

A block diagram of the model used in the simulation of the VWL system is given in FIG. 6.9.

Initially, counter CA is preset to minus K so that the contents of CA is the difference between K and the number of active channels. During each frame, the decoder examines the contents of CA. If an overflow is indicated, the decoder computes the maximum sample word length that allows samples from all active channels to fit into the frame of transmitted pulses. The sample word length is adjusted in the bit-masking logic under the control of the decoder output. Both the decoder and bit-masking logic consist of logic gates.

The bits that have been removed from a sample by frame overflow are modified to minimise the error that is introduced (see section 5.3) The most significant missing bit is replaced with a binary '0' and all remaining missing bits are replaced with a '1'.

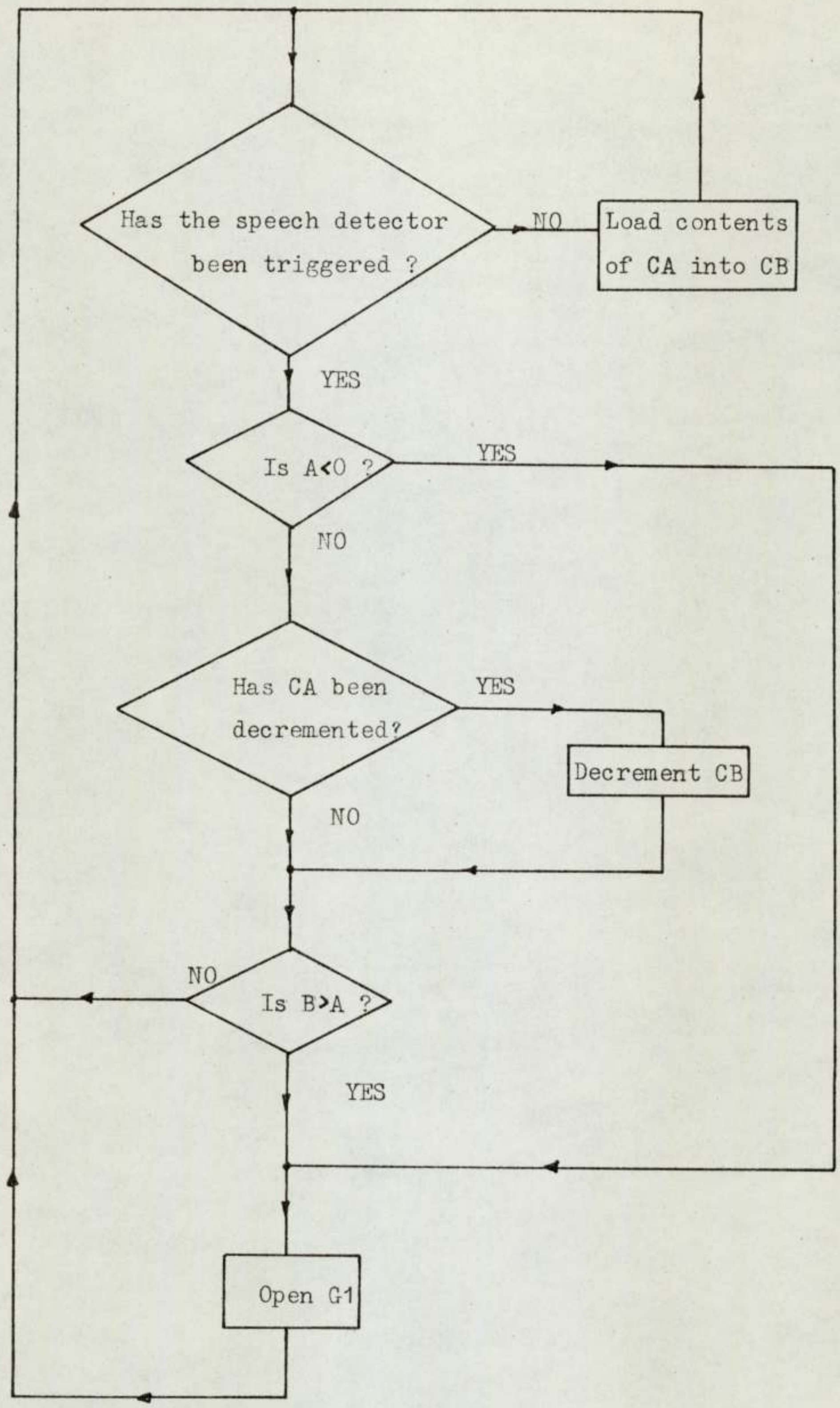


FIG. 6.8 Flowchart describing the operation of the ERROR GROUPING simulation

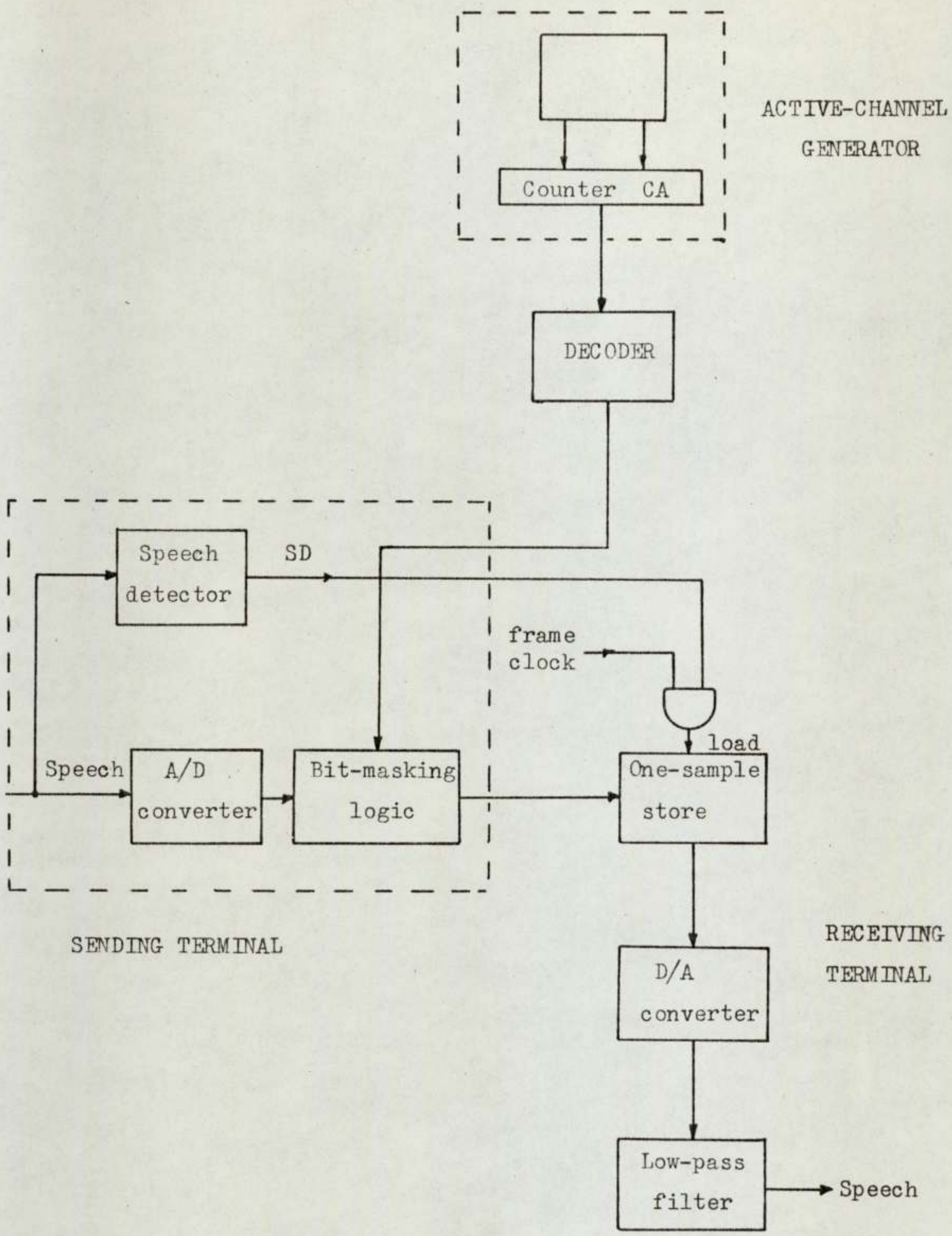


FIG. 6.9 Simulation of the VWL system

6.4 Practical design of the simulation model

This section describes the practical realisation of the principles described in section 6.3. A block diagram of the complete model is given in FIG. 6.10. This figure is basically a combination of FIGs. 6.3, 6.4 and 6.9.

The model consists of six basic sections; these being:

a) Speech-signal processing circuits (see section 6.4.1)

b) Feedback shift-registers (see section 6.4.2)

c) Active-channel generator (see section 6.4.3)

d) Overflow logic (see section 6.4.4)

e) Speech detector (see section 6.4.5)

f) Clock-pulse generator (see section 6.4.6)

The switches shown in FIG. 6.10 have the following functions:

Switch SWA varies the number of input channels by varying the number of stages in SRB.

Switch SWB selects the type of prediction (i.e. ZOP or ZAP).

Switch SWC selects the type of overflow-error distribution system.

Standard 74 series TTL integrated circuits were used for almost all the logic functions. The only exception is shift-register SRB, which is an MOS device. Where an input to a logic gate or other device is apparently left unconnected, this is equivalent to connecting it to a binary '1'.

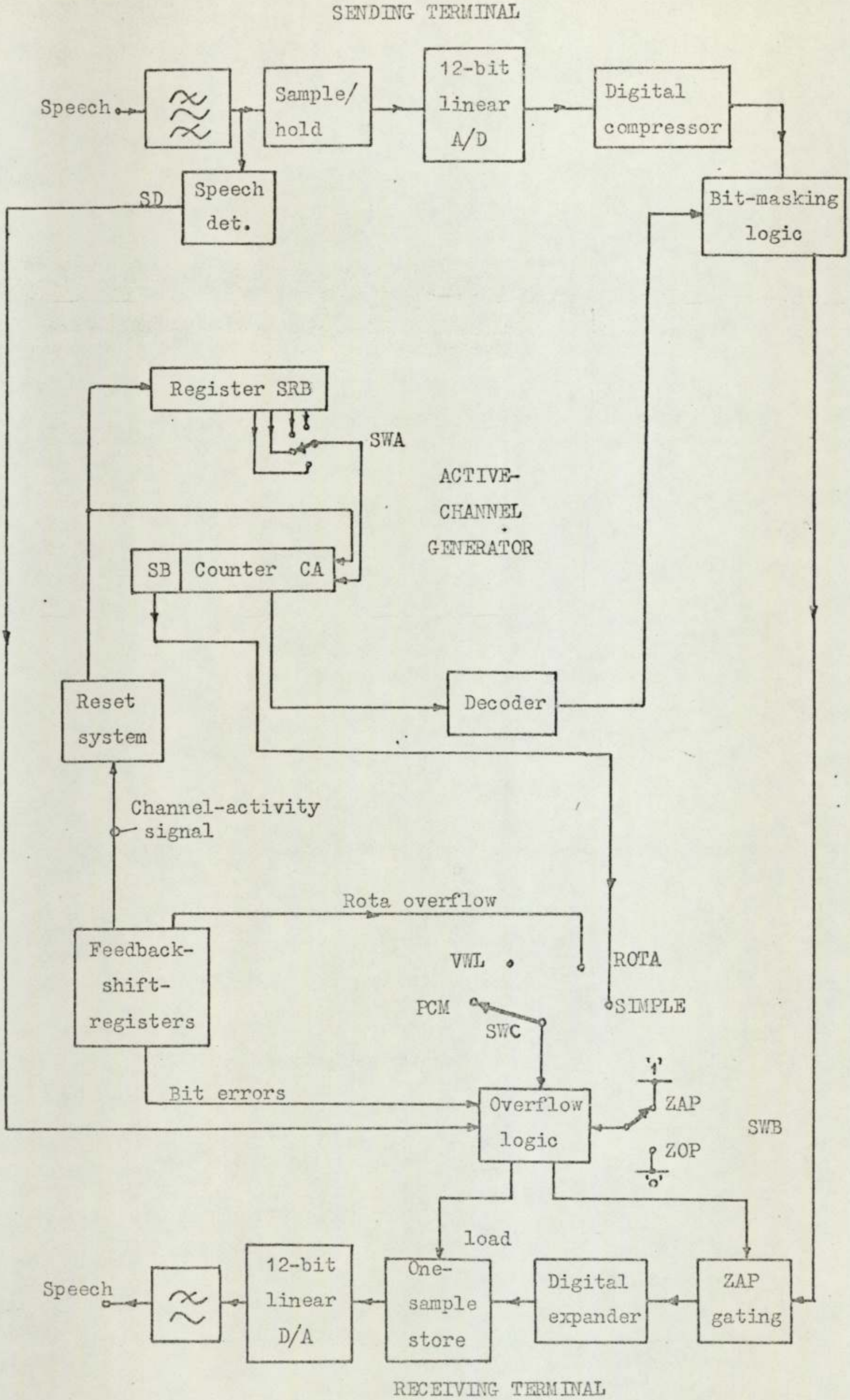
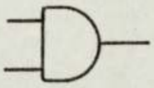
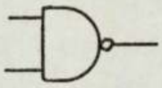
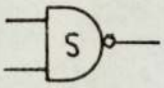
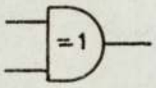
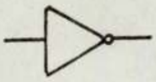
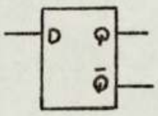
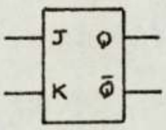
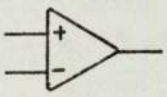
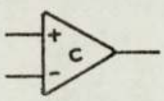


FIG. 6.10 The simulation model

A list of symbols used in subsequent diagrams is given below.

SYMBOL	FUNCTION
	AND gate
	NAND gate
	NAND gate with Schmitt-trigger input circuit
	Exclusive-OR gate
	Logic inverter
	D-type bistable
	JK bistable
	Operational amplifier
	Voltage comparator

A diagram illustrating the time position of all control pulses is given in FIG. 6.23.

6.4.1 Signal processing circuits

The signal processing circuits consist of two main parts.

a) Sending terminal

b) Receiving terminal

a) Sending terminal

Refering to FIG. 6.10, the incoming speech is initially band-pass filtered to the range 300-3400 Hz to simulate a practical telephone channel. The filter consists of a 3rd-order Butterworth high-pass filter followed by a 5th-order Butterworth low-pass filter. The filters are of the active-RC type and a detailed circuit diagram is given in FIG. 6.11.

After passing through a sample-and-hold (S/H) circuit the band-limited speech is linearly encoded in a 12-bit A/D converter. Both the S/H circuit and the A/D converter are proprietory items. The S/H circuit is an Analog Devices type SHA1A and the A/D converter is a Teledyne Philbrick type 4103.

An overall non-linear encoding is achieved by passing the 12-bit samples through a digital compressor. A circuit diagram of the compressor is given in FIG. 6.12. The compressor characteristic is segmented and is identical to the characteristic illustrated in FIG. 4.8 and described in section 4.4. The segments are chords to an A-law[19] curve, with A approximately equal to 87. A full description of the operation of the digital compressor is given in Appendix A4.

The compressed samples pass to the bit-masking logic. This circuit comes into operation only when the VWL system is being simulated. When any other overflow-error distribution system is in use, the samples

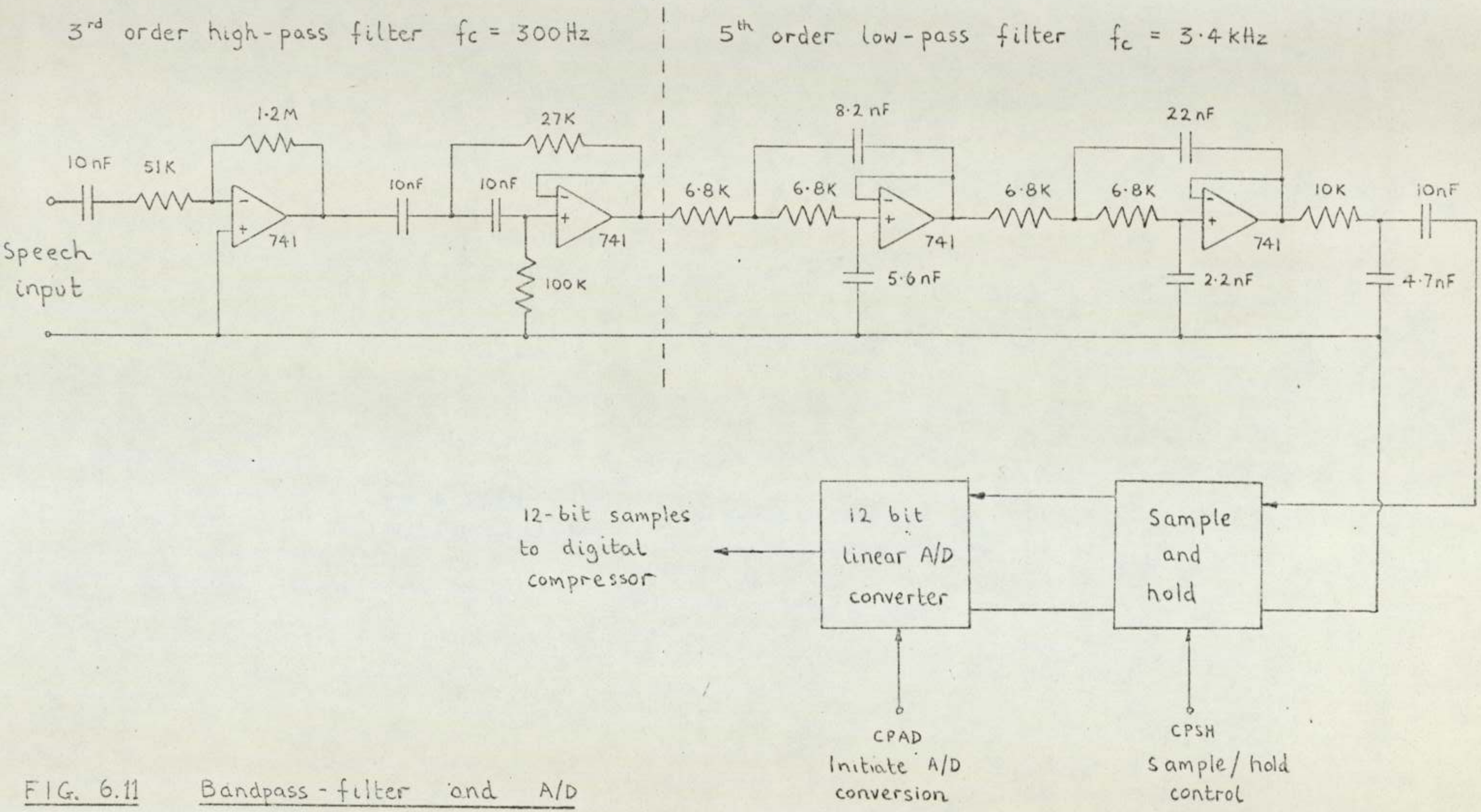


FIG. 6.11 Bandpass-filter and A/D converter

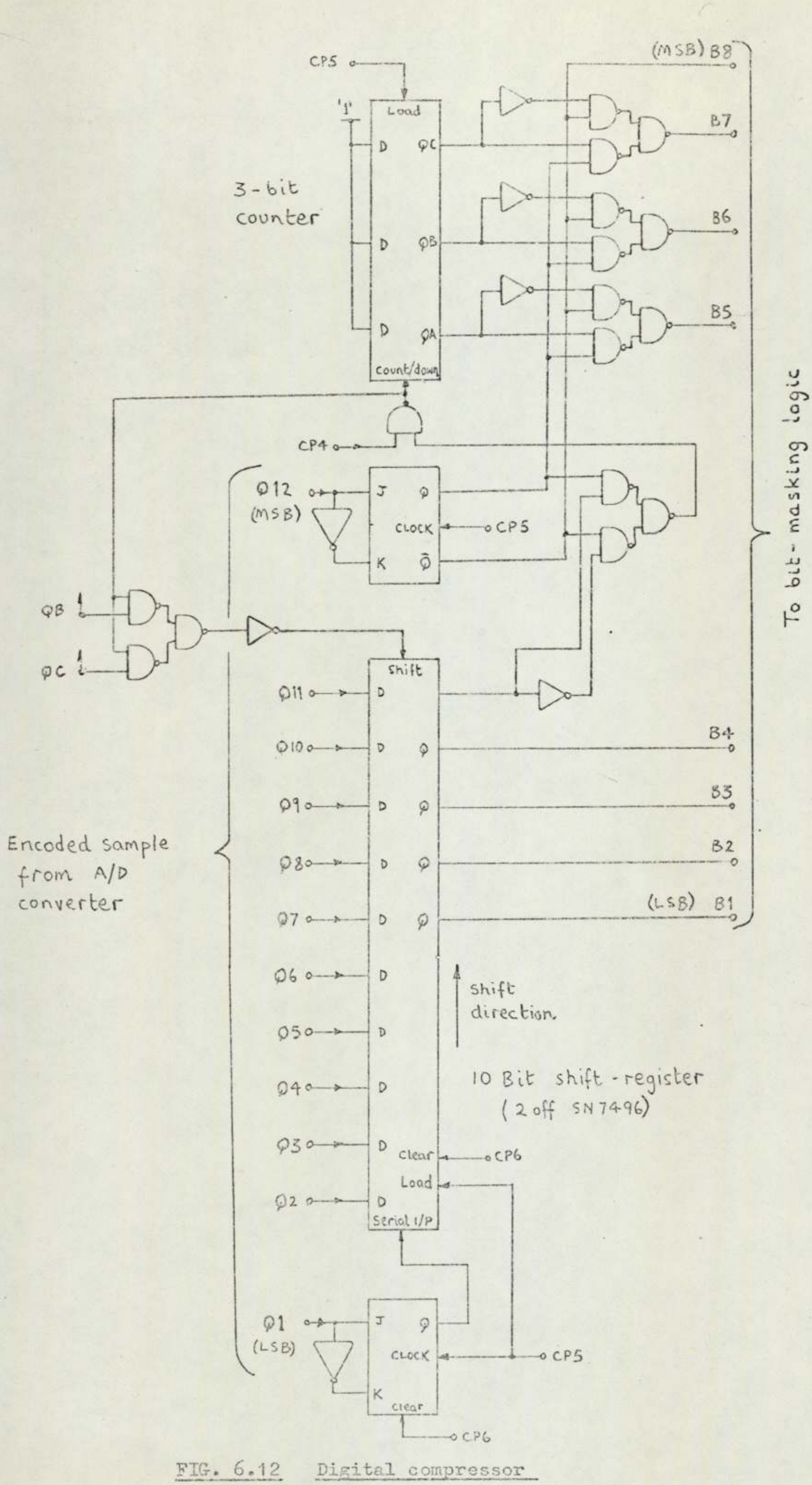


FIG. 6.12 Digital compressor

pass unmodified to the receiving terminal. The bit-masking logic modifies the sample word length when the frame has overflowed. Under the control of the decoder, this circuit removes and modifies the appropriate number of bits.

A circuit diagram of the bit-masking logic is given in FIG. 6.13. The gates on the left-hand side of FIG. 6.13 mask off bits under the control of the five lines from the decoder. The relationship between the bits removed and the signals on the decoder lines is summarised in the truth table below.

	Decoder lines				
Bits removed	7	6	5	4	3
-	0	0	0	0	0
B1	1	0	0	0	0
B1,B2	1	1	0	0	0
B1-B3	1	1	1	0	0
B1-B4	1	1	1	1	0
B1-B5	1	1	1	1	1

The gates on the right-hand side of FIG. 6.13 modify the missing bits. The most significant missing bit is replaced with a '0' and each remaining missing bit is replaced with a '1'. In a real system, missing bits would be replaced at the receiving terminal. However, in the model, the two operations of removing and modifying bits are grouped together for convenience.

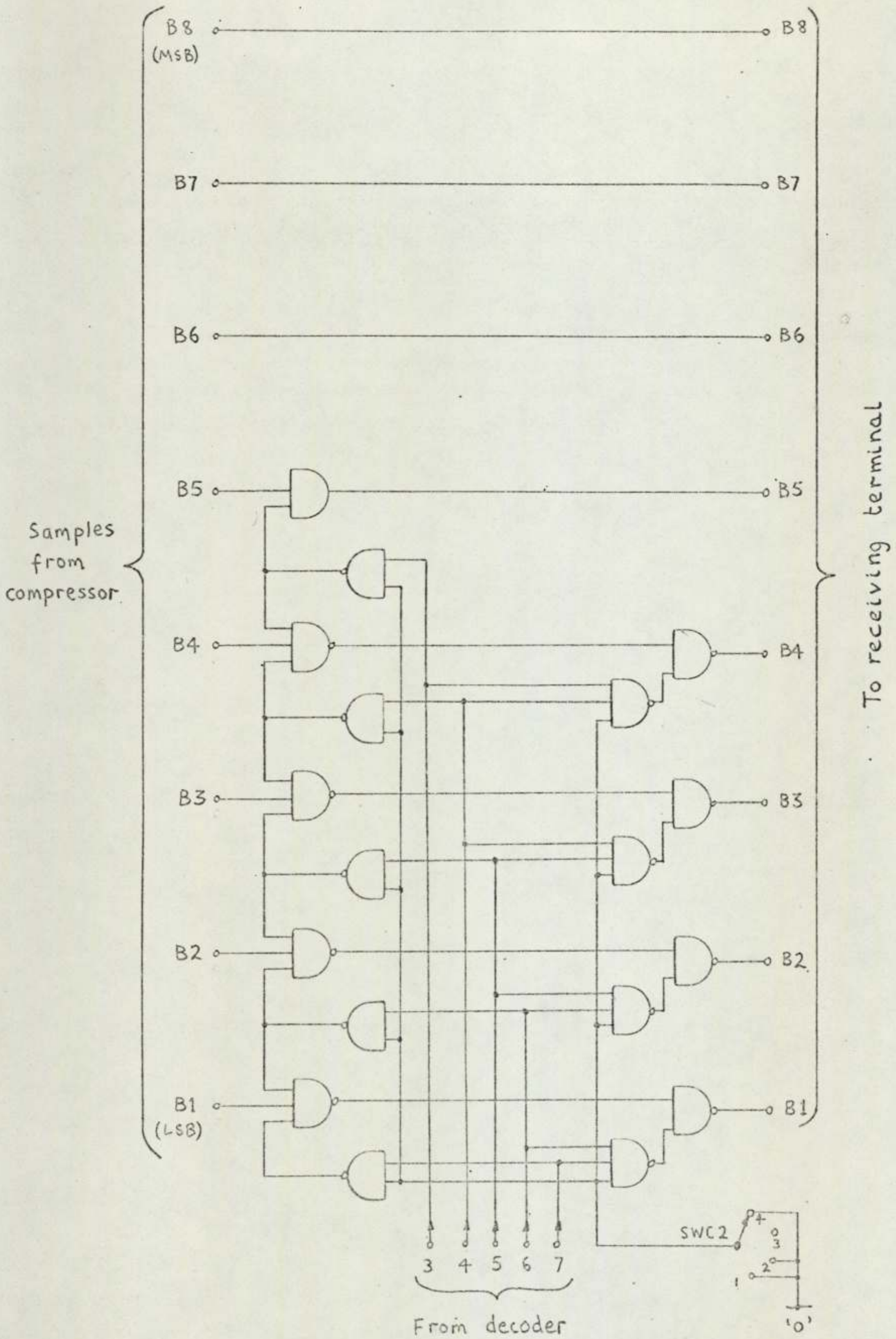


FIG. 6.13 Bit-masking logic

b) Receiving terminal.

The ZAP gating, and digital expander are detailed in FIG. 6.14. The ZAP gating is the vertical row of AND gates on the left-hand side of FIG. 6.14. These gates come into operation when the ZAP SIMPLE, and ZAP ROTA systems are being simulated. The gates are closed by a control signal from the overflow logic when a frame overflow occurs, or when the speech detector is not triggered. This simulates what happens at the receiving terminal of a real system (using ZAP) when no sample is received; a zero-valued sample is read out to the channel. When the ZOP systems and VWL system are being simulated the ZAP gates are permanently open.

The remainder of FIG 6.14 is the digital expander. This circuit, in conjunction with the D/A converter, gives an overall non-linear decoding. The transfer characteristic of the expander is the inverse of the characteristic of the compressor in the sending terminal. The combination of the two characteristics gives an overall linear relationship between the input and output signals of the transmission system. A full description of the operation of the expander is given in APPENDIX A5.

The receiving store, D/A converter and low-pass filter are detailed in FIG. 6.15. A sample is loaded into the store when a pulse is applied to the 'load' input. If a pulse is not applied, the content of the store remains unchanged. The store performs the same functions in the model as it would in a practical system. The D/A converter is a type 4005 made by Teledyne Philbrick. The output of the D/A converter is filtered to obtain the output speech signal. The low-pass filter is of the active-RC type and has a fourth-order Butterworth response.

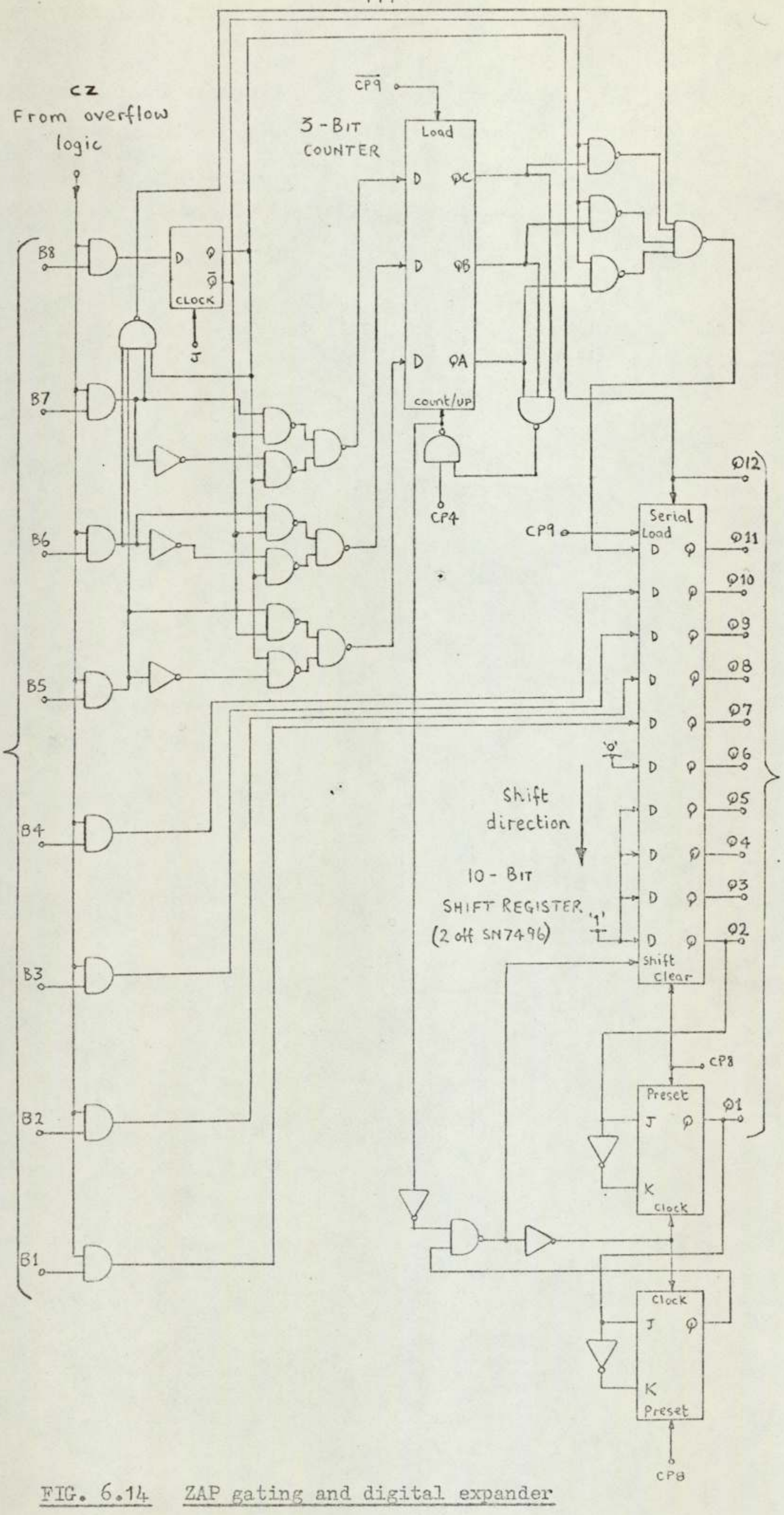


FIG. 6.14 ZAP gating and digital expander

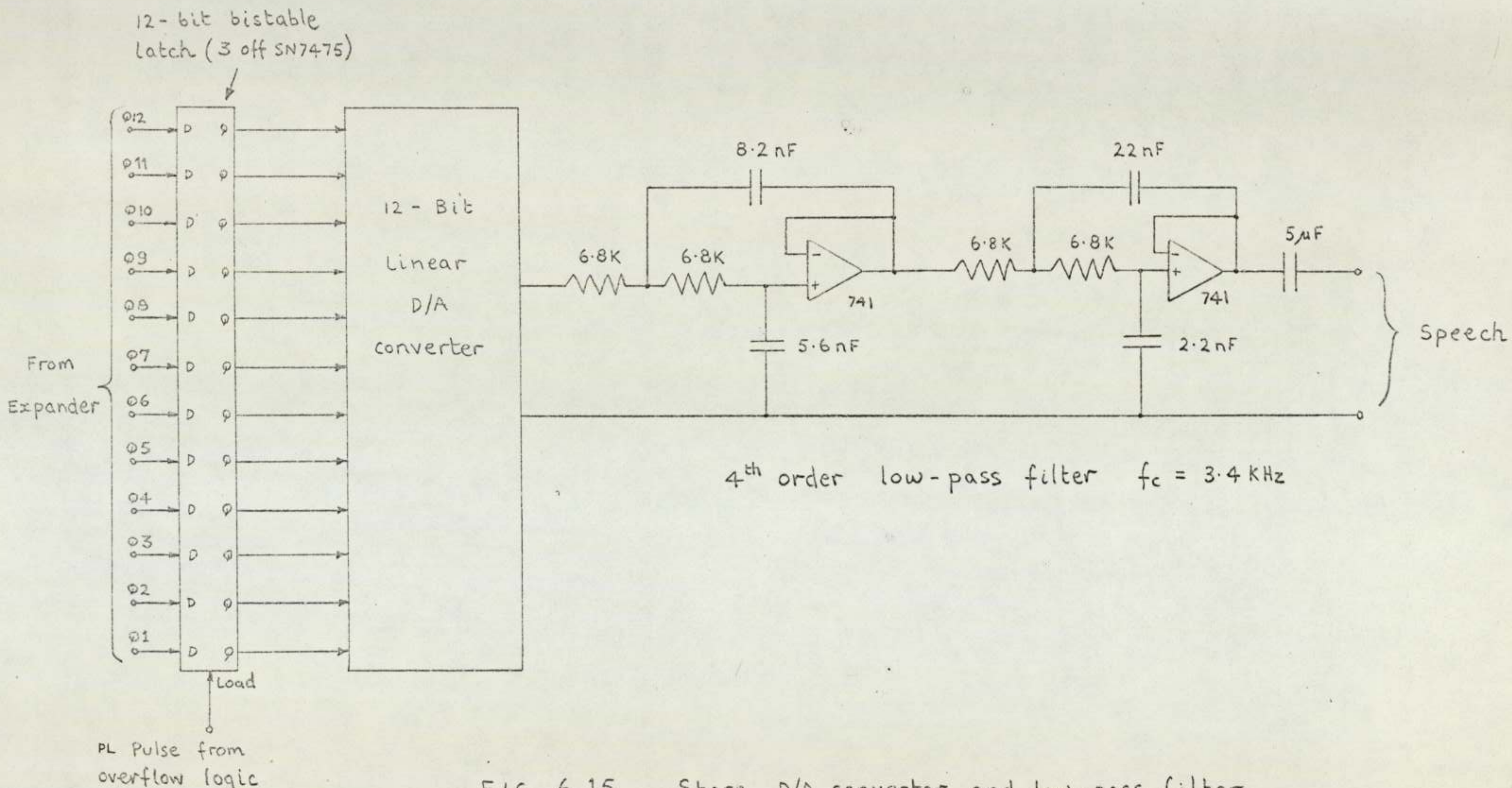


FIG. 6.15 Store, D/A converter and low-pass filter

6.4.2 Feedback shift-registers

The feedback shift-registers, shown in FIG. 6.16, are the heart of the error-generating circuits. They generate the following random signals:

a) Channel-activity binary signal for the active-channel generator.

b) Frame-overflow errors for the ROTA system.

c) Bit errors in the sample-address word.

Two feedback shift-registers are used to generate all of the above random signals. Each register is connected to generate pseudo-random binary signals in maximal-length sequences. The output from the shift-register stages are fed in pairs (one from each shift-register) into exclusive-OR gates. By correct combination of the shift-register outputs[22], the binary signal from an exclusive-OR gate is independent of any other exclusive-OR gate output. This procedure eliminates any statistical dependencies between generated error signals. It also eliminates local non-randomness in the error signals i.e. "bunching" (see section 6.3).

The outputs from the exclusive-OR gates are then fed into various logic gate combinations to generate the random signals mentioned above.

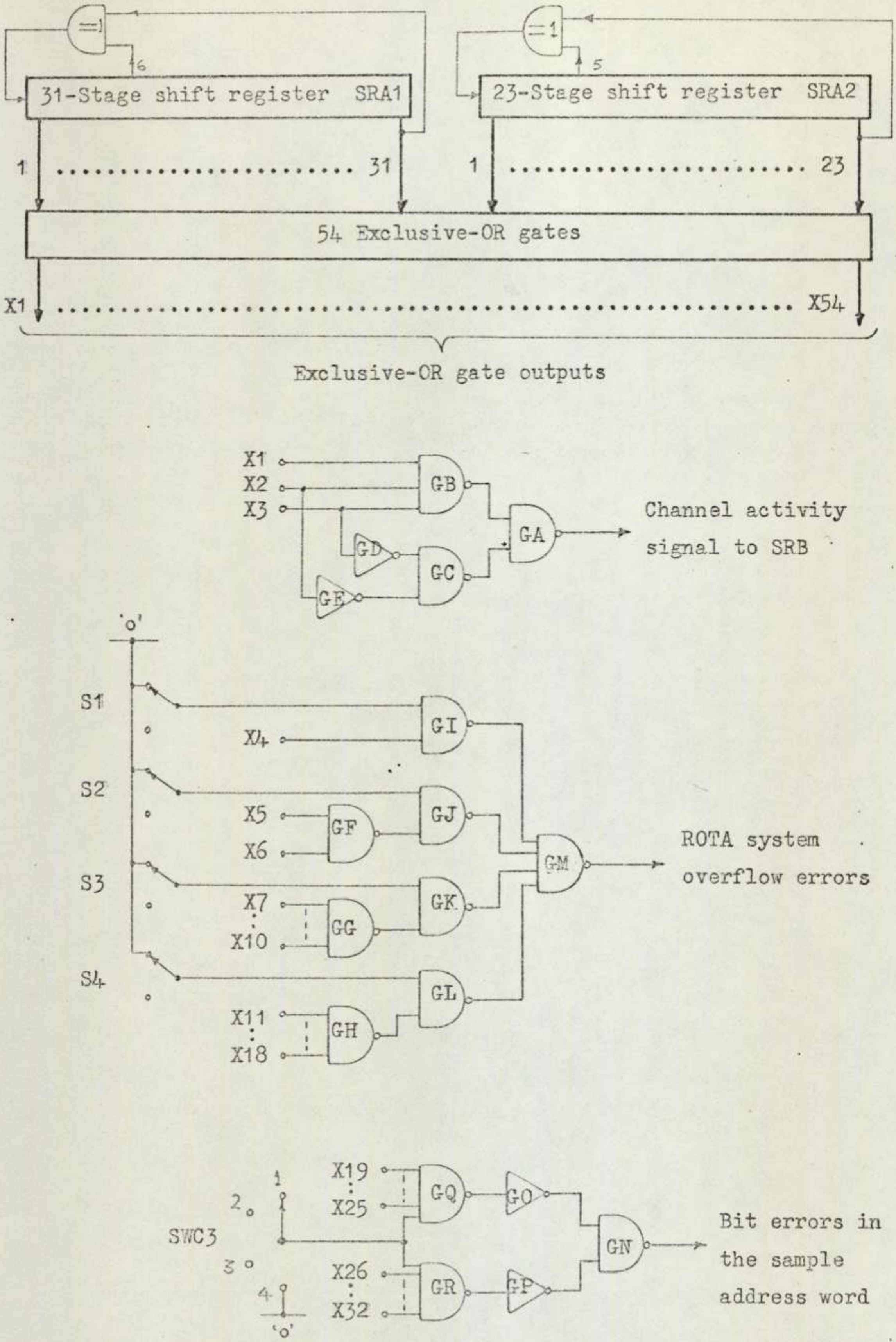


FIG. 6.16 Generation of random signals

a) Generation of the channel-activity signal

This signal is generated by gates GA to GE (see FIG. 6.16) in conjunction with three pseudo-random signals from the exclusive-OR gates. The channel-activity signal from GA is connected to the input of SRB in the active-channel generator.

The output from GA is a random signal with probability 0.375 of being a '1'. This signal represents the probability of a channel being active (i.e. α). This value of α was derived from the experimental work on 7-bit encoding described in chapter 4. However, since that work was carried out, 8-bit encoding has become the accepted standard and is used in the simulation model. Since the quantization is finer with 8-bit encoding, the value of α is slightly higher. However, measurements made by Sciulli and Campanella[12] for 8-bit encoding on practical telephone channels indicate that a value of 0.375 for α is a conservative figure.

Referring to FIG. 6.16, gate GB output becomes a '0' when all three inputs are at a '1'; this occurs with a probability of $(0.5)^3 = 0.125$. Similarly, gate GC output becomes a '0' with a probability of $(0.5)^2 = 0.25$. Since GB and GC detect mutually-exclusive events, the probability of GB <u>or</u> GC output being a '0'is 0.125 + 0.25 = 0.375.

The output of gate GA becomes a '1' when either of its inputs become a '0'. Hence, the probability of a '1' at gate GA output is 0.375.

b) Generation of the frame-overflow signal for the ROTA system

A random signal representing the occurrence of frame-overflow in the ROTA system is generated by gates GF to GM (see FIG. 6.16). The errors are generated by detecting when several random signals from the exclusive-OR gates are a binary '1'. If n such signals are being examined, the probability of an error is $(0.5)^n$.

The size of n (and hence the error rate) is controlled by switches S1 to S4, which open or close gates GI to GL . These gates, in conjunction with gates GF to GH, each detect the occurrence of the " all '1's" condition in a specific number of random signals. The number of inputs to gates GF, GG and GH is binary weighted to minimise the number of gates and switches required. There are 8 inputs to GH, 4 inputs to GG, etc.

Thus n can be varied between 1 and 15 by using the switches. This varies the probability of error between 0.5 and $(0.5)^{15}$ in multiples of 0.5.

c) Generation of bit errors in the sample address word

In an N-channel system there are n bits in the sample address word (pre-addressing). The probability of one bit-error in the sample address word can be calculated as follows. Let the probability of a bit-error in the transmission path be Pe. Also, assume that the occurence of an error is statistically independent of any previous errors. Then, the probability of one bit-error in the sample address word is given by :

$$P_{error} = \binom{N}{1} P_e (1-P_e)^{N-1} = NP_e (1-P_e)^{N-1}$$

If P_e is very small, then

$$P_{error} \doteqdot NP_e \qquad (6.3)$$

If the errors occurred in the sample address word, this would not be detected by the single-digit parity check at the receiving terminal. However, provided P_e is very small, the probability of two or more errors is negligible compared with probability of one error.

Referring to FIG. 6.16, the errors are generated by gates GN to GR. The technique used is identical to that used for generating ROTA overflow-errors. Several random outputs from the exclusive-OR gates are fed into GQ and GR. When all the inputs are a binary '1', the output from gate GN is a '0'; this indicates that an error has occurred. The output from GN is a '1' for any other combination of input signals. The switch SWC3 inhibits gates GQ and GR when it is in position 4 (PCM). This prevents errors from being generated at the output of GN.

The number of inputs to gates GQ and GR, to produce the required error probability, was calculated from equation (6.3). P_e was chosen as 10^{-6}, which is typical of practical digital transmission systems used for high-quality telephony. To simplify the design of the error-generating circuits, N was assumed fixed at a value of 60. This assumption is not strictly valid, since N is variable in the model. However, the range of variation used in the subjective tests was limited. Also, for N greater than approximately 60, the overflow noise increases very rapidly and masks any noise produced by bit errors.

6.4.3 Active-channel generator

A detailed circuit diagram of the active-channel generator and the decoder is given in FIGs. 6.17 and 6.18. The basic operation of these two circuits was described in sections 6.2 and 6.3.4.

The active-channel generator consists of three sections :

a) Shift-register SRB.

b) Counter CA.

c) Reset system.

a) Shift-register SRB

SRB is divided into six smaller shift-registers (see FIG.6.17). The total number of stages in SRB is controlled by the switches SWA1 to SWA5. Each switch is associated with one section of SRB. When a switch is in position 2, the associated register is connected in as part of SRB. When the switch is in position 1, the register is bypassed and forms no part of SRB. Thus, the number of stages in SRB can be adjusted between 32 and 94 in steps of 2. This varies N between 33 and 95.

It will be noted that the shift-registers are not switched directly in and out of circuit. The switches SWA1 to SWA5 provide a d.c. signal which is applied to a group of logic gates. These gates can be seen in FIG. 6.17 immediately below each shift-register. If the d.c. signal applied to these gates is a binary '0', then the <u>output</u> from the associated shift-register is directed to the next shift-register stage. This effectively connects the shift-register in series as part of SRB. However, if the signal applied to the gates is a '1', then the <u>input</u> to the shift-register is fed to the next shift-register stage. Thus, the shift-register is bypassed and forms no part of SRB.

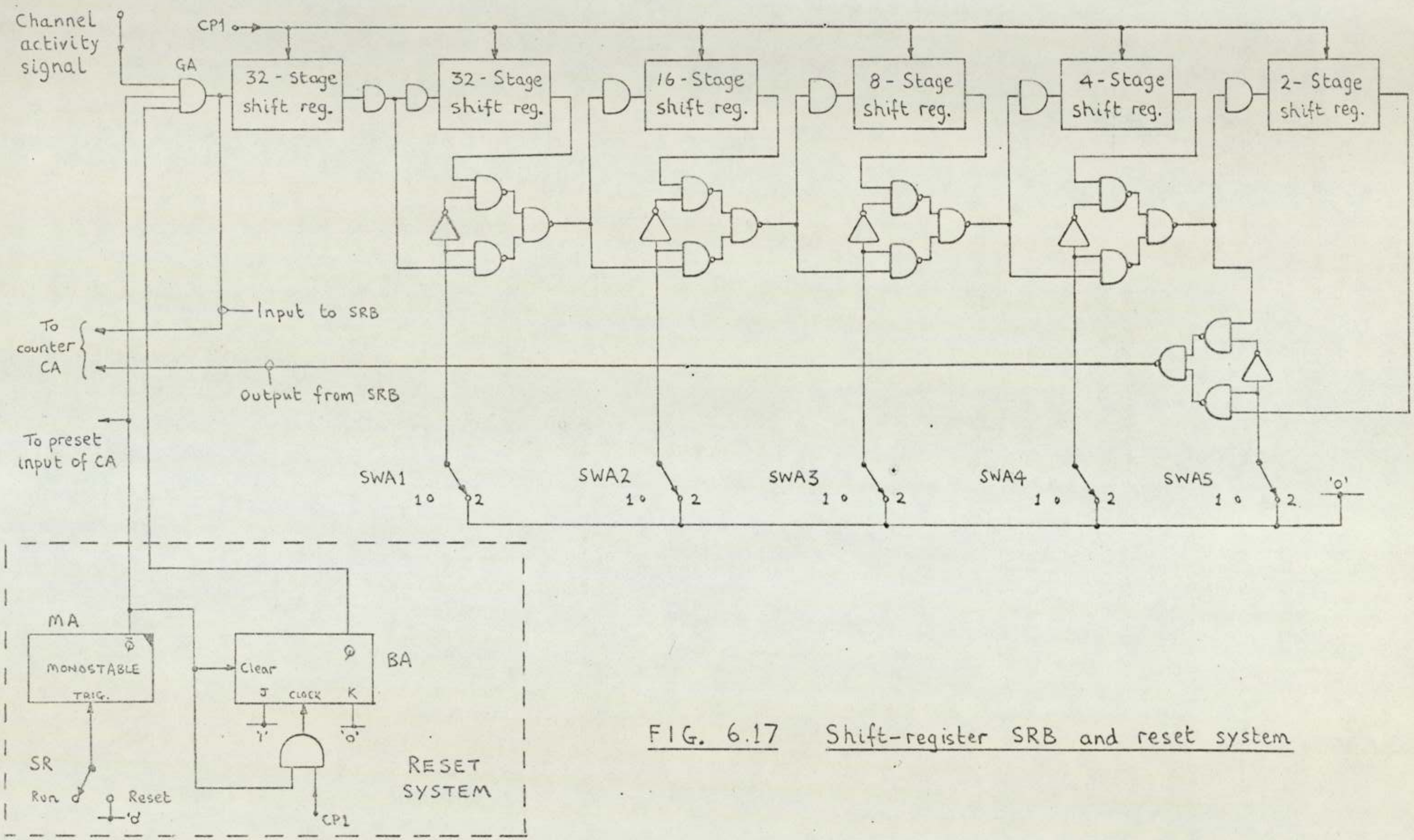


FIG. 6.17 Shift-register SRB and reset system

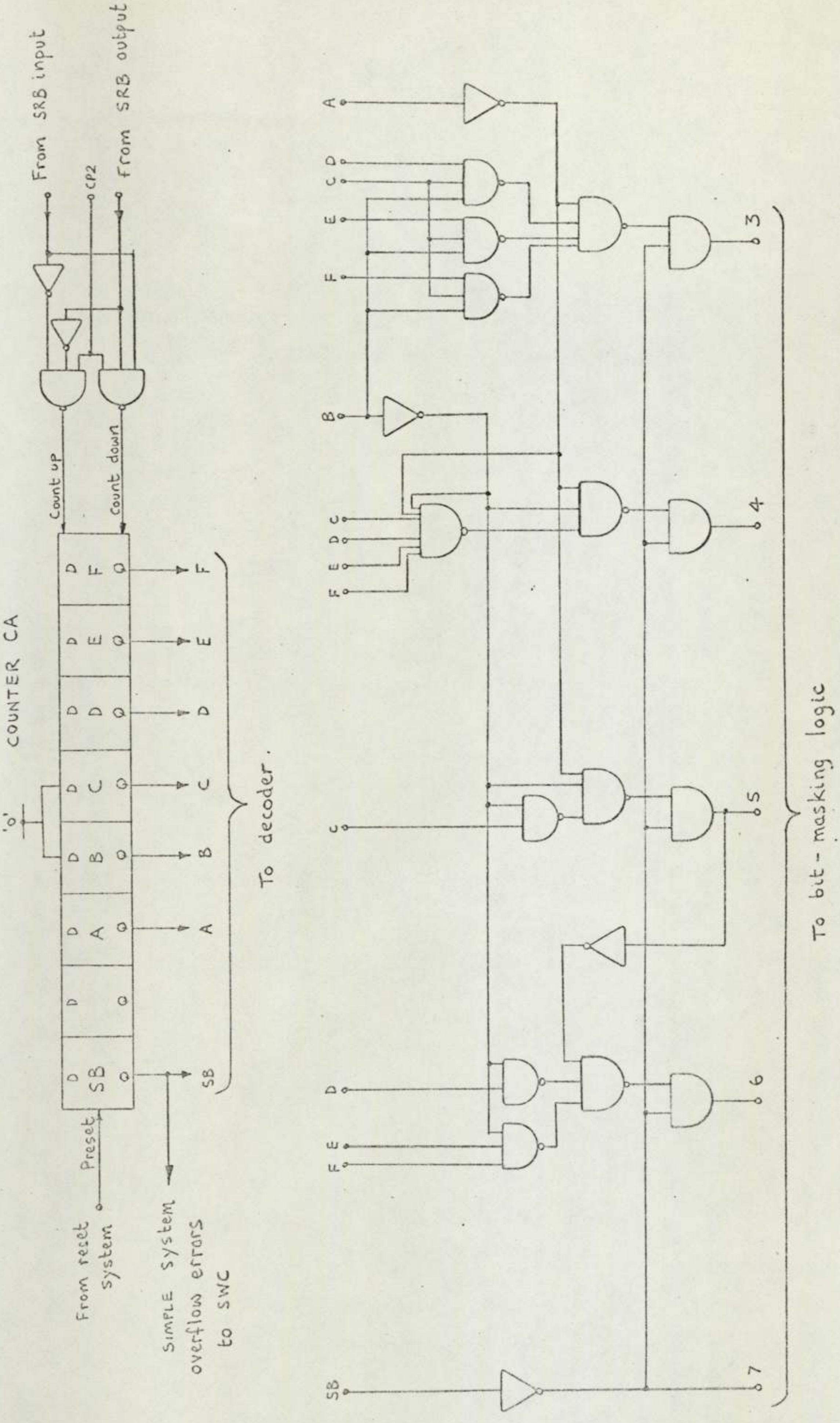


FIG. 6.18 Counter CA and decoder

b) Counter CA

The counter CA (see FIG. 6.18) is an 8-bit reversible counter, the contents of which can be preset to any initial value. The data inputs are permanently wired so that a pulse on the 'preset' input sets the contents of the counter to minus twenty five (i.e. minus K).

The logic gates to the right of CA ensure that the contents of SRB and CA remain synchronised. The frame pulse CP2 is "steered" to either the 'count/up' or 'count/down' input of CA or it is inhibited. The gates are controlled by the signals at the input and output of SRB. A truth table which describes the function of these gates is given below.

SRB input	SRB output	CP2 applied to
0	0	—
0	1	count/down
1	0	count/up
1	1	—

c) Reset system

The reset system initially synchronises the counter CA and shift-register SRB. This is necessary when the equipment is first switched on, or after the number of stages in SRB has been changed.

The system is operated when switch SR is briefly turned to the 'reset' position. This applies a signal to monostable multivibrator MA, which causes it to trigger. The $\overline{Q}$ output of MA becomes a '0', which closes gate GA. This causes binary '0's to be shifted into SRB. The timing period of MA is sufficiently long to ensure that SRB is completely filled with '0's.

The $\overline{Q}$ output of MA is also applied to the 'preset' input of CA. Thus, at the end of the reset cycle, there are no '1's in SRB and CA is preset to minus K.

While triggered, the $\overline{Q}$ output of MA also holds bistable BA in the 'clear' condition (i.e. Q = '0'). When MA releases, bistable BA remains 'cleared' until the next CP1 pulse 'sets' it. Thus, the Q output of BA holds GA closed for the period between the release of MA and the arrival of the next CP1 pulse. This ensures that the normal operation of SRB and CA commences at a convenient point in the frame period.

The decoder

The decoder (see FIG. 6.18) is used only when the VWL system is being simulated. This circuit is simply a combination of logic gates that examines the contents of CA. It computes the sample word length that is required to enable all samples to fit in the frame of transmitted pulses. There are five output lines from the decoder (marked 3,4,5,6 and 7) which are connected to the bit-masking logic. The relationship between the required sample word length and the signals on these lines is given in the table below.

Sample word length	7	6	5	4	3
8-bits	0	0	0	0	0
7 "	1	0	0	0	0
6 "	1	1	0	0	0
5 "	1	1	1	0	0
4 "	1	1	1	1	0
3 "	1	1	1	1	1

When the number of input channels N exceeds 66, it is possible that 2-bit encoding may be required. However, even when N is 95, the probability of the number of active channels exceeding 66 is negligible. For this reason no circuits were included in the decoder to accomodate the possibility of 2-bit encoding.

6.4.4 Overflow logic

The overflow logic, shown in FIG. 6.19, generates control signals for the ZAP gating and the one-sample store in the receiving terminal. The ZAP gating is used only when the model is simulating a system which uses zero-amplitude prediction. The type of prediction (i.e. ZAP or ZOP) is controlled by switch SWB. When SWB is in the 'ZAP' position, gate GB is opened. The control CZ is a '1' (ZAP gates open) when the speech detector is triggered (SD = '1') and the frame has not overflowed (OF='1'). If either SD or OF are a '0', then the ZAP gates are closed and a zero-valued sample is fed into the D/A converter. When SWB is in the 'ZOP' position, gate GB is closed and CZ is held at a '1'. Thus, the ZAP gates are held in the open state.

The control signal PL to the one-sample store in the receiving terminal is a pulse which, if undisturbed, occurs at frame frequency. If the pulse is present, the latest signal sample is loaded into the store. However, if PL is removed the store contents remain unchanged.

PL is generated from the frame-frequency pulse CP7. This is fed into gate GC together with the signal representing bit errors in the sample address word. When an error occurs, gate GC is closed. This removes pulse PL and causes a zero difference to occur in the receiving terminal.

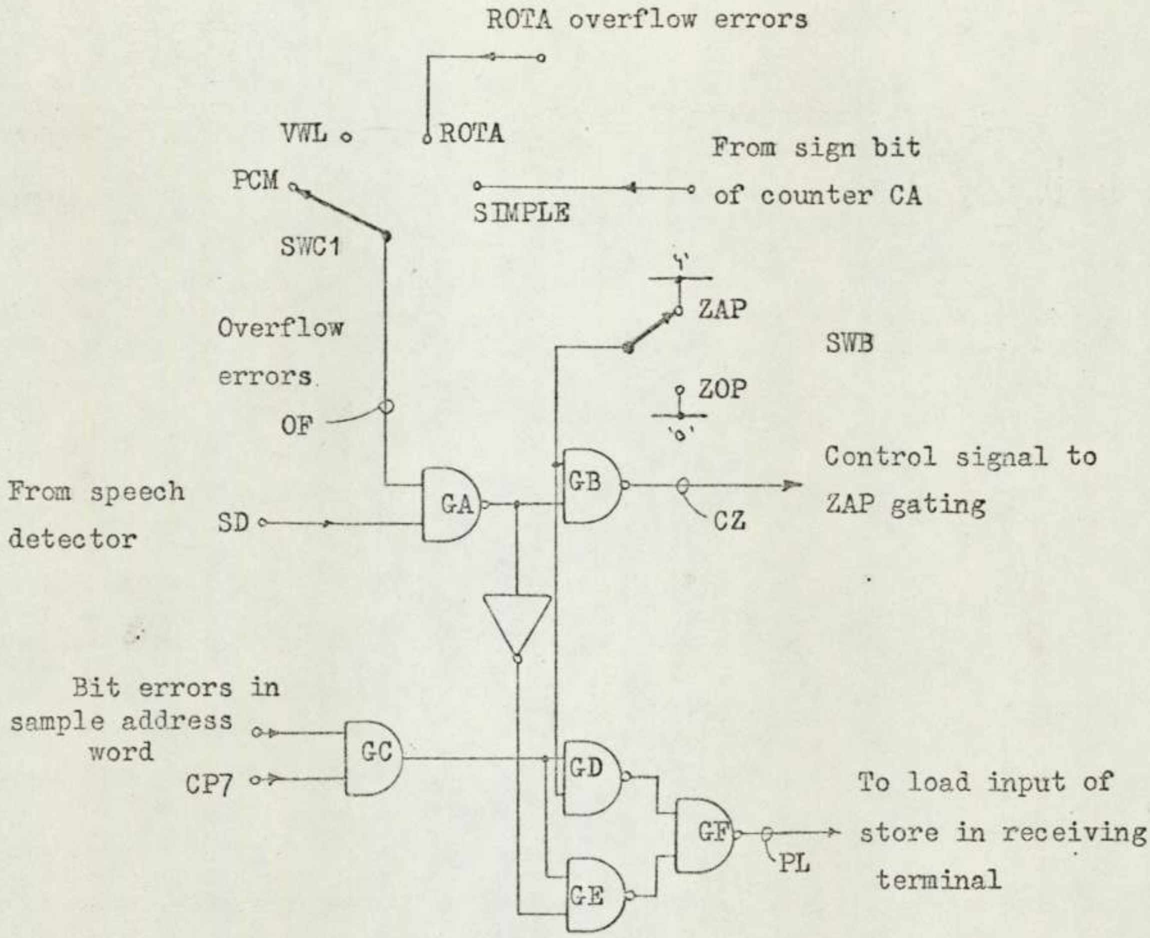


FIG. 6.19 Overflow logic

When SWB is in the 'ZOP' position, gate GD is closed and the pulses from GC have to pass via gate GE. Gate GE is closed (which removes PL) when either the frame has overflowed (OF = '0') or the speech detector is not triggered (SD = '0').

6.4.5 Speech detector

A block diagram of the speech detector is given in FIG. 6.20 and a circuit diagram in FIG. 6.21. The speech detector was designed using principles described by Fariello[18].

Speech is detected by two distinct mechanisms :

a) On an amplitude basis by the magnitude detector and sequence detector.

b) On a frequency basis by the sign-sequence detector.

The magnitude detector compares the analogue speech waveform with a fixed threshold voltage. When the threshold is exceeded, this is indicated to the sequence detector. If the threshold is exceeded in three consecutive frames, the sequence detector triggers the pulse generator; this indicates that speech has been detected. The inclusion of the sequence detector improves the immunity of the speech detector to noise. It prevents the detector from being triggered by occasional noise peaks.

The sign-sequence detector increases the sensitivity of the speech detector to the pure sibilant sounds (i.e. s and z). The frequency spectrum of these sounds is centred close to 4kHz; when present, they cause long sequences of alternately positive and negative samples from the A/D converter. These sequences are detected by the sign-sequence

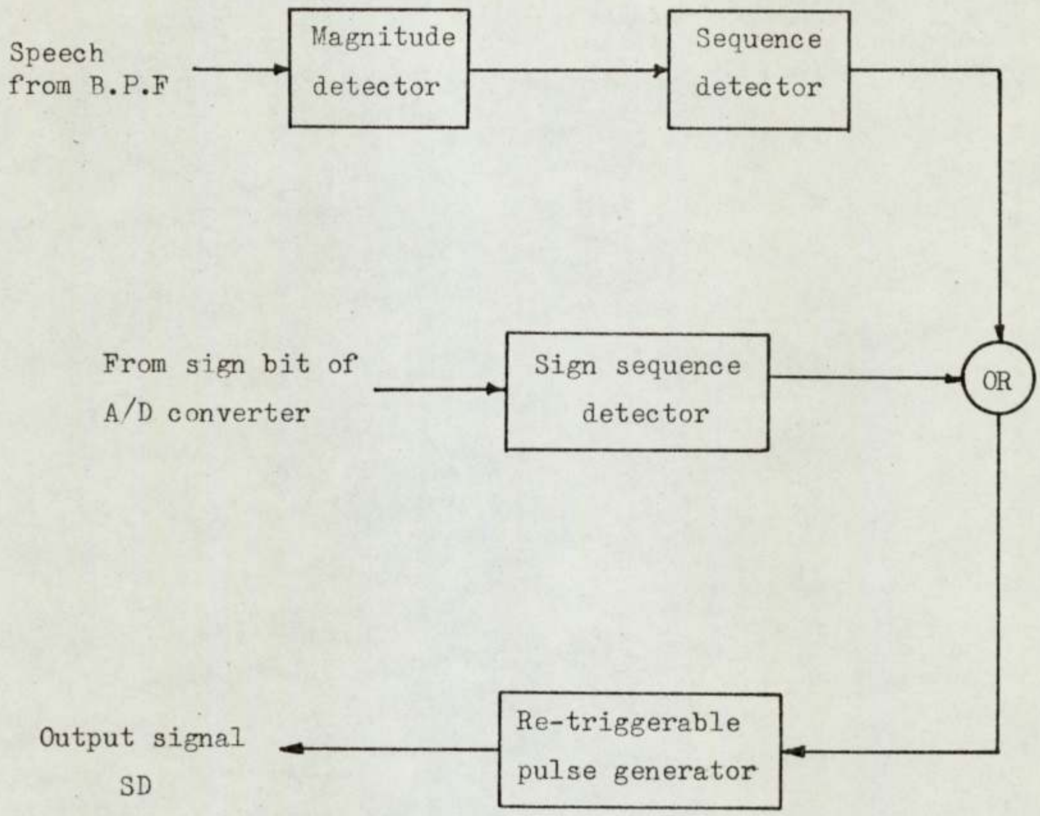


FIG. 6.20 Speech detector

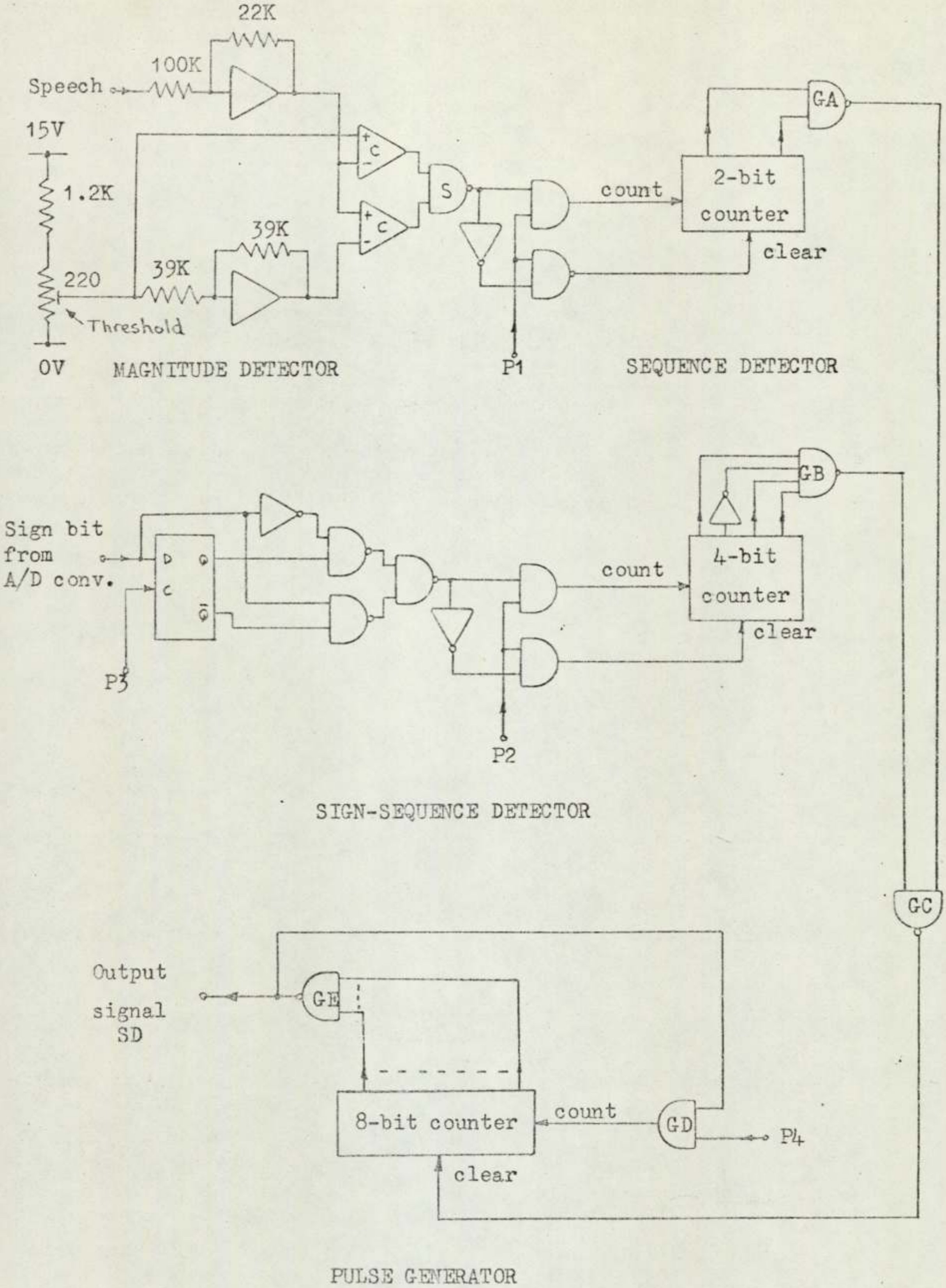


FIG. 6.21 Speech detector

detector. Noise, because of its random nature, very rarely produces these long sequences of sign reversals. Thus, noise triggering of the sign-sequence detector is negligible. The sign-sequence detector triggers the pulse generator after a sequence of 11 sign reversals.

In the magnitude detector, the analogue speech waveform is compared with the threshold voltage by the two comparators (see FIG. 6.21). One comparator checks the positive part of the speech waveform and the other comparator the negative part. If either comparator indicates that the threshold has been exceeded, the output of the Schmitt-trigger NAND gate goes to a '1'. The pulse P1 then increments the two-bit counter. If the threshold is not exceeded, P1 clears the counter. When the threshold has been exceeded in three consecutive frames, the contents of the two-bit counter is '11'. This causes the output from GA to go to a '0' which triggers the pulse generator via GC.

In the sign-sequence detector, the sign bit from the previous frame is stored in BA. This is compared with the sign bit in the current frame. If they differ (indicating a sign reversal), the pulse P2 is 'steered' to the 'count' input of the 4-bit counter. When eleven consecutive reversals occur, the contents of the counter cause the output of GB to go to a '0'; this triggers the pulse generator via GC.

The pulse generator consists of an 8-bit counter and two logic gates. The output from each stage of the counter is fed into gate GE. When the speech detector is not triggered, the counter is full (i.e. '11111111') and the output of GE (SD) is a '0'. The speech detector is triggered when a signal from GC clears the 8-bit counter. SD goes to a '1' which opens gate GD. The pulse P4 then increments the counter at 1ms. intervals. The time taken to fill the counter is 2^8ms unless the counter is cleared again during the timing cycle.

6.4.6 Clock pulse generator

The clock pulse generator provides all the control pulses for the various circuits in the simulation model. The pulses are all derived from a 1MHz master oscillator (see FIG. 6.22). The output of the oscillator is fed into a series of cascaded binary counters which act as dividers. The outputs from these dividers, lettered A to Q in FIG. 6.22, are then fed into logic gates to generate the required control pulses. A timing diagram showing the length and position of the control pulses is given in FIG. 6.23.

The frequency of the master oscillator is adjustable to allow the precise setting of the frame period (125μs).

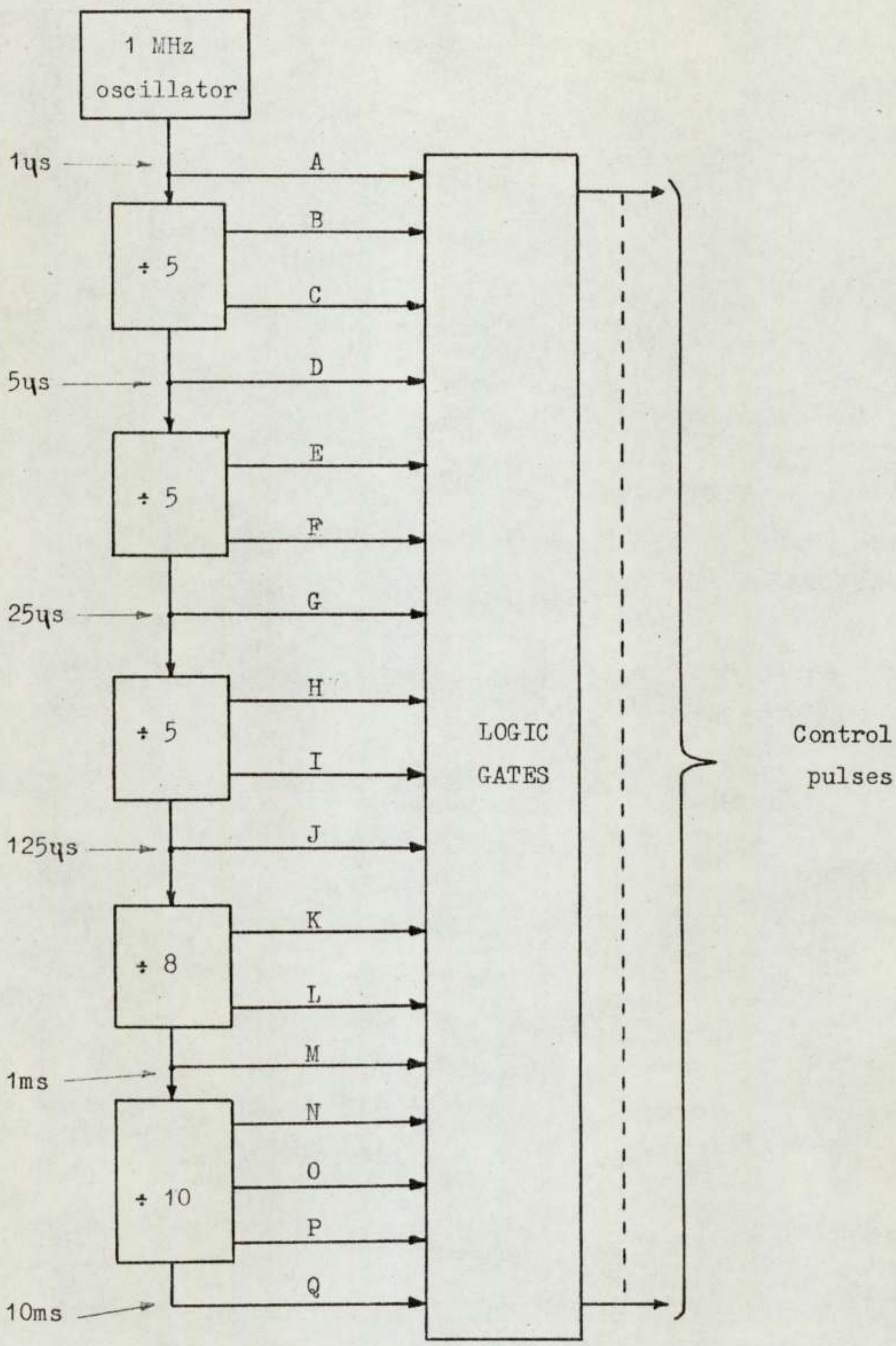


FIG. 6.22 Clock pulse generator

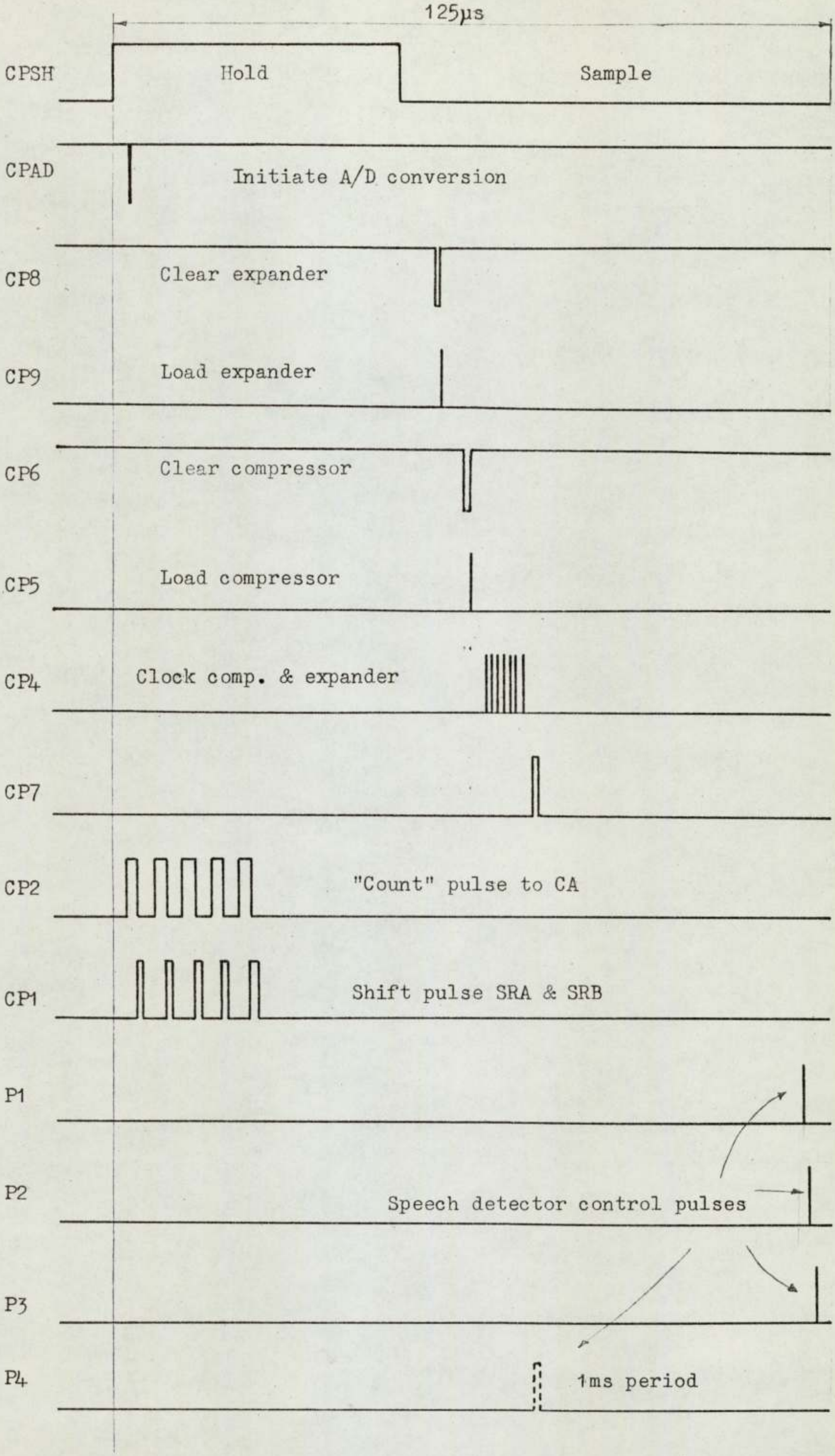


FIG. 6.23 Timing diagram of control pulses

6.5 Signal-to-overflow-noise ratio measurements

To confirm the theoretical calculations made in section 5.4, signal-to-overflow-noise ratio measurements were made on the simulation model.

Overflow-noise power is not easy to measure since overflow-noise only occurs when a signal is present. The method used for making the measurements is shown in FIG. 6.24. The problem of separating noise and signal was eased by using a sinusoidal signal. The noise power at the output of the model was measured by cancelling the signal in a summing amplifier. One input to the summing amplifier was the signal plus noise from the output of the model. The other input was the signal from the output of the low-pass filter in the sending terminal of the model; this signal is uncontaminated by overflow noise.

At a frequency of 1350 Hz, the signal inputs to the summing amplifier were in exact anti-phase. The resistor R was then adjusted for complete cancellation of the signal at the output of the summing amplifier. The power in the overflow noise was measured on the true r.m.s. reading voltmeter.

Signal-to-noise ratios were measured for the ZAP ROTA system and the VWL system. The results obtained are illustrated in FIG. 6.25, together with the theoretical values calculated in section 5.4. It can be seen that there is good agreement between the measured and theoretical values; although the measured values are consistently slightly greater. The probable explanation for these slight errors is as follows. In the theoretical analysis it was assumed that the baseband filters were ideal, with a cut-off frequency of 4kHz. Whereas, the filters in the model have a cut-off frequency of approximately 3.4kHz. Thus, the measured noise power is less than the theoretical prediction. The signal power is unaffected since all of the energy in the signal

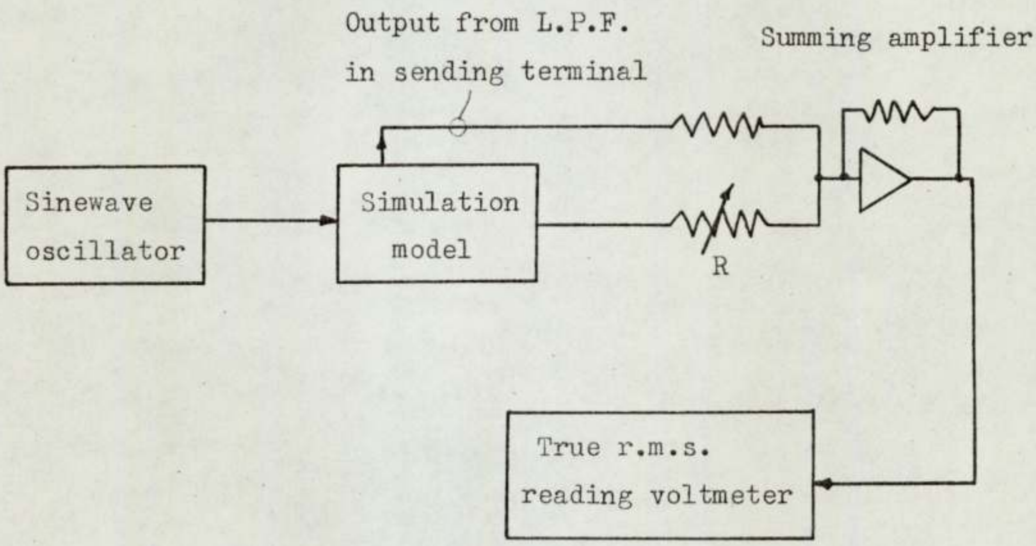


FIG. 6.24 Measurement of signal-to-overflow-noise ratio

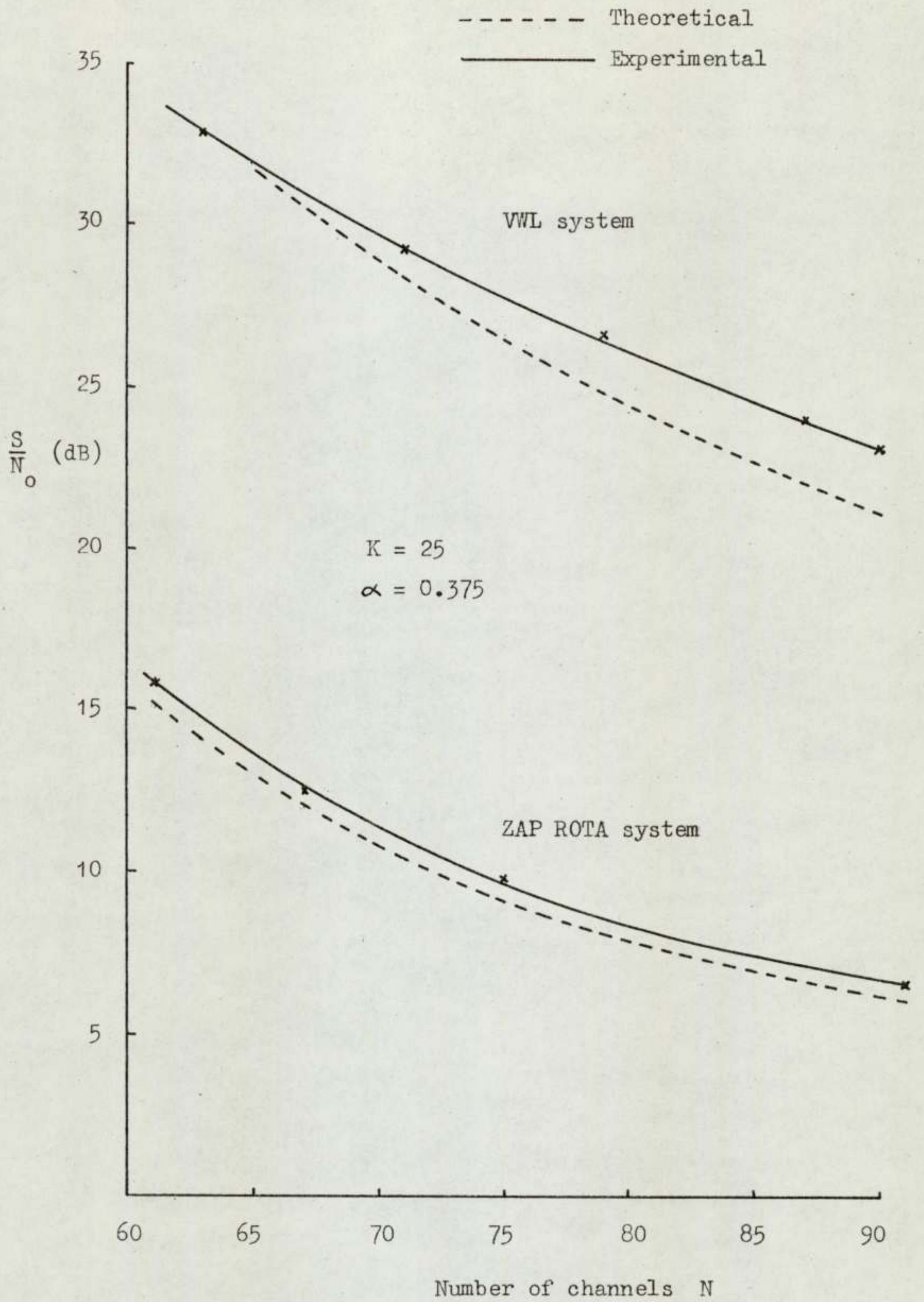


FIG. 6.25 Theoretical and experimental signal-to-noise ratios

is in the spectrum below 3.4kHz. This causes the measured value of signal-to-noise ratio to be slightly greater than the theoretical value.

C H A P T E R 7

SUBJECTIVE TESTS

SUBJECTIVE TESTS

7.1 Introduction

The performance of the frame-overflow-error distribution systems have been compared in terms of signal-to-noise ratio. This does not give a complete picture, however, since the distortions introduced are perceptually different. For example, the increase in quantizing noise introduced by the VWL system is different from the "clicking" noise caused by the ROTA system. For this reason, listeners opinions are also an important measure of performance. This is true for any voice communication system of course. In this chapter, listeners preference is used to compare the performance of the ROTA and VWL methods of overflow-error distribution.

The speech samples used in the evaluation were obtained by playing recorded sentences through the simulation model described in chapter 6. The output from the model was then re-recorded on magnetic tape. Then, listeners preference was measured as a function of the number of channels,using isopreference testing techniques.

7.2 Test procedure

The procedure outlined in this section is based upon IEEE recommended practice [23]. The object of an isopreference test is to measure a listener's opinion on the relative quality of two speech signals. This method basically assumes that speech quality can be measured based solely on listener's preference.

In an isopreference test, a listener compares a reference speech signal with a test speech signal. The test and reference signals are said to be isopreferent if a group of listeners are equally divided as to which signal they prefer. For the purpose of these tests the VWL system was designated the reference system.

A line connecting points representing isopreferent signals is called an isopreference contour. A typical isopreference contour is given in FIG. 7.1, where the variable parameters are N_R and N_T , and:

N_R = Number of channels to the reference system

N_T = Number of channels to the test system

To obtain the arbitrary unknown point A in FIG. 7.1, the procedure is as follows. The parameter N_R of the reference system is fixed at N_{R_0}. Then test signals, with N_T varying in a suitable range on either side of N_{T_0}, are paired with the reference. These pairs are then presented to the listener in a random order. The listener has then to decide which sample in each pair sounds better. If both samples sound equally good, the listener makes an arbitrary choice. The proportion of listeners preferring the reference signal is plotted against N_T and fitted with a smooth curve, as shown in FIG. 7.2. The 50 per-cent preference point defines the value of N_{T_0} in FIG. 7.1.

An alternative procedure to that outlined above is to keep N_T constant and vary N_R.

The isopreference testing technique was chosen after considering two other test procedures. These were:

a) RELATIVE PREFERENCE METHOD in which the quality of the test signal is measured against several reference signals representing different types of distortion.

b) CATEGORY JUDGEMENT METHOD in which the quality of the test signal is assigned to one of several categories such as "Excellent", "good" etc.

The relative preference method was not chosen because of its complexity compared with the isopreference method. The category judgement method was rejected, because it was thought that it would not give as sensitive a comparison between two speech signals as the isopreference method.

The principal advantages of the isopreference method are :

a) It is relatively simple to implement.

b) It compares the test and reference signals directly; this is important in this context, since we are mainly interested in comparing the subjective performance of the overflow-error distribution systems.

7.2.1 Obtaining speech samples

A block diagram of the method used to obtain the speech samples is given in FIG. 7.3. The tape on recorder A has short, phonetically balanced, sentences recorded on it (taken from ref. 23). The sentences were recorded in a quiet room on a FERROGRAPH type 632 tape recorder using a SHURE model 570 microphone. The two tape recorders shown in FIG. 7.3 are also FERROGRAPH type 632.

A preliminary listening test was carried out by the author to determine the approximate form of the isopreference contour. This aided the choice of a suitable range for N_T for each point on the isopreference contour.

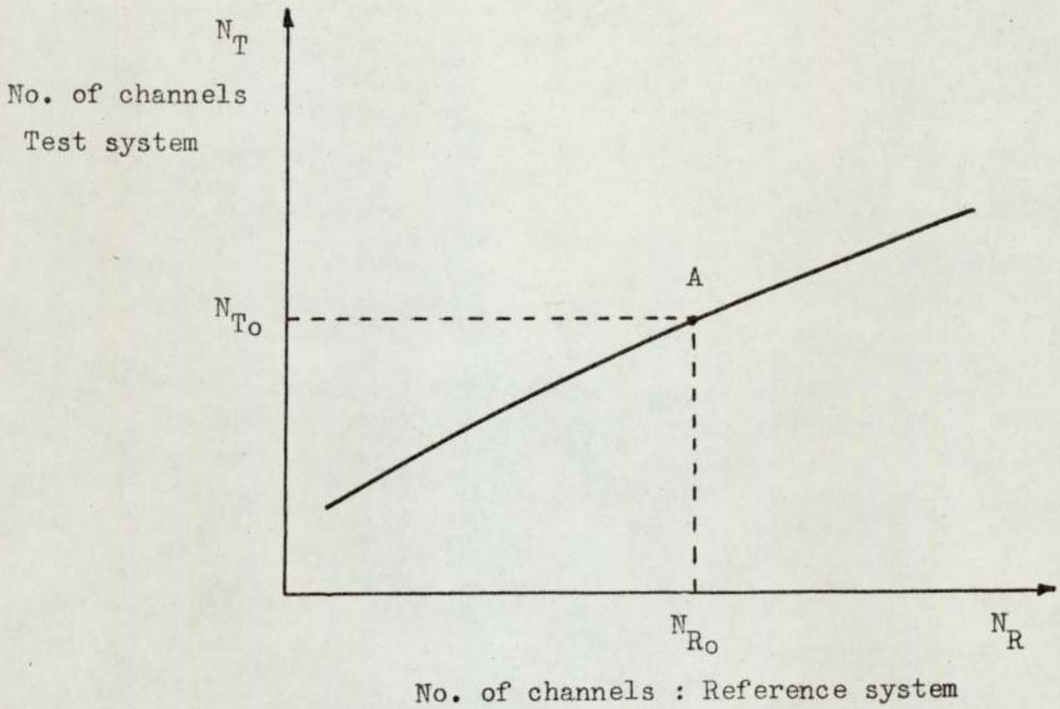


FIG. 7.1 A typical iso-preference contour

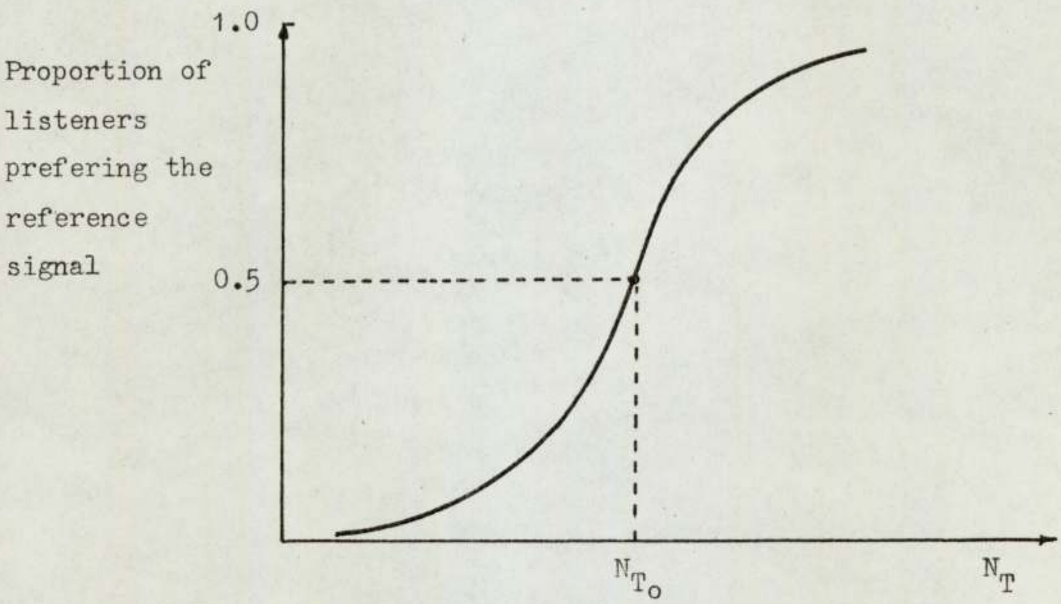


FIG. 7.2 Proportion of listeners prefering the reference signal

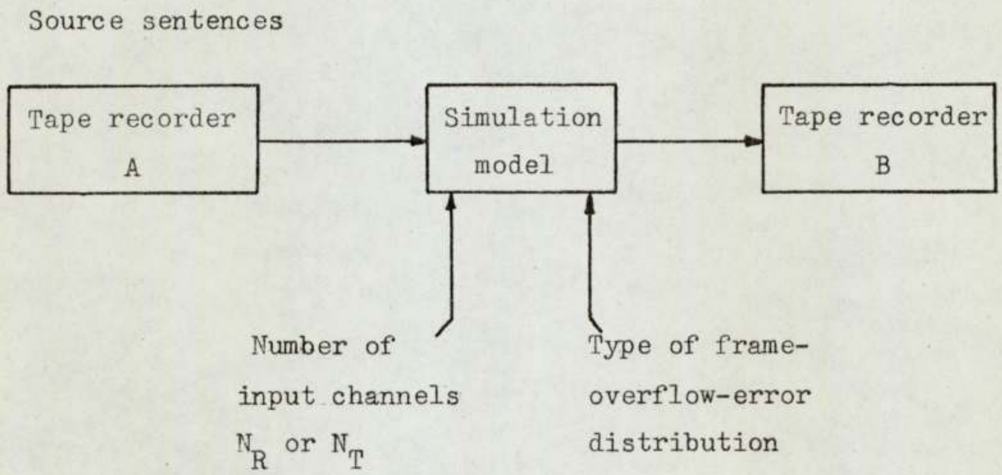


FIG. 7.3 Method of obtaining speech samples

Having decided upon a suitable range of sentence pairs, the order of them was randomised using a table of random numbers. The source sentences on tape recorder A, were then played back, one by one, through the hardware model and re-recorded on tape recorder B. The parameters of the model (No. of input channels and type of error distribution system) were adjusted for each sentence in accordance with the randomised list mentioned above. Each pair of sentences was repeated to give the listener a second chance to make a judgement. A gap of five seconds was allowed between each group of four sentences to give the listener time to record his preference.

7.2.2 Listeners and listening procedure

The listening group consisted of research students, laboratory technicians and secretaries. A total of twelve listeners were used for each test, although not necessarily the same twelve.

The listening tests were conducted in a quiet room. The test tapes were played back on a FERROGRAPH TYPE 632 tape recorder and BURR-BROWN TYPE K headphones were used for listening.

Before the initial speech tests, the volume control of the tape recorder was adjusted to give a speech volume that was comfortable to the author. Then each listener was asked to listen to two sentences played at the same volume. All listeners reported the volume as being suitable.

7.3 Results and discussion

The procedures detailed in section 7.2 were used to make the following tests :-

TEST A

Test system : ROTA system with zero-order prediction (ZOP) with N_T = 54, 57, 61 and 66

TEST B

Test system : ROTA system with zero-amplitude prediction (ZAP) with N_T = 51, 54 and 57

The isopreference contours obtained are illustrated in FIG.7.4. The horizontal solid bars indicate the 95% confidence limits which were calculated using small-sample theory. The sample variance (s^2) and sample mean were first estimated using a graphical method. Then the confidence limits were calculated from the formula :

$$\overline{X} + \frac{ts}{n} < u < \overline{X} - \frac{ts}{n} \qquad (7.1)$$

where - $\overline{X}$ = sample mean

n = sample size

u = population mean

t = From a table of t-distribution corresponding to 5% confidence.

The results obtained are slightly unsatisfactory in that they are not clearly defined. The cause of this can be attributed to one or more of the following reasons.

a) Insufficient number of listeners

b) The perceptual difference in distortions caused by the reference and test systems. For example, the increase in quantising noise produced by the VWL system is quite different from the "clicking" caused by the ZAP ROTA system.

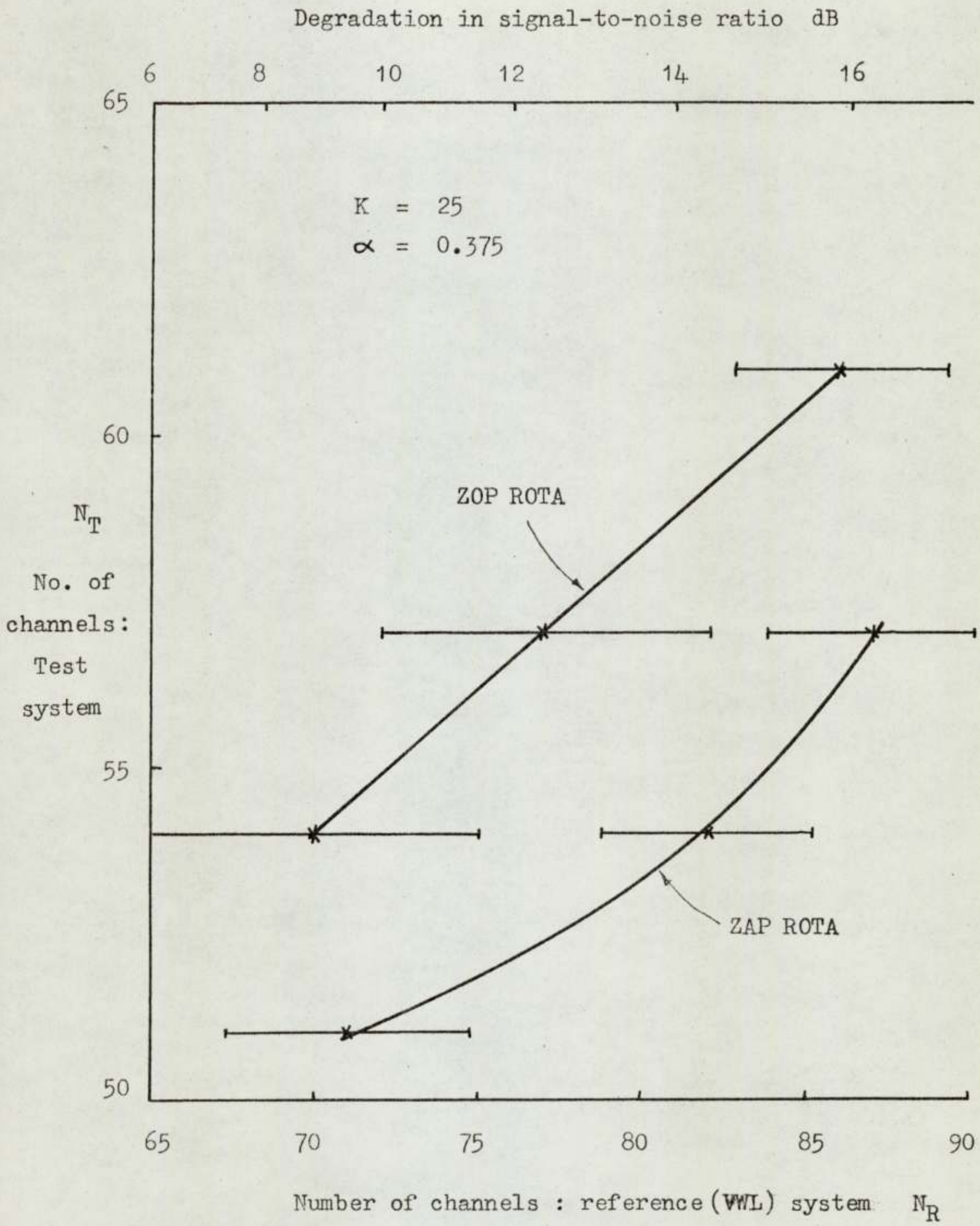


FIG. 7.4 Experimentally determined isopreference contours

Thus, a listener's judgement is swayed depending upon which type of distortion he prefers.

c) The shape and slope of the isopreference contour.

d) Using different source sentences for each test.

It is probable that a) and b) above are the major causes of the "spread" in the statistical data obtained.

To test the hypothesis proposed in d), test A was repeated using the same source sentence for every test. The results obtained are not significantly different from those obtained using different source sentences.

FIG. 7.4 shows the superior performance of the VWL system compared with both versions of the ROTA system. As an example from FIG. 7.4, 70 channels on the VWL system produces the same subjective quality as only 54 channels on the ZOP ROTA system and 51 channels on the ZAP ROTA system.

The scale at the top of FIG. 7.4 is the theoretical degradation in signal-to-noise ratio, caused by frame overflow in the VWL system, corresponding to the number of channels on the scale at the bottom. Degradation is defined as the amount by which the overall signal-to-noise is reduced when frame overflow occurs. The upper scale has been included to show the wide range of quality encompassed by the plotted curves. For N_R equal to 85, the noise introduced is very noticeable; whereas for N_R equal to 65, a discerning listener would not notice any increase in noise.

The degree of bandwidth economy that can be achieved depends upon the level of noise that can be tolerated. This, in turn, depends upon where the system is to be applied. Let us assume that the system is to be used in a high-quality telephone network, and that a signal-to-noise ratio of 30dB is acceptable (or equivalent subjective quality). Also, assume 8-bit encoding is used together with the same companding characteristic as in the simulation model. Under these conditions, the signal-to-quantizing-noise ratio with no frame overflow (i.e. conventional PCM) is 38dB [24]. Thus, a degradation of 38 - 30 = 8dB in signal-to-noise ratio is tolerable.

Referring to FIG. 7.4, 8dB of degradation corresponds to approximately 68 channels on the VWL system. Also from FIG. 7.4, for an equivalent subjective quality on the ZOP ROTA and ZAP ROTA systems, 53 and 50 channels respectively can be accommodated. The above data is summarised in TABLE 7.1 below, together with the bandwidth compression ratio R.

TABLE 7.1

Overflow-error distribution system	Number of channels N	Bandwidth compression ratio R
VWL	68	2.0
ZOP ROTA	53	1.65
ZAP ROTA	50	1.56

Subjectively equivalent performance of the VWL and ROTA systems for $S/(N_o + N_q) = 30dB$ (VWL system).

Theoretically, the ZOP ROTA system can accomodate approximately 47 channels for $S/(N_o + N_q) = 30dB$ (see section 5.4). Similarly, the ZAP ROTA system can accept approximately 46 channels. Comparing these figures with those given in TABLE 7.1 suggests that the noise produced by frame overflow in the ROTA systems is subjectively less disturbing than an equal amount of noise in the VWL system.

The value of R in TABLE 7.1 was calculated from

$$R = \frac{1}{\frac{1}{a} + \frac{K}{N}} .$$

This equation was derived in section 2.2.

The ideas used in deriving the data in TABLE 7.1 can be extended to construct curves of R as a function of signal-to-noise degradation, as shown in FIG. 7.5. Each point on these curves was derived in the same way as the data in TABLE 7.1. An important feature illustrated in FIG. 7.5 is that the VWL system increases its advantage over the two ROTA systems as the level of noise increases.

The values of R calculated above are the bandwidth compression ratios achieved compared with 8-bit PCM. However, the speech quality is degraded compared with PCM, so this method of calculating R is biased. As an example, suppose that when using the VWL system the SNR has been degraded by 6db. With PCM a SNR reduction of 6db is equivalent to removing one bit from each encoded sample. These ideas lead to the concept of effective bandwidth compression ratio (R_{eff}) where in the example quoted above :

$$R_{eff} = R. \frac{7}{8}$$

Similarly, for a 12db reduction in SNR :

$$R_{eff} = R. \frac{6}{8} \quad \text{and so on.}$$

This assumes that the increase in noise due to frame overflow in the VWL system, is subjectively the same as an identical increase in quantizing noise in simple PCM. R_{eff} is the bandwidth compression ratio of the proposed system compared with a PCM system which gives the same subjective quality.

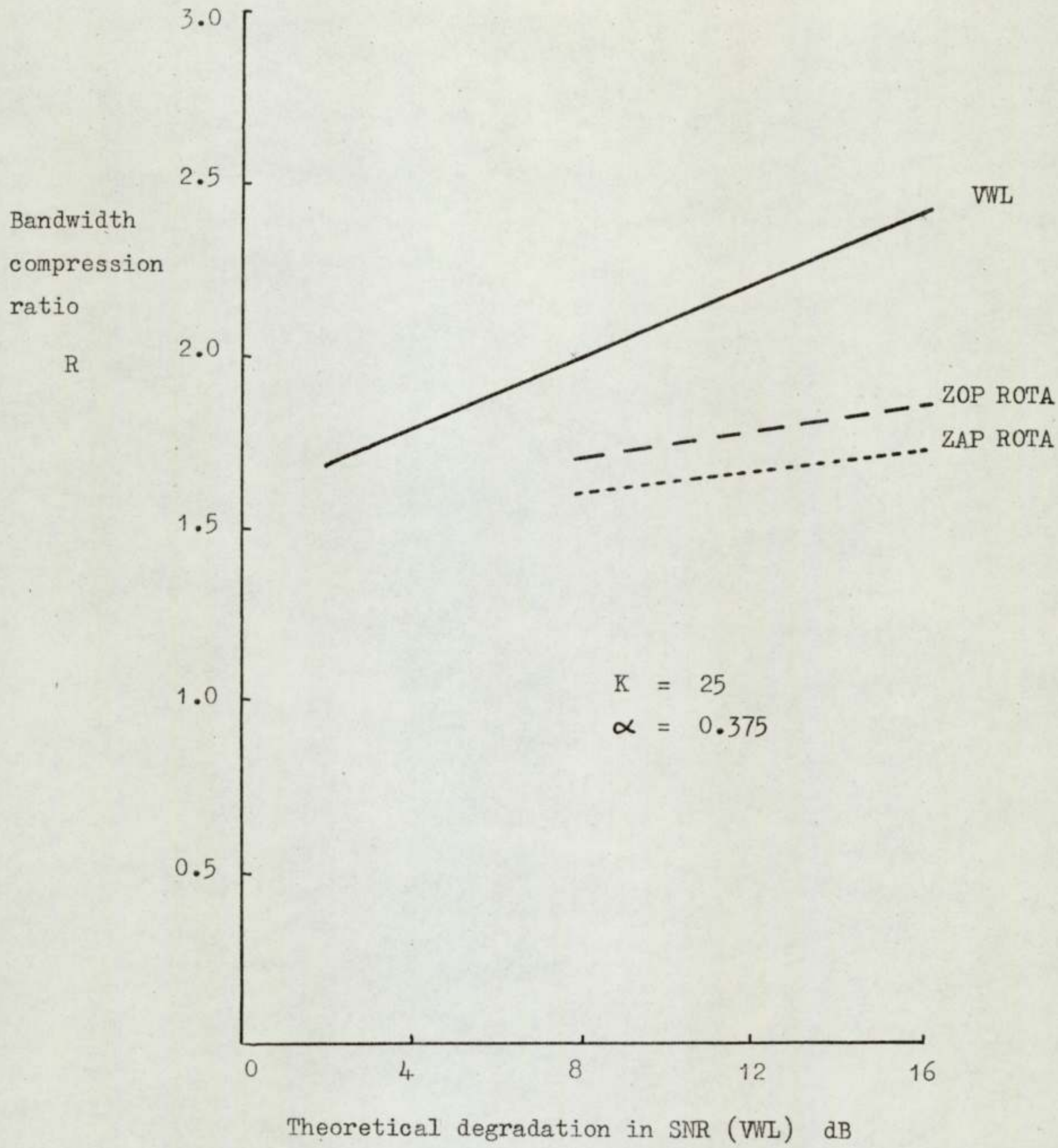


FIG. 7.5 Bandwidth compression ratio as a function of speech quality

This figure shows the variation of R with speech quality. The theoretical degradation in signal-to-noise ratio for the VWL system is used as a speech-quality reference.

Using the above ideas, curves of R_{eff} against degradation in SNR are given in FIG. 7.6. It can be seen that the performance of the VWL system (in terms of bandwidth compression) is fairly constant over a wide range of subjective quality. However, there is an optimum point at around 6db reduction in SNR. The performance of the two ROTA systems falls off rapidly with increasing noise and here again there is probably an optimum, somewhere just below 6db.

The concept of effective bandwidth compression ratio is useful, but it has a drawback ; it takes no account of the adaptive nature of the proposed system. The hardware model, on which the subjective tests have been made, was designed assuming all N-channels had a call on them. This is a worst-case condition and only occurs for short periods (i.e. the "busy hour"). Thus, the subjective quality, as measured in this chapter, is the worst that will occur. In a practical system the quality will be much better for most of the time, when the call-load is lighter. A simple PCM system does not adapt itself in this manner.

The SIMPLE system was not subjectively tested because of time limitations and also it is unlikely that this system will have any practical value. Also the ERROR GROUPING system could not be subjectively tested by the isopreference method because of the random nature of the speech clipping caused. For a short sentence the percentage speech clipping could vary from zero to some substantial value. Clearly this system needs several tens of minutes or even hours of speech to obtain a true measure of its subjective performance. However, the ERROR GROUPING system operates in a similar manner to T.A.S.I., the subjective performance of which is well documented [2,3].

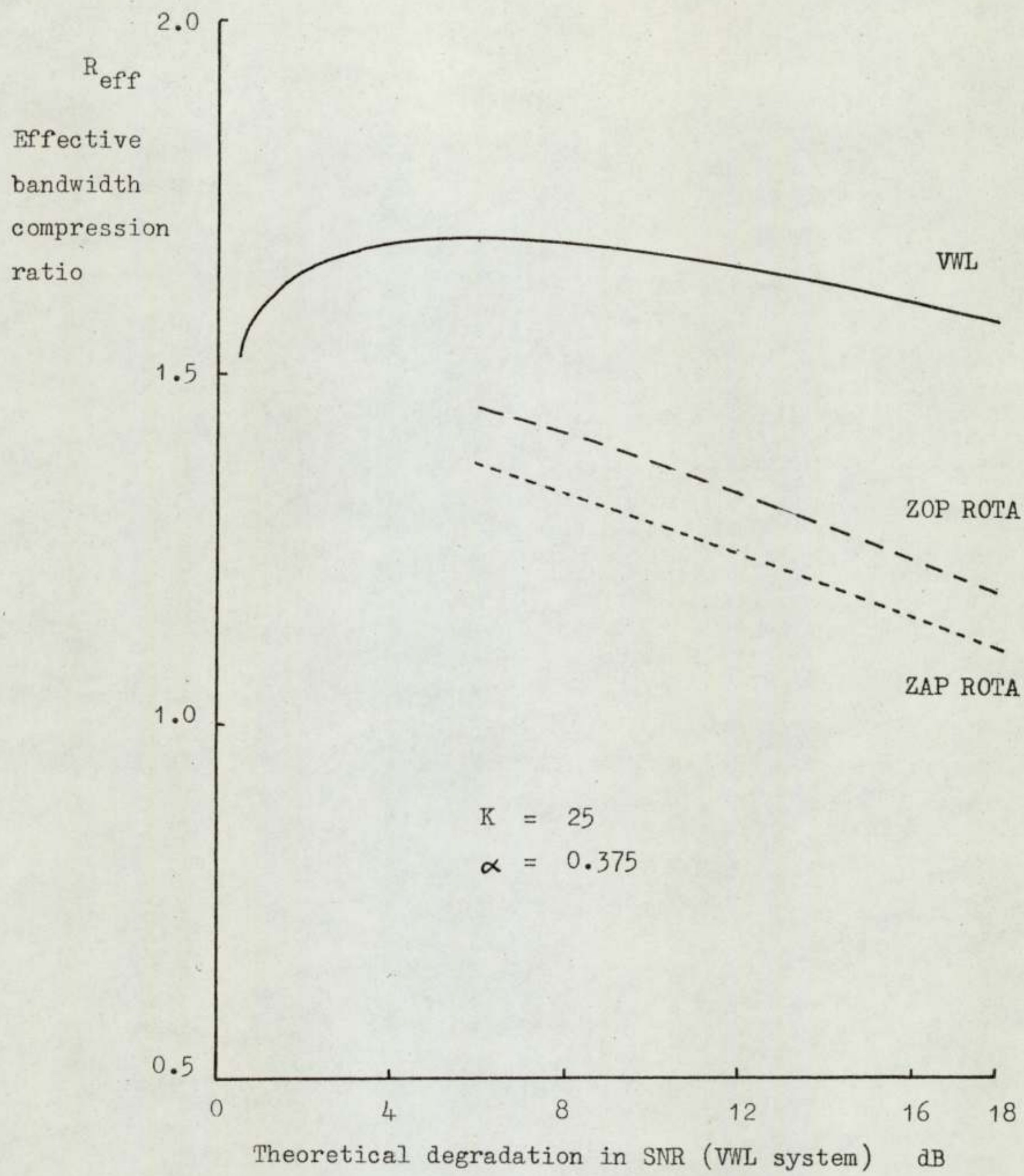


FIG. 7.6 Effective bandwidth compression as a function of speech quality

The theoretical degradation in SNR for the VWL system (or equivalent subjective quality), is the speech quality reference scale in this figure.

CHAPTER 8

COMPARISON OF THE PROPOSED MULTIPLEX TRANSMISSION SYSTEM WITH TIME ASSIGNMENT SPEECH INTERPOLATION (TASI) SYSTEMS

COMPARISON OF THE PROPOSED MULTIPLEX TRANSMISSION SYSTEM WITH TIME ASSIGNMENT SPEECH INTERPOLATION (TASI) SYSTEMS

8.1 Introduction

It is interesting to compare and contrast the proposed multiplex transmission system described in this thesis with Time Assignment Speech Interpolation (TASI) systems. Both systems achieve a bandwidth economy by eliminating redundant information from the speech signal. In a TASI system, a subscriber is connected to a transmission channel only when he speaks. In normal telephone conversations a subscriber speaks for less than half the time. This means that TASI can approximately double the number of channels for the same bandwidth.

The original TASI system[1-3] was applied to an analogue transmission system and was very complex. The TASI principle has more recently been applied to digital multiplex transmission systems[4,5]. The digital TASI system is far simpler than its analogue counterpart and recent advances in digital integrated circuit technology have made it easier to implement.

8.2 Redundancy removal

The TASI system only removes the redundancy associated with silent pauses in speech. The proposed system described in this thesis attempts to obtain greater bandwidth economy by eliminating redundancy in both silent pauses and active speech. In a digital TASI system speech is divided into silences and active speech by a speech detection device. All samples that occur in silences are classed as redundant and the remaining samples are non-redundant. Similarly, in the proposed system, samples that occur in silent pauses are classed as redundant.

However, there are additional samples that occur during active speech that are also classed as redundant.

8.3 Sample addressing

A digital TASI system, like the proposed system, has to transmit address information so that samples can be correctly routed at the receiving terminal. The technique used is very similar to the pre-addressing method[4]. For an N-channel system, an N-bit sample address word is transmitted. Each bit of the word is associated with one channel and indicates whether the channel is active or not. The time between changes in channel activity is very large compared with the frame period. Typically, a talkspurt is approximately one second in length and so spans 8000 frames. Consequently, it is not necessary to "update" the address information at the receiving terminal very frequently. For this reason, the transmission of each sample address word is distributed over several frames with a consequent saving of frame bits. This type of sample addressing causes the following effects however :

a) When a channel becomes active, there is a time delay before it is assigned a timeslot. As a result each talkspurt is clipped slightly. This effect places an upper limit on the number of frames over which the sample address word can be transmitted. However, it is not a serious limitation since a clip of several milliseconds can be tolerated.

b) Errors that occur in the sample address word during transmission persist for several frames (i.e. until the transmission of the next sample address word). However, the economy in frame bits makes it feasible to protect the sample address word against errors by trasmitting it in a redundant form [4].

In the proposed system, the channel activity can change frame by frame ; this makes it necessary to transmit address information during every frame. Thus, although the proposed system removes more redundancy from the speech signal than TASI, more address information is needed to make use of it.

8.4 Frame overflow

In a digital TASI system, when a channel becomes active it is given access to a timeslot if there is one available. If, due to frame overflow, there is no spare timeslot the channel waits until one becomes available. When a channel is given access to a timeslot, it retains this access until it becomes inactive again. Thus, frame overflow causes clipping of talkspurts. This is similar to the ERROR GROUPING method of frame-overflow error-distribution described in section 5.3.2.

The possibility of applying the overflow-error distribution methods described in section 5.3 to a digital TASI system was investigated. If the ROTA system were used, a frame overflow would cause one or more channels to lose access to a timeslot. Furthemore, these channels could not regain a timeslot assignment until the next sample address had been transmitted. Thus, these channels would lose several consecutive samples. The received speech would have segments "chopped" from it which would be most disturbing to the listener.

The VWL system, on the other hand, would not suffer from the same disadvantage. This system does not delete complete samples when the frame has overflowed. Instead the sample word length is reduced and there would be no "chopping" of the speech signal. Also, there is evidence that the application of the VWL technique to a digital TASI system would result in an improved bandwidth compression ratio. Consider a digital TASI system with the following parameters:

a) 25 timeslots per frame available for samples

b) 8-bit encoding

c) 8 bits per frame for sample addressing

d) The activity factor $\alpha = \tau_A = 0.4$

e) The VWL method of overflow-error distribution

From equation (5.16), for a signal-to-noise ratio of 30 dB during the busy hour, this system could accept 64 channels. The bandwidth compression ratio R is defined by:

$$R = \frac{\text{No. of bits per frame for PCM system}}{\text{No. of bits per frame for VWL-TASI system}}$$

Then,

$$R = \frac{64 \times 8}{25 \times 8 + 8} = 2.46$$

The value of R for a conventional TASI system is only 2.0.

CHAPTER 9

CONCLUSIONS

CONCLUSIONS

9.1 General conclusions

This thesis has described the investigation of a digital multiplex transmission system intended for telephony applications. The object of the system is to obtain a digit-rate reduction compared with conventional PCM, while maintaining good voice quality. The digit-rate is reduced by eliminating redundant samples from the sampled and quantized speech signal. A sample is defined as redundant when it is identical to the immediately previous sample; or, alternatively, when it is zero-valued. Since the number of samples transmitted per frame is less than the number of baseband channels, it is necessary also to transmit address information to enable the receiver to route each incoming sample to the correct output channel.

When the investigation was begun, this proposed system was believed to be original and a brief description of it was published [11]. When the work was nearly complete, a paper by Sciulli and Campanella [12] appeared which described a similar system developed in the COMSAT laboratories. This system is called SPEC (Speech Predictive Encoding Communication system). It is encouraging to find that the research laboratories of a major telecommunications organisation has independently been working on similar lines. However, it appears from the paper by Sciulli and Campanella that COMSAT have investigated only one of the modes of operating the system, and have not built and tested an experimental model. There are also differences in the theoretical analysis.

Clearly, for the system to be efficient in reducing the digit-rate, the probability of zero-difference (or zero-amplitude) must be relatively large. Since the number of samples that can be transmitted in each frame is less than the number of baseband channels, there is a possibility that the number of channels that are active (i.e. having non-redundant samples) may exceed the number of samples that can be transmitted. This is known as frame overflow. It was therefore necessary to measure the relevant speech statistics for typical telephone conversations in order to determine the probability of overflow occuring.

9.2 Measurements on conversational speech

The probability of zero-difference for conversational speech was measured with the aid of an A/D converter and a digital computer. It was noted during the experiments that noise in silent pauses significantly affected the results. Consequently, a speech detector was fitted in the speech-signal path to eliminate this effect. This device rejects noise in the absence of speech, but allows speech to pass. The need for such a device was also noted by Dorgello and van der Veer [7], and Sciulli and Campanella [12].

The probability of zero-difference can be divided into two fractions. First, there is the fraction of the total time that the talker is silent; in this case all of the samples are redundant. Secondly, there is the probability of zero-difference during active (detected) speech C. An analysis of several hours of recorded conversations that had been linearly encoded into a 7-bit code, yielded an average value of approximately 0.29 for C.

The probability of zero-difference is related to the instantaneous slope or rate of change of the speech waveform. The probability density function (pdf) of slope was obtained for several talkers, using a digital computer and A/D converter. The shape of the slope pdf is very similar to the instantaneous amplitude pdf i.e. a large central "peak" and long "tail".

9.3 Companding systems

Practical telephone digital transmission systems require some form of companding to accommodate variations in talker volumes. To observe the effect on the value of C, the following companding systems were simulated using a digital computer and A/D converter.

a) Instantaneous companding

b) Syllabic companding

It was discovered that both types of companding drastically reduced the value of C, compared to the linear encoding case. The average value of C obtained was 0.073 for instantaneous companding and 0.084 for syllabic companding.

On average, a telephone subscriber speaks for only 40% of the time [1]; then the probability, $P(d_0)$, of zero-difference for 7-bit encoding with instantaneous companding is $0.6 + 0.4C \doteqdot 0.629$. Similarly, the probability of zero-difference for syllabic companding is 0.626. Thus, the extra complexity of the syllabic companding system is not justified by a significant increase in $P(d_0)$.

The probability of zero-amplitude was also measured for both of the companding systems. The average value obtained was only slightly smaller than $P(d_0)$.

9.4 Sample addressing

To ensure the correct routing of samples at the receiving terminal, it is necessary to transmit additional "address" bits in each frame. Two methods of identifying transmitted samples were investigated; these were :

a) Full addressing, in which an address word is appended to each transmitted sample indicating the channel number.

b) Pre-addressing [7,12], in which each frame of transmitted samples is preceded by an address word that indicates those channels that are transmitting a sample and those that are not.

It was shown that for any number of channels greater than 4, method b) requires less address bits per frame than method a).

The pre-addressing method is more vulnerable to bit errors in the transmission path. However, provided the bit error rate is better than 1 in 10^5, a single parity check digit provides adequate protection for the sample address word (see section 2.1). Because of its economy in frame pulses, pre-addressing is the preferred method.

9.5 Frame overflow

An important feature of the proposed transmission system is how it caters for frame overflow. The importance of distributing the effects of frame overflow evenly throughout all channels was discussed in chapter 5. The following distribution methods were investigated:

a) ROTA system, in which overflow errors are distributed throughout all channels on a rota basis.

b) ERROR GROUPING system, in which overflow errors are grouped at the beginning of talkspurts, in a similar way to a TASI system.

c) VARIABLE WORD LENGTH system; when the frame has overflowed in this system, the word-length of samples is reduced so that all samples can be transmitted.

When the frame overflows, information is lost and cannot be retrieved. A formula was derived that relates the fraction of information lost by each channel, Φ, with the number of channels N, the number of timeslots K and the probability of a non-redundant sample α.

When frame overflow occurs, it introduces additional noise into the baseband channel. The ROTA and VWL systems were compared in terms of their noise performance under overflow conditions. A signal-to-noise ratio was derived for both versions of the ROTA system. This was compared with a formula derived for the VWL system by Dorgello and van der Veer [7]. These analyses demonstrated that both versions of the ROTA system (i.e. ZOP ROTA and ZAP ROTA) introduce far more noise than the VWL system, for the same number of channels.

Overflow noise increases with call-load, and reaches a peak during the "busy hour" when the call-loading is at its highest. It was shown that if a signal-to-noise ratio of 30 dB was acceptable during this period then, for a system transmitting 25 samples per frame :

a) The VWL system could accept 68 channels.

b) The ROTA system using zero-order prediction (ZOP) could accept 53 channels.

c) The ROTA system using zero-amplitude prediction (ZAP) could accept 50 channels.

Sciulli and Campanella have also derived a signal-to-noise ratio for the ZOP ROTA system . It was demonstrated in section 5.4 that their analysis gives a signal-to-noise ratio that is approximately 10 dB too high.

9.6 Simulation model

In order to make signal-to-noise ratio measurements, and subjective tests, a model of a practical transmission system was constructed. The model consists of a single speech channel with simulation of frame-overflow noise caused by the other channels. The model was designed using the following parameters:

a) 25 timeslots available for signal samples.

b) probability of a non-redundant sample $\alpha = 0.375$.

c) The number of channels variable from 33 to 95.

Both the ROTA and the VWL methods of overflow-error distribution were built into the model.

In the first half of chapter 6, statistical models were developed showing how each of the error-distribution systems could be simulated. The practical implementation of these models is described in the second half of chapter 6.

Signal-to-overflow noise ratio measurements were made on the model for the ZAP ROTA and the VWL systems, over a wide range of N. The results obtained agree very well with the theoretical values obtained in section 5.4.

9.7 Subjective tests

The simulation model was used to make a limited series of subjective tests on the ROTA and VWL systems. Isopreference testing techniques were used to compare the relative subjective performance of the two systems. The main conclusions to be drawn from the tests are :

a) The tests confirmed the superiority of the VWL system that was predicted by the signal-to-noise ratio analysis.

b) The VWL system increases its advantage over the ROTA system as the call-load increases.

c) For equal noise power, the noise produced by the ROTA system is, subjectively, slightly preferable to the noise produced by the VWL system.

d) For a 30 dB signal-to-noise ratio on the VWL system, or equivalent subjective quality on the ROTA system, the bandwidth compression ratios are:

VWL system $R = 2.0$

ZOP ROTA system $R = 1.65$

ZAP ROTA system $R = 1.56$

Thus, using the VWL system, the number of channels can be doubled compared with conventional PCM, for the same bandwidth.

9.8 Comparison with TASI systems

In chapter 8, the proposed system was compared with TASI systems. The main points were :

a) TASI systems eliminate redundancy from the silent pauses in speech only, whereas the proposed system eliminates redundancy from silent pauses and active speech.

b) In a digital TASI system, the sample addressing information occupies far fewer bits in each frame than the proposed system.

c) The VWL method of overflow-error distribution could be applied to a digital TASI system and improve the bandwidth compression ratio.

9.9 Suggestions for further study

(1) This thesis has been concerned with the theoretical investigation and simulation of the proposed multiplex transmission system. Clearly, an important aspect of future work is to confirm the results obtained so far by constructing and testing a complete system.

(2) The efficiency of the system (bandwidth compression) can be improved by decreasing the value of α. In this investigation, redundant samples have been eliminated by using zero-order (or zero-amplitude) prediction. By considering more than one previous sample (i.e. by increasing the order of prediction) it should be possible to reduce α. However, this would require more storage at both terminals and extra complexity.

(3) Another method of increasing efficiency is to reduce the number of bits required for sample addressing. Typically, the address information occupies 30% of the transmitted frame bits. It may be possible to use the statistical properties of the address information to achieve a data compression.

(4) In chapter 9, it was mentioned that the VWL method of frame-overflow error-distribution could be applied to a digital TASI system. Preliminary calculations have shown that this scheme gives a substantial improvement in bandwidth compression ratio compared with conventional TASI systems. The author intends to investigate the VWL-TASI system in the near future.

(5) It was mentioned in section 5.4 that the VWL system rarely uses all of the frame bits available for samples. By encoding some samples in b bits and some in b+1 bits when a frame overflow occurs, it is possible to use all frame bits. This improves the signal-to-noise ratio for the channels receiving the extra bit, but introduces the following problems :

a) Word synchronization at the receiving terminal.

b) Distributing the benefits evenly throughout all channels.

It would be interesting to discover the average bit wastage of the existing VWL system, also by how much the signal-to-noise ratio is improved by using the two word system.

APPENDIX A1 Computer programs used for probability measurements

The following two programs were used in the probability measurements described in chapter 3. The first program (S7) is written in FORTRAN 4, and the second program (CALC7) is written in MACRO.

```
C                   S7
        INTEGER TOTAL(128),RMS(128)
        DIMENSION X(128),U(128)
  110   READ(3,100) M,N
  100   FORMAT(I5/I5)
        IF(M) 109,109,111
  111   PAUSE1
        DO 101 I=1,128
        X(I) = 0
        U(I) = 0
  101   TOTAL(I) = 0
        DO 102 J=1,N
        CALL CALC(TOTAL,RMS)
        DO 103 K = 1,128
        Y = TOTAL(K)
        X(K) = X(K) + Y
        V = RMS(K)
        U(K) = U(K) + V
        RMS(K) = 0
  103   TOTAL(K) = 0
  102   CONTINUE
        SUM = 0
        SUMSQ = 0
        SQ = 0
        DO 104 L = 1,128
        SUM = SUM + X(L)
        Z = L - 1
        SQ = SQ + U(L)*(Z**2)
  104   SUMSQ = SUMSQ + X(L)*(Z**2)
        R = SQRT(SQ/SUM)
        STDEV = SQRT(SUMSQ/SUM)
        WRITE(2,112)
  112   FORMAT(4X,4HDATE//4X,12HCONVERSATION//4X,6HTALKER//
     1 4X,4HMIC.//4X,5HLEVEL,6X,11HRELATIVE TO//
     2 4X,12HENCODING LAW//4X,10H128 LEVELS//
     3 4X,13HSAMPLING TIME/)
        WRITE(2,105)SUM,STDEV,R
  105   FORMAT(4X,15HTOTAL SAMPLES =,F9.0,5X,27HRMS VALUE OF DIFF.
     1 SIGNAL =,F7.2//4X,28HRMS VALUE OF SPEECH SIGNAL =,F7.2/)
        WRITE(2,108)
  108   FORMAT(6X,5HDIFF.,3X,7HSAMPLES,6X,5HPROB.,5X,6HAMPLT.,
     13X,7HSAMPLES,6X,5HPROB.//)
        DO 106 I = 1,M
        J = I - 1
        PROB1 = X(I)/SUM
        PROB2 = U(I)/SUM
  106   WRITE(2,107)J,X(I),PROB1,J,U(I),PROB2
```

S7 continued.......

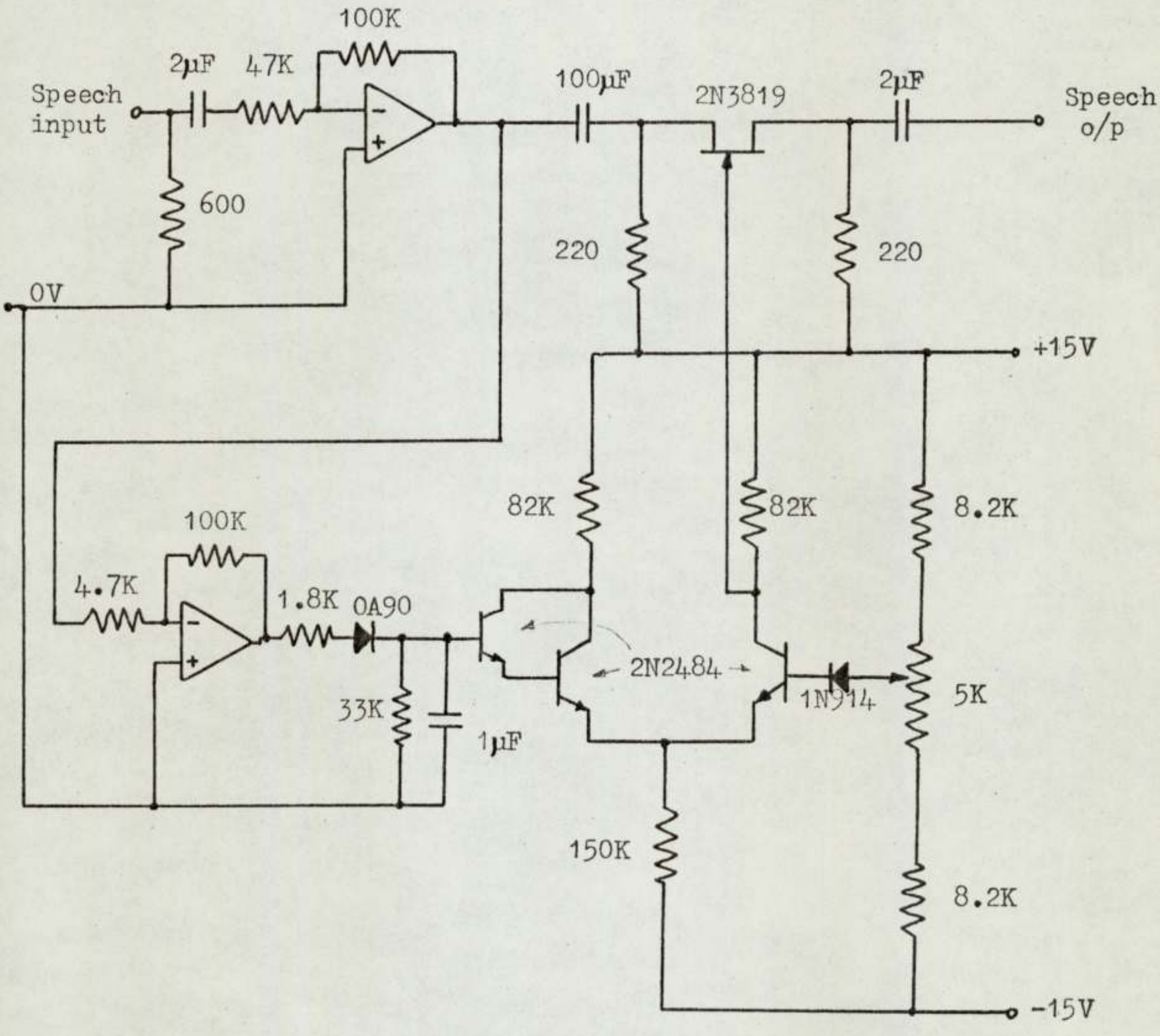
```
107   FORMAT(I1Ø,F1Ø.Ø,F12.5,I1Ø,F1Ø.Ø,F12.5)
      PAUSE2
      GO TO 11Ø
1Ø9   STOP
      END
```

```
          .TITLE CALC7
          .GLOBL CALC,.DA
ADSF=7Ø13Ø1
ADSC=7Ø13Ø4
ADRB=7Ø1312
ADSM=7Ø11Ø3
CALC      Ø
          JMS*     .DA
          JMP      .+3
BUFF1     Ø
BUFF2     Ø
          IOF
          CLA
          ADSM
          .TIMER 3ØØ,EXIT
START     IOF
          ADSC
          LAC      A
          JMS      ABSV
          TAD*     BUFF2
          DAC      STO1
          ISZ*     STO1
          LAC      A
          CMA
          TAD      (1
          TAD      B
          JMS      ABSV
          TAD*     BUFF1
          DAC      STO1
          ISZ*     STO1
          LAC      (-16
          DAC      COUNT
          ISZ      COUNT
          JMP      .-1
          NOP
          NOP
          ADSF
          JMP      .-1
          ADRB
          ION
          LRSS     13
          DAC      A
          NOP
          IOF
          ADSC
          LAC      B
          JMS      ABSV
          TAD*     BUFF2
          DAC      STO1
```

CALC7 continued.........

```
        ISZ*    ST01
        LAC     B
        CMA
        TAD     (1
        TAD     A
        JMS     ABSV
        TAD*    BUFF1
        DAC     ST01
        ISZ*    ST01
        LAC     (-16
        DAC     COUNT
        ISZ     COUNT
        JMP     .-1
        NOP
        NOP
        ADSF
        JMP     .-1
        ADRB
        ION
        LRSS    13
        DAC     B
        JMP     START
ABSV    Ø
        SMA
        JMP     .+4
        CMA
        TAD     (1
        JMP     .+4
        DZM     COUNT
        DZM     COUNT
        NOP
        JMP*    ABSV
EXIT    Ø
        DBR
        ION
        JMP*    CALC
A       Ø
B       Ø
ST01    Ø
COUNT   Ø
        .END
```

APPENDIX A2 Circuit diagram of the speech detector described in section 3.2



APPENDIX A3 Computer programs used in the simulation of the instantaneously and syllabically companded systems

A3.1 Syllabically companded system

The following two programs were used in the simulation of the transmit terminal of a pcm system using 7-bit encoding and syllabic companding. The first program (FSIM1) is written in FORTRAN 4, and the second program (SIM3) is written in MACRO.

```
C           FSIM1
C       SIMULATION OF A SYLLABICALLY COMPANDED SYSTEM WITH
C       SPEECH DETECTION
        INTEGER C(1024),D(512),A(512)
        DIMENSION DD(512),AA(512)
  1     READ(3,2)ID,IT,IL
  2     FORMAT(I5/I5/I5)
        PAUSE
        DETECT = Ø
        DO 3 I=1,512
        A(I) = Ø
        AA(I) = Ø
        D(I) = Ø
  3     DD(I) = Ø
        DO 4 J=1,7
  4     C(J) = -1
        DO 5 J=8,14
  5     C(J) = -2
        DO 6 J=15,28
  6     C(J) = -3
        DO 7 J=29,56
  7     C(J) = -4
        DO 8 J=57,112
  8     C(J) = -5
        DO 9 J=113,1Ø24
  9     C(J) = -6
        DO 1Ø K=1,IT
        CALL SIM(C,D,ID,IDET,A)
        DO 11 L=1,512
        AA(L) = AA(L) + A(L)
        A(L) = Ø
        DD(L) = DD(L) + D(L)
 11     D(L) = Ø
        DETECT = DETECT + IDET
 1Ø     IDET = Ø
        SUM = Ø
        DO 12 I=1,512
 12     SUM = SUM + DD(I)
        I = 1
        SPDET = DETECT/SUM
        WRITE(2,13)
 13     FORMAT(4X,4HDATE//4X,12HCONVERSATION//
```

FSIM1 continued.........

```
     14X,5HLEVEL,6X,11HRELATIVE TO//4X,12HENCODING LAW//
     24X,9H64LEVELS//)
        PRINT(2,)SUM
        PRINT(2,)ID
        PRINT(2,)SPDET
        WRITE(2,14)
  14    FORMAT(6X,5HDIFF.,3X,7HSAMPLES,6X,5HPROB.,5X,6HAMPLT.,
     13X,7HSAMPLES,6X,5HPROB.//)
        DO 15 I=1,IL
        J =I-1
        PROB1 = DD(I)/SUM
        PROB2 = AA(I)/SUM
  15    WRITE(2,16)J,DD(I),PROB1,J,AA(I),PROB2
  16    FORMAT(I1Ø,F1Ø.Ø,F12.5,I1Ø,F1Ø.Ø,F12.5)
        PAUSE2
        GO TO 1
        STOP
        END
```

```
               .TITLE SIM3
               .GLOBL    .DA,SIM
ADSF= 7Ø13Ø1
ADSC=7Ø13Ø4
ADRB=7Ø1312
ADSM=7Ø11Ø3
DAL1=7Ø51Ø1
SIM     Ø
        JMS*     .DA
        JMP      .+6
B1      Ø
B2      Ø
B3      Ø
B4      Ø
B5      Ø
        LAC*     B3
        DAC      DELTA
        IOF
        CLA
        ADSM
        .TIMER 3ØØ,EXIT
START   JMS      SUBR
        LAC A
        DAC      B
        LAC      C
        DAC      D
        JMP      START
SUBR    Ø
        IOF
        ADSC
        LAC      B
        ABS
        TAD      REF
```

SIM3 continued.........

```
SPA
JMP     .+4
LAC     REF
TAD     (-1
JMP     .+3
LAC     REF
TAD     (1
DAC     REF
LAC     REF
CMA
TAD     (1
TAD*    B1
DAC     STO
LAC*    STO
DAC     STO
LAC     STO
DAC     COUNT
LAC     B
JMP     .+2
LRSS    1
ISZ     COUNT
JMP     .-2
DAC     C
LAC     STO
TAD     (7
CMA
TAD     (1
DAC     COUNT
JMP     .+3
ABS
ABS
ISZ     COUNT
JMP     .-3
LAC     B
ABS
TAD     DELTA
SPA
JMP     .+4
LAC     (-2ØØØ
DAC     STO1
JMP     .+3
ABS
ABS
LAC     STO1
SZA
JMP     .+5
DZM     C
ABS
ABS
JMP     .+4
TAD     (1
DAC     STO1
```

SIM3 continued........

```
        ISZ     STO2
        LAC     C
        SMA
        JMP     .+4
        CMA
        TAD     (1
        JMP       .+4
        NOP
        NOP
        TAD*    B5
        DAC     STO
        ISZ*    STO
        LAC     C
        LLSS    6
        DAL1
        LAC     C
        CMA
        TAD     (1
        TAD     D
        SMA
        JMP     .+4
        CMA
        TAD     (1
        JMP     .+4
        NOP
        NOP
        NOP
        TAD*    B2
        DAC     STO
        ISZ*    STO
        ADSF
        JMP     .-1
        LAC     (-21
        DAC     COUNT
        ISZ     COUNT
        JMP     .-1
        ADRB
        LRSS    7
        DAC     A
        ION
        JMP*    SUBR
EXIT    Ø
        LAC     STO2
        DAC*    B4
        DZM     STO2
        DBR
        ION
        JMP*    SIM
A       Ø
B       Ø
C       Ø
D       Ø
STO     Ø
COUNT   Ø
REF     Ø
DELTA   Ø
STO1    Ø
STO2    Ø
        .END
```

A3.2 Instantaneously companded system

The following two programs were used in the simulation of the transmit terminal of a PCM system using 7-bit encoding and instantaneous companding. The first program (FSIM2) is written in FORTRAN 4, and the second program (SIM5) is written in MACRO.

```
C             FSIM2
C       SPEECH SLOPE ANALYSIS - 7-SEGMENT NON-LINEAR ENCODING
        INTEGER Z(2048),A(128),D(128)
        DIMENSION AA(128),DD(128)
        II = Ø
        DO 1ØØ M=1Ø25,1Ø4Ø
        Z(M) = II
        N = 2049 - M
        Z(N) = -Z(M)
  1ØØ   II = II + 1
        DO 1Ø1 M=1Ø4Ø,1Ø54,2
        DO 1Ø2 N=1,2
        I = M + N
        Z(I) = II
        K = 2Ø49 - I
  1Ø2   Z(K) = -Z(I)
  1Ø1   II = II + 1
        DO 1Ø3 M=1Ø56,1Ø84,4
        DO 1Ø4 N=1,4
        I = N + M
        Z(I) = II
        K = 2Ø49 - I
1Ø4     Z(K) = -Z(I)
1Ø3     II = II + 1
        DO 1Ø5 M=1Ø88,1144,8
        DO 1Ø6 N=1,8
        I = M + N
        Z(I) = II
        K = 2Ø49 - I
  1Ø6   Z(K) = -Z(I)
  1Ø5   II = II + 1
        DO 1Ø7 M=1152,1264,16
        DO 1Ø8 N=1,16
        I = M + N
        Z(I) = II
        K = 2Ø49 - I
  1Ø8   Z(K) = - Z(I)
  1Ø7   II = II + 1
        DO 1Ø9 M=128Ø,15Ø4,32
        DO 11Ø N=1,32
        I = M + N
        Z(I) = II
        K = 2Ø49 - I
```

FSIM2 continued........

```
110   Z(K) = -Z(I)
109   II = II + 1
      DO 111 M=1536,1984,64
      DO 112 N=1,64
      I = M + N
      Z(I) = II
      K = 2049 - I
112   Z(K) = -Z(I)
111   II = II + 1
      PAUSE1
113   READ(3,114)ID,IT,IL
114   FORMAT(I5/I5/I5)
      DO 115 IJ=1,128
      A(IJ) = 0
      AA(IJ) = 0
      D(IJ) = 0
115   DD(IJ) = 0
      DO 117 J=1,IT
      CALL SIM(Z,A,D,ID,IDET)
      DO 118 K=1,128
      AA(K) = AA(K) + A(K)
      DD(K) = DD(K) + D(K)
      A(K) =0
118   D(K) = 0
      DETECT = DETECT + IDET
117   IDET = 0
      SUM = 0
      DO 119 L=1,128
119   SUM = SUM + AA(L)
      SPDET = DETECT/SUM
      I = 1
      WRITE(2,120)
120   FORMAT(4X,4HDATE//4X,12HCONVERSATION//4X,6HTALKER//
     14X,4HMIC.//4X,5HLEVEL,6X,11HRELATIVE TO//
     14X,12HENCODING LAW//4X,10H128 LEVELS//
     14X,13HSAMPLING TIME/)
      PRINT(2,)SUM
      PRINT(2,)ID
      PRINT(2,)SPDET
      WRITE(2,122)
122   FORMAT(6X,5HDIFF.,4X,7HSAMPLES,6X,5HPROB.,5X,6HAMPLT.,
     13X,7HSAMPLES,6X,5HPROB.//)
      DO 123 I=1,IL
      J = I - 1
      PROB1 = DD(I)/SUM
      PROB2 = AA(I)/SUM
123   WRITE(2,124)J,DD(I),PROB1,J,AA(I),PROB2
124   FORMAT(I10,F10.0,F12.5,I10,F10.0,F12.5)
      PAUSE2
      GO TO 113
      STOP
      END
```

```
                .TITLE SIM5
                .GLOBL .DA,SIM
ADSF=7Ø13Ø1
ADSC=7Ø13Ø4
ADRB=7Ø1312
ADSM=7Ø11Ø3
DAL1=7Ø51Ø1
SIM     Ø
        JMS*    .DA
        JMP     .+6
B1      Ø       /Z
B2      Ø       /A
B3      Ø       /D
B4      Ø       /ID
B5      Ø       /IDET
        LAC*    B4
        DAC     DELTA
        LAC*    B1
        TAD     (2ØØØ
        DAC     ZA
        LAC*    B2
        DAC     AA
        LAC*    B3
        DAC     DA
        IOF
        CLA
        ADSM
        .TIMER 3ØØ,EXIT
START   JMS     SUBR
        LAC     A
        DAC     B
        LAC     C
        DAC     D
        JMP     START
SUBR    Ø
        IOF
        ADSC
        LAC     B
        TAD     ZA
        DAC     STO
        LAC*    STO
        DAC     C
        LAC     B
        ABS
        TAD     DELTA
        SPA
        JMP     .+4
        LAC     (-4ØØ
        DAC     STO1
        JMP     .+3
        ABS
        ABS
        LAC     STO1
        SZA
        JMP     .+5
        DZM     C
        ABS
        ABS
        JMP     .+4
        TAD     (1
        DAC     STO1
        ISZ     STO2
```

SIM5 continued.........

```
        LAC     C
        SMA
        JMP     .+4
        CMA
        TAD     (1
        JMP     .+4
        NOP
        NOP
        NOP
        TAD     AA
        DAC     STO
        ISZ*    STO
        LAC     C
        LLSS    6
        DAL1
        LAC     C
        CMA
        TAD     (1
        TAD     D
        SMA
        JMP     .+4
        CMA
        TAD     (1
        JMP     .+4
        NOP
        NOP
        NOP
        TAD     DA
        DAC     STO
        ISZ*    STO
        NOP
        NOP
        NOP
        NOP
        ADSF
        JMP     .-1
        ADRB
        LRSS    7
        DAC     A
        ION
        JMP*    SUBR
EXIT    Ø
        LAC     STO2
        DAC*    B5
        DZM     STO2
        DBR
        ION
        JMP*    SIM
A       Ø
B       Ø
C       Ø
D       Ø
STO     Ø
DELTA   Ø
ZA      Ø
AA      Ø
DA      Ø
STO1    Ø
STO2    Ø
        .END
```

APPENDIX A4 Description of the digital compressor used in the simulation model

This Appendix describes the operation of the digital compressor described in section 6.4.1. A circuit diagram is given in FIG. 6.12, and the positive half of the input/output transfer characteristic is given in FIG. A1. The characteristic is segmented and FIG. A1 also gives the range of input and output numbers encompassed by each segment.

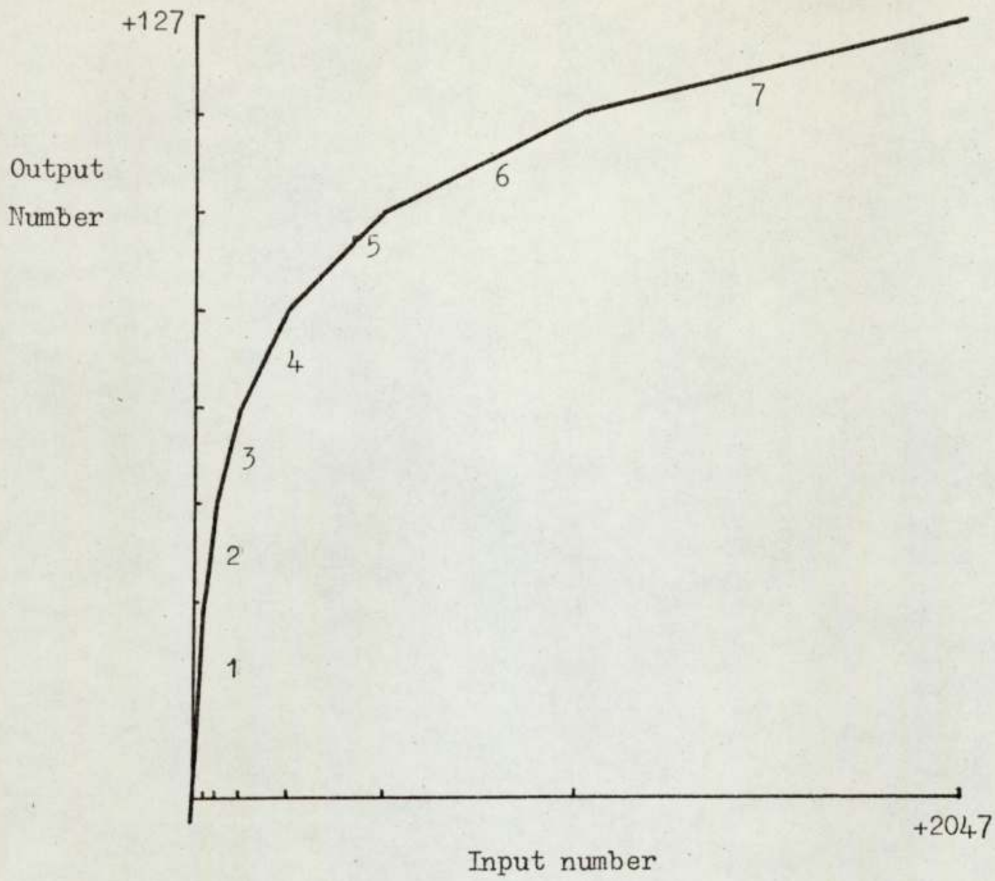
The compressor accepts a 12-bit linearly encoded sample at its input (in parallel) and gives an 8-bit compressed sample at its output (in parallel). Both input and output numbers are in symmetrical binary code. The output number comprises three parts :

a) A sign-bit; bit B8 is '0' for a positive number and '1' for a negative number.

b) An exponent: bits B5 to B7 indicate on which segment of the characteristic the number lies. All segments cover the same size range of output numbers, except the one that passes through the origin (i.e. segment number 1). This segment has two exponent codes associated with it (i.e. '000' and '001').

c) A mantissa; bits B1 to B4 indicate the position of the number on the segment.

The relationship between the input and output numbers has certain properties that are useful in executing the compression. These properties are shown in the table of numbers given in FIG. A1. For a positive number, the value of the exponent bits can be obtained by noting the number of '0's prior to the leading '1' in the input number



Seg. No.	Input number From	Input number To	Output number From	Output number To
7	01(0000)000000	01(1111)111111	0111(0000)	0111(1111)
6	001(0000)00000	001(1111)11111	0110(0000)	0110(1111)
5	0001(0000)0000	0001(1111)1111	0101(0000)	0101(1111)
4	00001(0000)000	00001(1111)111	0100(0000)	0100(1111)
3	000001(0000)00	000001(1111)11	0011(0000)	0011(1111)
2	0000001(0000)0	0000001(1111)1	0010(0000)	0010(1111)
1	00000001(0000)	00000001(1111)	0001(0000)	0001(1111)
	00000000(0000)	00000000(1111)	0000(0000)	0000(1111)

FIG. A1 The positive half of the compressor transfer characteristic and the range of numbers covered by each segment (excluding the sign-bit Q12). If there are 7 or more '0's, then bits B5 to B7 have the value '000'. Similarly, for 6 '0's bits B5 to B7 have the value '001', and so on. For a negative number, the value of the exponent can be obtained by noting the number of '1's to the left of the first '0'.

It can be seen from FIG. A1 that the slope of segment number 1 is unity. In other words, a change in the input number produces an equal change in the output number. Thus, on this segment, the mantissa (bits B1 to B4) change in unison with bits Q1 to Q4. The slope of the second segment is a half, so that the input number must change by two for a change of one in the output number. In this case bits B1 to B4 change in unison with bits Q2 to Q5. Similarly, on the third segment B1 to B4 are identical to bits Q3 to Q6, and so on for the other segments. Thus, the mantissa can be obtained directly from the input number by choosing the appropriate four bits. The mantissa and the corresponding bits in the input number are shown bracketed in FIG. A1. From this figure it can be seen that the required bits in the input number lie immediately to the right of the leading '1' (or leading '0' for a negative number).

The compressor consists of,basically, an 11-bit shift-register and 3-bit counter. There are also some logic gates that control clock pulses and provide other logic functions. The sequence of operations for the compression of a positive number is as follows :

(1) Clock pulse CP6 clears the shift-register.

(2) Clock pulse CP5 'loads' bits Q1 to Q11 into the shift-register, and the sign-bit Q12 into the upper bistable. CP5 also presets the counter to '111'.

(3) A burst of pulses (CP4) is applied (via gates) to the 'count/down' input of the counter and the 'shift' input of the shift-register. Each pulse decrements the counter and shifts the contents of the shift-register one stage upwards.

(4) This continues until the leading '1' appears in the top stage of the shift-register. This inhibits the application of any more pulses in the CP4 'burst' to the counter or the shift-register, The compression process is then complete.

(5) There are seven pulses in the CP4 burst, but only a maximum of six are applied to the shift-register. If, after six pulses, there is a '0' in the top stage of the shift-register (i.e. there is no '1' to the left of bits Q1 to Q4), the seventh pulse decrements the counter. The contents of the counter are then '000'; this indicates that the number lies on the lower half of the first segment.

(6) After compression the counter contains the exponent bits. The sign-bit of the output number is identical to the sign bit of the input number (they are shown underlined in FIG. A1). The mantissa bits are contained in the four stages of the shift-register immediately below the top stage.

The compression of a negative number follows a similar procedure to that outlined above, but has the following differences:

a) The compressor searches for the leading '0' in the input number (rather than the leading '1').

b) At the end of the compression sequence, the counter contains the exponent bits in inverted form (e.g. if the required exponent is '101', the counter contains '010'). Thus, for negative numbers, the exponent bits are obtained by inverting the contents of the counter. Referring to FIG. 6.12, this is done by the group of logic gates to the right of the counter.

APPENDIX A5 Description of the digital expander used in the simulation model

A circuit diagram of the digital expander is given in FIG. 6.14. The transfer characteristic is the inverse of that of the compressor in the sending terminal of the simulation model (see FIG. A1).

The expander is very similar to the compressor in that its main components are a shift-register and a counter. The sequence of operations are also similar except that they are in the reverse order.

The sequence of operations for the expansion of a positive number are as follows :

(1) Clock pulse CP8 clears the upper ten stages of the shift-register, and 'presets' the lower two stages to a '1'.

(2) The information on the 'data' (D) inputs of the shift-register is 'loaded' into the register, by CP9. The register then contains 'X B4 B3 B2 B1 0 1 1 1 1 1 1'. X is the contents of the top stage of the register. If the number is on the lower half of the first segment, X is a '0'; for any other part of the positive half of the characteristic, X is a '1'. B4 to B1 are the mantissa bits of the input (compressed) number. CP9 also loads the exponent bits (B5 to B7) of the input number into the counter.

(3) A burst of pulses (CP4) is applied (via gates) to the 'count/up' input of the counter and the 'shift' input of the register. Each pulse increments the counter and shifts the contents of the register one stage downwards. This

continues until either

a) The contents of the counter reaches '111'; this is detected by logic gates which inhibit the CP4 burst.

or b) After six pulses a '0' appears in the bottom stage of the register; this prevents further pulses being applied to the shift input of the shift-register.

(4) The expansion process is then complete. The sign-bit of the output number is the same as the sign-bit of the input number. The remainder of the output number is contained in the upper eleven stages of the shift-register.

The expansion of negative numbers follows an almost identical procedure. The only difference is that the exponent bits of the input number are inverted before they are loaded into the counter.

List of principal symbols

Symbols not listed here are defined as they appear in the text.

C	Probability of zero-difference in active (detected) speech
K	Number of timeslots per frame available for samples
N	Number of baseband channels
N_0	Frame overflow noise power
N_q	Quantizing noise power
$p(x)$	Probability density function
$P(a_o)$	Probability of sample being zero-valued
$P(d_i)$	Probability of a difference of i quanta between adjacent samples
$P(d_o)$	Probability of zero-difference between adjacent samples
R	Bandwidth compression ratio
R_{eff}	Effective bandwidth compression ratio
$R(\tau)$	Autocorrelation function
S	Signal power
x	Instantaneous amplitude of speech signal
$X(t)$	Analogue speech signal
x'	Instantaneous slope of the speech signal
y	Number of active channels
y_s	Simulated number of active channels
ZAP	Zero-amplitude prediction
ZOP	Zero-order prediction
α	Probability of a non-redundant sample
$\delta(\tau)$	Dirac impulse
Φ	Fraction of information lost by a channel due to overflow
τ_A	Speech activity factor
τ_s	Time between sampling instants

REFERENCES

1 Bullington K., Fraser J.M.: 'Engineering aspects of TASI', Bell Syst. Tech. J.,1959,38, pp. 353-364

2 Miedema H., Schachtman M.G. : 'TASI quality - effect of speech detectors and interpolation', Bell Syst. Tech. J., 1962, 41, pp. 1455-1473

3 Fraser J. M., Bullock D. B., Long N. G. : 'Overall characteristics of a TASI system', Bell Syst. Tech. J., 1962, 41, pp. 1439-1454

4 Amano K., Ohta C. : 'Digital TASI system in PCM transmission', International Communication Conference, June 1969, Boulder, Colo.

5 Nosaka K. : 'TTT system - 50 Mb/s PCM TDMA system with time pre-assignment and TASI features', International Conference on Digital Satellite Communication, 1969, IEE Conference publication No. 59, p. 83

6 Sciulli J. A. : 'A statistical characterization of digital voice signals for compression system applications', Proc. of UMR-M. J. Kelly Communications Conference, Oct. 1970

7 Dorgello A. J. C., van der Veer H. : 'Variable length coding increasing traffic capacity in PCM transmission systems', 1972 International Zurich Seminar on Integrated Systems, March 1972, paper C6

8 Flood J. E., Urquhart-Pullen D. I. : 'Time-compression-multiplex transmission', Proc. IEE, 1964, 111, pp. 647-668

9 Sekimoto T., Puente J. G. :'A satellite time-division multiple-access experiment', IEEE Trans. Comm., COM-16, 1968, pp. 581-588

10 Flood J. E., Urquhart-Pullen D. I. :'Time-assignment speech interpolation in time-compression-multiplex transmission', Proc. IEE, 1964, 111, pp. 675-683

11 Flood J. E., Whitehouse K. G. :'Digital multiplex telephone transmission system using statistical properties of speech', Electronics Letters, 24th Feb. 1972, 8, No. 4, p. 96

12 Sciulli J. A., Campanella S. J. :'A speech predictive encoding communication system for multichannel telephony', IEEE Trans. Comm., July 1973, COM-21, No. 7, pp. 827-835

13 Davenport W. B. :'An experimental study of speech wave probability distributions', J. of the Acous. Soc. of America, 1952, 24, No.4, p.390

14 Richards D. L. :'Statistical properties of speech signals', Proc. IEE, 1964, 111, No. 5, pp. 941-949

15 Purton R. F. :'A survey of telephone speech signal statistics and their significance in the choice of a PCM companding law', Proc. IEE, 1962, 109B, p. 60

16 Berry R. W. :'Speech volume measurements on telephone circuits', Proc. IEE, 1971, 118, No. 2, pp. 335-338

17 Watt-Carter D. E., Wheeler L.K. :'The Lincompex system for the protection of h.f. radio-telephone circuits', Post Off. Electr. Engnrs. J.,1966, 59, No. 3, pp. 163-167

18 Fariello E. :'A novel digital speech detector for improving effective satellite capacity', IEEE Trans Comm., 1972, COM-21, No. 2, pp. 55-59

19 Cattermole K. W. :'Principles of pulse code modulation', Iliffe Books Ltd., London, 1969, p. 134

20 Korn G.A. :'Random process simulation and measurements', McGraw-Hill, New York (USA), 1966

21 Davies W. D. T. :'Generation and properties of maximum-length sequences', Control, June July Aug. 1966, 10, pp. 302-4, 364-5, 431-3

22 Moss J. K., Simpson R. J., Tempest W. :'A pseudo-random pulse train generator with controllable rate for modelling of audiometric systems', The Radio and Electronic Engineer, 1972, 42, No. 9, pp. 419-424

23 IEEE : IEEE recommended practice for speech quality measurements, IEEE Trans., 1969, AU-17, No. 3, pp. 225-243

24 Cattermole K. W. : 'Principles of pulse code modulation', Iliffe Books Ltd., London, 1969, p 137

distributed to the four units. At this moment, when the pulse has entered the sampling lines, 40 samples are taken instantaneously. Fig. 2*a* shows the 40 samples reconstructed and displayed on an l.f. oscilloscope. The pulse amplitude is 40 mV at the sampling gates, or 4×40 mV before the power dividers. With a point distance $T_p = 45$ ps, we obtain a displayed risetime of about 165 ps. This risetime includes those of the pulse generator (70 ps), those of the 8 ns delay line and trigger pickoff (75 ps) and those of the power dividers (65 ps), which together cause about 122 ps risetime. Hence the sampling lines themselves demonstrate a risetime of 112 ps.

In Fig. 2*c*, the same pulse is displayed in the recurrent mode by a 12·4 GHz sampling scope, using a 60 ns delay cable to get a suitable pretrigger. The displayed risetime, about 170 ps, and the $\pm 10\%$ overshoot are mainly caused by the delay line. To demonstrate reliability and precision of the new instrument, about 100 single displays are written one over the other in Fig. 2*b*. The displayed noise level of the sampling gates is less than $\pm 1·5$ mV, and the additional noise within the fast transitions, caused by time jitter, is about ± 1 mV, resulting in a time jitter of ± 5 ps due to the slope.

Acknowledgments: The author is indebted to Prof. H Lueg for his encouraging and stimulating discussions. He also wishes to thank J. Herzog, T. F. Mentjes, F. W. Vorhagen and D. Schäfer for providing auxilliary devices.

RUDOLF SCHWARTE

Institut für Technische Elektronik, TH Aachen
5100 Aachen, Templergraben 55, W. Germany

28th January 1972

References

1 SELDMEYER, J. W., PATTEN, R. B., and FUSSEL, L.: 'Analyser for fast single events' *in* 'Nuclear electronics III' (International Atomic Energy Agency, 1962)
2 STANCHI, L.: 'Spatial sampling for fast single events', *IEEE Trans.*, 1969, NS, p. 107
3 CASHIN, P. M., and JOHNSON, D. A. H.: 'A novel method for analysing singly-occurring pulses with nanosecond resolution', *Radio & Electron. Eng.* 1969, 38

DIGITAL MULTIPLEX TELEPHONE TRANSMISSION SYSTEM USING STATISTICAL PROPERTIES OF SPEECH

Indexing terms: Multiplexing, Telephone networks, Speech, Digital circuits

A digital multiplex telephone transmission system is proposed, which should achieve a smaller bandwidth than conventional pulse-code modulation by using the statistical properties of speech waveforms. Each frame of transmitted pulses only contains a coded sample for a channel if the sample differs from the previous one from that channel. The frame can thus contain fewer time slots than there are channels, if there is a high probability of baseband signals changing by less than one quantising level during the frame period.

Multiplex communication systems normally provide a number of channels, each of which permanently provides an information-carrying capacity equal to the greatest demand that can be made by its baseband signal. For example, pulse-code-modulation (p.c.m.) systems can cater for the signal of every channel changing from the highest to the lowest quantising level in a single sampling period; for speech signals, this is clearly an extremely rare occurrence. It is theoretically possible to increase the number of channels which can be transmitted if redundancy can be eliminated from the baseband signals before multiplexing. This has been done to some extent in the time-assignment speech-interpolation (t.a.s.i.) system.[1] In this system, a baseband channel is only connected to a multiplex channel when the presence of speech is detected. Since, in a normal telephone conversation, speech is present in each direction for less than half the time, t.a.s.i. enables the number of channels provided by a multiplex system to be at least doubled.

New forms of multiplex systems are possible if the baseband signal of each channel is stored in an electronic memory for a short period before transmission.[2,3] Moreover, it should be possible to analyse the stored information to eliminate some redundancy prior to transmission.[4] This principle has already been applied to provide t.a.s.i. more economically than by conventional methods.[4-6] Further possibilities clearly exist.

A proposed multiplex system is shown in Fig. 1. At the sending terminal, the baseband signals of the channels are sampled sequentially (at the Nyquist rate), and the samples are encoded by an analogue–digital convertor, as in p.c.m. Each coded sample from a channel is written into the corresponding address in a store. A sample is only transmitted if it differs from the previous sample stored at the address for that channel. The frame of transmitted pulses can contain fewer time slots than there are channels if there is a high probability of the baseband signals of several channels changing by less than one quantising level in the frame period. This condition of 'inactivity' will be met, both by channels which are silent and channels whose signals have an instantaneous rate of change which is small (compared with that of a high-frequency component of maximum amplitude).

At the receiving terminal, the digits received are used to modify the sample value stored at the address corresponding to the channel. If no digits are received, the store reads out to the baseband channel the sample value received in the previous frame. Since the channel samples do not occupy fixed time slots in the frame, it is necessary to transmit additional information to cause each coded sample to be written into the correct address at the receiving store. This may be done in two ways:

(*a*) *Full addressing:* The word format for each transmitted sample contains two syllables; one gives the identity of the channel, and the other the magnitude of its sample.

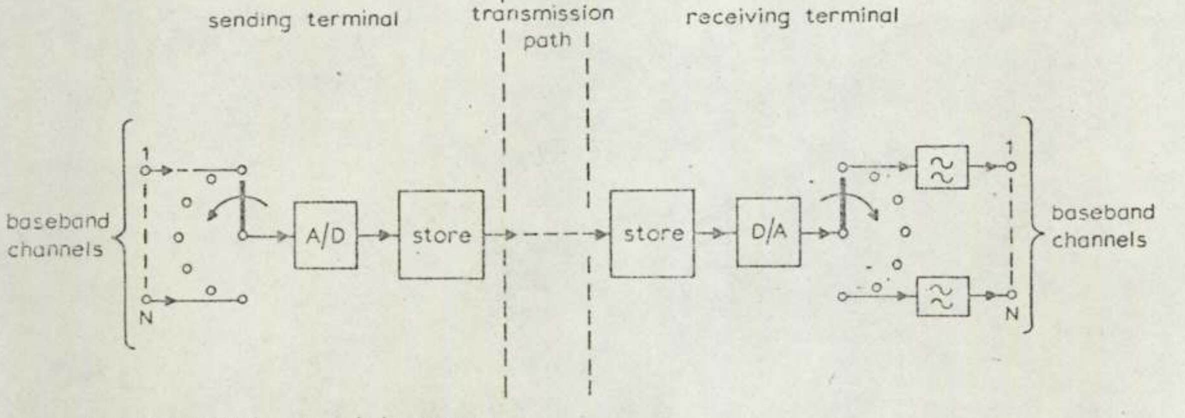


Fig. 1 *Proposed multiplex transmission system*

(*b*) *Preaddressing:* Each frame is preceded by an additional word containing one digit for each channel. If this digit is '1', the frame contains a sample for the corresponding channel; if it is '0', no sample is transmitted.

If the system contains N channels, method a requires $N\log_2{}^N$ bits, whereas method b only requires N bits, and therefore requires less bandwidth. An error in transmission of address information with method a only affects one channel, whereas, with method b, an error in one address bit affects the signals of all subsequent channels (unless a redundant code is used for the address word). However, an error only persists until the next sample is sent. If the error rate is small, the additional errors introduced by method b will also be small, and its lower bandwidth requirement makes this method preferable.

The economy in bandwidth obtainable from this system can be calculated easily (if digits required for synchronising are ignored), as follows: Let

$$R = \frac{\text{bandwidth required for } N\text{-channel system}}{\text{bandwidth required for } N\text{-channel p.c.m.}}$$

If method b is used for sending the address information, then

$$R = \frac{\text{number of bits per frame}}{\text{number of bits per frame for p.c.m.}}$$

$$= \frac{N + Ka}{Na}$$

$$= \frac{1}{a} + \frac{K}{N}$$

where

N = number of channels

K = number of time slots per frame ($K < N$)

a = number of bits per sample

To obtain a large bandwidth advantage, the ratio K/N should be small, but there will then be errors caused by samples being lost when the number of 'active' channels exceeds N. The value of K/N that can be achieved depends on the following:

(i) the permissible error rate

(ii) the number of digits per sample and the encoding law employed

(iii) the statistical distribution of the rate of change of speech voltages (in particular, the probability of the voltage changing by less than one quantum during the sampling period).

These factors are now being investigated by using a digital computer to make statistical analyses from speech recordings.

J. E. FLOOD
K. G. WHITEHOUSE

6th January 1972

Department of Electrical Engineering
University of Aston in Birmingham
Sumpner Building
19 Coleshill Street
Birmingham B4 7PB, England

References

1 BULLINGTON, K., and FRASER, K. M.: 'Engineering aspects of TASI', *Bell Syst. Tech. J.*, 1959, 38, p. 353
2 FLOOD, J. E., and URQUHART-PULLEN, D. I.: 'Time-compression-multiplex transmission', *Proc. IEE*, 1964, **111**, (4), pp. 647–668
3 SEKIMOTO, T., and PUENTE, J. G.: 'A satellite time-division multiple-access experiment', *IEEE Trans.*, 1968, **COM-16**, p. 581
4 FLOOD, J. E., and URQUHART-PULLEN, D. I.: 'Time-assignment speech interpolation in time-compression-multiplex transmission', *Proc. IEE*, 1964, **111**, (4), pp. 675–683
5 AMANO, K., and OHTA, C.: 'Digital TASI system in PCM transmission'. International communication conference, Boulder, Colo., June 1969
6 NOSAKA, K.: 'TTT system—50 Mb/s PCM-TDMA system with time-preassignment and TASI features'. International conference on digital satellite communication: *IEE Conf. Publ.* 59, 1969, p. 83

MICROWAVE ATTENUATION DUE TO RAIN AT 110 GHz IN SOUTH-EAST ENGLAND

Indexing terms: Electromagnetic-wave attenuation, Radiowave-propagation effects, Fading, Microwave links, Reliability

Fading due to precipitation has been recorded at 110 GHz for a period of one year over a 2·65 km path near Slough, England. The results show that frequencies as high as this could be used for relay links over short distances with acceptable reliability, despite the large attenuation due to heavy rain.

Introduction: It is well known[1] that attenuation by rainfall will impose severe restrictions on the widespread use of millimetre waves in communications. For example, calculations[2] show that a rainfall rate of 12·5 mm/h should produce an attenuation as high as 7 dB/km at 100 GHz, and experimental evidence[3,4] does not indicate that the actual value is likely to be much lower. Thus, in general, the use of such frequencies will be limited to single hops of a few kilometres.

Experiments are now being performed to examine the statistical characteristics of attenuation caused by precipitation at a frequency of 110 GHz (λ = 2·7 mm) over a typical link distance in south-east England. Some results, obtained from one year's observations, are presented here. Measurements of fading at such frequencies have been reported by workers in the USA,[5,6] but no published data are currently available for the UK.

Details of link: The transmitter and receiver are mounted 2·65 km apart in the Windsor–Slough area, approximately 30 km due west of central London. The transmitter (on the roof of a high building in Windsor) is at a bearing of 254° from the receiver site at the Radio & Space Research Station, Slough, and at an elevation of 1·3°. This gives ample ground clearance for the sidelobes of the receiving antenna.

The source is a reflex klystron producing a few milliwatts of c.w. power at a frequency near 110 GHz. The transmitting

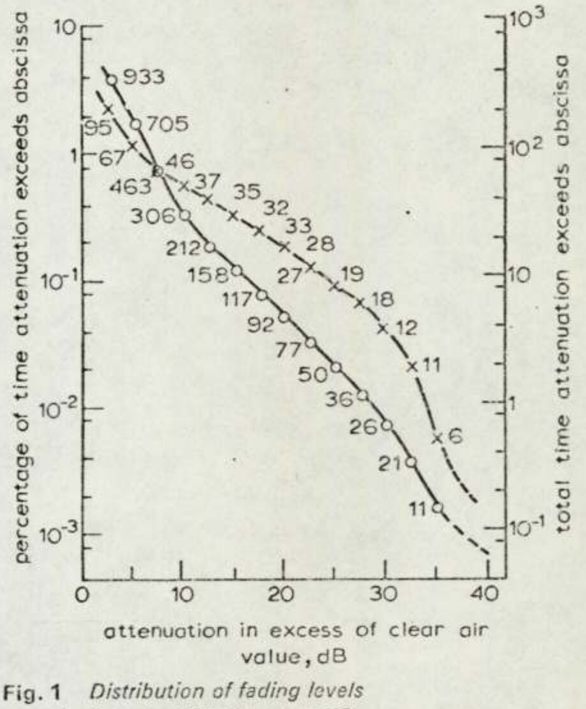


Fig. 1 *Distribution of fading levels*
○ ○ ○ 1st September 1970–31st August 1971
× × × Worst recorded calendar month (August 1971)
Numbers indicate occasions when attenuation level is exceeded

and receiving antennas are both 0·45 m-diameter dishes, and the receiver uses a harmonic-mixer crystal with a 55 GHz klystron as the local oscillator. The transmitter and receiver, including antennas, are each housed in temperature-controlled boxes with low-loss dielectric windows, and the windows themselves are well protected from direct rainfall by awnings.

The measured transmission loss $10\log_{10}(P_t/P_r)$ over the whole link, including antennas and feeds, is 46 dB, which agrees with the calculated estimate. Thus the received power