AUTOMATIC EQUALISATION OF TV TELEPHONE CHANNELS

by

Lee Thian Teck, B.E.(Hons.)

Submitted in partial fulfillment of the requirements for the degree of

Master of Engineering Science
in the
Faculty of Engineering

UNIVERSITY OF TASMANIA
HOBART

September, 1975.

ACKNOWLEDGEMENTS

The work contained in this thesis was undertaken in the Department of Electrical Engineering, University of Tasmania. I wish to thank the staff of the above Department for their help and cooperation, and my fellow post-graduate students for their useful discussions. I wish especially to express my appreciation for the guidance and encouragement given by my supervisors, Dr. J.P. Norton and Mr. G. The, throughout the duration of the work. Finally, I would like to thank the PMG Research Lab. in Melbourne for their financial assistance and Dr. A.J. Gibbs for his helpful interest.

I hereby declare that, except as stated therein, this thesis contains no material which has been accepted for the award of any degree or diploma in any university, and that, to the best of my knowledge, there is no copy of material previously published or written by another person except where due reference is made in the text.

CONTENTS

PAGE

Introduction

One possible way of implementing the TV telephone system is by transmitting the video signal over existing telephone cables. Not only is this attractive from an economical standpoint, it also appears to be technically reasonable. However, there are certain basic problems to be considered, particularly in reference to baseband analogue transmission over the local subscriber area.

For a start, it is expected that fixed equalisers will be required at regular intervals along the line to compensate for attenuation losses. These equalisers will have to be designed on the basis of some average channel characteristic. As a result, there will be uncorrected losses owing to deviations from the average characteristic for the different TV telephone connections. These and other channel imperfections will have to be compensated automatically.

This thesis is concerned with the design and testing of an automatic equaliser to serve the above-mentioned purpose. Chapter One outlines the basic principles of operation, together with the performance criterion used to assess the degree of optimality achieved. The mode of operation is discussed, as well as the different possible choices and arrangements of the basic equaliser function.

To provide further insight into the process, a mathematical study is given in Chapter Two. This includes considering the control algorithm for tap adjustments, the conditions for convergence, and the choice of step size to be used. The analysis is extended to cover adaptive operations, where the information for updating the tap settings is not exactly available but must be estimated on-line. In addition, ways of improving performance through modifying the basic equaliser structure (such as using a feedback configuration instead of the conventional feedforward arrangement) are also considered.

The theoretical investigation is carried further in Chapter Three. The effect of noise on equaliser capability and performance is discussed, along with the need for a defined standard of signal-to-noise ratio for acceptable transmission. This is followed by an evaluation of the effect of having a non-linearity in the feedback loop. The resulting system is studied, using some form of linearised analysis. Some reference to stability requirements is also given.

In order to confirm some of the theoretical findings of the previous chapters, and to demonstrate the potential of automatic equalisation in practice, a real-time experimental system was designed and assembled. Chapter Four outlines the work involved and the problems encountered. The choice of suitable test signals and their method of generation are discussed in some detail.

After the experimental equaliser had been built, various tests were conducted, using 500 yards of 4 lb P.I.U.T. cable to represent the TV telephone channel. The tests were designed to provide an estimate of some of the critical parameters and an assessment of system capability. The nature of the tests, with the results obtained, are given in Chapter Five.

The implications of the test results are discussed in Chapter Six. The areas requiring further investigation are identified. A preliminary discussion of the problems involved in integrating the system into the existing telephone network is also included.

The conclusion summarises the state of progress in evaluating the practicability of the automatic mean square equaliser for TV telephone applications.

1. The Concept of Automatic Equalisation

1.1 General

The purpose of a TV telephone service is primarily to realise face to face conversation. Secondary, but nevertheless important, uses include the display of drawings, photographs and various documents, and communicating with a computer, where the computer's responses are displayed visually on the set. The introduction of TV telephone service thus represents a new dimension in telephony.

There are considerable technical obstacles to be overcome to produce a practical, yet economic, system. A video cable network with special types of conductors to which all subscribers may be connected could conceivably be laid, but it would be quite uneconomic. The most practical solution would appear to be to use the existing telephone cable network. The existing loop plant consists of conventional wire pairs with each telephone line requiring one pair. The transmission of voice signals is over the frequency range of d.c. to 3.5 kHz. It is clear that modifications are needed if these same lines are to be used for video transmission over a much wider bandwidth. In the TV telephone system, one pair of lines will be needed for each direction of video transmission in addition to the pair used for the audio signal, making three pairs in all for each subscriber.

1.2 Impairment of TV telephone signals

The scope of this work will be confined to consideration of baseband analogue transmission over the local subscriber network, since this is where the major technical problems are likely to arise. The subscriber distribution area is characterised by light gauge cables, a multiplicity of gauges in a single connection, the use of cabinets and pillars, and the frequent use of tees and jointers in the continuing provision of new services. Such transmission irregularities introduce distortions to the transmitted video signal. The types of impairment include:

(a) Echoes

Deviations in the gain and phase characteristics of the transmission path produce a form of distortion that may be represented by a series of echoes (or ghosts) in the transmitted picture. The main sources of echoes are bridge taps, mixed gauges and frequent splicing. The occurrence of impedence mismatch at such places results in multiple reflections of the video signal, the net effect being the appearance of "ghosts" in the displayed image. Further, the cable is subjected to variations in ambient temperature which cause the gain and phase characteristics to change with time (1). This also causes echoes.

(b) Crosstalk

For the relatively short line lengths in the subscriber area, near-end crosstalk (defined as the coupling between two lines operated in opposite directions) is by far more significant than far-end crosstalk (coupling between two lines operated in the same direction). The extent of crosstalk coupling between two wire pairs depends also to a great degree on the configuration of the pairs. Pairs that are directly adjacent, i.e. in the same quad, influence one another more seriously than pairs that are further apart, i.e. in adjacent quads or in different basic units. Since in most situations, physical separation of the pairs is impractical, crosstalk coupling will arise. This places a limit on the spacing of repeaters, depending on the level difference that can be tolerated.

(c) Impulse Noise

Most calls involve connections through the exchange. In such cases, the opening of the contacts attached to a telephone pair carrying direct current may produce a train of transients of large amplitude containing energy distributed over several megahertz. The inevitable cross-

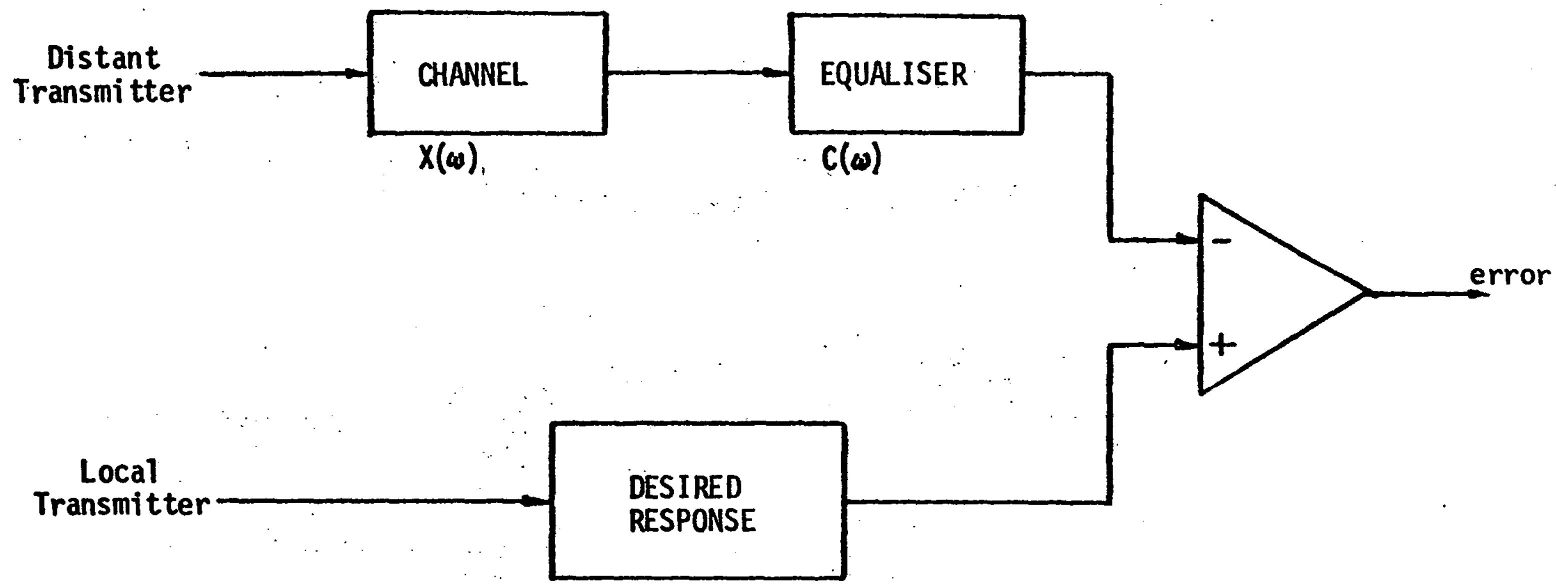


FIG.1.1 CHANNEL EQUALISATION

talk coupling allows some of this energy to be transferred from the telephone pairs to the TV telephone pairs in the same cable. The resulting interfering transients may reach amplitudes of the order of a volt and this together with their brief duration has led to the use of the term "impulse noise" to describe such disturbances. The impulse noise that occurs in practice is typically fairly complex, consisting of a series of separate rapidly-decaying oscillations. The duration of each may vary from 5 to 50 microseconds and may occur at intervals of 20 to 200 microseconds, and the entire transient may be of the order of a millisecond. Statistically, however, these transients are infrequent.

1.3 The Automatic Equaliser

As a result of the transmission irregularities existing in the local subscriber network, it is expected that the transmitted TV telephone signal will suffer distortion. Some form of compensating device or equaliser is thus necessary to remove the distortion and produce an acceptable response. In principle, if the channel to be equalised is known precisely, it is always possible to design an equaliser that will serve this purpose. However, in practice, a channel is random in the sense of being an ensemble of possible channels. Consequently, a fixed equaliser designed on average channel characteristics may not be adequate. An automatic equaliser which will remove the distortion whatever the actual characteristic of the channel may be is required.

As shown in Fig. 1.1, the automatic equaliser strives first to determine and then to synthesise the function $C(\omega) = {}^{1}/X(\omega)$ where $X(\omega)$ is the frequency response of the distorting channel. If the equaliser could perfectly synthesise ${}^{1}/X(\omega)$ (plus an arbitrary flat time delay), the distortion would be completely removed. In practice, exact equalisation is rarely attainable. Rather, the equaliser operates to minimise the discrepancy between a specified response and the actual response of the transmission medium automatically. The measure of

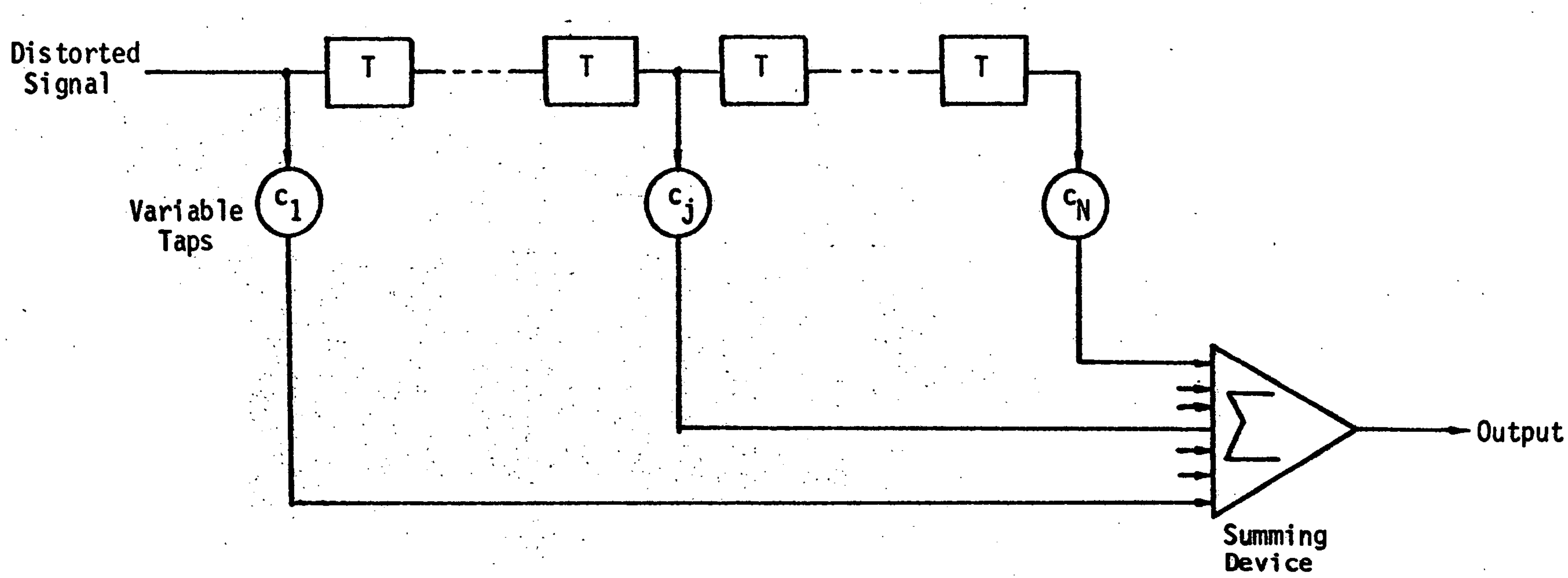


FIG.1.2 THE TRANSVERSAL EQUALISER

optimality is according to some suitably chosen performance criterion (see Section 1.5).

The device most commonly used for automatic equalisation is the transversal filter. It is simple in principle as can be seen from Fig. 1.2. The distorted signal enters a delay line and is picked off at equally spaced taps along the line. The signal from each tap is passed through an associated variable attenuator and all the attenuator output are then summed. The frequency characteristic of a N-tap transversal filter with tap gains c_n, n=1, ..., N spaced at T second intervals is given by

$$C(\omega) = \sum_{n=1}^{N} c_n \, e^{-jn\omega T}$$

(Where the channel is essentially low pass in nature, $T = \frac{1}{2B}$, the Nyquist period corresponding to a bandwidth B). The impulse response of the transversal filter is thus completely controlled by its tap settings. Accordingly, it can be used to provide amplitude compensation without phase distortion or vice versa. Its versatility lies in this ability to exercise independent control over both amplitude and phase response.

Alternatively, other sets of functions, of both first and second order, can be used as the basic equaliser function instead of the delay line. One such is the Laguerre set of networks (2) which can be easily realised. Mathematically, the impulse response of the nth Laguerre network is given by

$$l_n(t) = e^{\alpha t} \frac{d^n}{dt^n} \left(\frac{t^n}{n!} e^{-2\alpha t}\right) \text{ where } \alpha \text{ is a constant}$$

The corresponding transfer function, as expressed in the s-plane, is

$$L_n(s) = \frac{\alpha}{s+\alpha} \left(\frac{\alpha-s}{\alpha+s}\right)^n$$

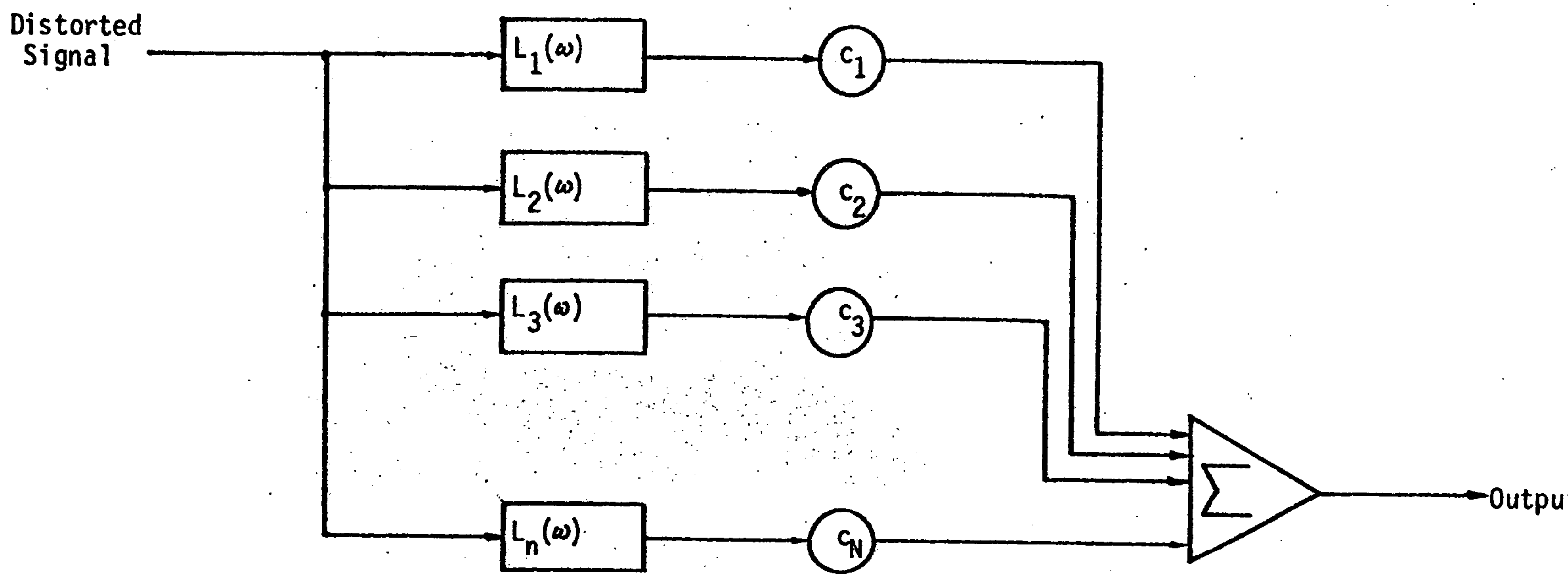


FIG.1.3 PARALLEL FORM OF EQUALISER

There appears to be no general answer yet as to the best form the transfer functions should take. In the context of this work, both the Laguerre networks and the delay line (using a rational function approximation to pure delay) were considered.

It is possible also to use a parallel arrangement as the equaliser as shown in Fig. 1.3. Like the serial arrangement, its output is given by the sum of weighted responses. It has the additional advantage of greater flexibility. On the other hand, for the serial layout, much of the filtering necessary for a particular response is performed by the preceding networks.

The operation of the equaliser, whether of the serial or parallel kind, involves continuously adjusting a set of tap gains to satisfy a chosen performance criterion. When the number of taps to be adjusted is large, the question invariably arises of interaction between the various settings. In general, interaction between parameters to be optimised tends to slow down the optimisation process. Kitamori (3) has shown that this is true for the equaliser configuration being considered, and that each tap setting turns out to be a function of the others as well.

This interaction can be reduced by using a set of networks whose outputs are orthogonal as the filter set for the equaliser. Rosenberger and Thomas (4) have suggested the use of a set of M filters having orthonormal impulse responses which are the first M members of a complete basis set, while Kitamori has proposed the use of the Laguerre set. It is not certain, however, if such a provision is really necessary since orthogonality is affected by the nature of the input signal, which is the output from the distorting channel in this case. Hence, because the response of the distorting channel is unknown, a priori, the desired orthogonality cannot usually be assured.

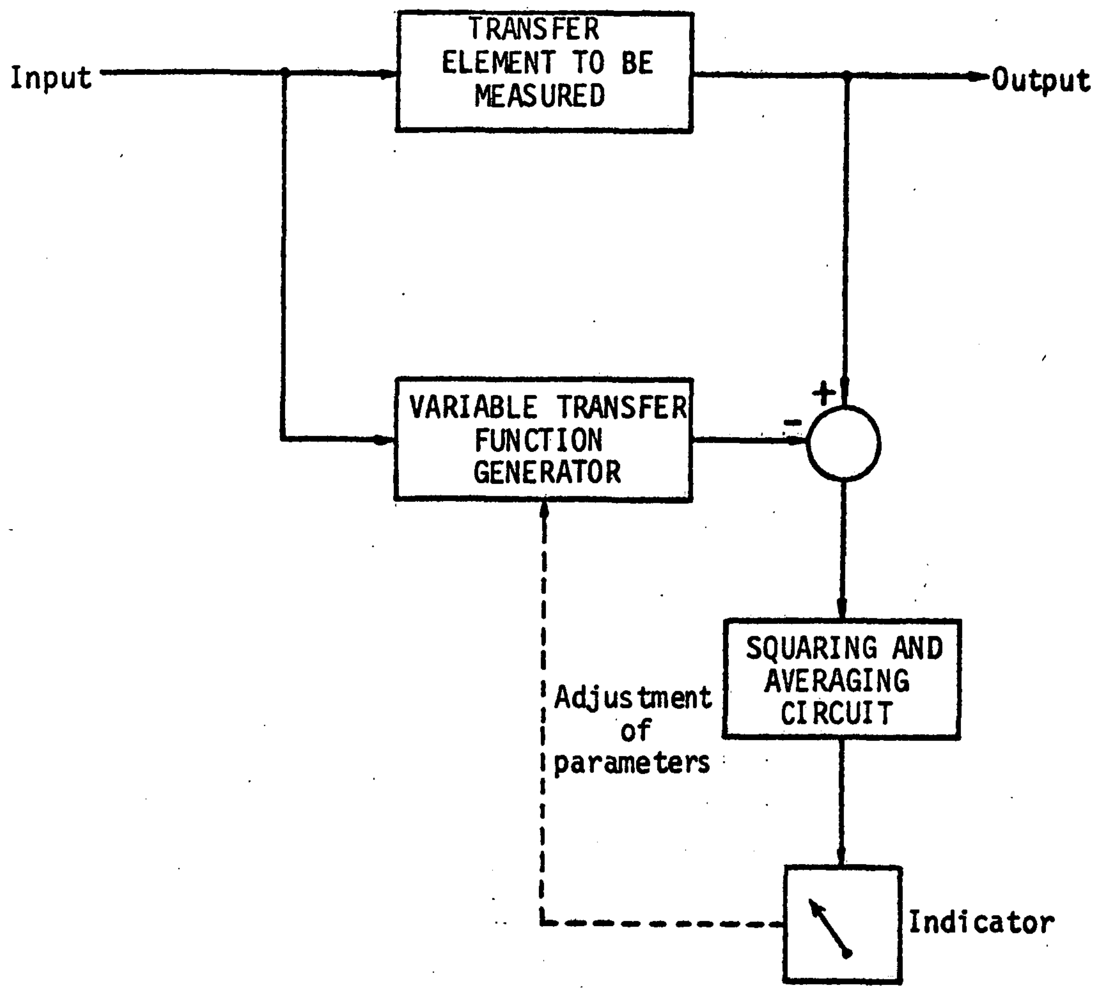


FIG.1.4 TRANSFER FUNCTION MEASUREMENT

1.4 Related Applications

In the course of equalisation, an automatic equaliser is called upon to perform a network synthesis. This suggests other areas of application utilising the same principle.

(a) Transfer function measurement

The conventional method involves the application of a step or sinusoidal input to the element whose transfer function is to be evaluated, and analysing the corresponding output. The transfer characteristics thus obtained are referred to as the step response and the frequency response respectively. There are a number of disadvantages with this conventional method:-

(i) the analytical form of the transfer function cannot usually be obtained easily.

(ii) the transfer characteristics obtained may be different from those under actual operating conditions, especially when the element has some non-linearity.

(iii) the test input so applied may disturb the operating system in which the element is involved.

A method which overcomes the above objections is the null method (3), using the arrangement shown in Fig. 1.4. The mean square of the difference between the outputs of the element and a variable transfer function generator, with the same input signal applied to both, is minimised by adjusting the variable parameters of the transfer function generator. The variable generator can take a number of forms,

e.g. first order $\dfrac{K}{sT + 1}$ or second order $\dfrac{K\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$

where K, T, ζ, ω_n are adjustable parameters.

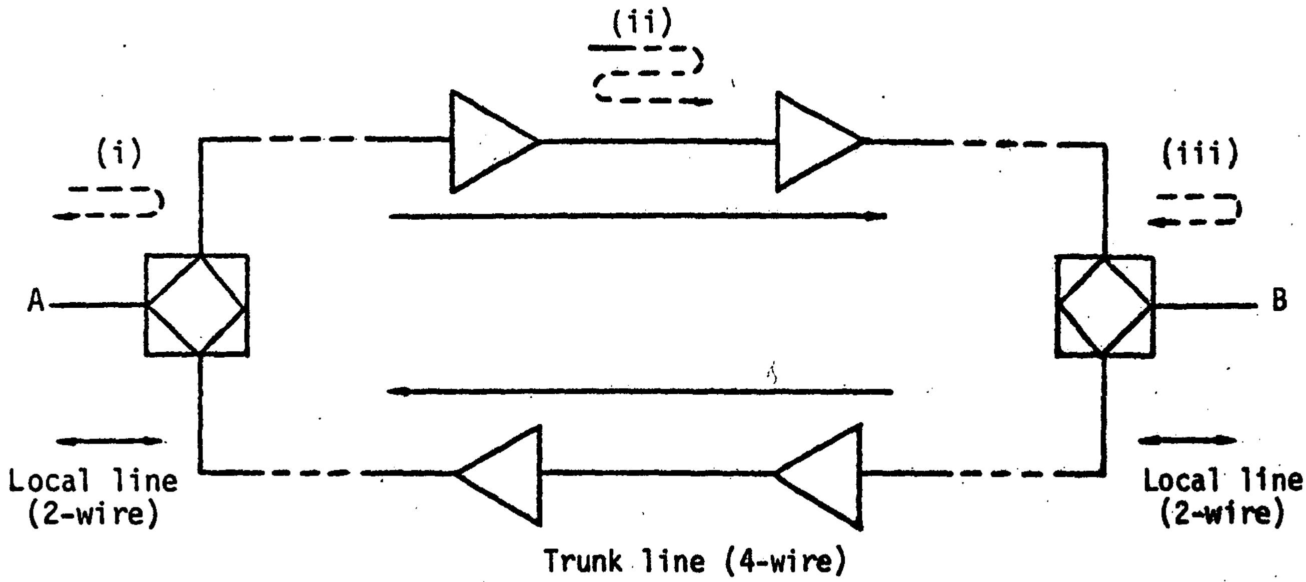


FIG.1.5(a) REFLECTIONS ARISING IN A LONG-DISTANCE TELEPHONE CALL

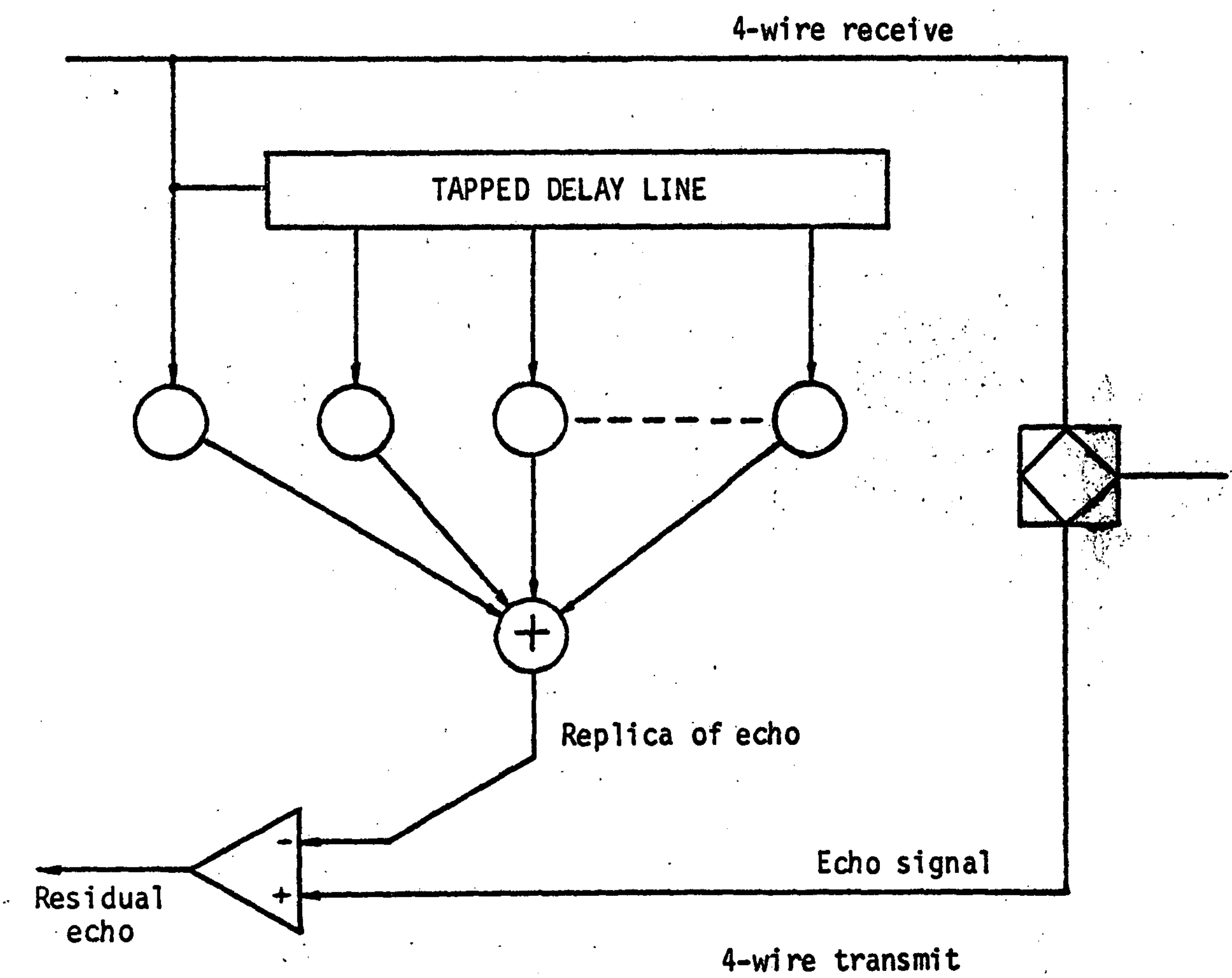


FIG.1.5(b) ECHO CANCELLATION (CONTROL LOOPS FOR TAP ADJUSTMENTS NOT SHOWN)

With this method, real time measurement is possible. Since no additional test input is needed, no unwanted disturbances are introduced into the system. Moreover, the transfer characteristics are now obtainable analytically in the form of a transfer function. Even if the element has some non-linearity, this method will satisfactorily linearise the transfer characteristics, taking account of the actual operating conditions.

(b) Echo cancellation

In the typical long-distance telephone conversation, three main types of reflections arise due to mismatch at the interfaces between the different transmission equipment. As shown in Fig. 1.5(a), where a transmission from A to B is assumed, these comprise:-

(i) reflected signals in the local line near A. They become audible at A practically underlayed and are known as side tones.

(ii) reflected signals that arise in the four-wire path. They suffer some attenuation owing to the repeated passage between the mismatch points.

(iii) reflected signals in the far local line near B. These return to A with a substantial delay and are heard as echoes.

In general, disturbances arising from reflections (i) and (ii) are unobjectionable. The echo delays that arise as a result of (iii) must however be taken care of. Typical echo delays are 20 milliseconds for a landline over 1000 miles and increase to 600 milliseconds for a call routed over a geo-stationary satellite. Such echoes are annoying or disturbing according to their level and delay. Previous techniques of control involving the introduction of attenuation into the echo path have proved to be unsatisfactory, especially for the longer delays.

The new technique of echo cancellation (2) uses principles identical to that of the equaliser to generate a replica of the echo. This is then subtracted from the return signal containing the unwanted echo. Synthesis of the replica is by means of a filter which, under the control of a feedback loop, adapts to the transmission characteristics of the echo path. Implementation is also possible such that the echo canceller is able to track any variations of the echo path that may occur during a conversation. The schematic of a typical echo canceller is shown in Fig. 1.5(b).

1.5 Performance Criteria

Since the process of equalisation involves matching the actual channel response to a specified response according to some measure of performance, the basic need is to establish a criterion for optimising the error of equalisation. One of two basic criteria is usually adopted - minimisation of either the absolute value or the mean square value of the error.

The first criterion uses a zero-forcing algorithm of the kind described by Lucky (5). The aim is to force the error to zero at a finite number of equispaced instants. Implementation is by passing both the error signal and the equalised output through infinite clippers to form the respective signum functions. The product of these functions is averaged and used to control the tap settings. Because only binary multiplication is involved, simple exclusive OR circuits can be used to form the product. The clipping operation introduced means, however, that some information has been discarded. As a result, the convergence properties are not as good as in the mean square error case. This type of equaliser can thus be used only for channels of relatively low dispersion.

The second criterion requires more complex hardware implementation. Because linear correlation of the error signal with the input signal is required, correlators are needed that must be accurate and have a bandwidth of operation compatible

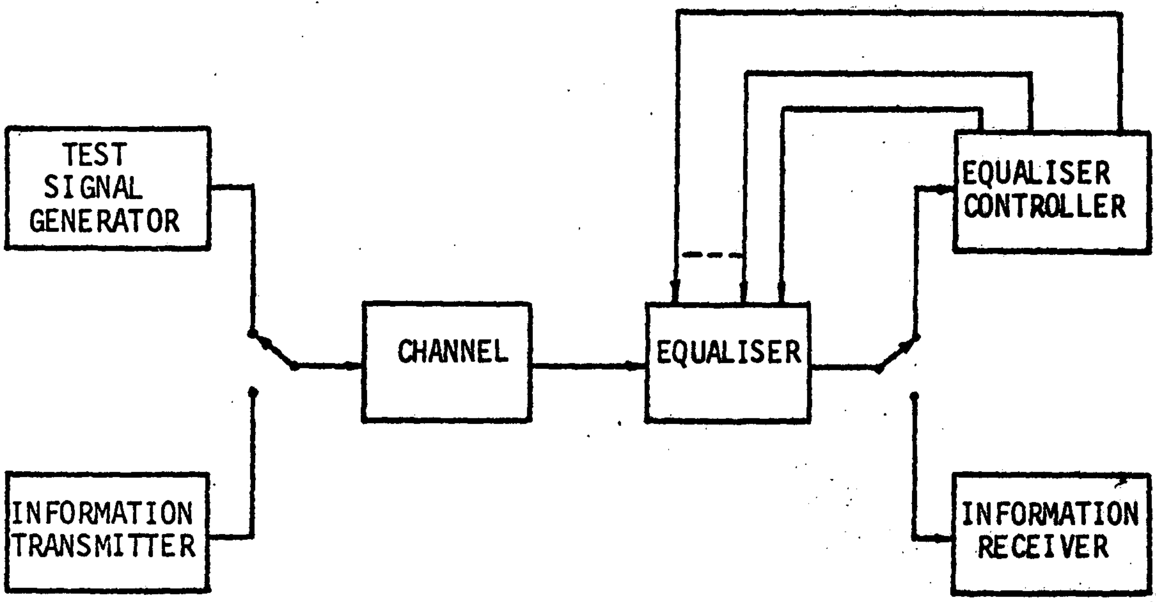


FIG.1.6(a) PRESET TYPE OF EQUALISER

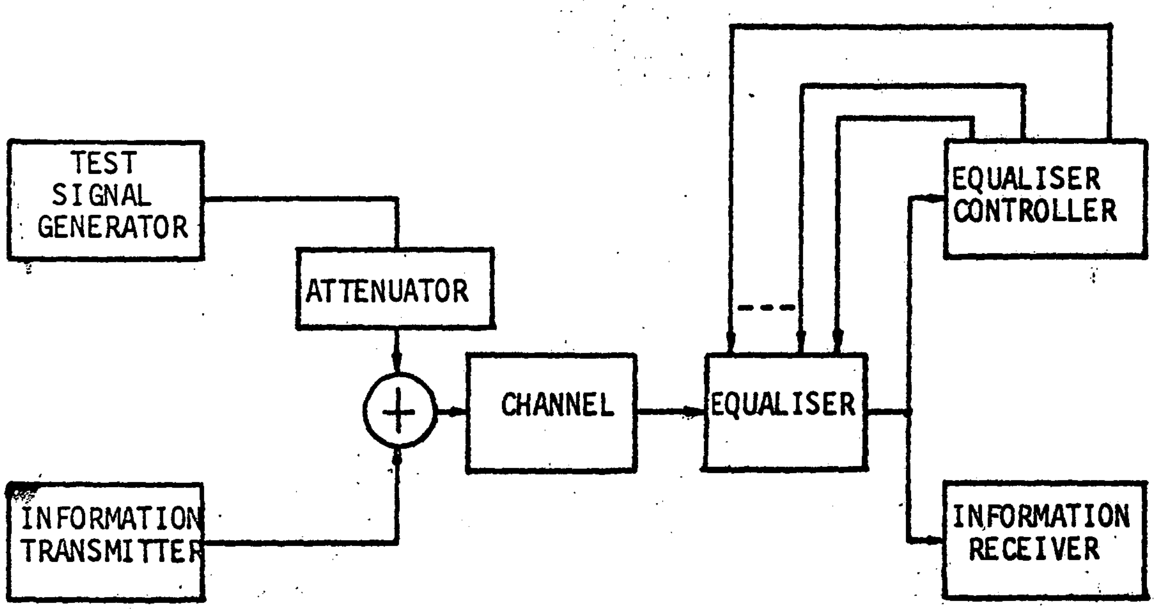


FIG.1.6(b) ONE FORM OF ADAPTIVE EQUALISER

with that of the channel. With the advance of integrated circuit technology, it is expected that this should not be a major problem. Moreover, the mean square error equaliser has far superior convergence properties. Convergence is assured for any initial distortion, and because of the quadratic nature of the error surface, convergence is to a unique minimum. Also, a number of iterative algerithms can be used successfully with the mean square error criterion.

1.6 Mode of Operation

Another way in which equalisers may be classified is according to the particular mode of operation. Firstly, equalisers may be operated in the training mode (6), in which case they are usually referred to as the present type. Adjustments are made in a training period prior to or during breaks in information transmission by using a pre-arranged test signal. The general set-up is given in Fig. 1.6(a). After the channel has been equalised, the test signal is removed. With the equaliser settings held constant, the equalised channel can then be used for information transmission. Most of the work done so far in the field of automatic equalisation is concerned with this type of equaliser. The same approach will be adopted for this investigation.

Secondly, equalisers may be operated in a tracking mode (7), in which case they are usually referred to as the adaptive type. No training period is required in this case, but equalisation is performed during the actual transmission of useful information. The equaliser settings are derived either from a low level reference signal sent with the data signal or from operating on the data signal itself. The controller operates continuously so that the channel is always being monitored. Thus if the frequency characteristic of the transmission path should change, the equaliser would immediately begin to compensate for the change. The adaptive equaliser can therefore be used for correcting time-varying distortions.

One form of adaptive equaliser has been suggested by Rudin (7). As shown in Fig. 1.6(b), the operation relies on the sending of a low level test signal simultaneously with the information-bearing signal. Despite its low energy level, it can be extracted by high performance correlation detectors and used to make the corrective adjustments. But because it appears in the high "noise" environment created by the information-bearing signal, steps must be taken to ensure that it is the test signal and not the "noise" that dictates the behaviour of the equaliser. Using a pseudo-random sequence as the test signal, Rudin has devised a scheme to ensure that the tap settings are independent of the "noise". The set-up is basically the same as that of the preset type except for the fact that the error signal is passed through a delay of T_D seconds before reaching the correlators. T_D is chosen to be equal to a multiple of the period of the pseudo-random sequence. Since the test signal and information signal are generated in completely different fashions, it is reasonable to assume that they are uncorrelated. Also, because T_D is selected large compared with the channel's dispersion, time translations of the "noise" are also independent. With these assumptions, and the periodicity of the test waveform, it works out that the equaliser responds only to the test signal and not to the high background "noise" present.

Initial experiments performed by Rudin have indicated the potential of such a technique. Two important problems that need further investigation are -

(i) the correlators are required to perform in a very hostile environment. They must be able to search out small components of a signal which is buried in "noise" and this requires tremendous dynamic range. Moreover, the long observation times that are needed impose stringent demands, such as the absence of any significant output drift, on the accuracy of the integrators that perform the cross-correlation.

(ii) the delay line of length T_D is critical to the satisfactory operation of the adaptive equaliser. Any irregularities in this long line will severely affect the performance, as the correlators have no means of compensating for even relatively small distortions that may occur in the line.

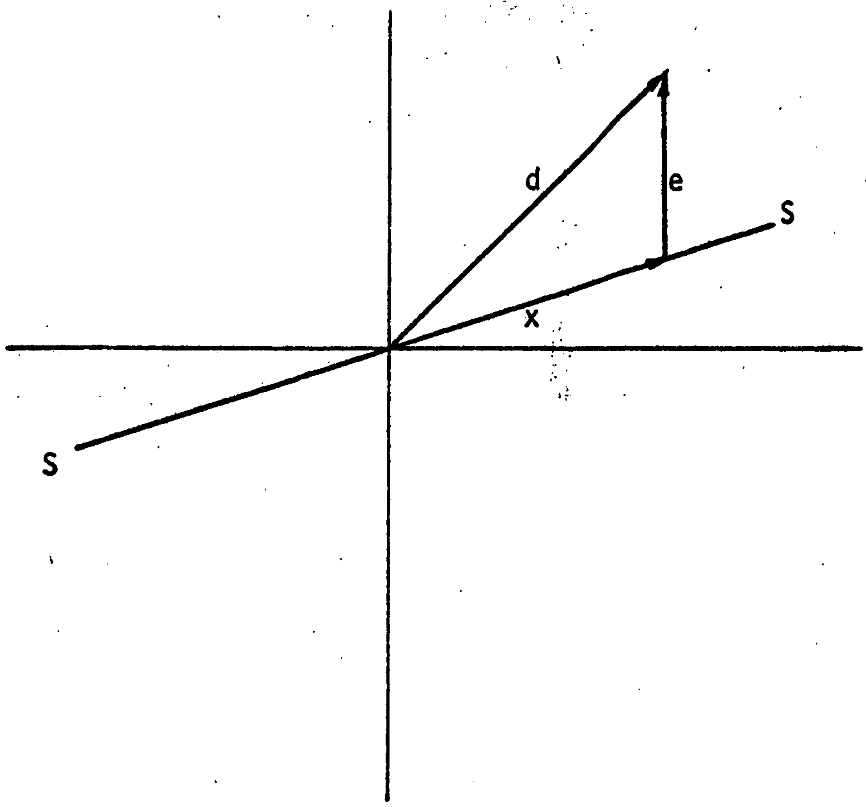
These two problems indicate that considerable advantage might well be obtained in an all-digital implementation. The only source of distortion then would be in the initial analogue-to-digital conversion. Delays of arbitrary length can easily be realised by shift registers, and correlators can be constructed to arbitrary accuracy. More investigation will be needed before such an adaptive equaliser can be practically implemented.

1.7 Hilbert Space Interpretation of Problem

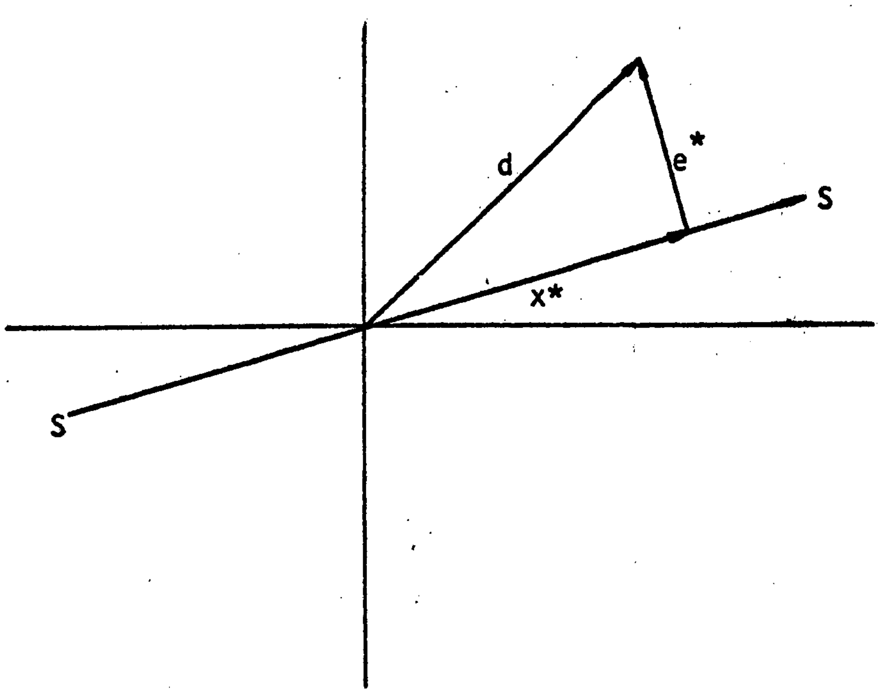
In the context of this work, "optimum" equalisation has been defined with reference to a mean square error criterion. Much work has been done in the past using such a criterion. They include systems which operate on either analogue or digital signals which are considered to be either deterministic or realisations of stochastic processes. Gibbs (8) has made the interesting observation that all these systems have an identical mathematical formulation (implying that they are mathematically indistinguishable) when working with the mean square error criterion. This unification is achieved by using the mathematical properties of the Hilbert space, which may be considered as a generalisation of an n-dimensional Euclidean space, but retaining the geometric properties of length and angle.

The geometrical interpretation of mean square minimisation can be understood by first looking at the following two-dimensional problem:-

"Given a fixed vector d in a plane and a fixed line S through the origin, find the vector x in S which is closest to d,



(a) Simple Problem



(b) Optimum Solution

FIG.1.7 TWO-DIMENSIONAL GEOMETRICAL PROBLEM i.e. minimise the length of the error vector e." (See Fig. 1.7)

The solution to this problem has three significant features:

(i) a solution always exists

(ii) the solution is unique

(iii) the optimum vector x^* is found by dropping a perpendicular from d to S i.e. $d - x^*$ is orthogonal (at right angles) to every vector in S.

In the equaliser problem, the vector d represents the desired response, the line S represents the subspace generated by the output of the transversal filter as its tap gains are varied, and the optimum tap settings give rise to x^*, which is the orthogonal projection of d on S. The control algorithm is that which causes the error to be orthogonal to the tap outputs (and therefore S). Derivation of the optimum settings is equivalent to the solution of a set of simultaneous linear equations by an iterative technique.

The Hilbert space formulation of the problem thus becomes:

"Given a vector d in the Hilbert space H and a closed subspace S of H find the vector x^* in S with the property that $\| d - x^* \| = \inf (\|d - y\| : y \in S)$ i.e. find the vector in S closest to d", where $\| d - x^* \|$ denotes the norm of the vector $(d - x^*)$.

As stated, this general abstract problem always has a unique solution, given by the orthogonal projection of the desired response on the subspace spanned by the transversal filter. The result can be directly applied to various physical problems by considering them as pertinent realisations of abstract Hilbert space.

To provide further insight into the process of automatic equalisation, the next chapter will deal with mathematical analysis of the equaliser operation and the control algorithm used to adjust the tap settings. Both the dynamic (rate of convergence) and steady state (residual error) performance will be assessed.

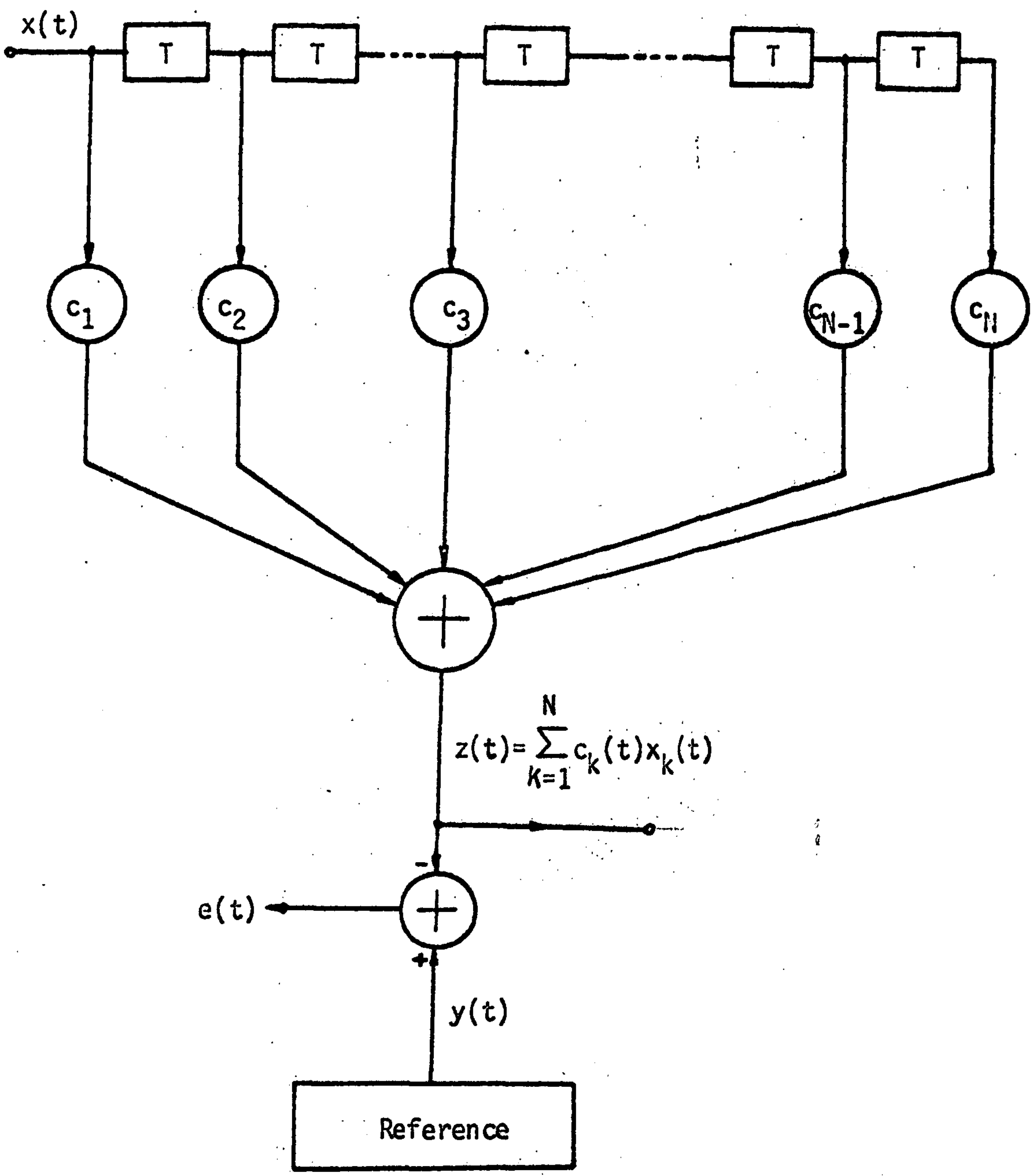


FIG. 2.1 DIAGRAM ILLUSTRATING BASIC PRINCIPLES OF OPERATION OF EQUALISER (CONTROL LOOPS NOT SHOWN)

2. Mathematical Study of the Equalisation Process

The flexibility of the transversal type equaliser has resulted in its frequent use in the automatic or adaptive equalisation of analogue or digital communication channels. Recent work has centred on the use of this device for specific channels, the use of various algorithms for setting the tap gains automatically, and the use of various performance criteria.

2.1 Basic Principles of Operation

Let $x(t)$ be the distorted signal to be equalised and let $y(t)$ be the reference signal, as shown in Fig. 2.1.

The output of the N-tap transversal filter is given by

$$z(t) = \sum_{k=1}^{N} c_k(t)\, x_k(t)$$

where $c_k(t)$ is the gain of the kth tap

and $x_k(t) = x(t - (k-1)T)$ for T = delay of each section of the transversal filter.

$$\therefore \text{Error } e(t) = y(t) - z(t) \qquad (2.1)$$

It is proposed to treat these continuous time functions as sampled functions with the sampling interval tending to zero in the limit. Though not mathematically rigorous, it is amenable to expression in vector-matrix form and allows for ease of manipulation. Equation (2.1) thus becomes

$\underline{e} = \underline{y} - \underline{z}$ where the underlined terms are column matrices. Under a mean square error criterion,

$$\varepsilon = (\underline{y} - \underline{z})^T (\underline{y} - \underline{z})$$

where the superscript T denotes transpose of the matrix.

$$= \underline{y}^T\underline{y} - 2\underline{z}^T\underline{y} + \underline{z}^T\underline{z}$$

$$= \underline{y}^T\underline{y} - 2\underline{c}^T\underline{b} + \underline{c}^T A\, \underline{c} \qquad (2.2)$$

where $A = \underline{xx}^T$, $\underline{b} = \underline{xy}$

A represents the signal + noise correlation matrix where it has been assumed that $\underline{x}$ comprises both the distorted signal and any additive noise. Note that A is symmetric. It can also be shown to be positive definite. The proof is as follows (9):-

$$\underline{c}^T A\, \underline{c} = \underline{c}^T \underline{x}\underline{x}^T \underline{c} = \sum_k (c_k x_k)^2$$

$$= \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} |X(\omega)|^2 \, |C(\omega)|^2 \, d\omega \qquad \text{by Parserval's}$$

theorem, where $X(\omega)$ is the received signal Fourier transform.

This implies that $\underline{c}^T A\, \underline{c}$ is non-negative for all vectors $\underline{c}$. But the received signal has finite energy and

$$C(\omega) = \sum_{k=1}^{N} c_k e^{-jk\omega T} \qquad \text{can have only isolated zeroes.}$$

Therefore, it can be inferred that $\underline{c}^T A\, \underline{c} > 0$ unless $\underline{c} = 0$, which proves that A is positive definite.

From equation (2.2), it can be seen that the mean square error is a quadratic function of the tap weights $\underline{c}$. Its gradient is given by

$$\nabla_{\underline{c}}\varepsilon = 2(A\,\underline{c} - \underline{b}),$$

which indicates that the gradient is a linear function of $\underline{c}$. The advantage of working with a quadratic performance measure lies both in this linear relation and in the freedom from relative minima.

For minimum mean square error,

$$\nabla_{\underline{c}}\varepsilon = 2(A\,\underline{c} - \underline{b}) = 0 \qquad (2.3)$$

Since A is non-singular, it has an inverse A^{-1}. Equation (2.3) can thus be solved to give

$$\underline{c}_{opt} = A^{-1}\underline{b} \qquad (2.4)$$

where $\underline{c}_{opt}$ is the optimum tap setting. For $\underline{c} = \underline{c}_{opt}$, i.e. at convergence, the mean square error is given by

$$\varepsilon = \varepsilon_{min} = \underline{y}^T\underline{y} - 2\,\underline{c}_{opt}^T\underline{b} + \underline{c}_{opt}^T A\,\underline{c}_{opt} \qquad (2.5)$$

using equation (2.2).

2.2 Tap Control Algorithm

The optimum tap setting $\underline{c}_{opt}$ is usually obtained in an iterative manner using some form of the gradient algorithm as opposed to solving the simultaneous equation $A\underline{c}=\underline{b}$ directly.

Assuming that the gradient of ε with respect to the tap weights is available, the taps are adjusted according to the recursion

$$\underline{c}_{i+1} = \underline{c}_i - \frac{1}{2}\alpha_i \nabla_{\underline{c}}\varepsilon \qquad (2.6)$$

where the subscript i denotes the ith iteration and takes the values 0,1,2, ..., and α_i's are suitably chosen constants.

In the gradient algorithm, the tap weights are moved in a direction opposite to that of the gradient (the significant feature of using a gradient algorithm is that the gradient can be conveniently evaluated without any knowledge of the actual error surface itself. Moreover, no derivative computation is needed).

$\therefore$ Substituting for $\nabla_{\underline{c}}\varepsilon$ from equation (2.6),

$$\underline{c}_{i+1} = \underline{c}_i - \alpha_i\,(A\underline{c}_i - \underline{b}) \qquad (2.7)$$

If the cross-correlation between the equalisation error $\underline{e}$ and the input to the equaliser $\underline{x}$ is computed and used to adjust the tap settings,

$$\begin{aligned}\underline{c}_{i+1} &= \underline{c}_i + \alpha_i\,(\underline{x}\,\underline{e}) \qquad (2.8)\\ &= \underline{c}_i + \alpha_i\,\underline{x}\,(\underline{y} - \underline{z})\\ &= \underline{c}_i + \alpha_i\,\underline{x}\,\underline{y} - \alpha_i\,\underline{x}\,\underline{x}^T\underline{c}_i\end{aligned}$$

The iterations have converged when $\underline{c}_i = \underline{c}_{i+1} = \underline{c}_{opt}$

i.e. when $\underline{x}\,\underline{y} = (\underline{x}\,\underline{x}^T)\,\underline{c}_{opt}$

$$\therefore\quad \underline{c}_{opt} = (\underline{x}\,\underline{x}^T)^{-1}\underline{x}\,\underline{y} \qquad (2.9)$$

Note that this is the same as the ordinary least squares estimate (10) to minimise $\underline{e}^T\underline{e}$. Moreover, it has been defined in the previous section that $A = \underline{x}\,\underline{x}^T$ and $\underline{b} = \underline{x}\,\underline{y}$. This means that equation (2.9) is equivalent to $\underline{c}_{opt} = A^{-1}\underline{b}$, which in turn is the result established in equation (2.4). It can thus be seen that one way of implementing the gradient algorithm is by determining the cross-correlation between the equalisation error and the output of the distorting channel, and using this to update the tap settings. Also, the least squares approach can be used to give an identical mathematical formulation of the optimum tap setting.

At convergence, the cross-correlation $\underline{x}\,\underline{e}$ becomes

$$\begin{aligned}
\underline{x}\,\underline{e} &= \underline{x}\,\underline{y} - \underline{x}\,\underline{z} \\
&= \underline{x}\,\underline{y} - \underline{x}\,(x^T\underline{c}_{opt}) \\
&= \underline{x}\,\underline{y} - \underline{x}\,\underline{x}^T\,(\underline{x}\,\underline{x}^T)^{-1}\underline{x}\,\underline{y} && \text{from equation (2.9).} \\
&= 0
\end{aligned}$$

Therefore, the error sequence becomes orthogonal to the input sequence at convergence. This result is intuitively obvious since when all the cross-correlation coefficients become zero, no further adjustment of the taps can improve the equalisation (see equation (2.8)), and the optimum condition is reached.

Before proceeding further with the analysis, it is necessary to introduce a tap gain error vector defined by

$$\underline{p} = \underline{c} - \underline{c}_{opt}$$

$\therefore$ From equation (2.6), it follows that

$$\begin{aligned}
\underline{p}_{i+1} &= \underline{c}_i - \alpha_i\,(A\underline{c}_i - \underline{b}) - \underline{c}_{opt} \\
&= \underline{c}_i - \underline{c}_{opt} - \alpha_i A\,(\underline{c}_i - \underline{c}_{opt}) && \text{since } A\,\underline{c}_{opt} = \underline{b}
\end{aligned}$$

$$\therefore\ \underline{p}_{i+1} = (I - \alpha_i A)\,\underline{p}_i \tag{2.10}$$

where I = unity matrix.

From equation (2.2),

$$\begin{aligned}
\varepsilon_i &= \underline{c}_i^{\,T} A\,\underline{c}_i - 2\,\underline{c}_i^{\,T}\underline{b} + \underline{y}^T\underline{y} \\
&= (\underline{p}_i = \underline{c}_{opt})^T A\,(\underline{p}_i + \underline{c}_{opt}) - 2(\underline{p}_i + \underline{c}_{opt})^T\underline{b} + \underline{y}^T\underline{y} \\
&= \underline{p}_i^{\,T} A\,\underline{p}_i + 2\underline{p}_i^{\,T} A\,\underline{c}_{opt} - 2\underline{p}_i^{\,T}\underline{b} + \varepsilon_{min}
\end{aligned}$$

using equation (2.5)

$$= \varepsilon_{min} + \underline{p}_i^T A\, \underline{p}_i \text{ since } \underline{p}_i^T A\, \underline{c}_{opt} = \underline{p}_i^T A\, A^{-1}\underline{b} = \underline{p}_i^T \underline{b}$$

from equation (2.4)

$\therefore$ Mean square error at ith iteration is given by

$$\varepsilon_i = \varepsilon_{min} + \underline{p}_i^T A\, \underline{p}_i \tag{2.11}$$

The term $\underline{p}_i^T A\, \underline{p}_i$ represents the excess mean square error. In practice, ε_{min} cannot usually be attained and there will be an excess error even at convergence.

2.3 Convergence of Tap Gain Error

An explicit condition for the convergence of the iterative algorithm described in the previous section can be obtained by considering equation (2.10).

$$\underline{p}_{i+1} = (I-\alpha_i A)\, \underline{p}_i$$

$$= (I-\alpha_i A)\, (I-\alpha_{i-1} A) \;....\; (I-\alpha_0 A)\, \underline{p}_0$$

For a constant step size parameter i.e. $\alpha_i = \alpha =$ constant,

$$\underline{p}_{i+1} = (I-\alpha A)^{i+1}\, \underline{p}_0$$

According to Lytle (11), $\underline{p}_{i+1}$ will converge to zero if and only if $(I-\alpha A)^{i+1}$ converges to zero, which is equivalent to the requirement that all the eigenvalues of $(I-\alpha A)$ be less than one in magnitude,

i.e. $|1-\alpha\lambda_k|<1$ where (λ_k) is the set of eigenvalues of A and are all real since A is a real, symmetric matrix.

i.e. $-1<1-\alpha\lambda_k<1$

Thus, for convergence, the step size factor α must satisfy the relation

$$0<\alpha<2/\lambda_{max} \tag{2.12}$$

where λ_{max} is the maximum eigenvalue of A.

For α within this range, the iterations converge at least geometrically for any starting point at a rate given by ζ, the spectral radius of $(I-\alpha A)$, where the spectral radius of a matrix is defined to be the magnitude of the largest eigenvalue of the matrix.

$$\therefore \quad \zeta = \max_{k} \; |1-\alpha\lambda_k|$$

Convergence is fastest if ζ is minimised.

Since λ_k lies in the interval $(\lambda_{min}, \lambda_{max})$, the function $|1-\alpha\lambda_k|$ assumes its maximum at one of these end points for every α. The best choice of α, i.e. the one for which max $(|1-\alpha\lambda_{min}|, |1-\alpha\lambda_{max}|)$ is smallest, is obtained when (12, p.225)

$$1-\alpha\lambda_{min} = -(1-\alpha\lambda_{max})$$

$$\text{i.e.} \quad \alpha = \frac{2}{\lambda_{min} + \lambda_{max}} \tag{2.13}$$

Equation (2.13) represents the optimum value of α to be used for a constant step size algorithm. With this value of α,

$$\begin{aligned} \zeta &= 1-\alpha\lambda_{min} \\ &= 1-\left(\frac{2\lambda_{min}}{\lambda_{min} + \lambda_{max}}\right) \\ &= \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} \qquad (2.14) \\ &= \frac{r-1}{r+1} \quad \text{for} \quad r = \frac{\lambda_{max}}{\lambda_{min}} \end{aligned}$$

The factor r must therefore be reduced in order to speed convergence. It can thus be seen that the rate of convergence is related to the spread of the eigenvalues. To obtain rapid convergence of the tap error $\underline{p}$ to zero, the eigenvalues of A must have values which are close to one another.

In most practical situations, however, the eigenvalues of A are not known. Equation (2.14) can still provide useful information if a set of bounds can be derived for these eigenvalues, based on the spectrum of the received signal and noise (9).

Let $\lambda_{min} \geqslant m = \min_{|\omega| \leq \pi/T} (|X(\omega)|^2)$ and

$\lambda_{max} \leqslant M = \max_{|\omega| \leq \pi/T} (|X(\omega)|^2)$

where $X(\omega)$ is the received signal Fourier Transform

$$\therefore \quad r = \frac{\lambda_{max}}{\lambda_{min}} \leqslant \frac{\lambda_{max}}{m} \leqslant \frac{M}{m}$$

It follows that $\zeta = \frac{M-m}{M+m}$

Therefore, for the best choice of α, convergence proceeds at least at a rate given by the geometric factor $(M-m)/(M+m)$. Thus, useful information regarding convergence can be obtained without knowledge of the actual channel characteristics.

2.4 Consideration of Step Size

From equation (2.12), it is clear that the choice of the step size factor has a direct influence on the convergence of the tap gain error vector. Therefore, it is proposed next to consider further the bounds on the step size to ensure convergence.

2.4.1 Constant Step Size

A has been defined to be the received signal + noise correlation matrix. In any practical channel, noise will always be present. Channel noise affects the system performance in two ways:

(i) It introduces a direct error by its presence in the final output.

(ii) It introduces an error in the setting of the tap gains. In the presence of noise, the tap weight corrections (gradient estimates) contain undesired random components. A consequence of this use of noisy gradient estimates rather than the true value of the gradient components is that it gives rise to an excess mean square error at the output, referred to earlier. As a result, the tap weights no longer converge to the optimal value but instead, approach some neighbourhood of a sub-optimal setting. It then fluctuates randomly about this setting.

With regards to (ii), the amplitude of the random fluctuations increases with α. It has been pointed out that increasing α brings about an increase in the convergence rate. It is now clear, however, that the excess mean square error is also increased as a result. There is thus a practical upper limit on the value that α can take that is normally much smaller than that given in equation (2.12). This upper limit can be found by considering equation (2.11):

$$\varepsilon_i = \varepsilon_{min} + \underline{p}_i^T A\, \underline{p}_i$$

Let $(\underline{u}_k)$ be a set of orthonormal eigenvectors of A and let Q be a NxN matrix whose kth column is the eigenvector $\underline{u}_k$. Q is thus a modal matrix of A and is orthogonal which implies that

$Q^T = Q^{-1}$ or $Q^TQ = I$.

Also, let D be a NxN diagonal matrix whose kth diagonal element is the eigenvalue λ_k.

Under this notation, the symmetric matrix A can be represented in the form

$$A = QDQ^T$$

The proof follows simply from the fact that $AQ = QD$.

Premultiplying both sides by Q^{-1} then gives

$$\begin{aligned} A &= QDQ^{-1} \\ &= QDQ^T \quad \text{since } Q^T = Q^{-1} \end{aligned}$$

Equation (2.11) thus becomes

$$\begin{aligned} \varepsilon_i &= \varepsilon_{min} + \underline{p}_i^T\, QDQ^T \underline{p}_i \\ &= \varepsilon_{min} + \underline{s}_i^T D\, \underline{s}_i \qquad \text{where } \underline{s}_i = Q^T \underline{p}_i \\ &= \varepsilon_{min} + \sum_{k=1}^{N} \lambda_k s_i^2 \qquad (2.15) \end{aligned}$$

Proakis and Miller (13) have shown that the average value of the excess mean square error (i.e. the second term in equation

(2.15)) works out to be

$$\Delta\varepsilon = \frac{\alpha}{2}\,\varepsilon_{min}\,\mathrm{tr}(A) \qquad (2.16)$$

where $\mathrm{tr}(A)$ = trace of $A = \sum_{k=1}^{N} \lambda_k$

In other words, the excess mean square error is dependent upon α, the step size factor. The choice of α thus involves a trade off between the speed of convergence (fast convergence for larger α) and the amount of excess mean square error (small error for smaller α) that can be tolerated. From equation (2.16),

$$\alpha = \frac{\Delta\varepsilon}{\varepsilon_{min}} \quad \frac{2}{\mathrm{tr}(A)}$$

Consider $\frac{\Delta\varepsilon}{\varepsilon_{min}} < 1$ (the reasonableness of this choice will depend on the convergence speed and the size of the varying errors, i.e. on the relative importance of convergence speed and jitter about the optimum).

This implies that

$$\alpha < \frac{2}{\mathrm{tr}(A)} < \frac{2}{\lambda_{max}} \qquad (2.17)$$

since $\mathrm{tr}(A) = \sum_k \lambda_k > \lambda_{max}$.

Thus choosing α according to equation (2.17) still gives a value that is within the range required for convergence of the tap gain error vector.

2.4.2 Variable Step Size

The conflicting requirements between the dynamic (rate of convergence) and steady state (mean square error attainable) performance arise from the use of a fixed step size parameter α. Mark (14) has proposed a variable step size procedure, with decreasing step sizes as the iterative algorithm converges, to overcome this problem, i.e. $|\alpha(i)|$ is made a decreasing function of i. In this proposal, the ratio $\alpha(i)/\sqrt{\varepsilon(n)}$ is kept constant. The initial choice $\alpha(0)$ is then given by

$$\alpha(0) = \alpha(s/s)\sqrt{\frac{\varepsilon(0)}{\varepsilon(s/s)}}$$ where s/s denotes the steady state value.

$$= \sqrt{\frac{\varepsilon(0)}{\varepsilon(s/s)}} \; \frac{\Delta\varepsilon}{\varepsilon_{min}} \; \frac{2}{tr(A)}$$ from equation (2.16).

For $\frac{\Delta\varepsilon}{\varepsilon_{min}} = 0.05$, and a 90% reduction in the mean square error from initial to steady state conditions,

$$\alpha(0) = \frac{0.1}{tr(A)} \; \sqrt{10}$$ which is within the range prescribed for convergence.

Provided $0<\alpha(0)<\frac{2}{\lambda_{max}}$, other choices of $\alpha(0)$ are possible. However, if the choice is too large, there is the danger that the system may be driven too hard. This would produce large ripples in the transient response which could require a long time to settle.

2.5 Convergence of Excess Mean Square Error

Instead of considering the dynamic behaviour of the tap gain error vector, as has been done so far, convergence can also be studied in terms of the actual mean square error. From equation (2.11), it is seen that ε_i, the mean square error corresponding to the ith iteration, consists of two terms, the first term ε_{min} being irreducible. Further insight into the convergence of the equalisation process can be obtained by looking at the second term $\underline{p}_i^T A\, \underline{p}_i$ (15).

It was shown in the previous section that A, the received signal + noise correlation matrix, is expressible as $A = QDQ^T$. Substituting for A in equation (2.10),

$$\underline{p}_{i+1} = Q\,(I-\alpha_i D)\,Q^T \underline{p}_i$$

Doing the same successively for $\underline{p}_i$, p_{i-1}, ...,

$$\underline{p}_{i+1} = Q\,(I-\alpha_i D)\,Q^T Q\,(I-\alpha_{i-1} D)\,Q^T \ldots Q^T Q\,(I-\alpha_0 D)\,Q^T \underline{p}_0 .$$

Since $Q^TQ = I$, this simplifies to

$$\underline{p}_{i+1} = Q \left(\prod_{j=o}^{i} (I-\alpha_j D) \right) Q^T \underline{p}_o \qquad (2.18)$$

$\therefore$ The excess mean square error term becomes

$$\underline{p}_i^T A \, \underline{p}_i = (\underline{p}_o^T Q \, (\prod_{o}^{i-1} (I-\alpha_j D)^T) Q^T)(QDQ^T)(Q(\prod_{o}^{i-1} (I-\alpha_j D))Q^T \underline{p}_o)$$

$$= \underline{p}_o^T Q \, (\prod_{o}^{i-1} (I-\alpha_j D)^T) \, D(\prod_{o}^{i-1} (I-\alpha_j D)) \, Q^T \underline{p}_o \qquad (2.19a)$$

$\prod_{o}^{i-1} (I-\alpha_j D)$ is diagonal (since D is the diagonal matrix of eigenvalues) and therefore symmetric.

The matrix multiplication in (2.19a) thus yields,

$$\underline{p}_i^T A \, \underline{p}_i = \sum_{k=1}^{N} f_k(i)$$

$$\text{where } f_k(i) = (\underline{p}_o^T \, \underline{u}_k)^2 \, \lambda_k \, \left(\prod_{j=0}^{i-1} (1-\alpha_j \lambda_k)^2 \right) \qquad (2.19b)$$

Since A is positive definite, all its eigenvalues will be positive,

i.e. $\lambda_k > 0$ for all k

This in turn implies that $f_k(i) \geq 0$ for all k.

Hence, from (2.19b), $\underline{p}_i^T A \, \underline{p}_i$ converges to zero if and only if $f_k(i)$ all converges to zero, which in turn is true if and only if the terms $(\prod_{o}^{i-1} (1-\alpha_j \lambda_k)^2)$ all converge to zero. The convergence of the mean square error has thus been shown to be dependent upon two sets of parameters: the step size factors $\alpha_o, \ldots, \alpha_{i-1}$ and the eigenvalues of A $\lambda_1, \ldots, \lambda_N$.

Gitlin et al (17) has also studied the quadratic term $\underline{p}_i^T A \, \underline{p}_i$ but his approach is basically different:

$$\underline{p}_{i+1}^T A \, \underline{p}_{i+1} = \underline{p}_i^T A \, (I-\alpha_i A)^2 \underline{p}_i,$$

using equation (2.10) and noting that A is symmetric.

$$= (1-2\alpha_i A + \alpha_i^2 A^2) \, \underline{p}_i^T A \, \underline{p}_i \qquad (2.20)$$

To study this recursion, let λ_{min} and λ_{max} denote the minimum and maximum eigenvalues of the symmetric, positive definite matrix A.

$$\therefore\ \lambda_{min}\ \underline{p}^T\underline{p} \leq \underline{p}^T A\ \underline{p} \leq \lambda_{max}\ \underline{p}^T\underline{p} \tag{2.21}$$

Applying (2.21) twice to equation (2.20), and remembering that A is of the form $\underline{x}\ \underline{x}^T$

$$\underline{p}_{i+1}{}^T A\ \underline{p}_{i+1} \leq (1-2\alpha_i\lambda_{min} + \alpha_i{}^2\lambda_{max}{}^2)\ \underline{p}_i{}^T A\ \underline{p}_i \tag{2.22}$$

Equation (2.22) indicates that $\underline{p}_{i+1}{}^T A\ \underline{p}_{i+1}$ will approach zero as i becomes large provided

$$\gamma_i \triangleq (1-2\alpha_i\lambda_{min} + \alpha_i{}^2\lambda_{max}{}^2) < 1 \tag{2.23}$$

This can be guaranteed by choosing

$$\alpha_i \leq \frac{2\lambda_{min}}{\lambda_{max}{}^2} \tag{2.24a}$$

(Note that $\dfrac{2}{\lambda_{max}}\ \dfrac{\lambda_{min}}{\lambda_{max}} \leq \dfrac{2}{\lambda_{max}}$ since $\dfrac{\lambda_{min}}{\lambda_{max}} \leq 1$.

Therefore selecting α according to equation (2.24) still satisfies the bounds specified for convergence of the tap gain error).

The rate of convergence will be optimised in the sense that the step size parameter α will be chosen to minimise γ_i for each i. Differentiating (2.24a), the optimum value α_{opt} is obtained to be

$$\alpha_{opt} = \frac{\lambda_{min}}{\lambda_{max}{}^2} \tag{2.24b}$$

(This is less than or equal to the value $\alpha_{opt} = \dfrac{2}{\lambda_{min} + \lambda_{max}}$ as derived by Gersho (9) and given in equation (2.13)).

Equation (2.24b) indicates that α_{opt} is a constant, independent of i, the number of iterations, and is half the maximum permissible step size. With this value of α,

$$\underline{p}_{i+1}{}^T A\ \underline{p}_{i+1} \leq \left(1 - \left(\frac{\lambda_{min}}{\lambda_{max}}\right)^2\right)\ \underline{p}_i{}^T A\ \underline{p}_i \tag{2.25}$$

Hence convergence is exponentially bounded and the rate is dependent upon the ratio of the minimum and maximum eigenvalues of A. If the eigenvalues of A are the same (this would be the case when the inputs to the taps are orthonormal, see Section 2.7 for elaboration of this point), then $\lambda_{min} = \lambda_{max}$, and convergence is achieved in one step. This agrees with the observation of Chang (15) who used this result as a basis for the design of a rapidly converging equaliser structure.

In summary, it is seen from equations (2.14) and (2.25) that the spread of the eigenvalues (i.e. the ratio of λ_{max} to λ_{min}) influences the convergence of both the tap gain error vector $\underline{p}$ and the excess mean square error term $\underline{p}^T A\, \underline{p}$. Ungerboeck (16) has shown that the convergence of the excess mean square error is also dependent upon the number of taps N. This can be expected from an examination of equation (2.16), which shows the direct relationship between the two. Since, for most practical cases, the mean square error tolerable is the dominating factor, the much tighter bound of equation (2.17)

$$0 < \alpha < \frac{2}{tr(A)} \qquad \text{where } tr(A) = \sum_{k=1}^{N} \lambda_k$$

should be imposed on the allowable value of α (rather than $0 < \alpha < \frac{2}{\lambda_{max}}$ which is the bound obtained from consideration of tap gain fluctuations).

2.6 Analysis of Adaptive Operation

The underlying assumption in the previous sections is that the gradient of ε with respect to the tap weights $\underline{c}$ is available, and consequently, the tap adjustments are done in a completely deterministic fashion. This will be true if isolated test pulses are used for the adjustments. But, as it sometimes occurs in practice, the exact gradient of the mean square error is not available and must be estimated from the received data. In this case, a slightly different approach is required, with the results obtained from assuming a training mode providing a convenient reference. The aim is to establish that the taps converge in some probabilistic sense to the optimum settings.

Gitlin et al (17) has undertaken such an analysis and found that

$$\underline{p}_{i+1}^T A \underline{p}_{i+1} \leq (1-2\alpha_i \lambda_{min} + \alpha_i^2 \lambda_{max}^2) \underline{p}_i^T A \underline{p}_i + \alpha_i^2 \sigma^2 \quad (2.26)$$

where σ^2 is the variance of the "background noise". (The use of this term should not be taken to imply that it is due solely to additive noise in the channel). It is interesting to note the similarity of equations (2.20) and (2.26). The random nature of the tracking mode algorithm is clearly indicated by the σ^2 term.

To study the convergence properties of the mean square error, it is required to find an optimum step size that minimises some upper bound on $\underline{q}_i \triangleq \underline{p}_i^T A \underline{p}_i$ at each iteration. Differentiating the RHS of (2.26) with respect to α_i and setting the derivative to zero gives this optimum step size as

$$\alpha_{i\ opt} = \frac{\lambda_{min} \underline{q}_{i\ opt}}{\sigma^2 + \lambda_{max}^2 \underline{q}_{i\ opt}} \quad (2.27a)$$

for i = 1, 2, ...

It follows that $\alpha_{i\ opt} \geq 0$. Rewriting equation (2.27a),

$$\underline{q}_{i\ opt} = \frac{\sigma^2 \alpha_{i\ opt}}{\lambda_{min} - \lambda_{max}^2 \alpha_{i\ opt}}$$

and for this to be consistent with $\underline{q}_{i\ opt} \geq 0$, the term in the denominator must be positive. This will be so if

$$\alpha_{i\ opt} \leq \frac{\lambda_{min}}{\lambda_{max}^2}$$

By substituting equation (2.27a) into the RHS of equation (2.26), an iterative bound on the optimum step size can be obtained. The result is given by

$$\underline{q}_{i+1\ opt} \leq (1 - \lambda_{min} \alpha_{i\ opt}) \underline{q}_{i\ opt} \quad (2.28)$$

Combining (2.27) and (2.28), and simplifying,

$$\alpha_{i+1\ opt} \leq \alpha_{i\ opt} \left(\frac{1 - \lambda_{min} \alpha_{i\ opt}}{1 - \lambda_{max}^2 \alpha_{i\ opt}^2}\right) \quad (2.29)$$

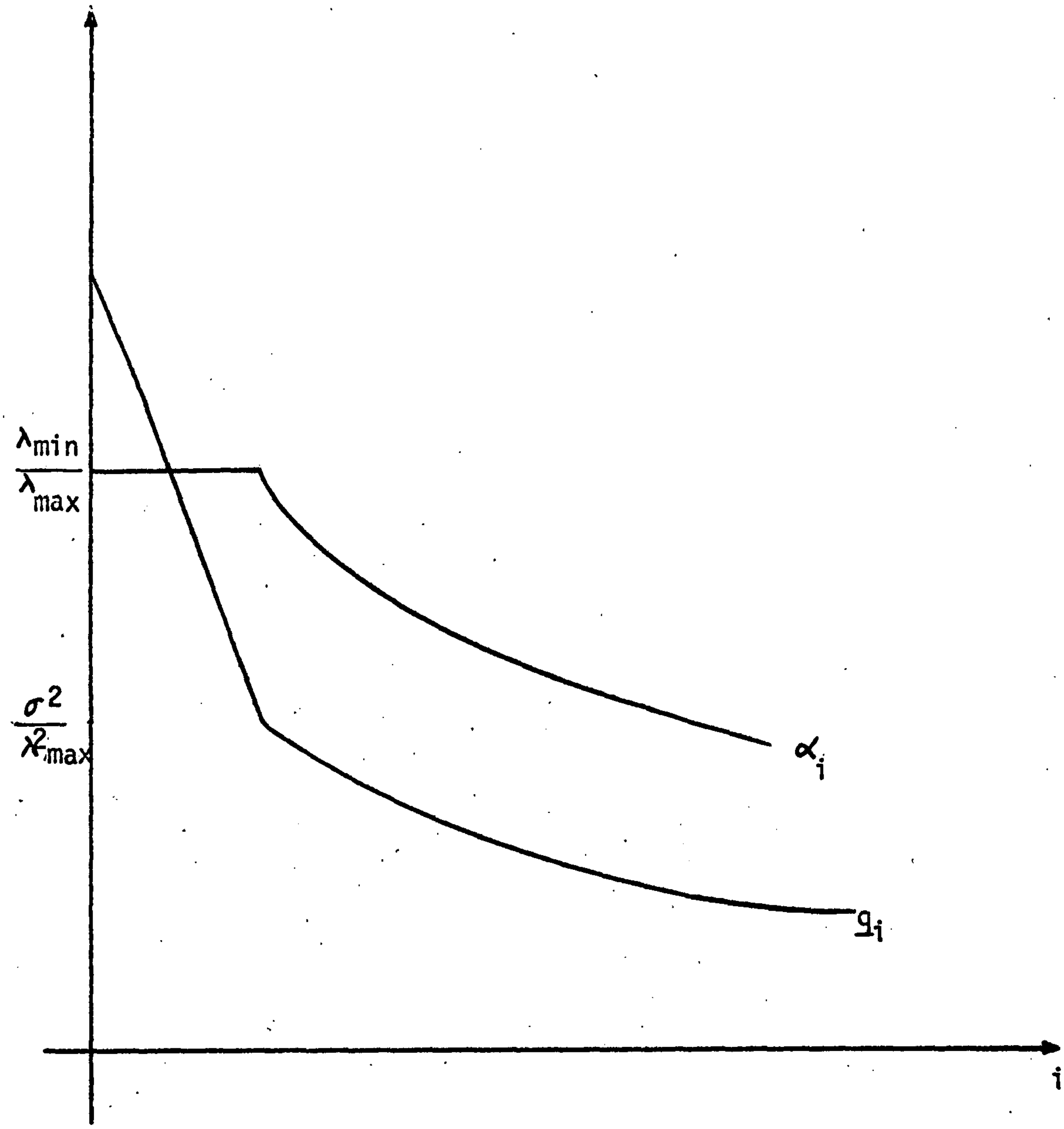


FIG.2.2 VARIATION OF STEP SIZE AND MEAN SQUARE ERROR IN ADAPTIVE OPERATION

which is valid in the range $0 \leq \alpha_{i\ opt} \leq \frac{\lambda_{min}}{\lambda_{max}^2}$ from the results of equation (2.27).

In this range, the bracketed term in (2.29) is less than unity (but greater than zero). This means that the optimum step size will decrease monotonically with i. An exact study of (2.29) appears to be difficult. It is possible, however, to assess the behaviour of $\alpha_{i\ opt}$ for small and large i. For small i, it is reasonable to assume that the initial error will usually be much larger than the "background noise", i.e. $\lambda_{max}^2 \underline{g}_{i\ opt} \gg \sigma^2$. With this assumption, equation (2.27a) simplifies to

$$\alpha_{i\ opt} \simeq \frac{\lambda_{min}}{\lambda_{max}^2} \qquad \text{(compare with equation (2.25))}$$

Combining this with equation (2.28),

$$\underline{g}_{i+1\ opt} \leq \left(1 - \left(\frac{\lambda_{min}}{\lambda_{max}}\right)^2\right) \underline{g}_{i\ opt}.$$

Thus, the mean square error initially delays at an exponential rate while the optimum step size is a constant. This will be the mode of behaviour until the stage when the "large initial error" assumption no longer holds. The dynamic behaviour is then governed by equation (2.29).

For large i, it turns out that $\underline{g}_i$ is proportional to the step size, and both approach zero in a 1/i manner.

The result is summarised in Fig. 2.2. The large initial (constant) step size ensures rapid convergence to a region near the optimum. The algorithm then switches to the smaller (decreasing) step size which provides more precise convergence by reducing fluctuations about the optimum point.

2.7 Other Control Algorithms and Implementations

2.7.1 Alternative Algorithms for Tap Adjustment

The discussion so far has assumed the use of a gradient algorithm of some kind for the adjustment of the tap settings. The most commonly used is the steepest descent method which involves a one-dimensional search for the minimum. It is by no means the only algorithm applicable for iterative adjustment of the taps. De and Davies (18) have, for instance, proposed the use of a conjugate gradient algorithm which converges to the solution in a finite number of steps.

Another algorithm of interest, that employs variable step sizes designed to minimise the error after a specified number of iterations, has been proposed by Schonfeld and Schwartz (19). The optimum set of step sizes (α_i) is obtained by minimising the mean of the tap gain error $\underline{p}_i$.

From equation (2.10), $\underline{p}_i = (I-\alpha_{i-1}A)\ \underline{p}_{i-1}$

It follows that at the Mth step,

$$\underline{p}_M = P_M(A)\ \underline{p}_o$$

where $P_M(A) = \prod_{i=1}^{M} (I-\alpha_{i-1}A)$ which is a polynomial of degree M in A.

It was shown earlier, in equation (2.13), that the best value of α for a fixed step size algorithm is $\alpha = \frac{2}{\lambda_{min}+\lambda_{max}}$. This value is optimum only locally, in the sense that it gives the optimum step size to go from one iteration to the next. In the proposed new approach, the local behaviour of the error norm for the (M-1) intermediate iterations is disregarded. Instead, only the error vector at the Mth iteration is considered.

$$\|\underline{p}_m\| = \|P_M(A)\ \underline{p}_o\| \leq \|P_M(A)\| . \|\underline{p}_o\|$$

where the norm of a vector $\underline{x}$ is defined as its Euclidean length:

$\|x\| = (\sum_k x_k^2)^{\frac{1}{2}}$ and the norm of a symmetric matrix A as

$\|A\| = \max_{|\underline{q}|=1} |A\underline{q}| = \max_k |\lambda_k(A)|$. Following this definition, the norm of $P_M(A)$ is given by its largest (in magnitude) eigenvalue. But since $P_M(A)$ is a polynomial in A, the eigenvalues of $P_M(A)$ are simply the values of the polynomial $P_M(x)$ when evaluated at the eigenvalues of A. Therefore, the problem is reduced to finding the set of step sizes generating the polynomial operator $P_M(x)$ such that

$$\|P_M(A)\| = \max_{\lambda_k(A)} |P_M(\lambda_k)| \qquad \text{is a minimum.}$$

The eigenvalues of A are, of course, not known in advance since the channel characteristic is not known. But it is reasonable to assume that their upper and lower bounds are known, i.e.

$$\lambda_l \leq \lambda_k(A) \leq \lambda_u$$

The optimisation problem can now be re-stated in terms of the upper bound

$$\max_{\lambda_k(A)} \left| P_M(\lambda_k) \right| \leq \max_{\lambda_l \leq x \leq \lambda_u} \left| P_M(x) \right|$$

Thus, instead of minimising the maximum magnitude of the polynomial over the discrete set (λ_k), the procedure rather is to minimise for the worst distribution of the λ_k's over the range (λ_l, λ_u). Such a minimum when found will be independent of any given initial setting. Moreover, it can be shown that $\|P_M(A)\| < 1$, which means that the algorithm will converge for any dispersive channel.

Among the class of polynomials of degree M, the polynomial of least magnitude in the interval (-1,1) is the Chebyshev polynomial. Therefore, selecting $P_M(x)$ to be a Chebyshev polynomial, the optimum step sizes work out to be

$$\alpha_i = 2((\lambda_u + \lambda_l) - (\lambda_u - \lambda_l) \cos (\tfrac{(2i+1)\pi}{2M}))^{-1} \quad i=1, 2, \ldots.M.$$

It is clear that for M=1, $\alpha_0 = \dfrac{2}{\lambda_l + \lambda_u}$ which is the optimum fixed step size derived in equation (2.13).

Comparing the rates of convergence, it appears that the new algorithm gives an improvement and is especially suited for highly dispersive channels. The improvement, though, is guaranteed only at the end of every M iterations and not necessarily for iterations in between. In fact, computer simulations have shown that the output error may even increase for the (M-1) intermediate steps. Because of this, the new algorithm is useful only in a training mode. A second order search algorithm, that is an extension of the first order scheme just described, has also been suggested (20). This has been designed to provide the minimum norm of the tap gain error at each iteration. Convergence is thus always monotonic, and the algorithm may be used in a tracking mode as well.

2.7.2 Alternative Implementations of the Automatic Equaliser

Some of the more interesting alternative algorithms that have been used for iterative adjustment of the taps have been noted. While they appear attractive from the point of improved convergence rate and smaller number of iterative steps needed, they also require more complex hardware implementation. The extent to which the degree of equalisation required for a particular communication channel justifies the cost of using more complex hardware has yet to be established.

In addition to investigations on algorithms other than steepest descent that might conceivably be used to improve equaliser performance, efforts have also been made to modify the equaliser structure whilst retaining the use of the steepest descent algorithm. One such work has been carried out by Chang (15).

From equation (2.19b), it is seen that the excess mean square error is made up of the sum to N terms of the error component $f_k(i)$. Prior to the (i+1)th iterative step, the kth error component is

$$f_k(i) = (\underline{p}_o{}^T\underline{u}_k)^2 \lambda_k \left(\prod_{j=o}^{i-1} (1-\alpha_j\lambda_k)^2\right)$$

After the (i+1)th adjustment, this becomes

$$f_k(i+1) = (\underline{p}_o{}^T\underline{u}_k)^2 \lambda_k \left(\prod_{j=o}^{i} (1-\alpha_j\lambda_k)^2\right)$$

This implies that $\quad f_k(i+1) = (1-\alpha_i\lambda_k)^2 f_k(i)$

The (i+1)th adjustment will reduce each of the error components by a large factor if $\quad (1-\alpha_i\lambda_k)^2 \ll 1 \quad$ for all k

i.e. if $\quad \alpha_i \cong \frac{1}{\lambda_k} \quad$ for all k

If the λ_k's are very different, α_i cannot satisfy this simultaneously, and hence the (i+1)th adjustment cannot reduce each of the error components by a large factor. This result suggests that reducing the difference between the eigenvalues of A, the signal + noise correlation matrix, should bring about an improvement in convergence rate. But the difference in the eigenvalues is due, in the first instance, to correlation between the various tap inputs. Therefore, if the inputs are made orthonormal, A becomes an identity matrix and all the eigenvalues are equal. It would then be possible, at least theoretically, to minimise the mean square error in a single adjustment.

The input vectors (x_i) can be orthogonalised by the Gram-Schmidt technique to form the orthogonal set (w_i) where each (w_i) is a linear combination of the (x_i).

i.e. $\underline{w} = \underline{x}\, V$ where V is an upper triangular square matrix.

For systems where the amplitude distortions in the channel are not severe, the new equaliser can be implemented by simply adding a preset weighting matrix to the conventional transversal equaliser. With this modification, the convergence rate can be improved without changing the value of the minimum mean square error attainable and the convexity of the adjustments, and without complicating the gain-control adjustment loop. Quan and Gibbs (21) have suggested a similar approach but using the orthogonalisation procedure in conjunction with a Seidel iteration.

So far, the schemes considered have all been proposed in time domain terms. Walzman and Schwartz (22) have, on the other hand, proposed an automatic equaliser that utilises Rosen's gradient projection method to optimise parameters in the discrete frequency domain. The algorithm is shown to converge in the mean for any channel, even in the presence of noise. Convergence of the tap gain error is according to the inequality

$$\|\underline{p}_i\| \leq \rho^i \|\underline{p}_0\|$$

where ρ = spectral radius of $(I-\alpha A)$. This is similar to the result obtained in Section 2.3, which shows that both time and

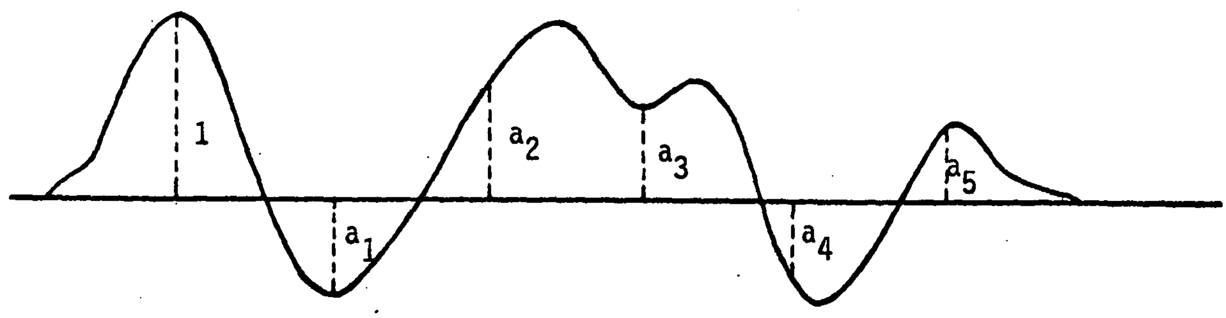


FIG.2.3(a) IMPULSE RESPONSE OF A NON-IDEAL CHANNEL

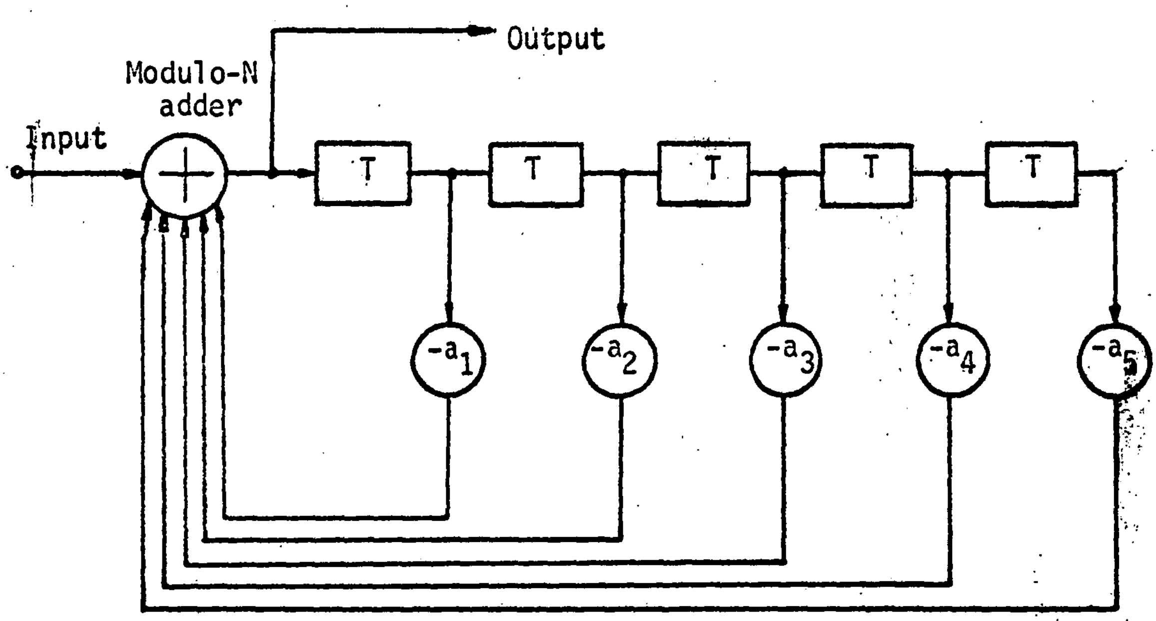


FIG.2.3(b) FEEDBACK TRANSVERSAL FILTER

frequency domain equalisation exhibits geometric convergence. There is, however, a very important difference. For the time domain case, the eigenvalues of A are difficult to compute, especially where the order of A is large. In fact, determination of the eigenvalues may often involve more computation than iterative optimisation of the parameters. The procedure usually adopted is to find bounds on the eigenvalues and use these as estimates. Because the eigenvalues are not known exactly, the value of α obtained, as in equation (2.13), is suboptimal.

On the other hand, for the frequency domain case, A turns out to be a diagonal matrix (this follows from the properties of the Discrete Fourier Transform). This means that the eigenvalues are found exactly by inspection. The step size parameter α can therefore always be optimally chosen, which accounts for the more rapid convergence of frequency domain equalisers over their time domain counterparts.

2.8 Feedback and Non-linear Transversal-Type Equaliser

The conventional transversal equaliser consists of a delay line with feed-forward taps. In principle, an infinite number of such taps is required for exact equalisation of any general communication channel. This can be seen by considering the sampled impulse response A(z) of a non-ideal channel as shown in Fig. 2.3a. Using the z-transform,

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots. + a_5 z^{-5}$$

If a transversal filter of sampled impulse response P(z) is connected in series with the channel, the output at the sample times is given by the product of the two polynomials, i.e. P(z)A(z). For exact equalisation at the sample times, it is required that P(z)A(z) = 1, or $P(z) = \frac{1}{A(z)}$. Hence, an infinite number of taps, with settings equal to the coefficients obtained from expanding $\frac{1}{A(z)}$ in negative powers of z, are needed.

Tomlinson (23) has suggested the use of feedback

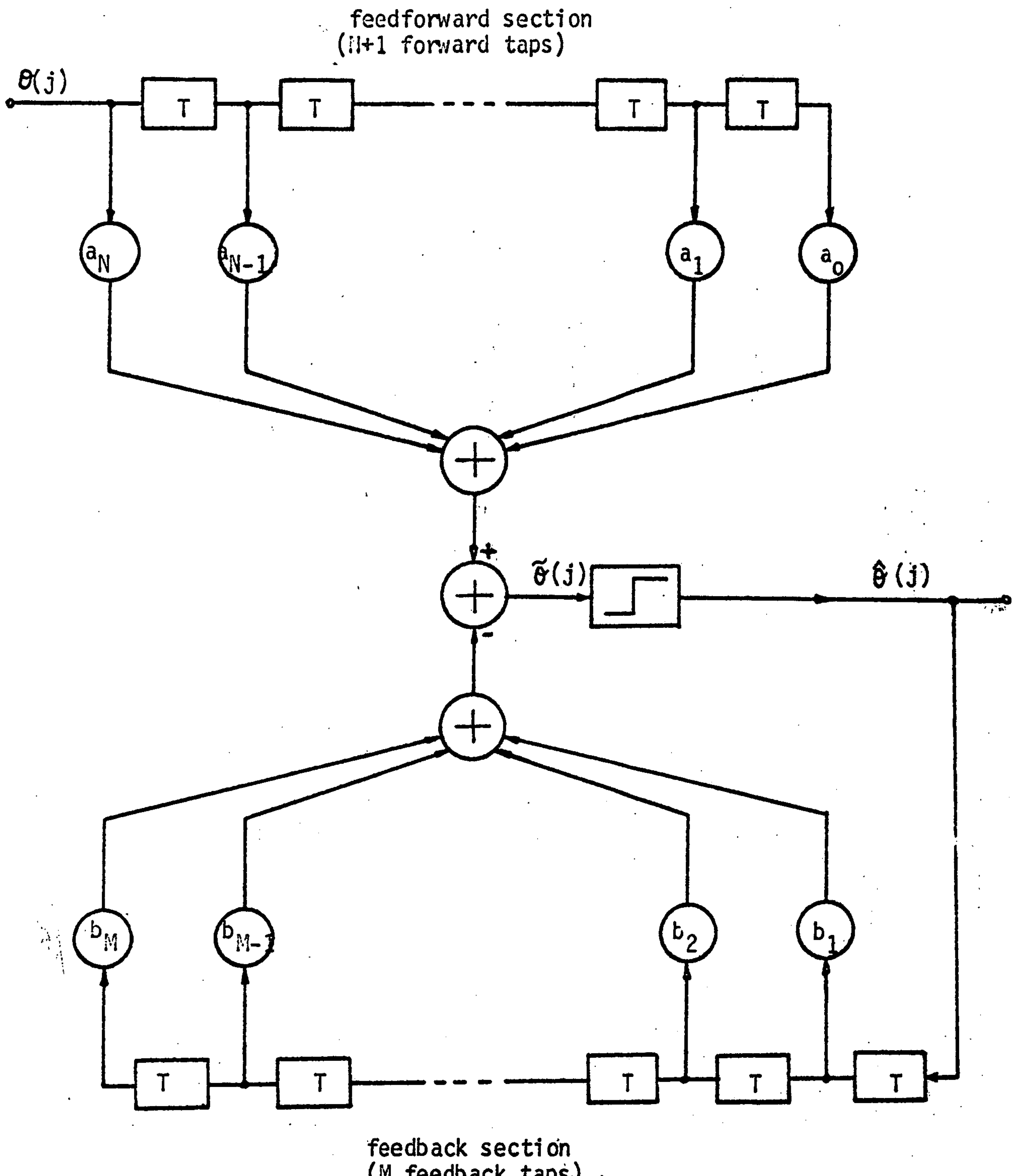


FIG.2.4 NON-LINEAR TRANSVERSAL TYPE EQUALISER

connections to overcome this drawback. The ideal response $P(z) = \frac{1}{A(z)}$ can be achieved by using a finite length transversal filter with feedback connections as shown in Fig. 2.3b. The tap settings are $-a_1$, $-a_2$, ..., $-a_5$. Such an arrangement as it stands, is liable to be unstable. By using a modulo-N adder in place of the ordinary summer, the output can be maintained between $\pm \frac{N}{2}$. The filter will then always be stable. In this way, the modulo inverse filter can be used for equalisation simply by making its successive feedback coefficients equal in amplitude to the successive samples of the impulse response of the channel.

A logical extension is the type which consists of a combination of feed-forward and feedback transversal filters. For digital transmission, a non-linear element is needed in the feedback path to convert the output from the feed-forward section to a final decision. Thus, past decisions are utilised in making a decision on the present signal element. The general arrangement is as shown in Fig. 2.4 (Note that for analogue transmission, the limiter at the output is not needed).

George, et al (24) have described the use of such a decision feedback equaliser to recover a sequence of digits transmitted at a high rate over a noisy, dispersive linear channel. At the output of the conventional equaliser, there is error caused by both undetected digits and previously detected digits. If the previous decisions are correct, they can be used to subtract coherently the error caused by the previously detected digits. This is achieved by passing the past decisions through the feedback tapped delay line. The remaining error due to undetected digits and the effect of additive noise is minimised in the usual manner by the feed-forward delay line. For such an equaliser, any errors at the output will tend to occur in bursts, since a decision error in the feedback section will propagate more incorrect decisions. Simulation studies have since indicated the ability of the equaliser to recover spontaneously from this condition. Moreover, the improved capability over that of the conventional linear type far outweighs this setback.

The equaliser makes an estimate $\tilde{\theta}(j)$ of $\theta(j)$, the digit that is sent at time $t = jT$, and then converts this estimate to a final decision $\hat{\theta}(j)$ with a non-linear circuit. The error is defined to be $e(j) \triangleq \tilde{\theta}(j) - \theta(j)$, and the aim is to minimise the mean square value of the error. Instead of calculating and manually adjusting the tap gains after determining the channel characteristics, the equaliser can be made to work adaptively. This is achieved by using the past decisions for adaptation as well, again assuming that these are correct.

For the decision feedback equaliser, the mean square error remains a quadratic function of the tap gain errors. There are, therefore, no "locally optimum" settings, and a hill-climbing adaptation can readily be made to converge close to the optimum. This suggests that the forward taps can be adjusted by using the cross-correlation between the error and the input signal, just as for the conventional equaliser. The feedback taps, however, are updated by using the cross-correlation between the error and the decisions $\hat{\theta}(j)$.

Alternatively, both sets of taps may be adjusted by using the one set of measurements - the cross-correlation between the error and the decisions. In this case, the taps no longer converge to the correct value which minimises the mean square error. For certain channels, though, this second algorithm has been found to give a better performance because of the smaller error in the cross-correlation measurement.

In the last few sections, alternative algorithms for controlling the tap settings as well as alternative implementations of the equaliser structure have been considered. While these proposals may result in improved performance, they also involve more complex hardware. It remains to be established if the degree of equalisation required for the TV telephone channel warrants the extra expenditure. For the scope of this work, therefore, the conventional equaliser structure will be used in conjunction with the steepest descent algorithm.

Since noise will be present in all practical systems, the next chapter will go on to consider, in greater detail, the effect of noise on equaliser performance. Also, the effect of introducing a non-linearity into the error loop, thereby creating non-linear feedback, will be discussed.

3. Effects of Noise and Non-Linear Feedback

3.1 Noise in the Channel

For a TV telephone network using existing telephone cables, the noise is introduced into the baseband loops primarily through crosstalk coupling from other transmission paths in the same cable. The extent of coupling depends on both the proximity of the pairs in question and on the difference in their transmission levels.

The noise may be broadly classified under two categories (25):

(i) random noise interference, which consists of the sum of thermal noise and all those other interferences appearing in the signal which are at too low a level to be separately identifiable. The amplitude distribution of this type of noise is nominally gaussian within the dynamic range of the channel. The spectrum may vary widely, depending on the characteristics of the transmission systems through which the signal has passed.

(ii) switching or impulse noise. In spite of its high amplitude, it appears to contribute little to the total random noise power because of its short duration. It may, however, disturb the equilibrium condition of the equaliser. Impulse noise is characterised by sampling the noise on the idle channel at a fixed rate, and counting samples of amplitude exceeding a given threshold. The result is expressed as the estimated probability of the noise exceeding the threshold. For the TV telephone channel being investigated, no such measure of actual statistics is yet available.

3.2 Effect of Noise on Equaliser Performance

The analysis given in chapter two has assumed an additive, white gaussian noise, with zero mean, which is uncorrelated with the received signal. Under these conditions, the noise has been seen to affect the equaliser's performance in two ways:

(i) it introduces a direct error by its presence in the final output.

(ii) it corrupts the gradient estimates used for the tap adjustments, resulting in an error in the gain settings. Consequently, the system converges to some neighbourhood of a sub-optimal setting, about which random fluctuation occurs. The amount of noise that is tolerable before the error between the optimal and sub-optimal settings becomes objectionable needs to be investigated.

Each of these effects will be considered in turn.

Firstly, it is possible to reduce the direct error by proper design of the equaliser. Provided the signalling statistics are known, a priori, the equaliser can be designed to minimise the total error consisting of both the component resulting from channel distortion and the component resulting from noise. Lucky and Rudin (6) have suggested that the mean-square equaliser will achieve this if the test signal used for the equalisation has a spectral density identical to that of the signal to be transmitted over the equalised channel. Thus, if the power spectrum of the information signal is known beforehand, it is simply a matter of choosing the proper spectral weighting function for the test signal. For the TV telephone system, it is specified that the channel response should have a high frequency roll-off to suppress the effect of the energy components above the 1MHz upper limit (see Section 4.7). In view of this, the spectral weighting function to be used need not be flat up to 1MHz, but can have a lower cut-off frequency. In order to observe the effect on equaliser performance of varying the spectral weighting, a weighting function with variable cut-off is used in this work.

Secondly, regarding the use of noisy gradient estimates, Gersho (9) has found that the error between the optimal and sub-optimal settings is small for low noise levels. Even where the noise level is fairly substantial, the sub-optimal setting may still be acceptable, provided the error surface given by $\mathcal{E}(\underline{c})$ is "shallow" in a large neighbourhood of the minimum. In this case, a fairly large departure of $\underline{c}_{\text{sub-opt}}$ from $\underline{c}_{\text{opt}}$ may still correspond to a small increase in mean-square error only. Alternatively, this effect of noise can be minimised by increasing the signal-to-noise ratio. Chang (15) has observed that with a signal-to-noise ratio of 30 dB, the use of noisy estimates for the tap adjustments has only a minor influence on the accuracy of the equaliser. To decide if this is also the case for the TV telephone channel, it is necessary to consider the expected signal-to-noise ratio of the video signals that will be transmitted over the channel.

3.3 Expected Signal-to-Noise Ratio for Video Transmission

No generally accepted standard has yet been laid down concerning the minimum signal-to-noise ratio required for satisfactory video transmission in TV telephone operation. Any such specification will have to take into account not only the optical conditions governing picture reproduction but also the characteristics of the human eye (26). Given the contrast range of about 1:50 of conventional picture tubes, the human eye is still just able, under favourable conditions, to perceive variations in brightness due to a noise level which is about 52 dB below the maximum video signal level. The subjective impression of the interference, however, depends not only on its level, but also on its frequency of occurrence. Thus, in general, the eye is more sensitive to low frequency interference than to high frequency ones. This is because the latter tends to produce patterns with very narrow bright-dark zones that come close to the limit of resolution of the displayed picture. Moreover, the prescribed high frequency roll-off serves also to suppress the noise at this end of the spectrum. These factors together mean that a lower signal-to-noise ratio will be adequate in the upper frequency region.

In the light of these considerations, Ebel (27) has suggested a value of 50dB for the range of lower video frequencies (up to about 300 kHz), and a lower value thereafter. Depending on the extent to which the extra expenditure on noise

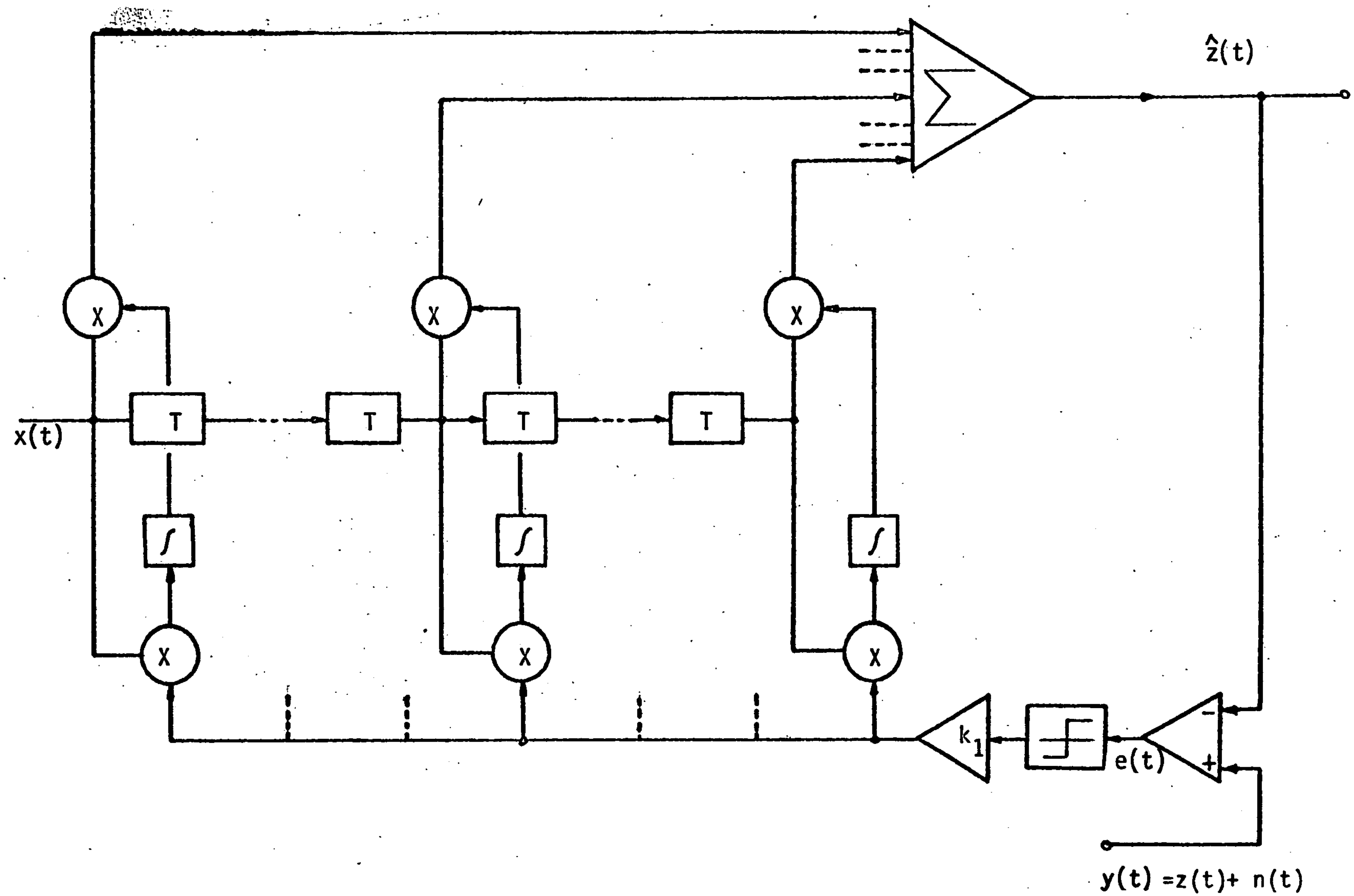


FIG.3.1 ADAPTIVE CONTROL LOOP WITH CLIPPER

reduction is justifiable, a value of about 17 to 35 dB at the upper cut-off frequency may be sufficient. But in view of the need to minimise the effect of noise corruption of the equaliser settings, as indicated in the previous section, it is preferable to select the higher value.

3.4 Effect of Non-Linearity in Feedback Loop

As mentioned in Chapter One, a problem that parallels that of adaptive equalisation of transmission channels is the adaptive cancellation of echoes in long-distance telephone circuits. A paper by Sondhi (2) dealing with this problem has suggested the introduction of a non-linearity in the feedback loop as a means of improving system performance. This modification will be considered in this section and will be extended to include any possible usefulness for equalisation purposes.

3.4.1 Application in Adaptive Echo Cancellation

In Fig.3.1, $x(t)$ denotes the speech input from the far-end, and $y(t)$ denotes the return signal from the near-end echo path and comprises the echo $z(t)$ and the noise $n(t)$. The principle of operation relies on the synthesis of a replica $\hat{z}(t)$ of the echo by the N-tap transversal filter. The control loop uses the error, given by

$$e(t) = z(t) - \hat{z}(t) + n(t)$$

to improve the estimate $\hat{z}(t)$ continuously and thus cancel out the unwanted echo. The adaptive loop can be described by a set of simultaneous non-linear differential equations of the first order:-

$$\frac{d}{dt}\underline{c} = K_1 F(e(t))\,\underline{x} \qquad (3.1)$$

where $\underline{c}$ = matrix of tap gain co-efficients

K = positive constant of proportionality

In equation (3.1), the same procedure has been adopted of writing the continuous time functions in vector-matrix form as in Chapter Two. The non-linearity F is specified to be any monotonic, non-decreasing odd function. Because of the introduction of F, the system is no longer mean-square in the sense understood so far.

With this set-up, it is found that the uncancelled echo e(t) converges monotonically to zero under ideal, noiseless conditions. In practice, noise will always be present, but convergence can still be achieved as long as the magnitude of e(t) is large compared to that of the noise n(t). If this is not the case, the system can only be considered to converge in an average sense. Even then, it is only true for the quasi-stationary case where the feedback constant K_1 is small (and hence the tap gains are only slowly-varying). The final state is one in which the tap co-efficients fluctuate about their optimum setting.

Even if convergence is assumed under the conditions stipulated above, it remains an extremely difficult problem to estimate the convergence rate. This depends upon three factors:-

(i) the statistics and level of the incoming signal and noise. Generally, increasing the signal level serves to speed up convergence.

(ii) the choice of the non-linearity F. This will have a profound influence on the dynamic behaviour of the system. One particular non-linearity which has been suggested is the infinite clipper with

$$F(e) = \mathrm{sgn}(e) = \begin{cases} +1 & \text{for } e>o \\ -1 & \text{for } e<o \end{cases}$$

(iii) the value of the positive constant K_1. For rapid convergence, K_1 should be chosen to be as large as possible. Unfortunately, this conflicts with the requirement that K_1 be made small for maximum immunity to noise. There appears to be no theory at present for calculating the optimum value, and further investigation is needed. Because of the dependence of the convergence rate on the signal level as well (factor (i)), it is clear that the value of K_1 can be adjusted to give a certain speed only for some average level of $x(t)$.

A completely general solution to the problem of determining the convergence rate has yet to be found. Sondhi (2) has obtained an estimate of the mean rate for the noiseless case by specifying a low value for K_1, and taking the input to be a stationary, gaussian process with zero mean.

In a similar analysis, Rosenberger and Thomas (4) have also had to impose certain restrictive assumptions on the input signal and noise in order to obtain a particular solution. The input is assumed to be gaussian, with zero mean and bandlimited, with a rectangular power density spectrum. Likewise, the noise is assumed to be bandlimited to the same frequency, and statistically independent of the input. It is further specified that the echo path should contain no signficant non-linearities and be time-invariant.

Allowing for these imposed assumptions, observations using the infinite clipper have indicated a lesser dependence of the time constant of the adaptive loop on the signal level. Thus, without the clipper, the time constant of convergence is proportional to the signal power and a 20 dB change in signal level changes the time constant by a factor of 100. On the other hand, with the clipper introduced, the same change in signal level changes the time constant by only a factor of 10. Moreover, in the event of a sudden burst of noise occurring when the system is near equilibrium (i.e. when $e(t)$ has been reduced to a very low value), the rate at which the system moves away becomes more or less independent of the noise level. It would appear that the introduction of the clipper has brought about an improvement in the overall system capability, providing the effects of noise are tolerable.

Both analyses confirm that there will always be an uncancellable residual, even for the non-linear system. The residual arises from the presence of noise and the assumption that the echo to be cancelled is exactly representable by the

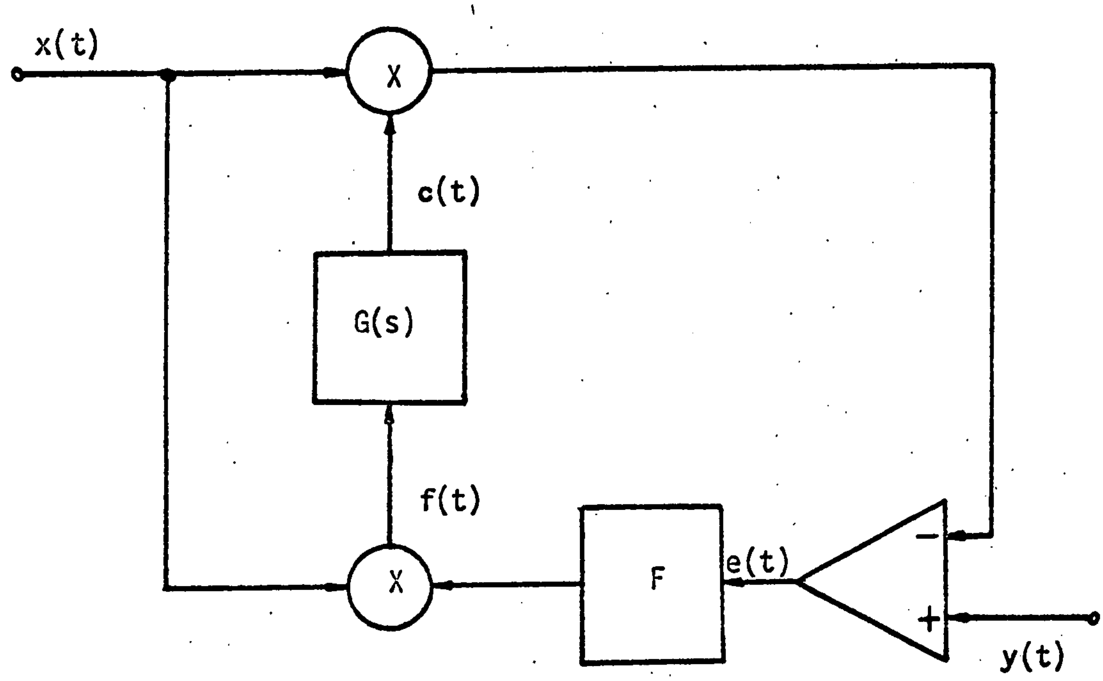


FIG.3.2 BLOCK SCHEMATIC OF SIMPLE MODEL

output of a transversal filter with a finite number of taps, which have constant (or slowly varying) gains. Rosenberger and Thomas (4) have defined an incompleteness or truncation factor I as a measure of this residual. I is a function of both the environment, i.e. the echo path transfer function over the bandwidth of the input signal, and the canceller configuration, i.e. the filter set used and the number of taps. The maximum achievable cancellation is given by $S = -10 \log_e I$, and represents a theoretical upper limit of system capability.

3.4.2 Application in Adaptive Equalisation

The apparent ability of the clipper to improve the performance of the echo-canceller has prompted investigation of its usefulness for the equalisation process. Furmage (28) has made some preliminary studies, using a time-scaled model of the television channel as the system to be equalised. Using a delay line with ten taps, and operating under a mean-square error criterion, it was found that improvement was obtained in both the static (residual error) and dynamic (convergence rate) performance.

(i) Static behaviour:

Initially, it was found that for received signals with low amplitudes, the equaliser would not converge. The output integrators merely drift off until saturation of the amplifiers occurs. This malfunction was subsequently traced to the effect of the d.c. offset error of the correlators. It can also be theoretically predicted by analysing the simple loop of Fig.3.2. If the loop gain is small, the a.c. component of the correlator output can be neglected. The d.c. component is then given by

$$\begin{aligned} f_{dc}(t) &= x(t)\; e(t) \\ &= x(t)\; y(t) - c(t)\; x^2(t) \end{aligned}$$

$$\text{since } e(t) = y(t) - c(t)\; x(t)$$

It is clear that in the presence of any d.c. offset introduced by the correlator, there will be an error in the value of the tap setting c(t). The error can be reduced by better zeroing of the associated multiplier, and also by raising the levels of the received and reference signals. The improvement achievable in the latter case is limited, especially if the other parts of the system are not to be saturated first. The limitations can be overcome by amplifying the error signal instead. The clipper immediately following the difference amplifier provides this desired amplification. It was found thereafter that the equaliser converges even for low level received signals.

(ii) Dynamic behaviour:
As in the observations with the echo canceller, it was found that the convergence rate was altered from a quadratic to a linear dependence on the amplitude of the received signal. The clipper thus reduces the equaliser's sensitivity to input level variations. Moreover, there was an improvement in the overall settling time. Furmage quoted a reduction by a factor of ten.

It would appear then that the use of the clipper introduces possible advantages in adaptive equalisation of communication channels. In addition to giving an improved overall system capability, the non-linear system can be implemented with simpler circuitry than that considered so far. Thus, the four-quadrant precision multipliers used for correlating the error signal with the received signal may be replaced by simple transistor switches.

On the other hand, the use of the clipper has the disadvantage of causing increased variations of the correlator outputs about their mean values. It remains to be established if, under these circumstances, the effects of noise will be tolerable.

3.5 Analysis of Non-Linear Equaliser Loop

As a first step in establishing the validity of the results outlined in the previous section, some form of theoretical analysis will be attempted. The analyses published so far have all involved restrictive assumptions owing to the non-linear nature of the system. These assumptions have been found necessary in order to obtain a reasonably uncomplicated solution. Similarly, in this case, a more general analysis cannot easily be attempted owing to:

(i) the shortage of information on the signal and noise spectra and probability distributions

(ii) the presence of the multipliers and the clipper in the equaliser loop

Therefore, it is the intention here to resort to linearised analysis, with reference particularly to small-signal operating conditions. Though it is acknowledged that this approach may be limited in applicability, nevertheless, it may afford some understanding of the behaviour of the system.

It is proposed to consider first the simplified system shown in Fig.3.2. To linearise the system, it is necessary to find a way of representing the clipper in the error path. This can be conveniently achieved by applying the describing function concept (29). A sinusoidal input is assumed, with the describing function defined as the complex ratio of the amplitude of the fundamental component of the output of the non-linearity to that of its input. The validity of the describing function approximation rests implicitly on the assumption that the linear portion of the system acts as a low pass filter, i.e. all terms except the fundamental can be neglected. It will be demonstrated, in the course of the analysis, that this condition is satisfied for the loop under investigation.

The non-linearity in this case, i.e. the clipper, has a characteristic with odd symmetry and hence it introduces no phase shift. The describing function is therefore a real function given by $F = \frac{4K_1}{\pi E}$, where K_1 = loop gain

$e(t) = E \sin \omega t$

∴ From Fig.3.2., assuming that the output of the difference amplifier is sinusoidal, Output of correlation multiplier, $f(t) = x(t)\,F\,e(t)$

$$= x(t)\,\frac{4K_1}{\pi}\sin\omega t$$

If $x(t) = X\sin\omega t$, then

$$f(t) = \frac{4K_1X}{\pi}\sin^2\omega t$$

$$= \frac{2K_1X}{\pi}(1 - \cos 2\omega t)$$

$$\therefore \quad \text{Tap setting } c(t) = \int_0^T f(t)\,dt$$

$$= \frac{2K_1X}{\pi}\left(t - \frac{1}{2\omega}\sin 2\omega t\right)_0^T \qquad (3.2)$$

From equation (3.2), it is clear that the high frequency term has been attenuated in passing through the integrator. It follows that if x(t) should contain frequency components other than the fundamental, these will be similarly attenuated by the integrator. This low pass filtering action satisfies the assumption upon which valid use of the describing function is based.

Equation (3.2) also indicates that, with this non-linearity, the tap setting is dependent upon the loop gain and the level of the received signal. This agrees with the observations of Sondhi and Furmage as outlined in Sections 3.4.1 and 3.4.2 respectively.

By applying the describing function method on the simplified equaliser system, initial results have been obtained which agree with that of previous investigations. Nevertheless, it is difficult to really establish the mathematical legitimacy of the above approach. An attempt has been made by Bass (30) to provide a firm mathematical foundation for the use of the describing function concept. He worked out several necessary and sufficient conditions for the existence, in a given dynamical system, of periodic motion described as "π symmetric". As the name suggests, these are undamped vibrations which possess odd symmetry relative to an angular translation of π radians. The use of the describing function tacitly assumes that only vibrations of this type exist in the system. Although his results are quite complete and rigorous, they are very difficult to apply.

Moreover, the solution obtained with the describing function method is only a "first approximation" to the answer. Extensions to the basic method exist which will correct for the presence of unattenuated higher order components. No such further study has been attempted however owing to the limitation of time.

3.6 Stability Considerations

Another area that requires investigation is the stability of the equaliser system. The study of the stability of a non-linear system is often not an easy exercise. Linear techniques such as the Routh-Hurwitz criterion, the root locus method, or Bode's phase and gain margin would fail to give proper results except under those conditions in which the non-linear system can be approximated by a linear one. Such conditions exist only for small perturbations about a stable operating point. If a new operating point is chosen, the results, in general, would no longer be valid.

A more general criterion based on considerations of global stability is required. One such is Lyapunov's second method (32). Lyapunov connected the stability of a system with the presence of a V-function, the time derivative of which satisfies certain conditions. Thus, a system is asymptotically stable if a positive-definite function V, with time derivatives which are definite and of opposite sign from V, exists. Narendra and McBride (33) have successfully applied this method in analysing the stability of the equaliser system in the absence of the clipper. They used the equation

$$\frac{\partial}{\partial t} c_i(t) = K_1 x_i(t)\, e(t)$$

$$= -K_1 x_i(t) \sum_k c_k(t)\, x_k(t) + K_1 x_i(t)\, y(t)$$

Or, adopting the matrix notation of Chapter Two,

$$\dot{\underline{c}} = -K_1\underline{x}\,\underline{x}^T\underline{c} + K_1 y(t)\underline{x} \qquad (3.3)$$

By considering the non-negative Lyapunov function

$$V = \underline{c}^T\underline{c} > 0 \qquad \text{for all } \underline{c} \neq 0,$$

it can be shown that the homogeneous part of equation (3.3) is asymptotically stable for positive K_1. It follows that $\underline{c}$ cannot increase without limit if both $y(t)$ and $\underline{x}$ are bounded. Thus, the equaliser system is stable even if a relatively large value of K_1 is used to ensure rapid convergence.

With the introduction of the infinite clipper into the equaliser loop, the problem becomes more difficult. In general, methods have been derived which will give the Lyapunov function for a linear system in a rather straightforward manner. But much effort may have to be expended before a Lyapunov function for a non-linear system can be found. In this particular case, a partial solution would be to consider Booton's representation of the equivalent clipper gain (31), given by

$$K_B = \frac{1}{\sigma_e\sqrt{2\pi}}\int_{-\infty}^{\infty} \bar{e}\,\bar{F}(\bar{e})\,\exp\left(-\frac{\bar{e}^2}{2}\right) d\bar{e} \qquad (3.4)$$

where $\bar{F}(\bar{e})$ = normalised form of $F(e)$

σ_e^2 = variance of input to clipper.

For the equaliser system being considered, equation (3.4) gives

$$K_B = \pm\frac{K_1}{\sigma_e}\sqrt{\frac{2}{\pi}}$$

Since σ_e changes as the system converges, it follows that the effective gain of the clipper will also change. Therefore, it seems that using Booton's method, the clipper may be taken to be just a variable gain element which introduces no phase shift. By treating this variable gain K_B as part of the overall loop gain, the resulting system can be shown to be stable, using the approach of Narendra and McBride as before.

As has already been pointed out, the non-linear nature of the system, in particular the presence together of both the multipliers and the clipper, renders any attempt at producing a general theoretical analysis a difficult task. More elaborate methods than those that have been considered thus far, such as perturbation methods or extensions to Lyapunov's second method, may have to be used. But although such concentrated studies would be interesting from a theoretical point of view, they are not likely to yield much more insight into the operation of the system. This is especially true with the present shortage of information on the signal and noise spectra and probability distributions. In view of these limitations and also the shortage of time, no further theoretical study was pursued. Instead, attention was focussed on testing the system experimentally. The next chapter thus reports on the design of the experimental equaliser system.

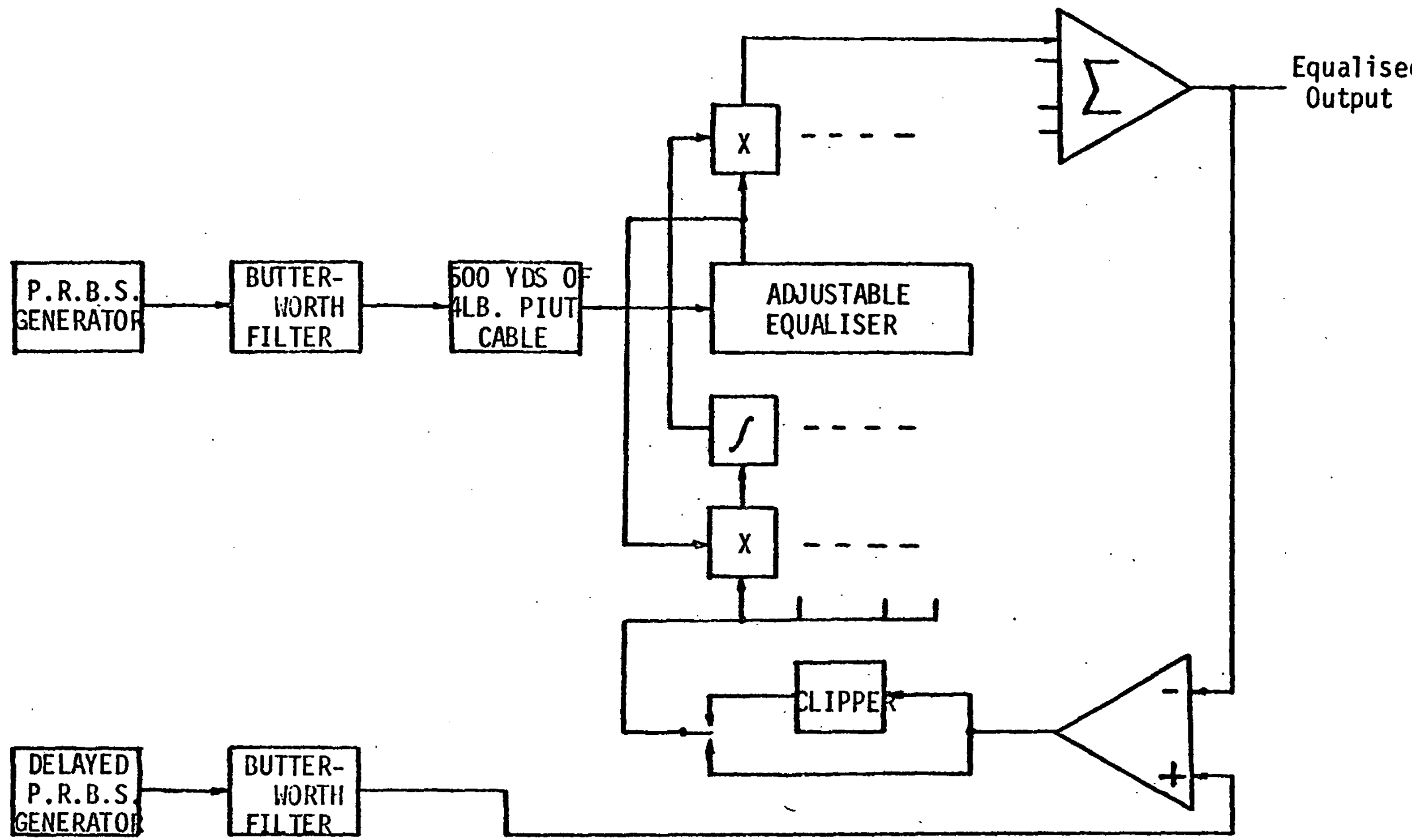


FIG.4.1 BLOCK DIAGRAM OF EXPERIMENTAL EQUALISER

4. Design of the Experimental System

4.1 General

In order to confirm some of the theoretical findings of the previous chapters, and to demonstrate the potential of automatic equalisation in practice, a real time experimental system was designed and constructed. The block diagram, as shown in Fig.4.1, is similar to that used in a previous investigation (28) on the automatic equalisation of television channels (but time-scaled with a bandwidth of 5kHz). The essential difference is that the TV telephone channel is the subject of interest instead, and all circuits have to be realised for operation in real time corresponding to a nominal bandwidth of 1 MHz. Hence, the circuits used previously were either modified or redesigned to meet the extended bandwidth requirements. This chapter consists of an outline of the work involved and some of the problems encountered. The actual hardware details can be found in the Appendix A.

4.2 Basic Implementation

Equalisation is achieved in a training mode. During this interval, a test signal is passed through the channel and operated on by the equaliser. The output is then compared with a properly synchronised but undistorted version of the same signal. The correlators then calculate the various cross-correlation coefficients (between the error and the output of each tap) for all the delay line taps. The polarity of the coefficient gives the direction of change required for optimisation, while the magnitude indicates the corresponding rate of change. The attenuators, which are capable of both positive and negative weights, then automatically adjust the settings to achieve mean square error minimisation.

In the selection of an appropriate delay line, three parameters must be established: The bandwidth, tap spacing, and number of taps (6). In the present scheme, the entire bandwidth of the channel is to be equalised. Hence, the usable bandwidth of the delay line must be at least equal to the channel's bandwidth. This in turn means that the spacing between the taps must be less than or equal to the Nyquist interval i.e. the reciprocal of twice the bandwidth. The third parameter, i.e. the number of taps needed, depends on the nature and degree of the distortion likely to be found in the channel and on the precision of equalisation required. In theory, an infinite number of taps is required for exact equalisation, as seen from Section 2.8. As this is not practicable, a truncation error will be incurred by the use of a finite number of taps.

That the number of taps is an important parameter in the selection of an appropriate delay line has been confirmed by some of the theoretical findings of Chapter Two. For instance, from equation (2.11), it is seen that the mean square error is given by the sum of two terms: an irreducible term arising from noisy measurements and finite truncation, and a quadratic term giving the excess mean square error. The latter term has been found to be directly dependent on the number of taps used (see equation (2.15)). Each additional tap increases through its tap gain fluctuations the expected excess mean square error. This accumulation of errors in the tap coefficients may often more than offset the added correcting ability of the longer delay lines. There is as yet no method of determining the optimum number of taps to be used for a particular application.

While considering the number of taps to be used, it is perhaps worth noting the selection of the position for the main or reference tap. This is the tap that is used as the timing reference in synchronising the remote and local p.r.b.s. generators so that near optimum use is made of the transversal equaliser (see Chapter Six). The common practice is to arbitrarily designate the centre-tap to be the reference tap. However, the impulse response of the channel to be equalised is often such that another position is more desirable. In these cases, the final mean square error obtained will not correspond to the true minimum for a transversal equaliser with the given total number of taps. For instance, it has been found empirically (6) that most disturbing echoes lag the undistorted impulse. It would appear then that a lower residual distortion may be obtained by shifting the reference tap to a position about two-thirds down the line.

The exact optimum position for a given channel can be found by trying out all possible positions and choosing the one which yields the minimum mean square error. Clearly, this is not the most satisfactory procedure. Newhall et al (34) have described an algorithm which provides a partial solution to this problem. It involves finding the approximate inverse of the system to be equalised. If $H(z)$ is taken to represent the sequence of equispaced samples of the impulse response of the given system, then the algorithm calculates the Laurent's series expansion of $\frac{1}{H(z)}$ in positive and negative powers of z. Truncation of the convergent series and multiplication by a delay factor yields a stable and casual $H'(z)$, which is the approximate inverse of $H(z)$. Its coefficients correspond to the tap gains of the transversal equaliser. The algorithm is applicable in all cases except when $H(z)$ has a root on the unit circle. In such a case, the two-sided expansion of $\frac{1}{H(z)}$ is non-converging. Beyond this limitation, the algorithm yields automatically the optimum position for the reference tap. For certain specific channels that were considered, it has been shown to result in an improvement in the value of the residual error.

4.3 Method of Tap Control

Having selected the appropriate parameters for the design of the delay line, it remains to decide upon the method whereby the tap settings are to be varied to minimise the mean square error. The discrete method, as implemented by Arnon (35) requires only simple circuitry in its realisation. The operations to be performed are:-

(i) averaging the output of each correlator
(ii) sampling to determine the polarity of the cross-correlation values

(iii) adjusting the digitally controlled attenuators in discrete increments in the direction determined by the measured polarities.

The continuous method, on the other hand, requires more complex circuitry but results in improved static and dynamic performance. The main difficulty lies in the realisation of the variable gain circuits for continuous control of the tap settings. High precision analogue multipliers are needed. Moreover, there must be provision of some form of hold to maintain the tap settings at their optimum values during transmission of the actual TV telephone signal. It is expected that equaliser adjustments will be performed during the vertical blanking interval when no picture signal is being transmitted. The proposed interval (36) between adjustment is thus 15 or 30 milliseconds, depending on whether these are made once every field or frame. For a 2:1 interlace pattern, two fields are required to complete one frame. The hold circuits must thus be capable of maintaining the tap settings for at least this length of time.

As there do not appear to be major problems using integrated circuitry, the continuous method is adopted for the experimental system so as to achieve better equalisation performance.

4.4 Modification of Existing Circuits

Part of the work of building the system involved modifying the circuits that were used in the time-scaled investigation of automatic equalisation of television channels. These include the analogue multipliers, the delay line sections, the limiter, the set of Laguerre networks, and the summing and difference amplifiers. Basic to the modification was the selection of an operational amplifier that would satisfy the new bandwidth requirements. A Fairchild μA 749C dual in-line operational amplifer was used in place of the μA 709 of the time-scaled model. All other components were frequency scaled as required, with further impedance scaling in some cases so as to give practical values.

In the course of the testing, it was found that the μA 749C suffers from slew rate limiting near the upper frequency limit (at 500 kHz and above). Output slewing results when the rate of change of voltage required from the amplifier exceeds the specified slew rate. The peak undistorted sinusoidal output that can be obtained at a given frequency is given by

$$V_p = \frac{\rho}{2\pi f} \qquad \text{where } \rho = \text{slew rate in volts}/\mu\text{sec}, \quad f = \text{frequency in MHz}$$

Thus, in this case, for f = 1 MHz and ρ = 2 volts/μsec at unity gain, V_p works out to be only about 0.3 volts. The slew rate can be increased by proper frequency compensation, depending on the particular external circuit configuration being used. By this means, the peak undistorted output at 1 MHz was increased to about 2 volts.

There remained certain other circuits to be designed and built in order to meet the different requirements of the TV telephone channel. These include:

(i) a pseudo-random binary sequence generator which provides the test signal. Provision must also be made to generate a variable delayed version to serve as the reference signal.

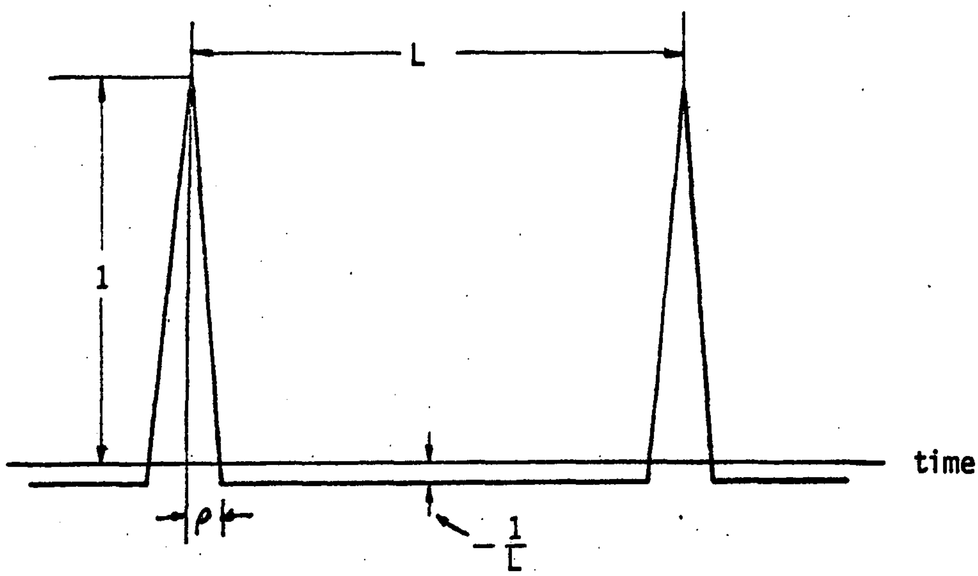
(ii) a sixth order Thomson filter to approximate the desired TV telephone channel response.

(iii) a fixed equaliser to compensate for most of the attenuation loss of the cable at the higher frequencies, thus leaving the automatic equaliser to correct only for any residual loss.

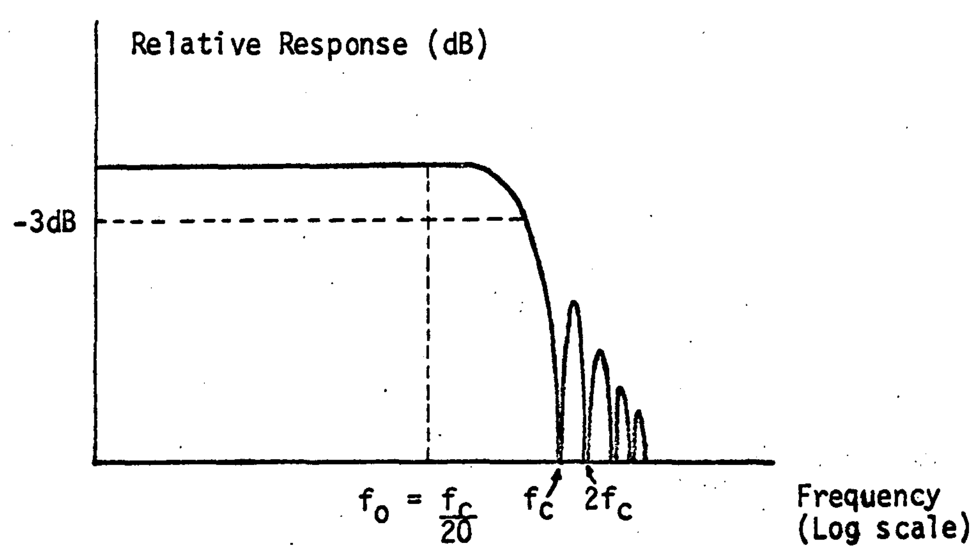
(iv) a square wave generator to drive the shift registers that made up the pseudo-random signal generator.

(v) a set of ten integrators (one for each tap of the delay line), as the analogue computer used in the previous tests on the time-scaled model was no longer adequate.

The rest of this chapter will consider the factors involved in the design of these circuits.



(a) Autocorrelation Function



(b) Envelope of Line Spectrum

FIG. 4.2 PSEUDO-RANDOM BINARY SEQUENCE CHARACTERISTICS

4.5 The Test Signal

For an N-stage shift register with feedback connections, the sequence obtained depends on both the feedback connections and the initial loading of the shift registers. For a given number of stages in the register, there is a maximum ($L=2^N-1$) to the number of bits which occurs before the sequence repeats itself. A maximal length pseudo-random binary sequence has the following properties:-

(i) For n stages, the sequence is repetitive with period given by $L\rho$ where ρ = period of driving clock. The odd number of bits L is due to the fact that the state of all zeroes at the output of each stage is prohibited. (For this case, the output of the exclusive OR gate will always be at zero and the state of the shift register will remain unchanged indefinitely).

(ii) The autocorrelation function of the sequence is as shown in Fig.4.2(a) and is periodic with the same period as the sequence. Provided that the time of measurement is taken over a complete period of the sequence, the autocorrelation function approximates in a known manner that of white noise.

(iii) The power spectrum, as shown in Fig.4.2(b), is a line spectrum, with frequency interval between spectral lines given by $\frac{f_c}{L}$ where f_c = clock frequency. The envelope has the shape of $\left(\frac{\sin x}{x}\right)^2$. The amplitude of the spectral lines is substantially constant up to $f_c/20$, with the reduction in power being only 0.0358 dB at this frequency. The 3 dB bandwidth is at 0.443 f_c.

Because of these properties and its reproducity, the pseudo-random binary sequence is particularly suitable as a test signal in control and communication systems. By suitable filtering, a good approximation to a signal with Gaussian distribution can be obtained (37). The suitability of the pseudo-random sequence for application in the field of automatic equalisation has been investigated by Chang and Ho (38). They considered the shortest possible sequence period, i.e. where the period is equal to the length of the delay line. The reasons for making this choice are:-

(i) Adjustment is made only at the end of each training sequence period and it is desired to make the largest number of adjustments during a fixed training time.

(ii) When the period of the pseudo-random sequence is shorter than the length of the transversal equaliser, the analysis is difficult since A^{-1} may not exist.

In general, the statistics of the true information transmitted will be different from those of the training sequence. Hence, the optimum equaliser setting derived using the latter may not necessarily be optimum when used for information transmission. Chang and Ho have found that any loss in optimality resulting from the use of a pseudo-random sequence is not significant. Also, the degree of degradation decreases as the sequence period is increased.

The design of a pseudo-random binary sequence (p.r.b.s.) generator is governed by two factors:-

(i) the choice of the driving clock frequency. This depends upon the highest frequency of interest in the system under test and how flat the power spectrum of the p.r.b.s. is required to be up to this frequency. It has been mentioned that the bandwidth suggested for the TV telephone channel is 1 MHz. However, other considerations, such as are outlined in Section 4.7, have indicated the desirability of an overall response that is flat to about 200 kHz and then rolling off to 20 dB at 1 MHz (refer to Fig.4.5). In view of this, a clock frequency of 5 MHz was selected. With this value, the power spectrum of the p.r.b.s. is substantially flat up to 250 kHz with a drop of only 0.036 dB at this frequency.

(ii) the choice of the shift register length N. The value of N affects the repetition period of the sequence, and hence the separation of the spectral lines, though it does not affect the shape of the envelope. A large value of N thus leads to an improved signal since it increases the number of spectral lines within the passband. A small value of N, on the other hand, allows time for the transients within the system to decay, without unduly increasing the training period. In practice, the value chosen is determined firstly, by the time interval available within the TV telephone signal for insertion of the test signal, and secondly, by the range of distortion expected and the degree of equalisation desired. With regard to the second point, Lucky and Rudin (6) have noted that the equaliser may react to the properties of the test signal rather than to that of the channel if the period of the autocorrelation function of the p.r.b.s. is less than the length of the expected dispersion in the channel. This has been confirmed by Furmage's investigation (28).

It is still not possible at this stage to determine the optimum parameters to be used in this particular application. The shift register was designed to consist of ten stages, with the outputs of stages seven and ten modulo-two added and fed back to the input of the first stage. The resulting p.r.b.s. is a 1023-bit sequence ($L=2^{10}-1$), with a period of $1023 \times \frac{1}{5} = 204.6$ microseconds.

4.6 The Reference Signal

It is expected that the received test signal will be time-delayed due to the channel distortion. For proper synchronisation with the locally-generated reference signal, allowance must be made for this delay. In actual implementation, it is expected that some form of timing arrangement will be introduced to achieve this desired synchronisation automatically. In the course of this work, however, the simple procedure was adopted of manually varying the delay of the local sequence with respect to the received test signal such that the residual error is minimised. The p.r.b.s. generator was thus designed with a second output to provide a variable-delayed version of the original sequence.

It is known (39) that a sequence delayed by an arbitrary number of bits up to (2^N-1) can be obtained by adding the outputs of the various stages of the original generator in a network of modulo-two adders. The rate at which the new sequence is generated is determined by the rate $f_c = \frac{1}{t_c}$ at which the clock pulses are shifted. The delay obtainable thus ranges from zero to $(2^N-1)t_c$ seconds, in increments of t_c seconds.

Let D be a linear algebraic operator defined to represent the delay by one clock interval. Thus $D^N X$ represents X delayed by N bits.

The feedback equation is given by

$$X = a_1 DX \oplus a_2 D^2 X \oplus \quad \ldots\ldots \quad \oplus D^N X$$

where $a_i = \begin{cases} 1 & \text{if output from ith stage (is included in} \\ 0 & \text{(is not the feedback signal.} \end{cases}$

and $\oplus$ denotes modulo-two addition.

In this case, the feedback equation is

$$X = (D^7 \oplus D^{10})X$$

i.e. $0 = X \oplus D^7 X \oplus D^{10} X$

Modulo-two adding $D^{10}X$ to both sides,

$$D^{10}X = X \oplus D^7 X \qquad (4.1)$$

since $D^{10}X \oplus D^{10}X = 0$

Repeated multiplication by D gives

$$\begin{aligned}
D^{11}X &= D^1X \oplus D^7X \\
D^{12}X &= D^2X \oplus D^8X \\
D^{13}X &= D^3X \oplus D^9X \\
D^{14}X &= D^4X \oplus D^{10}X \\
D^{15}X &= D^5X \oplus D^{11}X = D^5X \oplus D^1X \oplus D^7X
\end{aligned}$$

These equations give the outputs of the stages to be modulo-two added to give the original sequence delayed by a desired number of bits. Continuing in this way to phase L, where $L = 2^N-1$, a phase table can be constructed. Part of this is shown in

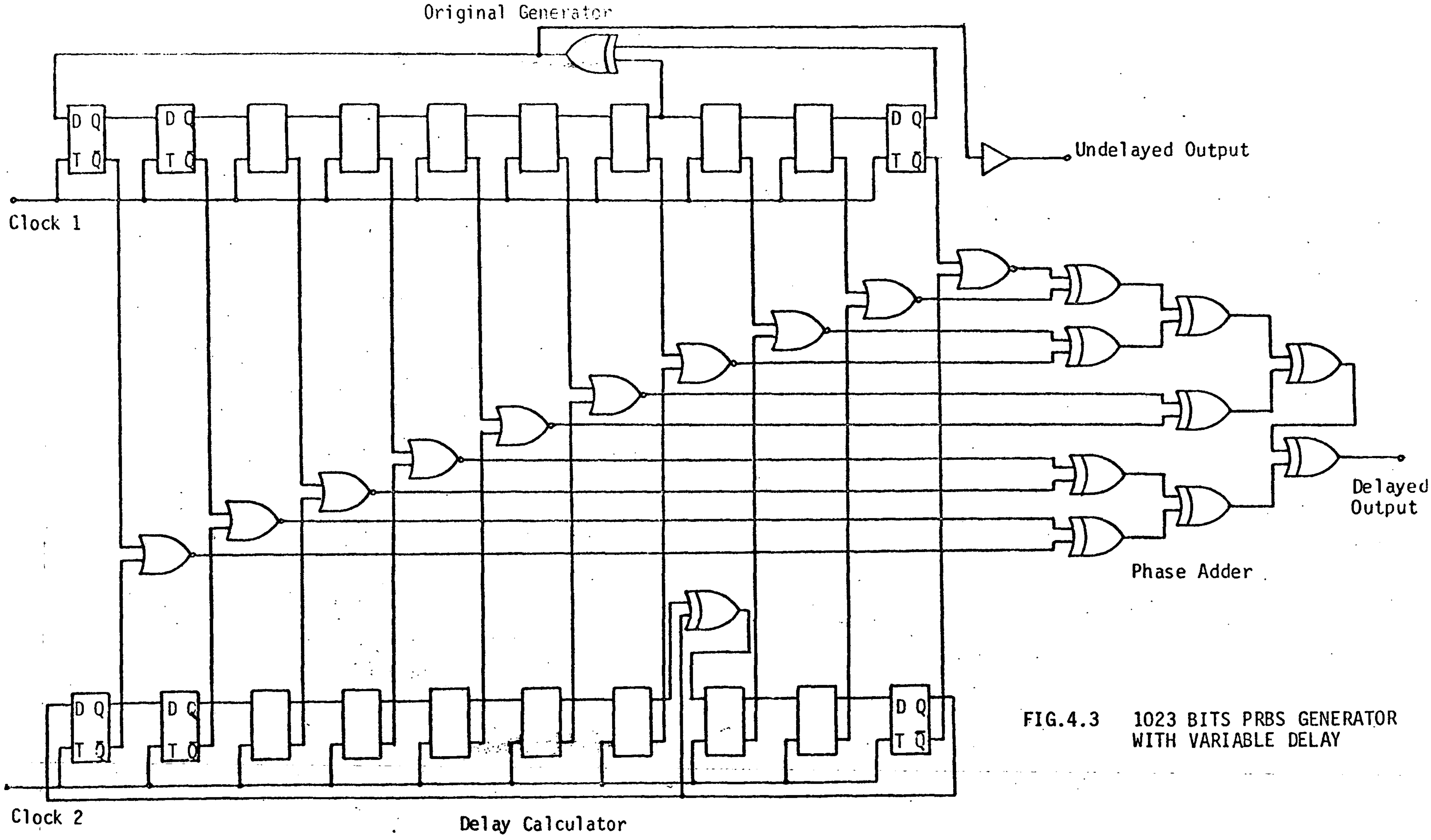


FIG.4.3 1023 BITS PRBS GENERATOR WITH VARIABLE DELAY

Appendix B. Examining the phase table shows that multiplying equation (4.1) by D is equivalent to shifting the contents of each column one place to the right. The feedback operation is equivalent to modulo-two adding the outputs of stages 7 and 10 and feeding into stage 8. This means that the process of generating the equations of the phases may be automated. As illustrated in Fig.4.3, a programmer or delay calculator can be used to indicate automatically which outputs are to be added. By including a stage-identifying circuit (e.g. one which identifies when the outputs of all stages are at logical one), the appropriate shift pulses can be supplied to drive the delay calculator, thus providing a sequential change of delay.

New techniques have been suggested (40) whereby the propagation delay time through the feedback path may be greatly reduced, allowing the speed capability of the flip flops to be fully utilised. These techniques apply only to the following cases:-

(i) where feedback is from two stages, one of which must be the first stage

(ii) where feedback is from four or more stages

In this case, where feedback is from the seventh and tenth stage, the conventional arrangement has to be used. With the TTL logic elements used, the propagation delay time of the modulo-two adder (12 nanoseconds) is not expected to have too adverse an effect on system performance.

4.7 Model Response of the TV Telephone Channel

The cathode ray tube that will be fitted in each TV telephone set is inherently capable of displaying signals at much higher frequencies. It is essential, therefore, to prevent either components of the camera signal or interference at frequencies above 1 MHz from reaching the display. It has been suggested (25) that for actual TV telephone implementation, an overall frequency which is down 20 dB at 1 MHz and more beyond that will be adequate to meet this requirement. At first sight, it would appear that the maximum resolution within the specified

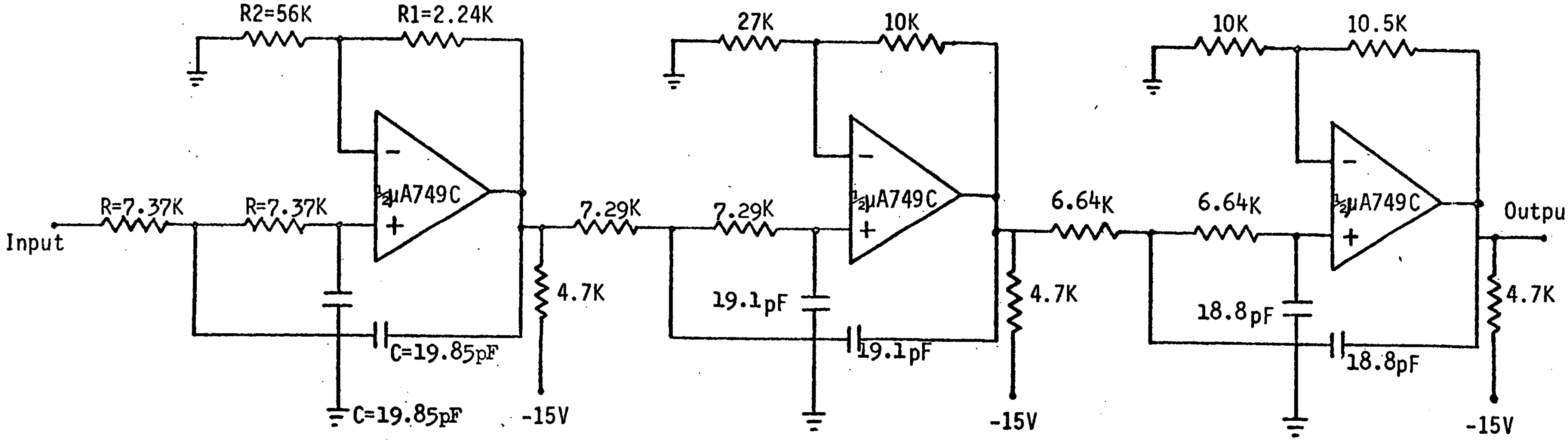


FIG.4.4 SIXTH ORDER THOMSON FILTER

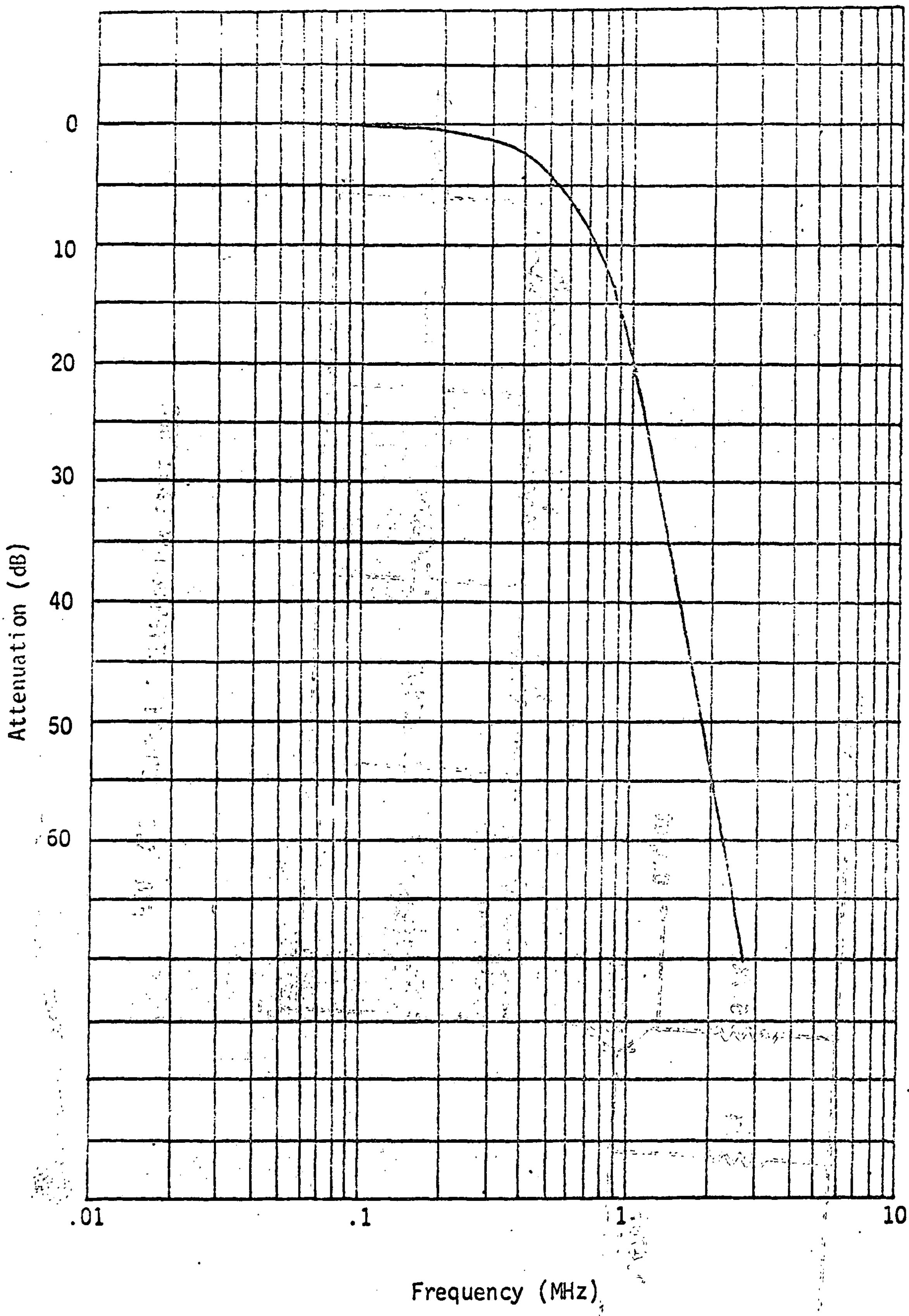


FIG.4.5 ATTENUATION CHARACTERISTIC OF SIXTH ORDER THOMSON FILTER

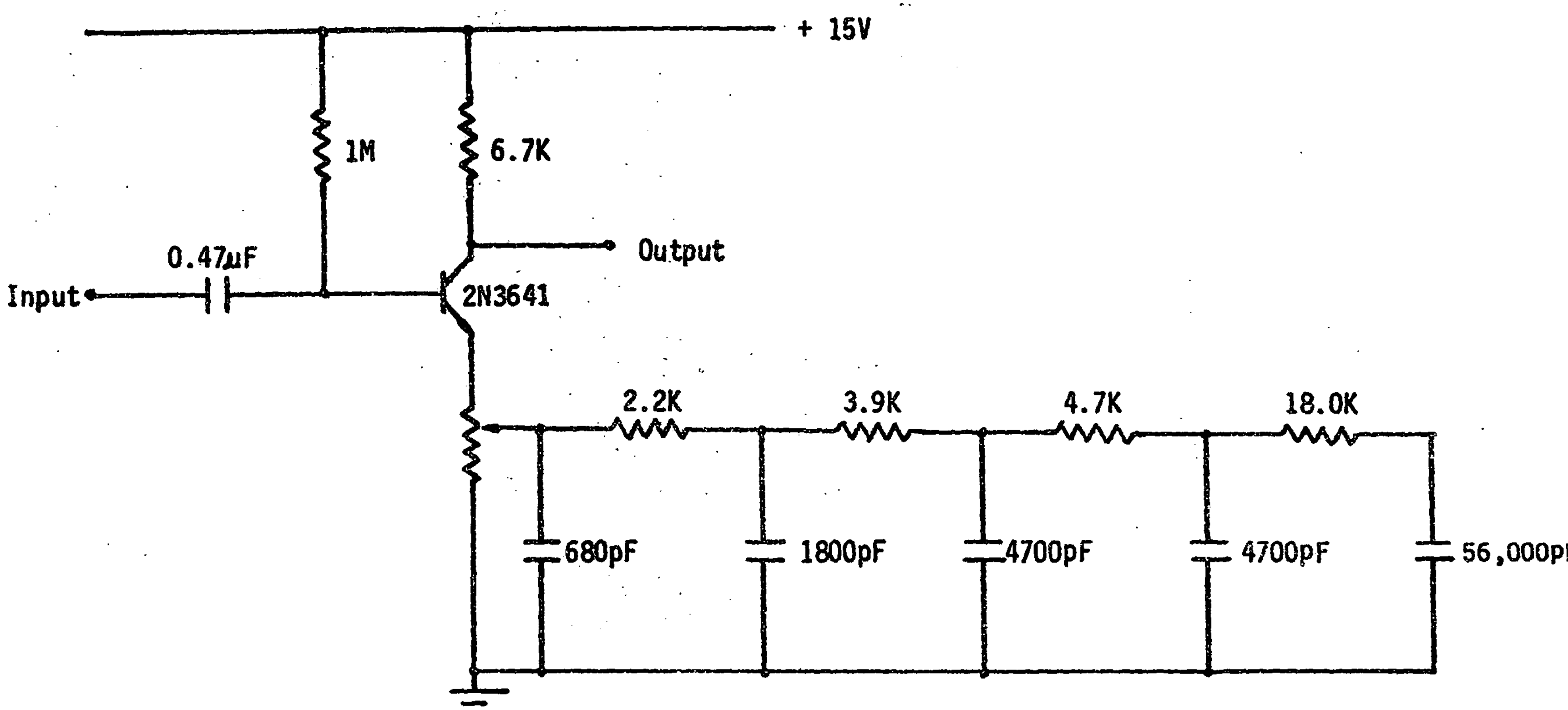


FIG. 4.6 BODE-TYPE ATTENUATION EQUALISER

band can be obtained by using a phase-equalised, sharp cut-off filter that produces 20 dB suppression at the band edge. Such a filter, however, will produce subjectively unacceptable ringing in the resultant picture. To obtain a rapid roll-off in frequency response without ringing, a filter whose impulse response is approximately a Gaussian density function should be used.

It is known that a Thomson filter approximates the ideal linear phase characteristic of a Gaussian filter. The six pole Thomson filter, as shown in Fig.4.4 was thus used to approximate the desired response. All components used were matched to within 1% of the specified values. Each of the three sections has a transfer function given by

$$\frac{V_{out}}{V_{in}} = \frac{k_0}{(\frac{s}{\omega_0})^2 + (\frac{s}{\omega_0 Q})+1} \qquad \text{where } \omega_0 = \frac{1}{RC}$$

$$Q = \frac{1}{3-k_0}$$

$$K_0 = 1 + \frac{R_1}{R_2}$$

The filter characteristic is given in Fig.4.5.

4.8 Compensation of Cable Attenuation Loss

It is envisaged that, in practice, attenuation equalisers will be installed at pre-determined intervals along the cable to compensate for most of the attenuation loss. This would mean that the automatic equaliser need only correct for any residual loss, as well as for any variations due to temperature effect and deviation from average cable characteristics. In order to adhere to actual operating conditions, an attenuation equaliser was also incorporated in the experimental system that was tested.

The attenuation equaliser used is of the Bode type (41). The circuit configuration is given in Fig.4.6.

$$\text{Gain} = \frac{V_{out}}{V_{in}} = \frac{R_L}{Z_E}$$

where Z_E is the driving-point impedance of the RC network. The poles and zeroes of Z_E interlace on the negative real axis of

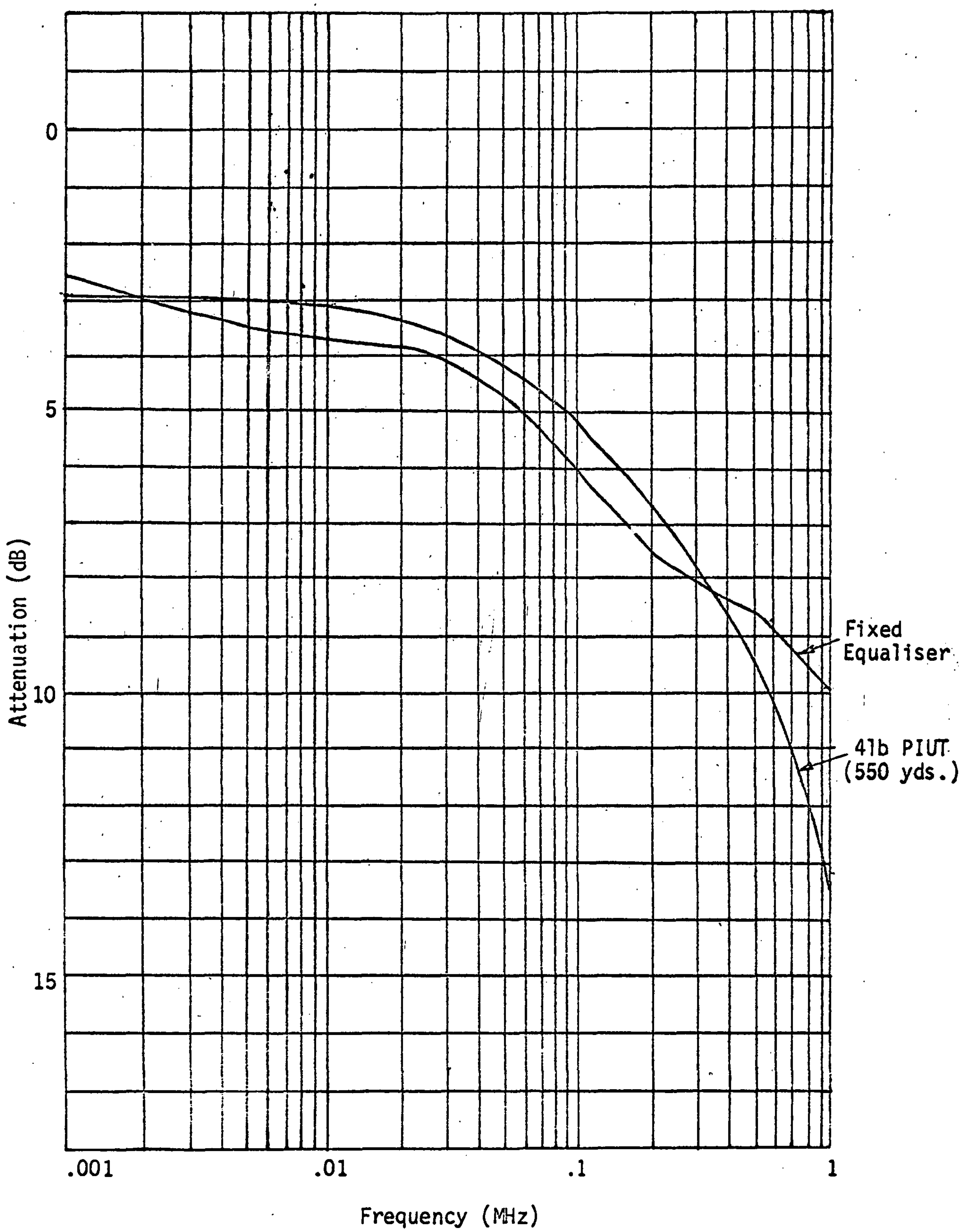


FIG.4.7 ATTENUATION CHARACTERISTIC OF FIXED EQUALISER

the complex frequency plane, the singularity closest to the origin being a pole. As a result of this property, the equaliser gain increases monotonically with frequency, with a maximum slope of 20 dB per decade change in frequency.

In Fig.4.6, the components used were matched to those specified, and trimming capacitors were included to allow for better tolerances in the values of C. The potentiometer can be used to adjust the equaliser characteristic so as to allow for use with different lengths of the same type of cable (4lb. P.I.U.T. in this case). When the wiper is at the emitter end, a maximum length of cable is equalised (corresponding to an attenuation slope of 20 dB/decade). The characteristic for equalising 550 yards of cable is plotted in Fig. 4.7. The residual error left to be compensated by the automatic equaliser, is given in Fig.4.8.

For correct use of the attenuation equaliser just described, the cable must be terminated in its characteristic impedance. If this is not, there will be a loss due to mismatch at the terminations. This loss will depend on both the frequency and the physical length of the cable being equalised. Moreover, the relationship with respect to length is not a linear one. Hence, the loss cannot be included in the design of the equaliser. Terminating networks that simulate the characteristic impedance of the cable may be required. In the tests carried out, only the nominal value of 120 ohms was used as the termination.

As noted, the purpose of including the attenuation equaliser was merely to simulate the conditions under which the system is expected to operate. Consequently, care was not taken to build one that will compensate for as much of the attenuation loss as is possible. For actual utilisation, more careful design (42) could be adopted.

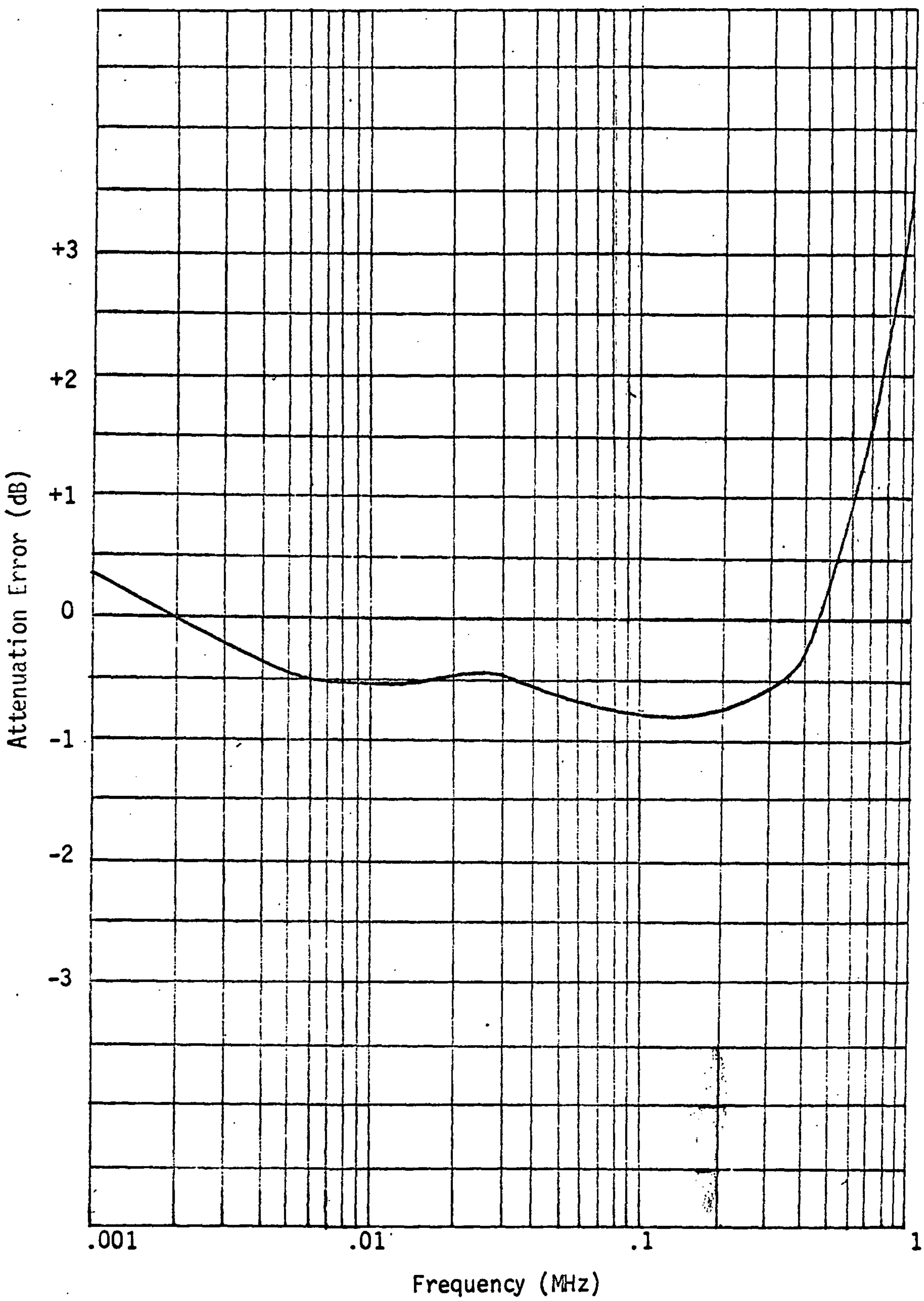


FIG.4.8 ATTENUATION ERROR TO BE CORRECTED BY ADAPTIVE EQUALISER

5. Performance of the Experimental System

5.1 General

Based on the principles outlined in the previous chapter, a real-time experimental equaliser was built. Both the Laguerre set of functions and the delay line were used as the basic equaliser set to afford a means of comparison. Various tests were carried out to determine the capability of these equalisers. These included investigating the effect of -

(a) introducing a non-linearity (the infinite clipper) into the error path.

(b) simulating the presence of noise in the received signal.

Tests were also devised to provide an assessment of some of the parameters that contribute towards optimum equaliser performance. The number of taps required and the synchronisation of the received and reference signals were considered.

These tests and their results will be discussed in this chapter.

5.2 System Parameters

The experimental set-up used is as shown in Fig.4.1. Both the received and reference signals were generated by the same pseudo-random binary signal generator, with a bit rate of 5 Megabits per second. These were filtered to approximate gaussian signals and then fed into the equaliser. For the delay line equaliser, nine sections were required, corresponding to ten taps with a spacing of 0.5 microseconds between taps. In the case of the Laguerre equaliser, an additional input section was needed and the cut-off of each section was fixed at two-thirds the band-width of interest (= 4.19×10^6 radians per second). A drum of 4lb. PIUT cable, 500 yards in length, was used to represent the TV telephone channel.

5.2.1 Delay between received and reference Signals

It was found that the degree of equalisation achievable was poor when the received and reference signals were transmitted simultaneously. This was not unexpected since, in passing through the channel, the former had suffered both amplitude and phase distortion. Consequently, it was no longer in synchronism with the undistorted local reference. To improve the equaliser performance, it would be necessary to delay the transmission of the latter deliberately so that both become delayed by the same amount.

This optimum condition can only be achieved if the amount of delay introduced by the imperfect channel is known. However, it will not be of a fixed value, but will vary from one channel to the next. So, at best, optimisation can only be effected based on some average channel characteristic. This problem of synchronisation will be an important consideration in the actual implementation of the system and will be discussed more fully in the next chapter.

For the present investigation, the reference signal was delayed by different amounts relative to the received signal. The variation was made in steps determined by the spacing of the equaliser taps. This was done because the process of physically varying the delay in transmission of the reference signal was equivalent to moving the main (or reference) tap of the equaliser down the line. That in turn meant the creating of more pre-echoes and less post-echoes.

For the transversal equaliser, it was found that minimum residual error was obtained when the fourth tap was selected as the reference tap. This was achieved by delaying the reference signal by 1.5 microseconds relative to the received signal. Similar tests on the Laguerre equaliser indicated that four pre-echoes were needed. In each case, the residual error was estimated to be less than 5% of the desired or reference signal.

5.2.2 Number of Taps

Another parameter affecting equaliser capability is the number of taps. It has been mentioned in chapter 4 that the length of the equaliser required depends on the nature and degree of distortion likely to be found in the channel, and the precision of equalisation desired. It is clear that an optimum solution cannot yet be arrived at in the absence of information on the former and definite standards laid down for the latter.

An attempt was made, nevertheless, to gauge the extent to which the choice of the number of taps was critical to equaliser capability. With the received and reference signals properly synchronised using the results of the previous section, the transversal and the Laguerre equalisers were allowed to converge, with only seven taps instead of ten but with all other parameters unchanged. A deterioration in performance of less than 1% was observed. This appears to indicate that, within the limits of the experimental conditions, the use of more than seven taps will only bring about a marginal improvement in system capability.

Further experimentation with six taps or less resulted in an increasingly significant residual error. The use of seven taps would thus seem to represent a lower limit for the particular channel being equalised. More detailed investigation would need to be conducted to establish this conclusively. However, for the remaining tests that were carried out on the experimental system, only seven taps were employed.

5.3 Effect of Non-Linearity in Error Path

The theoretical evaluations of chapter 3 have pointed to the desirability of introducing a non-linearity into the feedback loop of the equaliser. To confirm some of these predictions, an infinite clipper (or hard limiter) was built and incorporated into the system, at the output of the difference amplifier producing the error signal. The performance of this "modified" equaliser was monitored and compared to that of the original mean-square equaliser.

It was observed that the presence of the clipper had the effect of speeding up convergence, by a factor of about 10.

By adjusting the level of the clipper between ±5 Volts and ±10 Volts, the gain of the loop was varied. This resulted in an inversely proportional variation in the settling time (defined as the time for the residual error to fall to within 5% of its initial value). These observations concurred with those made by Sondhi, in his work on the adaptive echo canceller (2). The same improvement was observed when the clipper was incorporated into the Laguerre equaliser. The settling time obtained was of the order of a few hundred microseconds.

For both mean square equalisers, the settling time was also found to be dependent upon the time constant of the cross-correlators feeding the taps. In other words, the choice of the RC product for the set of integrators is important. This was investigated by keeping C constant, and using different values for R. The speed of convergence that was obtained decreased as the value of R was increased. In fact, the smaller the RC product, the more rapid was the convergence process and the shorter the settling time became. For instance, by reducing the value of RC from 95 to 0.95 milliseconds, a corresponding improvement in settling time from 20 to 4 seconds was observed. This effect became less pronounced when the clipper was introduced, since then the settling time was only a few hundred microseconds and any further improvement was not easily noticeable.

5.4. Effect of Noise in Received Signal

In particular, the equaliser will be required to operate under noisy conditions, since noise is present to varying degrees in all communication channels. The effect of noise in the received signal must therefore be evaluated.

Most of the work with adaptive equaliser systems have been concerned with the effect of Gaussian noise. Lucky and Rudin (6) have demonstrated theoretically that the mean-square equaliser will be able to reduce the sum of the residual error and the additive, uncorrelated noise. This is confirmed by Furmage's results (28) with his time-scaled TV system. In the absence of time to carry out more exhaustive tests, it was decided to assume that gaussion noise would not pose a major problem.

A more serious obstacle is the effect of impulse or switching noise in telephone channels. Any complete study, however, will not be possible since no attempt has been made to characterise the statistics of the impulse noise actually present. It is anticipated that one effect of such impulses would be to disturb the equilibrium of the equaliser. Isolated impulses were, therefore, generated and added to the input to the equaliser when it had nearly converged. The impulses had durations ranging from 10 to 100 microseconds and amplitudes of up to 5 volts. With the amplitude and duration as parameters, the ability of the system to tolerate such extraneous peaks was assessed.

It was found that the equilibrium of the mean-square equaliser was affected when the disturbing impulse was more than 15 dB above the received signal. The mean-square error increased by about 20 dB in 0.5 seconds, the rate increasing as the level of the disturbance was increased. The situation improved with the introduction of the clipper, and a higher level of noise (of the order of 25 dB) was tolerable. A change of only about 3 dB in 2 seconds was observed in the latter case.

5.5. Effect of Choice of Basic Equaliser Function

For the particular channel considered, the test results indicate that the Laguerre equaliser performs as well as the transversal equaliser. This differs from the observation of Furmage, who found the Laguerre equaliser to be less satisfactory, requiring more than ten taps for good performance. The difference between the two findings arises from the different types of channel that were equalised. The channel that Furmage investigated had an impulse response time-spread of 170 microseconds (main pulse) as compared to 2 microseconds for the simulated TV telephone channel. In addition, the time scales of the delay lines and laguerre networks that he used were 200 times greater.

5.6. Accuracy of Experimental Results

The accuracy of the test results is dependent upon a number of factors, one of which is the correct choice of parameters and components for the different circuits. This factor will be discussed in this section.

5.6.1. Choice of Multipliers

Good linearity of the multipliers is essential to give a faithful reproduction of the signals applied at the inputs. Before the multipliers are used, therefore, careful adjustments are needed to eliminate as much of the d.c. offset errors as possible. The gain, on the other hand, may be of lower tolerance since any gain discrepancies will be corrected by the feedback loop. In view of this, in cases where zero offset and accurate gain cannot be achieved simultaneously, the former was attained at the expense of the latter.

Other than good linearity, a second important requirement is zero a.c. offset. Theoretically, the choice of the continuous method of tap control means that infinite resolution of tap settings is possible. But in practice, the accuracy of the settings depends on the presence of any a.c. components in the outputs of the cross-correlators feeding the taps. Such signals, when present, cause the settings to fluctuate about their optimum values. For the multipliers used, the fluctuations were able to be limited to about 1% of the actual tap values through proper nulling, and were thus not significant.

5.6.2. Choice of Operational Amplifiers

The μA749C dual in-line operational amplifier had inadequate high-frequency performance. Its slew rate limited the accuracy of the different circuits (see Section 4.4.). For frequencies above 500kHz, the signals were distorted. An upper limit had to be imposed on the magnitudes of the received and reference signals to avoid saturation of the amplifier outputs.

The circuits could have been redesigned, using another operational amplifier with a higher slewing capacity. Some were in fact so modified, e.g. the summing and difference amplifiers. But time did not permit any extensive modification. Instead, a compromise was accepted and the upper video frequency of interest was reduced from 1MHz to 400kHz. The results so obtained can still be taken as representative of real-time operation, since the proposed TV telephone standard does not require the equalised response to be flat up to 1MHz. The actual 3 dB bandwidth specified is 400kHz (see Section 4.7.).

The difficulties experienced with the different circuits restricted the number of tests that could be usefully carried out. Nevertheless, the limited results were helpful in the preliminary assessment of the equaliser. The next chapter will consider the implications of these results for actual implementation. Additional tests will also be suggested.

6. Feasibility of Automatic Equalisation of the TV Telephone Channel

6.1. Implications of Test Results

From the preliminary test results reported in Chapter Five, it would appear that automatic equalisation of the TV telephone channel represents a workeable concept. The performance of the experimental equaliser was encouraging, within the scope of the limited tests that were carried out.

One direct inference that can be drawn from the test results is the desirability of having non-linear feedback. For the non-linearity that was considered, the settling time was significantly reduced. However, the precise nature of its effect is still not known. The study by Furmage (28) suggests that it acts as a variable gain device, effectively increasing the gain of the feedback loop during convergence.

It has also been confirmed that optimum use of the equaliser is achieved when the received and reference signals are properly synchronised to account for the delay introduced by the channel. In the actual implementation, there must be proper synchronism, whatever the particular TV telephone connection being equalised. This topic will be discussed later in the chapter.

The test results also indicate the possibility of difficulties being associated with disturbing impulses. Initial observations suggest that the equaliser may not converge if impulse noise above a certain magnitude is coupled to the received signal just prior to entering the equaliser. But the exact extent of such a setback can only be conclusively determined if the properties and distribution of the impulse noise present in telephone channels are known.

Overall, the test results indicate the comparable capability of transversal and laguerre equalisers. A lower limit of seven taps was found to provide sufficiently good equalisation. It should be emphasised though that the observations were valid only for the particular channel used i.e. adjacent pairs in the same quad of a 500 yards length of 4 lb PIUT cable. Further investigation would be needed to establish the optimum size and type of equaliser. The relevance of such work would again be conditional upon knowing the extent of distortion and the range of variation from channel to channel existing in practice.

6.2. Areas Requiring Further Investigation

From the discussion in the previous section, some factors requiring further study emerge. Experimentally, there are some immediate tests to be carried out to consolidate or improve on preliminary findings. Theoretically, there are some analytical methods to be considered to give a better understanding of the equalisation process. These will be outlined in Sections 6.2.1 and 6.2.2 respectively.

6.2.1. Further Tests

(i) Equaliser Circuits. By proper choice of components to eliminate such restrictions as slew rate limiting and saturation of output waveforms, the performance of the different circuits can be improved. Similarly, the design of the fixed equaliser of Section 4.8. can be corrected to better simulate the operating conditions.

(ii) Non-linearity in Feedback Loop. From Chapter Three, it has been seen that the choice of the non-linearity F significantly affects the performance of the system. It may be worthwhile, therefore, to introduce a deadband (when the residual error has become small) in place of the infinite clipper. This will reduce fluctuations in the tap settings during convergence, thereby minimising the settling time. The deadband can be achieved by simply opening the loop at the place where the clipper would otherwise be. Alternatively, another function could be used as F. One such example is described by the equation

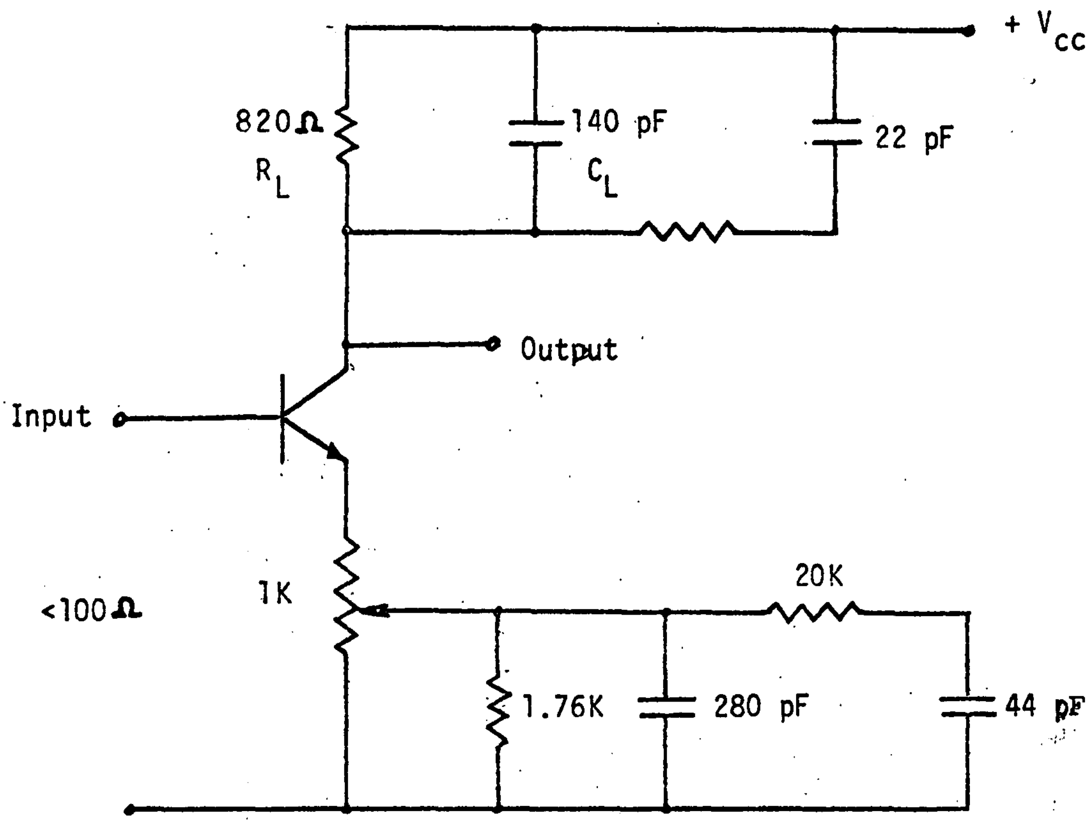
$$F(e) = \begin{cases} e & \text{for } -e_0<e<+e_0 \\ e_0 & e>+e_0 \\ -e_0 & e<-e_0 \end{cases}$$

(iii) Channel being equalised. In the tests, the TV telephone channel was represented by 500 yards of cable. This choice assumes that the fixed equalisers will be spaced at that interval along the line. Brown (1) and others have suggested, however, that the spacing could be extended to about a mile. The actual optimum value will depend largely on the crosstalk properties of the lines. Closer spacing will reduce the possibility of impulse noise coupling through near-end crosstalk paths from audio to video pairs. The effect of using a longer cable as the channel should hence be investigated. In addition, different pairs within the same cable can be used, varying from adjacent pairs to pairs in adjacent layers.

(iv) Changes in temperature. It is known that the frequency characteristic of the telephone line varies with temperature, the deviation being greatest at the higher frequencies. An additional loss increasing with frequency is thus incurred when the line is at a higher temperature than the ambient, for which the fixed equaliser was designed.

Sargeant and Deans (43) have measured the deviations for typical lines. For 500 yards of 4 lb PIUT cable, the variation is observed to be less than ½ dB for ± 10°C variations from the nominal temperature of 20°C. It increases for further changes in temperature, but is still less than ½ dB, even at the upper frequency limit of 1MHz. Such uncompensated losses become significant as the length of the line used increases. Over the local subscriber area, where the connections are typically about one to two miles long, the cumulative error may come to about 2 or 3 dB.

The equaliser must cope with the error introduced by temperature effects. Tests need to be devised in which changes to the channel characteristic are simulated to approximate those due to temperature variations. The circuit designed by MacGregor (44) can be used. As shown in Fig. 6.1., it simulates the temperature dependence of the insertion loss of 2½ miles



Notes : 1. Transistor f_T > 100 MHz

β_N > 100

2. Values of R_L and C_L can be adjusted to take into account loading

Fig. 6.1 Circuit to Simulate Temperature Dependance of Cable.

of 4 lb PIUT cable. For frequencies over several hundred kilohertz, this variation of insertion loss will give results identical to that for attenuation variation. The simulated temperature, given by the potentiometer setting, is

$$\tau = 40(1 - x^2)$$

where τ = temperature, measured in degrees Centigrade

x = function of potentiometer resistance between wiper and ground.

6.2.2. Further Analysis

In Chapter Three, a linearised analysis was attempted, appropriate to small signal conditions. To obtain a more general solution, it would be necessary to use non-linear techniques. Possible approaches are:

(i) Stability analysis (45). This involves the application of LyapunoV's method, or extensions of it, due to Popov and others.

(ii) Phase plane method (46). This is a graphical technique with no intrinsic limitation in accuracy since all non-linear characteristics are treated without modification or linearisation. It is usually restricted to second order systems. Further study is required to establish its adequacy here, particularly with regards to multiplicative non-linearity.

6.3. Problems in Actual Implementation

Thus far, the performance of the automatic equaliser has been evaluated in isolation. Little reference has been made to its compatibility when integrated into the existing telephone network. Without going into details, some of the relevant problems are highlighted in this section.

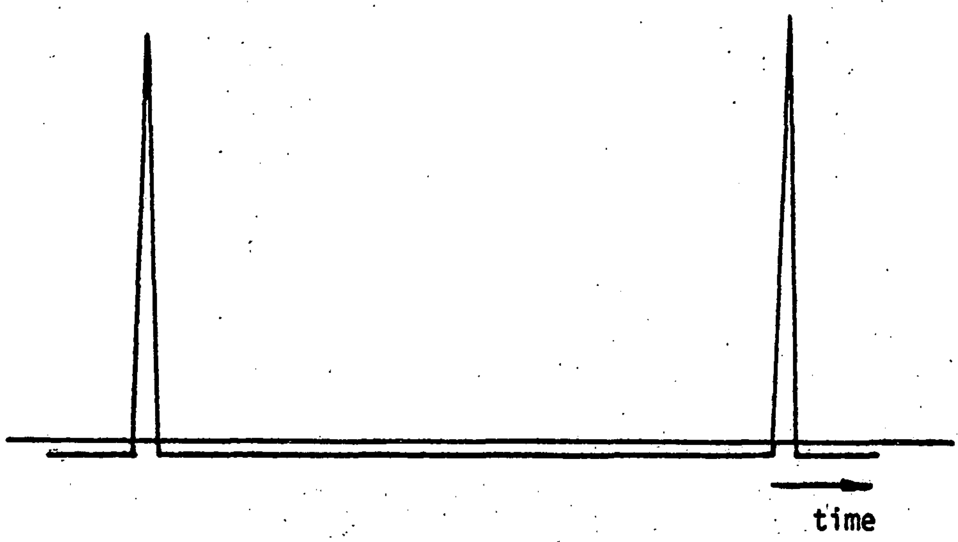


Fig. 6.2.(a) Autocorrelation Function of Pseudo-Random Test Signal for Ideal Channel

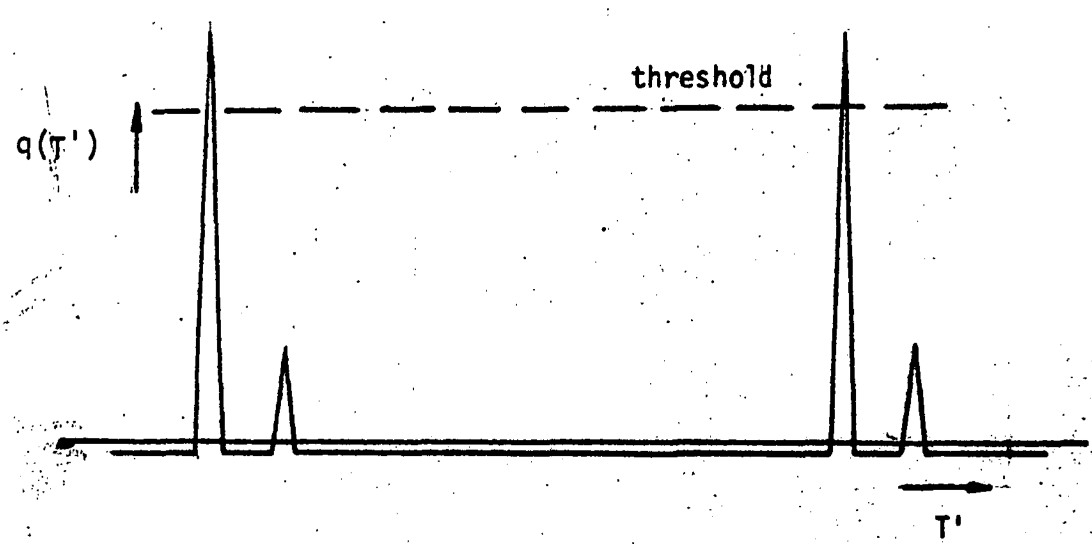


Fig.6.2.(b) q(T') Function for Distorted Channel.

6.3.1. Synchronisation of Received and Reference Signals

Equalisation is achieved by minimising the mean-square difference between the received signal and an identically-generated but undistorted local reference. The two signals must be properly synchronised. The difficulty in establishing correct timing is that the telephone line will generally vary in length and characteristic from one connection to the next. Any scheme of synchronisation must have automatic provision for such variations.

The objectives are:

(a) providing the clock frequency for driving the local reference generator

(b) establishing the proper phase for this clock, and synchronising the two random sequences.

Lucky and Rudin (6) have proposed a technique that is applicable to any general-purpose communication channel. Two pilot tones are added to the received signal, one each at the upper and lower edges of the band of interest. By suitably combining these tones, the required clock frequency is obtained.

The second objective is achieved by estimating the arrival time T' of the received signal x(t). A maximum likelihood estimate can be obtained by adjusting T' until

$$q(T') \equiv \frac{1}{t_1}\int_0^{t_1} y(t + T')x(t)dt,$$

which is the cross-correlation of the received and the reference signals, is a maximum. For large observation times t_1, this can be accurately determined despite the noise component in x(t). In fact, when t_1 is very large, the shape of q(T') would approach that of the autocorrelation function of the pseudo-random test signal for an ideal channel. The function is shown in Fig. 6.2.(a).

It is known from Wheeler's paired echo theory (47) that the effect of linear distortion in a bandlimited channel can be represented in terms of pairs of echoes of the time-domain impulse response. Therefore, for a distorted channel, the $q(T')$ function might have the appearance of Fig. 6.2.(b), which illustrates the systematic contribution of a single echo introduced by the channel. In such a case, the search for the absolute maximum of $q(T')$ can be accomplished in two successive modes. In the first mode, the correct "spikes" of $q(T')$ is found, and in the second mode, the maximum of the "spike" is determined.

For mode one, coarse alignment is obtained by cross-correlating the two signals and comparing the result with a fixed threshold. The threshold is pre-determined empirically so that only the large spikes, corresponding to the undistorted pulse, exceeds it. The phase of the timing signal is continuously increased until $q(T')$ reaches this level. After that, the phase is locked. In practice, if the same signal format as the television system is used for TV telephone operation, sufficient timing information may already be present. This could be utilised to obviate the necessity of the coarse alignment.

For mode two, the maximum is found by partial differentiation with respect to T', and setting the result to zero. The equation is given by

$$\frac{\partial q(T')}{\partial T'} \equiv \frac{1}{t_1} \int_0^{t_1} y'(t+T')x(t)dt$$

where $y'(t)$ is the time-derivative of $y(t)$. The expression is only valid for a small region over which $q(T')$ can be assumed to be convex. Practically, the partial derivative is generated by another cross-correlator.

6.3.2. Transmission of Test Signal for Equalisation

The operation of the automatic equaliser also relies upon the transmission of a pseudo-random test signal. For the preset equaliser, this is sent prior to or during breaks in the transmission of the information signal. For the adaptive type, on the other hand, both signals are added and then transmitted simultaneously. In either case, there is the problem of integration of the test signal into the TV telephone signal format. It must be gated in at the sending end and retrieved subsequently at the receiving end.

As discussed in Section 1.6., only the preset type is being considered here. The test signal can hence be introduced into time slots that are not being used for video transmission. For the television system, there are vertical blanking intervals occurring between each field of scan. By using these blanking intervals, the equaliser tap settings can be periodically updated as the channel characteristic changes slowly with time.

The length of the vertical blanking interval depends on the picture standards. Cagle (36) has proposed the following parameters:

line repetition rate	8 KHz
frame rate	29.9625Hz
resolution ratio	1:1
aspect ratio	1.1:1

Additionally, he recommends a two-to-one interlacing pattern. This means that the picture is scanned twice, the second time between the first set of scanning lines. In this way, two fields are required to complete one frame.

With these parameters, 267 lines are transmitted per frame. This includes those lost in blanking, so there are effectively only 250 active lines. The empty time slot available corresponds to the time for 8 lines, which is 1000 microseconds, if adjustments are made one per field.

To avoid introducing undesirable effects into the picture, all equaliser operations must be confined to the blanking intervals. This means allowing time for any transients to decay, and taking into account the horizontal sync. pulses normally also transmitted in the blanking intervals. These factors dictate the length of the pseudo-random sequence that can be used. The sequence selected, consisting of 1023 bits with a period of 204.6 microseconds (see Section 4.5), fits comfortably into the 1000 microsecond blanking slot. Since the average settling time obtained using the clipper is about 400 micro-seconds, the equaliser operations would be completed within the slot.

CONCLUSION

Preliminary studies have been made on some aspects of automatic equalisation of TV telephone channels. The technical feasibility of equalising baseband analogue signal, transmitted over a short length of telephone cable, has been demonstrated. By incorporating a non-linear device, the capability of the automatic equaliser has been further increased, with satisfactory performance even in a limited noise environment.

In the context of the work, it has firstly been assumed that baseband analogue transmission is used. This assumption had not been substantiated but was adopted on the basis of economic and technical possibilities. Other methods of transmission could well turn out to be more technically satisfactory. For instance, carrier transmission with amplitude modulation, which leaves the audio frequency free for voice transmission, or pulse code modulation techniques could be utilised, with definite advantages. However, considerable extra cost would be involved in installing modulation and repeater equipment both at the exchange and at the subscriber premises. Such can only be offset against the cost of one or two wire pairs. It remains to be established whether the tradeoff is justifiable.

Secondly, it has not been possible to assess accurately the degree of optimality of the equaliser used. This is because the precision of equalisation required, and the nature and degree of distortion to be expected in the channel are not fully known. The former depends on the TV telephone picture standards, which have yet to be officially formulated. The latter depends on the availability of samples of the impulse responses of various connections likely to be encountered in practice. With these information, a complete study can then be made of

(a) the size of equaliser required (i.e. the optimum number of taps to be used)

(b) the type of equaliser required (i.e. the possibility of using other basic equaliser functions and the extent of tradeoff between sophistication of algorithm and complexity of hardware)

(c) optimum equaliser utilisation (i.e. proper synchronisation of local reference with received signal)

It may even be possible to establish a fixed weighting of the tap gains that will give optimum system performance.

Finally, all the measures of performance so far are esentially objective. They could well be quite different from subjective assessments. Some kind of distortion measure that is related quantitatively to the actual picture observed on the screen is needed. The method of pulse and bar testing adopted for television could perhaps be examined for its possible application in the TV telephone situation.

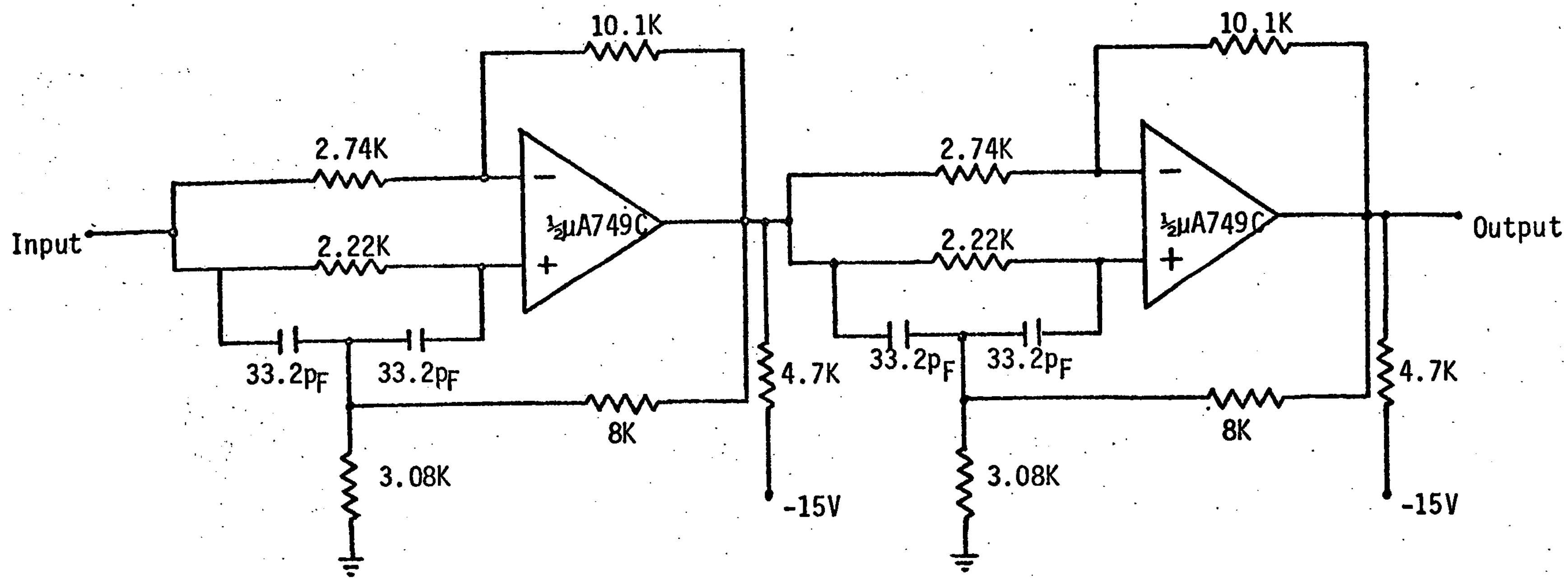


FIG. A.1 0.5 MICROSECOND DELAY LINE SECTION

APPENDIX A

This appendix contains further details of some of the hardware that was used in the experimental equaliser. The configurations of the circuits, together with component values and tolerances, are given.

A.1. Delay Line Section

The active circuit used to provide a rational function approximation to pure delay is based on suggestions proposed by Potter (48), but with the elimination of any inductive components. Potter has shown that a network with normalised pole and zero positions given by

$$s = -1.6 \pm j\,1.0 \quad \text{and} \quad s = +1.6 \pm j\,1.0$$

respectively will give a constant delay of $\frac{1}{2}$ T, with a tolerance of 0.5% over the required frequency range. Two such circuits are therefore needed to provide each section of delay T. For an N-tap delay line, N-1 delay sections are needed.

The choice of the value for T, the tap spacing, is critical to the adequacy of the transversal equaliser to correct for any expected channel distortion. For a channel which is essentially low-pass in nature, $T = \frac{1}{2W}$ where W = highest frequency of interest. Thus, in this case, for a 1MHz system, T was designed to be 0.5 microseconds.

The circuit configuration is as given in Fig.A.1. Each half-section is a second order network with transfer function of the form

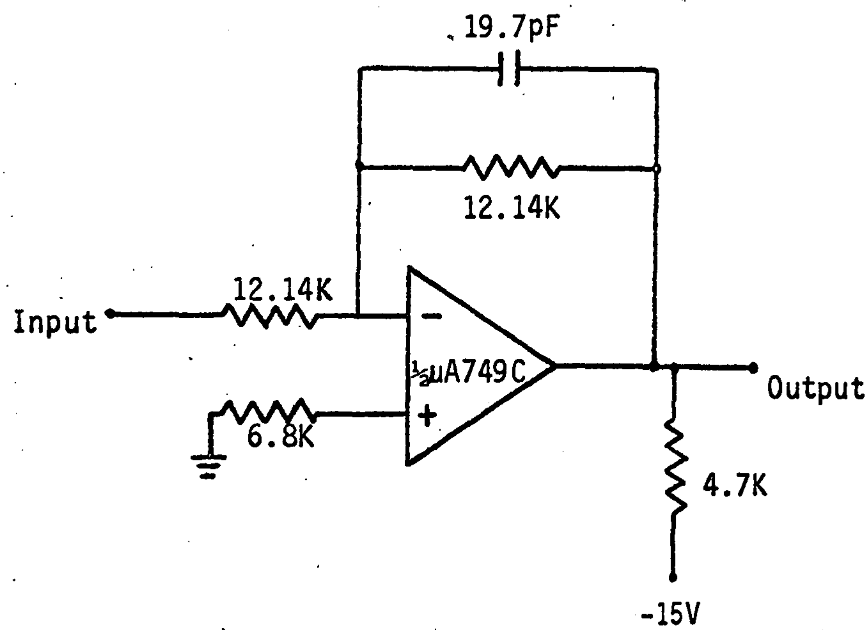
$$\frac{V_{out}}{V_{in}} = \frac{\left(\frac{s}{\omega_0}\right)^2 - 2\zeta\frac{s}{\omega_0} + 1}{\left(\frac{s}{\omega_0}\right)^2 + 2\zeta\frac{s}{\omega_0} + 1}$$

(For circuit analysis, refer to (28)). All components were matched to within 1% of the values specified.

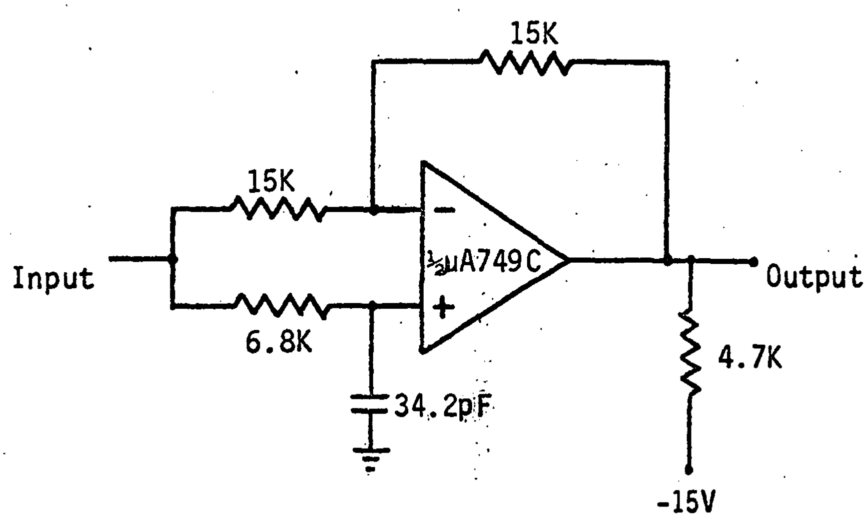
A.2. Laguerre Network

Another set of function was investigated as the possible basic equaliser function to be used instead of the delay line. It is the Laguerre set of first order networks. The impulse response of the nth Laguerre network is given by

$$l_n(t) = e^{\alpha t} \frac{d^n}{dt^n} \left(\frac{t^n}{n!} \; e^{-2\alpha t}\right)$$



(a) Transfer Function $\frac{\alpha}{s+\alpha}$



(b) Transfer Function $\frac{s-\alpha}{s+\alpha}$

FIG.A.2 LAGUERRE NETWORKS

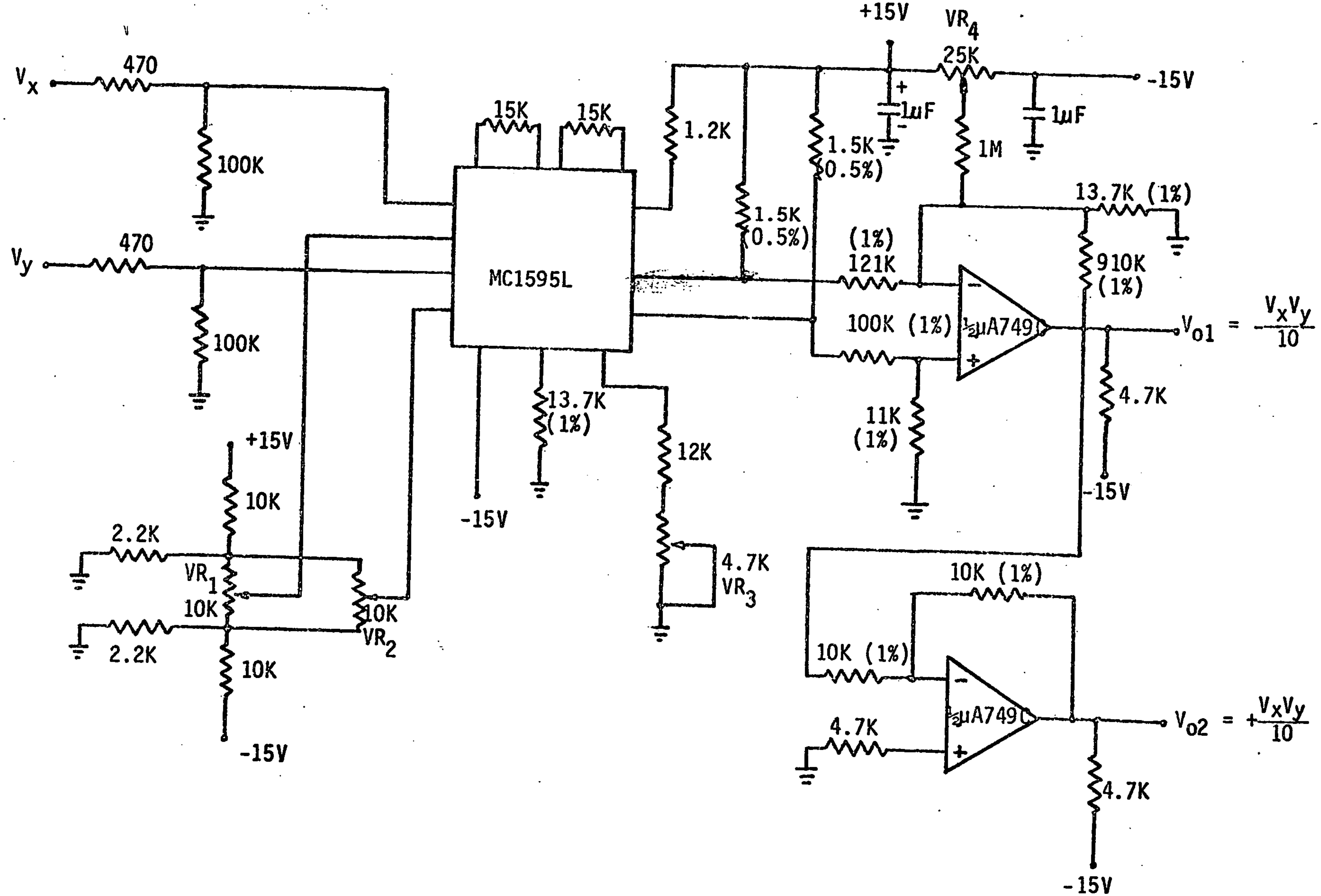


FIG.A.3 INTEGRATED CIRCUIT MULTIPLIER

with the corresponding transfer function

$$L_n(s) = \frac{\alpha}{s+\alpha}\left(\frac{\alpha - s}{\alpha + s}\right)^n$$

Synthesis is possible by passive circuitry, such as simple RC ladder networks. Active circuits were used however to avoid any problems of impedance mismatch.

The circuit arrangement is shown in Fig.A.2. The design was based on a value of α equal to two-thirds of the bandwidth of the system to be equalised. Just as for the delay networks, the components were closely matched to the values specified. For a ten-tap equaliser, nine sections of transfer function $\left(\frac{\alpha - s}{\alpha + s}\right)$ were built, together with the input network.

A.3. Analogue Multipliers

Adopting the continuous method of tap control as discussed in Section 4.3, analogue multipliers with good precision and wide bandwidth were needed. In all, two sets of N multipliers (N = number of taps in the equaliser) were required. Sondhi and Presti (49) have indicated that the multipliers used for controlling the tap settings have to satisfy more stringent requirements than the corresponding ones used in correlating the error signal with the received signal. The outputs of the former must be strictly proportional to the received signal.

To meet these specifications, integrated circuit multipliers with good static accuracy (within 1% and 2% of the maximum output for the X and Y inputs respectively) were used. Fig.A.3 shows the multiplier connections, with op-ampl. level shift. The variable potentiometers R1 and R2 are for ac offset adjustments of the two inputs, R3 for scale factor adjustment, and R4 for dc offset compensation. Of these adjustments, the most critical is the last one since, as seen from Chapter 3, any dc offset error will directly affect the output of the correlation multipliers. Thus, a multi-turn potentiometer was used for R4 to allow for more precise zeroing of the dc level.

It was found, during the course of the investigation, that the μA 749C's slew-rate limiting resulted in the positive output being

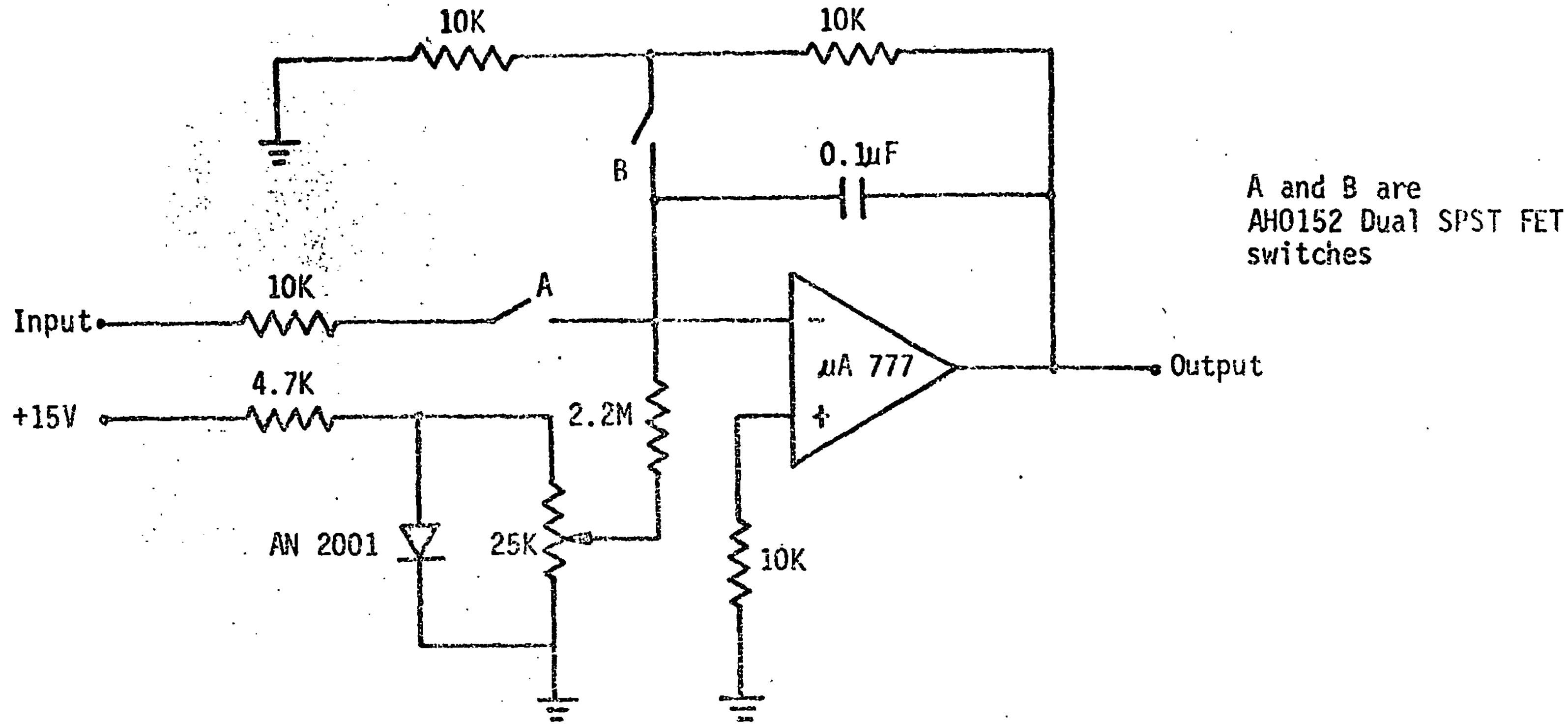


FIG.A.4 INTEGRATOR

incorrect for the higher frequencies of operation (above 100 kHz). As the situation could not be adequately improved even with application of different external compensations, only the negative output was used for each of the multipliers.

A.4. Integrator

The operation of correlation of the error signal with the received signal requires one integrator to be associated with each multiplier. The circuit used is shown in Fig.A.4. Since equalisation will only be performed during breaks in the transmission period, the integrator must be capable of holding its output when required. Provision is therefore made for operating the integrator in the hold mode, and for correcting any drift in the output (by varying the 25K potentiometer) that might affect the gain setting of the taps.

The RC product of the integrator should be chosen such that integration is performed over several periods of the pseudo random sequence. Lucky and Rudin (6) have selected RC $\simeq$ 4 x (p.r.b.s. period). The value of RC used will directly affect the gain of the feedback loops, and hence the speed of convergence. In this case, for a 1023-bit sequence, driven at a clock frequency of 5MHz,

$$\text{Period} = \frac{1023}{5} \text{ microseconds}$$

$$= 204.6 \text{ microseconds}$$

The value of RC used $= \dfrac{10 \times 10^3 \times 0.1}{10^6}$ s = 1 millisecond

A.5. Clipper Circuit

The μA 710 comparator was used to provide the limiting action in the investigation of the effect of non-linear loops. A number of circuit additions were necessary as seen in Fig.A.5. The discrete, matched pair in front of the comparator serves the purpose of increasing the gain as the μA 710 does not provide sufficient gain for the switching action. Any offset voltages present can be balanced out by the 47K variable potentiometer. In addition, allowance must be made for the fact that the input current of the μA 710 is often high enough to cause significant errors due to loading of the signal and reference-voltage sources. This is taken care of by the introduction of the transistor pair (SE4902).

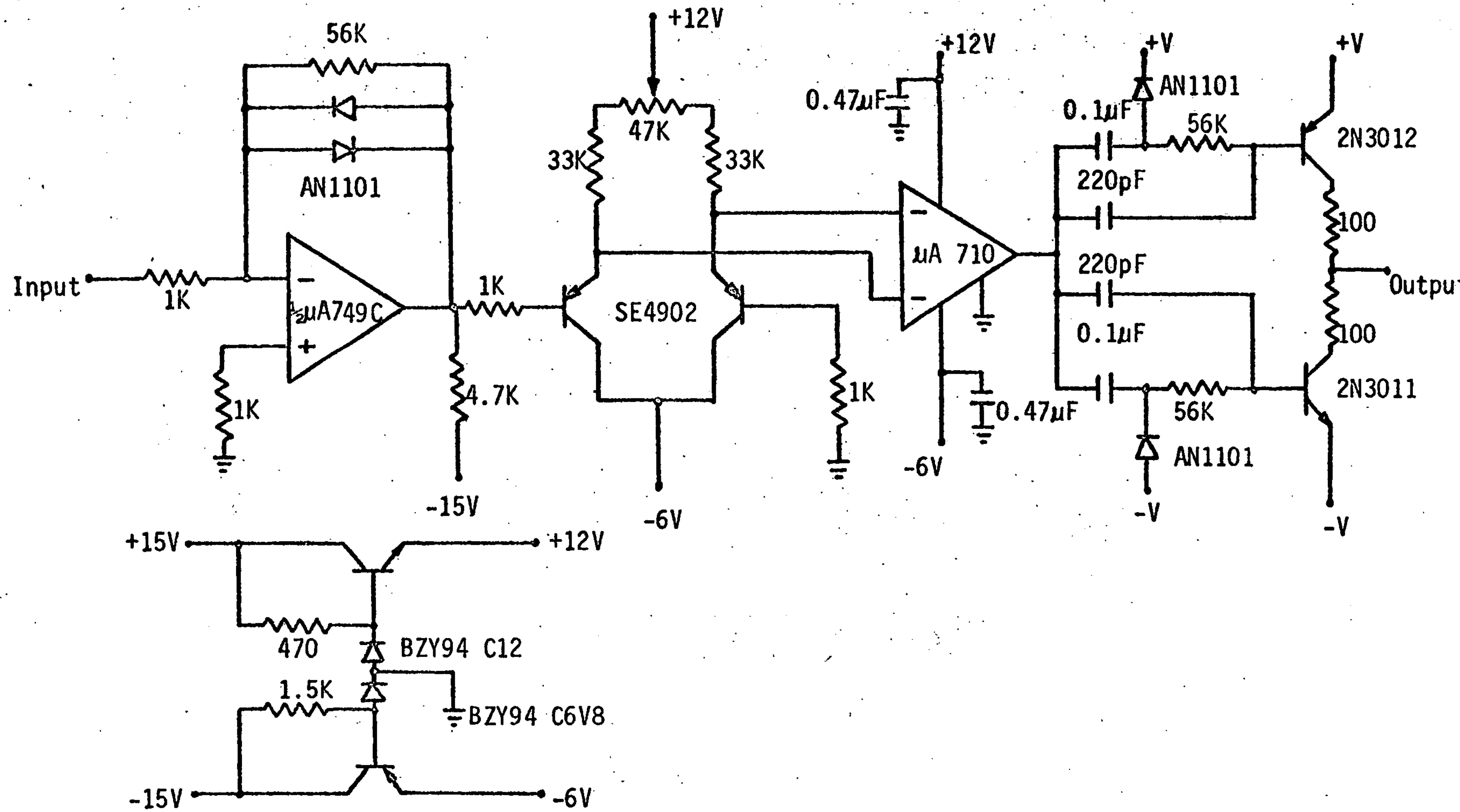


FIG.A.5 CLIPPER CIRCUIT

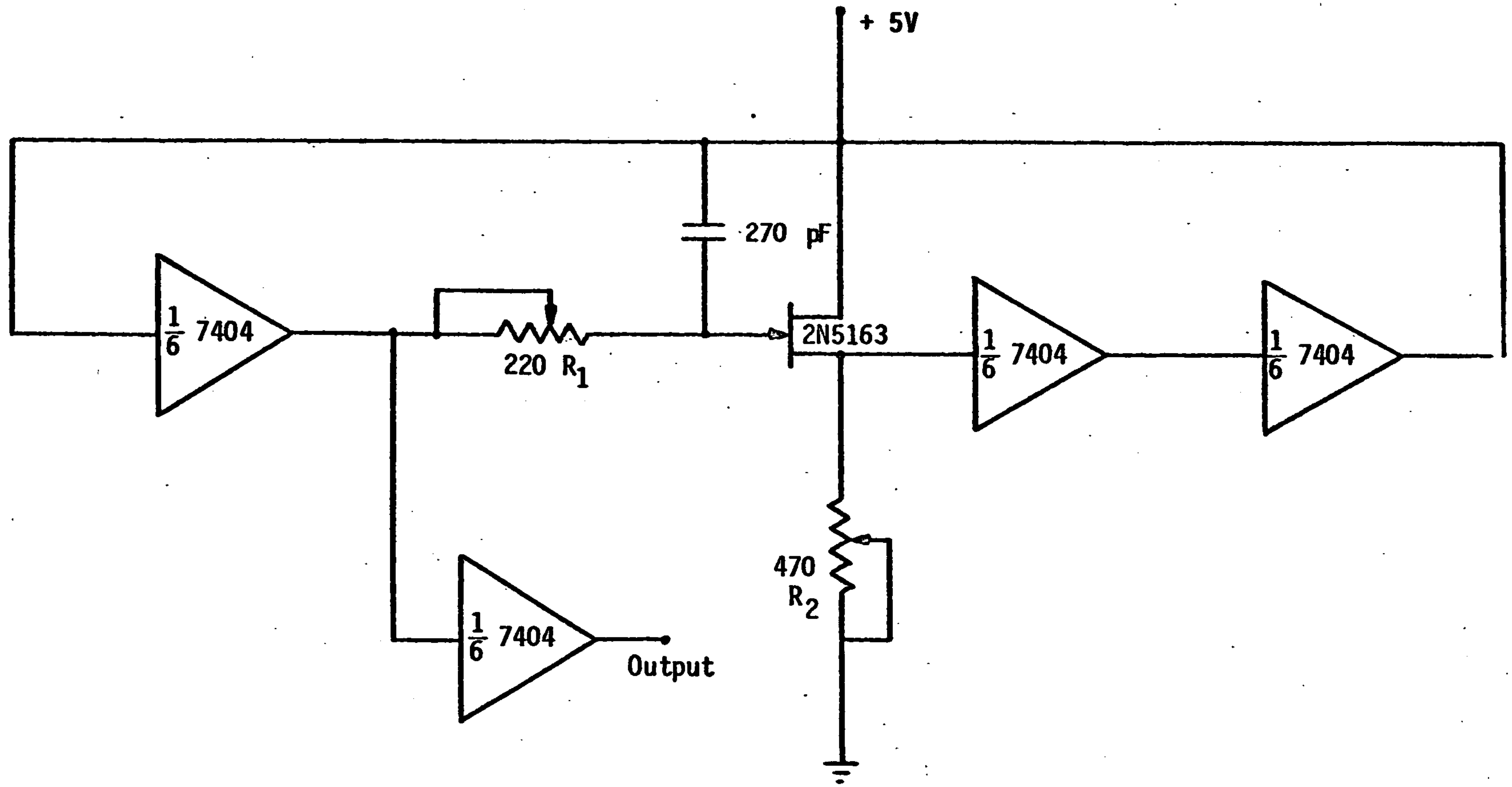


FIG.A.6   5MHz SQUARE WAVE GENERATOR

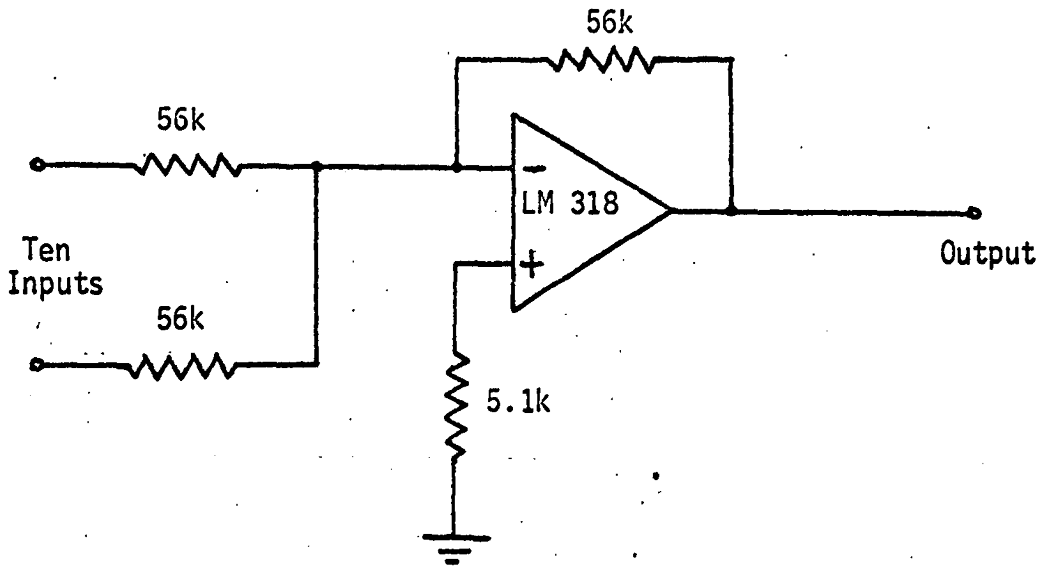


FIG.A.7(a) SUMMING AMPLIFIER

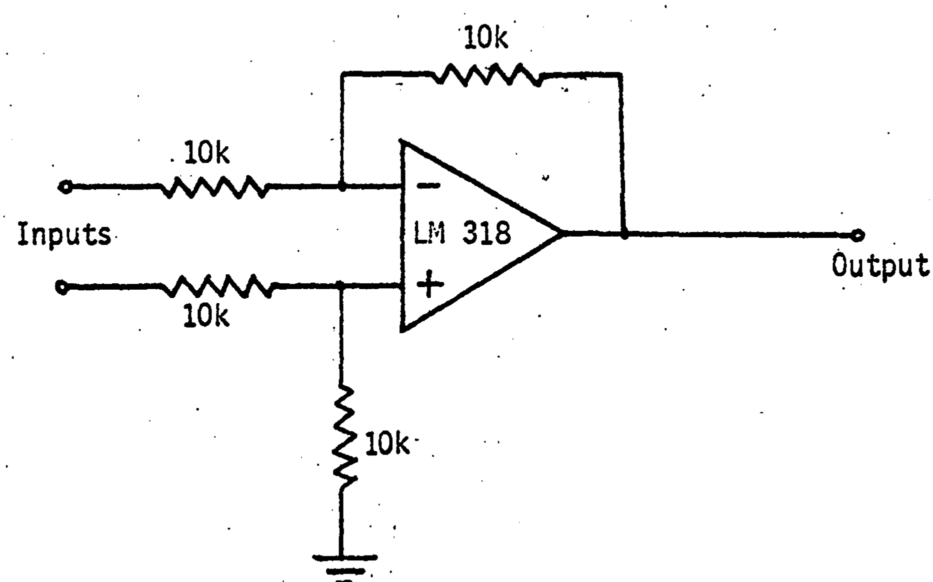


FIG.A.7(b) DIFFERENCE AMPLIFIER

The output circuit is necessary to provide a symmetrical output swing, since the output of the μA 710 is between the levels +3.2 volts and -0.5 volts. This output circuit is powered by an independent adjustable supply, so as to minimise the effects of supply transients within the system. The rest of the clipper circuit is powered by the simple regulator arrangement, which is also shown in Fig.A.5.

A.6. Driving Clock for the P.R.B.S. Generator

The shift registers that generate the pseudo random binary sequences require an external clock to provide the pulses for shifting the contents of the registers. The actual clock frequency to be used is determined by a number of factors, as has been outlined in Section 4.5. In this investigation, a value of 5MHz was chosen.

As a commercially manufactured square wave generator satisfying the frequency requirements was not available during the course of the project, a design given in (50) has been used. The circuit, as shown in Fig.A.6, uses three inverters to provide the gain and regenerative feedback necessary to sustain oscillations. The field effect transistor performs as a source follower. For the frequency chosen, the values of R_1 and C required can be obtained from

$$f_c = \frac{1}{2R_1C}$$

The potentiometer R_2 allows adjustment of the duty cycle of the output pulse train.

A.7. The Remaining Circuits

Fig.A.7 gives the circuit arrangements for the summing and difference amplifiers. In each case, the input and feedback resistors were matched to an accuracy of within 1% of the specified values. The capacitor at the output of the summing amplifier serves to prevent any dc drift of the tap settings from affecting the performance of the correlation circuits. The LM 318 operational amplifier with a much higher slew rate of 50 volts/microseconds was used in place of the μA 749C.

A set of inverting amplifiers was also needed to give the correct feedback polarity for the adaptive loops. A stabilised power supply using two μA 723 precision voltage regulators was also built to provide the ± 15 volts required for most of the circuits.

APPENDIX B

As mentioned in Section 4.6, the original sequence delayed by a desired number of bits can be obtained by modulo-two adding the outputs of the appropriate stages of the original generator. The phase table in this appendix shows which stages are to be so added to give a required delay. The table is obtained by modulo-two adding the outputs of stages 7 and 10 and feeding into stage 8. The delay increments are in steps of 0.2 microseconds (corresponding to a clock pulse rate of 5MHz).

Delay \ Stage	1	2	3	4	5	6	7	8	9	10
0	0	0	0	0	0	0	1	0	0	1
1	1	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	0	0	0	1	0	0
9	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1
11	1	0	0	0	0	0	0	1	0	0
12	0	1	0	0	0	0	0	0	1	0
13	0	0	1	0	0	0	0	0	0	1
14	0	0	0	1	0	0	0	1	0	0
15	0	0	0	0	1	0	0	0	1	0

A '1' indicates that the output of that stage is to be included in the modulo-two addition. For the generator built, a delay of up to 1023 steps can be obtained.

BIBLIOGRAPHY

1. Brown, J.M. "The Picturephone System : Baseband Video Transmission on Loops and Short-Haul Trunks".

   Bell System Tech.J. 50, No.2. (1971), 395-425.

2. Sondhi, M.M. "An Adaptive Echo Canceller".
   Bell System Tech.J. 46, No.3. (1967), 497-511.

3. Kitamori, T. "Applications of Orthogonal Functions to the Determination of Process Dynamic Characteristics and the Construction of Self-Optimising Control Systems".

   Automatic and Remote Control : Proceedings of the First International Congress of I.F.A.C. Edited by J.F. Coales, London : Butterworths, 1961. Vol. II, 613-618.

4. Rosenberger, J.R.; and Thomas, E.J. "Performance of an Adaptive Echo Canceller Operating in a Noisy, Linear, Time-Invariant Environment".

   Bell System Tech.J. 50, No.3 (1971), 785-813.

5. Lucky, R.W. "Automatic Equalisation for Digital Communication".
   Bell System Tech.J. 44, No.4. (1965), 547-588.

6. Lucky, R.W.; and Rudin, H.R. "An Automatic Equaliser for General-Purpose Communication Channels".

   Bell System Tech.J. 46, No.9. (1967), 2179-2208.

7. Rudin, H.R. "A Continuously Adaptive Equaliser for General-Purpose Communication Channels".

   Bell System Tech.J. 48, No.6. (1969), 1865-1883.

8. Gibbs, A.J. "Generalised Mean-Square-Error Minimisation with Application to Automatic and Adaptive Systems".

   Aust. Telecommun. Res. 6, No.2. (1972), 30-38.

9. Gersho, A. "Adaptive Equalisation of Highly Dispersive Channels for Data Transmission".
Bell System Tech.J. 48, No.1. (1969), 55-70.

10. Astrom, K.J.; and Eykhoff, P. "System Identification - A Survey". Automatica 7, No.2. (1971), 123-162.

11. Lytle, D.W. "Convergence Criteria for Transversal Equalisers". Bell System Tech.J. 47, No.8. (1968), 1775-1800.

12. Forsythe, G.E.; and Wasow, W.R.
Finite Difference Methods for Partial Differential Equations.
New York : Wiley & Sons Inc., 1960.

13. Proakis, J.G.; and Miller, J.H. "An Adaptive Receiver for Digital Signalling through Channels with Inter-Symbol Interference".
IEEE Trans. on Inf. Theory. IT-15, No.4. (1969), 484-497.

14. Mark, J.W. "Variable-Gradient Method for Adaptive Equaliser". Electronics Letters. 7, No.21. (1971), 636-638.

15. Chang, R.W. "A New Equaliser Structure for Fast Start-Up Digital Communication".
Bell System Tech.J. 50, No.6. (1971), 1969-2014.

16. Ungerboeck, G. "Theory on the Speed of Convergence in Adaptive Equalisers for Data Transmission".
IBM J. of Res. and Dev. 16, No.6. (1972), 546-555.

17. Gitlin, R.D.; Mazo, J.E.; and Taylor, M.G. "On the Design of Gradient Algorithms for Digitally Implemented Adaptive Filters".
IEEE Trans. on Circuit Theory. CT-20, No.2. (1973), 125-136.

18. De, S.; and Davies, A.C. "Convergence of Adaptive Equaliser for Data Transmission".
Electronics Letters. 6. No.26, (1970), 858-861.

19. Schonfeld, T.J.; and Schwartz, M. "A Rapidly Converging First-Order Training Algorithm for an Adaptive Equaliser".
IEEE Trans. on Inf. Theory. IT-17, No.4. (1971), 431-439.

20. Schonfeld, T.J.; and Schwartz, M. "Rapidly Converging Second-Order Tracking Algorithms for Adaptive Equalisation".
IEEE Trans. on Inf. Theory. IT-17, No.5. (1971), 572-579.

21. Gibbs, A.J.; and Quan, A. "Control Algorithm for Minimum Mean-Square-Error Automatic Systems".
Unpublished.

22. Walzman, T.; and Schwartz, M. "Automatic Equalisation Using the Discrete Frequency Domain".
IEEE Trans. on Inf. Theory. IT-19, No.1. (1973), 59-67.

23. Tomlinson, M. "New Automatic Equaliser Employing Modulo Arithmetic".
Electronics Letters. 7, Nos.5/6. (1971), 138-139.

24. George, D.A.; Brown, R.R.; and Storey, J.R. "An Adaptive Decision Feedback Equaliser".
IEEE Trans. on Commun. Technol. COM-19, No.3. (1971), 281-292.

25. Crater, T.V. "The Picturephone System : Service Standards".
Bell System Tech.J. 50, No.2. (1971), 235-269.

26. Guntersdorfer, S.; Hexberle, W.; and Lueder, R. "Technical Problems of a Video Telephone for the Subscriber Lines of a Public Telephone Network".
Siemens: Rep. on Teleph. Engng. 4, No.1. (1968), 49-56.

27. Ebel, H. "Basic Technical and Economical Aspects of Videophones".
IEEE International Convention Digest 1970, 252-253.

28. Furmage, S.G. "Adaptive Equalisation of Television Systems". Uni. of Tasmania : M.Eng.Sc. Thesis 1972.

29. Clay, R. Non-linear Networks and Systems. New York : Wiley & Sons Inc., 1971, 186-192.

30. Bass, R. "Mathematical Legitimacy of Equivalent Linearisation by Describing Functions".

Automatic and Remote Control : Proceedings of the First International Congress of I.F.A.C. Edited by J.F. Coales, London : Butterworths, 1961. Vol. II, 895-905.

31. Booton, R.C. "Non-linear Control System with Random Inputs". IRE Trans. on Circuit Theory. CT-1 (1954), 9.

32. Kazda, L.F. "Techniques for Analysing Non-linear Control Systems". Adaptive Control Systems : Proceedings of Symposium in Garden City, Long Island, N.Y. Oct. 1960. Edited by Caruthers, F.P.; and Levenstein, H. Pergamon Press : New York, 1963, 21-32.

33. Narendra, K.S.; and McBride, L.E. "Multiparameter Self-Optimising Systems Using Correlation Techniques".

IEEE Trans. on Automatic Control. AC-12, No.1. (1967), 53-59.

34. Newhall, E.E.; Qureshi, S.U.H.; and Simone, C.F. "A Technique for Finding Approximate Inverse Systems and its Application to Equalisation".

IEEE Trans. on Commun. Technol. COM-19, No.6. (1971), 1116-1127.

35. Arnon, E. "An Adaptive Equaliser for Television Channels". IEEE Trans. on Commun. Technol. COM-17, No.6. (1969), 726-734.

36. Cagle, W.B.; Stokes, R.R.; and Wright, B.A. "The Picturephone System: 2C Video Telephone Station Set".

Bell System Tech.J. 50, No.2. (1971), 271-312.

37. Roberts, P.D.; and Davies, R.H. "Statistical Properties of Smooth Maximal Length Linear Binary Sequences".
Proc. IEE. 113, No.1. (1966), 190-196.

38. Chang, R.W.; and Ho, E.Y. "On Fast Start-up Data Communications Systems using Pseudo-random Training Sequences".
Bell System Tech.J. 51, No.9. (1972), 2013-2027.

39. Curry, R.F. "A Method of Obtaining All Phases of a Pseudo-random Sequence".
Proc. of Nat. Electronics Conf. 22 (1966), 518-523.

40. Quan, A.Y.C. "Generation of Fast Pseudo-random Binary Sequences at High Bit Rates".
Aust. Post Off. Res. Lab. Rep. No. 6707 (1972)

41. Thiele, A.N. "Variable Equalisers for Short Video Cables".
Proc. of IRE Aust. 28, No.7. (1967), 215-231.

42. MacGregor, I. "Design of a Variable Attenuation Equaliser and the Associated Impedance Simulating Network for Unloaded Cables".
Aust. Post Off. Res. Lab. Rep. No. 6650 (1971)

43. Sargeant, V.K.; and Deans, N.M. "Variation of Cable Parameters with Temperature, 4 lb PIUT and 2/40 PEIQC". Aust. Post Office Res. Lab. Rep., No. 6671 (1972).

44. McGregor, I "Simulator for Temperature Dependence of Insertion Loss of 2½ mile of 4 lb PIUT". Unpublished work (1973).

45. Barbashin, G. "Construction of Lyapunov Functions for Non-linear Systems". Automatic and Remote Control Proceedings of the First International Congress of I.F.A.C. Edited by J.F. Coales, London: Butterworths, 1961. Vol. II, 1637-1640.

46. Gibson, J.E. Non-linear Automatic Control.
New York: McGraw Hill Book Co. Inc., 1963, 237-290.

47. Wheeler, H.A. "The Interpretation of Amplitude and Phase Distortion in terms of Paired Echoes". Proc. of I.R.E. 27(1939), 359-385.

48. Potter J.B. "A Laboratory VSBSC Data Transmission Facility". Dep. of Electl. Engng. Res. Rep., No. 1/68. Univ. of Melbourne (1968).

49. Sondhi, M.M.; and Presti, A.J. "A Self-Adaptive Echo Canceler" Bell System Tech. J 45, No. 10 (1966), 1851-1854.

50. Sibert, R. "Broadband Pulse Generator Uses Small Timing Capacitance" Electronics 44, No. 17 (1971), 80.