304-699: Research Project

Supervised by Dr. Peter Kabal

ECHO CONTROL TECHNIQUES

IN THE

TELEPHONE NETWORK

presented by

Pedro R. Higuerey Perez

Stud. No. 7720619

This report is submitted to the Department of Electrical Engineering in partial fulfillment of the requirements for the Degree of Master of Engineering (Option B).

McGill University

Department of Electrical Engineering,

October, 1979.

ABSTRACT

Conventional echo suppressors offer inadequate performance on circuits exhibiting long propagation delay and/or low echo return loss (ERL), due to an increased incidence of double-talk and annoying echo. Echo cancellation is an alternative technique, with the potential of overcoming the echo problem on low ERL circuits and maintaining echo control during double talk. This report describes the design of an echo canceller with an adaptive filter. A description of the echo control problems encountered in telephone network is first made, then the echo suppressor device is described and its disadvantages mentioned. Finally, the description of an echo canceller is made. The proposed device uses nonuniform coding which reduces its cost and size. In addition, it works well not only during double talk but also during phase roll, which are among the most adverse circuits conditions an echo canceller may encounter. As an important fact, when satellite trunks are equipped with echo cancellers the quality of the circuit equals that of domestic terrestrial circuits.

Dedicado a Pedro A. y Augusta R.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author would like to express his appreciation and deep gratitude to Professor P. Kabal who acted as a project advisor and provided valuable guidance during the course of this project report.

Special thanks are also due to M. Nunez as well as to all the graduate students and friends who provided moral and technical aid for this project.

The financial assistance of the "Fundacion Gran Mariscal de Ayacucho" during my M. Eng. program is gratefully acknowledged.

Finally, the author wishes to thank Ms. J. Barbot for her typing, in spite of a busy schedule.

CHAPTER 1

INTRODUCTION

1.1 Basic Concepts

In the early 1920's long-distance telephone communications via cable circuits used voice-frequency loaded cables having a low propagation velocity. Only a few hundred miles of this type of cable exhibited delays in the tens of milliseconds.

The two directions of transmission for long-distance communications were then, as now, carried over physically separated cable pairs. This separation of direction is called four-wire transmission. A commercial telephone instrument transmits and receives over the same pair of wires. Although such two-wire transmission is entirely satisfactory for local telephone calls, for long-distance connections it is necessary to convert from two-wire transmission to four-wire transmission. This was, and still is, accomplished by a hybrid transformer circuit.

To prevent energy in the receive direction from entering the send path, the net impedance and the impedance seen on the telephone side of the hybrid circuit must be very closely matched. The latter, however, varies from one telephone connection to another because of such differences as length, wire gauge, and number of telephone extensions. Echo, by definition, is a signal, usually reduced in strength, returning to its source after a time delay. Although reflections also occur on local and short-haul connections, the very short delay (<10 msec.) renders the echo imperceptible. However, on long-haul circuits of over 1500 miles, for example, it can no longer be ignored and therefore requires special attention.

Present echo control measures consist of increasing loss allocation in four-wire circuits up to approximately 1850 miles and applying echo suppressors to longer circuits. In many instances, echo suppressors introduce speech distortion in the form of chopping of parts of words. In other instances, they do not provide adequate echo control on long propagation delay circuits due to the inherent conflict in design requirements for simultaneously blocking echo and passing speech. The requirements become too severe when echo return loss is insufficient. To overcome these problems, echo cancellation methods have been developed. Echo cancellation was proposed in the early sixties; preliminary experiments showed that it was, at least potentially, a superior method of echo control. It first appeared that this technique was not economically feasible for large-scale implementation. However, subsequent developments have proved that economy can be realized without sacrificing the basic advantage of the cancellation technique.

Echo cancellation is a method of echo suppression on telephone lines, which eliminates certain disadvantages of the differential echo suppressors used at present, i.e., the occurrence of speech mutilations and unsuppressed echoes experienced when both parties speak simultaneously.

In order to facilitate further discussion, it is convenient to define the following terms:

a) The echo path is the voice-frequency band-limited path between the receive and send sides of the four-wire circuit. This circuit may consist of several links in tandem and always includes a hybrid circuit and local two-wire telephone connection.

b) The echo return loss (ERL) is the ratio (dB) of the echo signal to the nominal signal level, as experienced in an actual telephone circuit without any form of echo protection.

c) The echo return loss enhancement (ERLE) is the apparent increase in ERL resulting from the action of the echo canceller.

d) The over-all echo return loss (OERL) is simply the sum of the echo return loss and the echo return loss enhancement, i.e.,

$$\mathrm{OERL} = \mathrm{ERL} + \mathrm{ERLE}$$

1.2 Review of the Literature

The problems of echo and propagation delay and the use of conventional echo suppressors in telephone communications have been extensively treated in the literature[1-9].

The problem of delay became acute during the 1920's and 1930's because voice-frequency circuits 500 miles and more in length were being set up for the first time in loaded cable. The propagation speed of these facilities was much slower than that of the open wire used previously, and the delay became sufficient to cause objectionable echoes.

D.L. Richards of the British Post Office performed some of the first investigations in the field of echo-free delay[10]. Richards was interested in the changes in the dynamics of conversations on circuits having appreciable delay. R. Riesz and E. Klemmer at Bell Laboratories approached the problem in a different way[4,8]. They wanted to determine the reactions of users to delay in normal telephone conversations. Krauss and Bricker[11] and Gould[12] performed other tests. In testing for the effects of delay alone, researchers have been restricted to the laboratory; when echo is permitted, studies are also made on regular telephone calls. In the early 1960's, studies were in progress in many companies. These tests showed that the then extant echo suppressors had to be modified before they could be used for telephone calls on long-delay circuits. Moreover, these studies resulted in specifications for new echo suppressors specifically designed for long-delay circuits.

Various experimental versions of these new echo suppressors were used in the subsequent studies on working telephone circuits. Two major studies were conducted by Bell Laboratories in which actual telephone subscribers participated[13]. One of these versions used in synchronous satellite communication was implemented using the technique of a transversal filter to synthesize a replica of the echo[14].

In 1966, Miura, Sato and Nagata[15] of Nippon Electric Company described work on a digitally implemented, impulse response echo canceller, referred to as the "blockless echo suppressor". The method appeared to work quite well, but was limited to time-invariant echo paths. In 1967, Sondhi[16] reported on work performed at Bell Laboratories on an adaptive model of an echo cancellor. Since then, much work have been done subsequently at Bell Laboratories[17-20], at COMSAT Laboratories[21-23], and by workers in Australia[24-25], Germany[26], France[27] and Japan[28]. INTELSAT has recently conducted a world-wide field trial of two echo canceller prototypes, one developed by NEC and the other by COMSAT[22]. Objective observations of their performance on operational satellite circuits have verified that the designs are very robust and are able to cope with all echo-path conditions encountered; including nonlinearities introduced by syllabically companded carrier systems and limited time variance. Subjective evaluations by means of call back interviews have shown that when circuits with low loss and less than average ERL are equipped with echo cancellers they exhibit a significant improvement in quality relative to similar circuits equipped with echo suppressors. In most cases about 25 dB ERLE was provided by the canceller, with greater than 15 dB ERLE attained within the first three words of a conversation.

Bell Laboratories had developed an echo canceller integrated circuit which costs about one-tenth the price of present echo-canceller circuits. These characteristics makes it feasible to be used into the giant telephone network.

1.3 Outline of the Report

Chapter II of this report defines the echo control problems encountered in modern telephone network. The effects of echo on the users of the telephone network and the different parameter to evaluate the telephone conversation are briefly described.

Chapter III describes the echo suppressor. First, the functional description and the characteristics of a simple suppressor are considered. Then, the comparison between the full and the split echo suppressors is made. Finally, the tone disabler section (for the transmission of data) of the echo suppressor is considered.

Chapter IV, addresses the main reason of this project. It examines all the details concerning the design of an echo canceller. First, the mathematical foundation is presented, along with the description of the high-speed digital processing used in the convergence algorithm. The convergence time is found to be faster using non-linear encoding. This type of encodings allows also a reduction in hardware size. Then, different parameters affecting the performance of the echo canceller are considered. Problems such as "phase roll", double talk detector and stability varies their effects accordingly to these parameters. Each of these problems is analyzed separately.

Chapter V contains a summary of the project, as well as some suggestions for improvements in the echo canceller are made.

CHAPTER II

ECHO CONTROL CONSIDERATIONS

When we speak, we like to hear what we are saying, but only if we hear it right away. Studies of subjective reaction to echo in the telephone network have shown that it is difficult to carry on a conversation when one's own voice returns with a delay of more than a few tens of milliseconds. Delayed echoes as much as 40 dB below the outgoing voice level will cause some speakers to characterize the transmission quality as unsatisfactory, particularly since the bulk of our telephone experience is with connections on which distant echoes have been suppressed.

2.1 Source of Delay

"Geostationarity" satellites orbit at an altitude of about 40.000 Kms. Consider an international telephone call via such a satellite. The speech must travel from one telephone through domestic long-distance facilities to the earth station, then up to the satellite, down to the distant earth station, and again through long-distance facilities to the telephone of the party called. The distance between earth stations via the satellite is 83.340 Km; and although signals travel at the speed of light, it takes 300 msec. to reach the distant earth station. Delay in the domestic facilities might add another 20 msec at each end of the circuit producing a total delay of the order of 320 msec. Twice this time must elapse before one party can possibly hear the response of the other.

2.2. Source of Echo

Any impedance mismatch in a transmission system will reflect energy back toward the source and be a potential cause of echo. If there is a mismatch at each end of a transmission line, the energy will be repeatedly reflected back and forth until dissipated by the line. Consider a telephone circuit as shown in Fig. 2.1. If the impedance of the balancing network N is equal at all frequencies to the impedance of the two-wire line, the energy from the incoming four-wire branch is equally divided between the line and the network and the conversion is accomplished without echo. The "balance" in the hybrid circuit is then considered perfect.

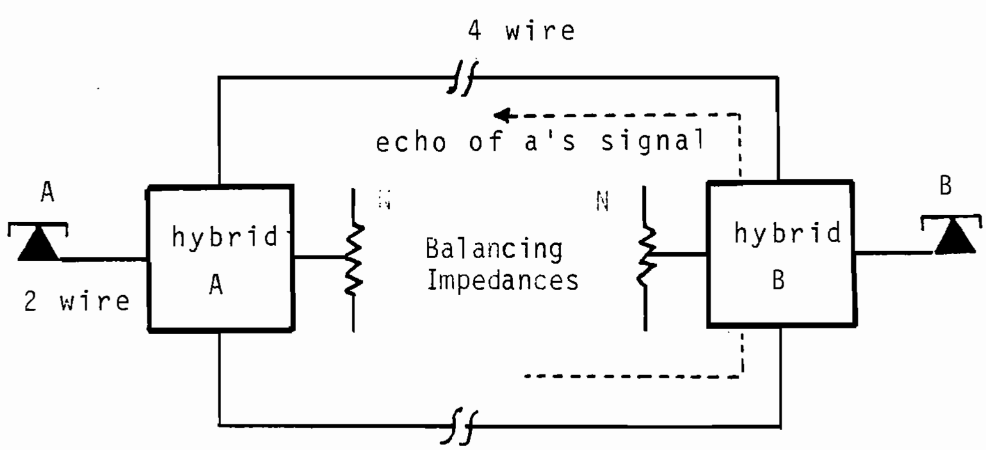

Figure 2.1 : Model for long-distance telephone circuit. The four-wire carrier section may possibly include a satellite link.

If there is any mismatch between the line and network impedances, the balance will be less than perfect and some of the energy will be transferred to the outgoing branch of the four-wire system.

In addition to the lack of impedance matching there will be time delay between the speakers, as mentioned in section 2.1.

It has been found[3] that echo becomes increasingly objectionable as the delay is increased. This is because close-in echoes tend to be masked by sidetone speech, but the masking effect decreases rapidly after speech ceases.

2.3 Effects of Delay and Echo

The disturbing effect of echo on the talker increases with both the level of the echo and with the delay associated with the echo. Since delay is fundamentally a function of circuit connection length and can not easily be reduced and since it is presently uneconomical to improve the loop return losses, the most obvious method of reducing echo effects has been to insert loss increasing with distance. This approach was implemented by the Via Net Loss (VNL) plan which has been applied to communication networks. Naturally, since the loss reduces signal level at the same time as reducing echo, there is a limit to the loss which may be inserted. The VNL plan is applied to trunks in connection for which the round trip delay is 45 msec. or less.

For connections with round trip delays exceeding 45 msec., echo suppressors have been considered essential. Figure 2.2 represents the relation between round-trip delay and the minimum loss required to attenuate the echo sufficiently to provide a non-objectionable echo.

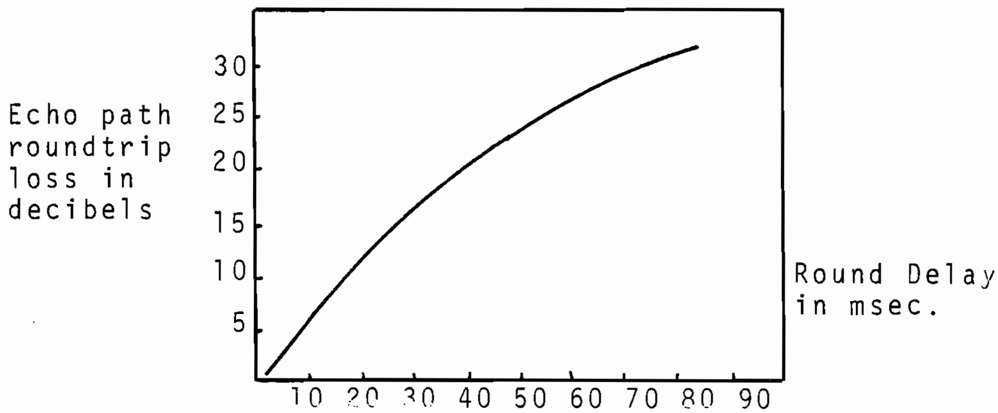

Figure 2.2: Average listener judgement of minimum echo attentuation necessary for just tolerable talker echo conditions on modern telephone sets. (After D.H. Hamsher[31].)

2.4 "Grade of Service" Concept

Figure 2.2 has been obtained using a "grade of service" concept. It has been developed based on extensive experience and testing both in the laboratory and in actual operation. Its purpose is to establish objectives for performance in terms of circuit quality determined by customer opinion within the constraints of the telephone plant. Grade of service is a combination of two imputs. On the one hand, subjective experiments determine the telephone user's reactions to conditions which may be encountered in the telephone circuit. The results of these experiments are in the form of opinion curves for the condition of interest. The other input consists of a survey of the telephone plant which provides distributions of its objective measurable performance. The distribution of the opinion data is combined with the distribution of the plant performance to obtain a grade of service distribution from which a performance objective may be selected. An example of such an objective is a grade of service which results in the rating of "good" or "better" for 95 percent of all telephone circuits. Once this has been established, it is possible to allocate objectives to each of the measurable variables, such as noise, crosstalk, echo and received volume.

Figure 2.3 is an analytical model of opinion that shows how variations in loss, noise and echo in the transmission path affect customer opinion about the quality of transmission. The contours represent the percentage of telephone users who rate telephone service as "good" or "better" for varying values of overall loss and noise.

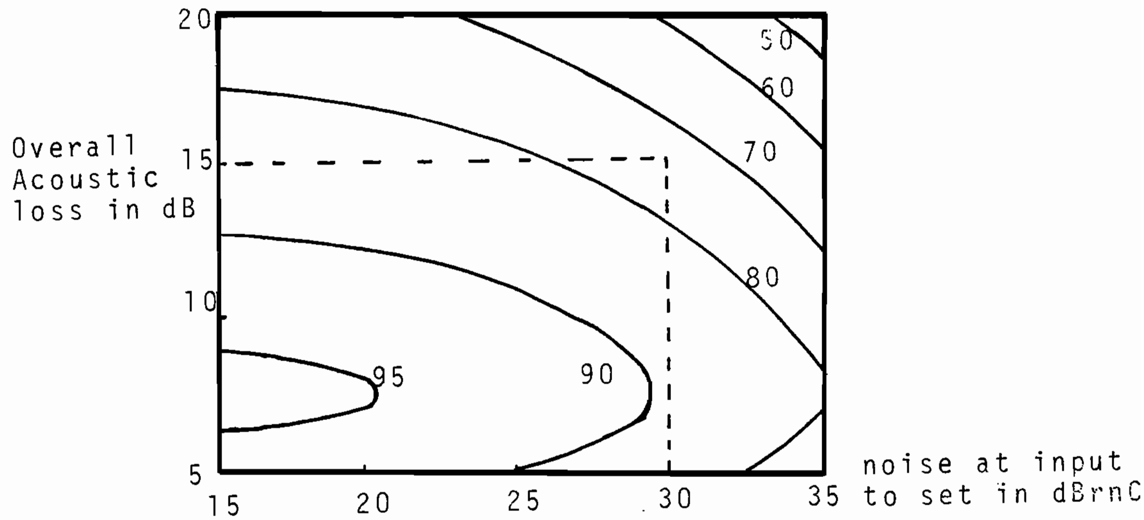

Figure 2.3 : Loss-Noise grade of service . Percent "good" or "better".

This model was obtained by Bell Laboratories using the results of subjective tests. This grade of service model indicates that there is a preferred range for loss, where the received speech is neither too loud nor too soft. When circuit noise was low, opinion on the quality of transmission depended almost completely on the amount of loss. High values of noise reduced opinion ratings for all values of loss. For example, with a loss of 15 dB and noise of 30 dBrnC (dashed lines), 75 percent of telephone customers rate the transmission as "good" or "better". If either loss or noise is increased slightly, this percentage soon drops below 70. Note that the highest customer ratings are for loss close to 7.5 dB and noise held under 20 dBrnC. If loss and noise for the primary signal path are held constant at the values shown by the dashed lines, opinion is then a function of loss and delay in the echo path, as shown in Figure 2.4. For example, for a 30 dB loss and a delay in the echo path of 100 msec., only 20 percent of telephone users find the quality of transmission "good" or "better". Note that this percentage rises only to about 75 percent for even the shortest echo path delays. The reason is that, even when there is no echo delay, only about 75 percent of telephone users consider transmission "good" or "better" at the constant levels of noise and loss assumed for this model. The highest percentage of "good" or "better" ratings occurs when the echo path loss is sufficiently high to make the echo unnoticeable.

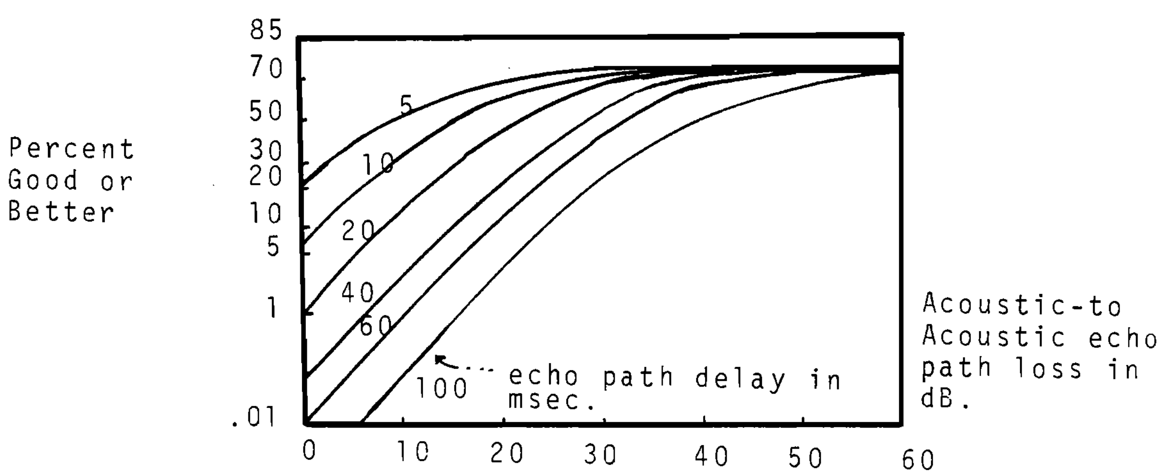

Figure 2.4 : Loss-Noise-Echo grade of service.
Acoustic Loss = 15 dB. Noise at set =30 dBrnC[39].

The effects of echo and propagation delay have probably been responsible for more subjective experiments than any other single parameter in the telephone plant.

2.5 Further Considerations

To provide some additional insight into the complexities of the echo control problem, the interactive effects of propagation delay, sidetone, circuit loss, and circuit noise on subscriber reaction to echo will be briefly considered.

Sidetone is caused by the mouth-to-ear coupling of a telephone handset. Since it involves no delay, it is a degenerate form of echo. However, it is an important parameter for two reasons. First, it gives the telephone user the impression that the circuit is "alive". Tests at Bell Laboratories[32] have indicated that the preferred sidetone level is quite high, i.e., comparable to the volume which the talker would like to receive from the distant party. Hence, it masks the weaker echo signals which occur with very short delay during active speech. The second function of sidetone is to provide a feedback mechanism which assists the talker in controlling the level of his own speech. This directly influences the echo control problem, since the echo level increases as the transmit speech becomes louder and vice-versa. As the delay increases, the telephone user begins to hear a disturbing hollow effect due to echo. For even larger delays, the echo occurs during silent intervals of speech where it becomes quite noticeable and annoying. Further increases in delay may make it impossible to carry on a conversation. Noise tends to mask echo in the telephone system. On a very high-quality circuit, the dynamic range of the human ear permits low-level interference such as echo to be heard. Terrestrial circuit noise is additive and proportional to circuit mileage. This is advantageous in terms of the echo control problem because of the masking effect of noise on low-level echo. However, current international practice tends toward lowering the noise objectives and hence increasing the importance of proper echo control[33]. Circuit loss is also used to control echo. Since echo becomes more apparent with increasing delay, it is reasonable to expect that, as the delay increases, increased loss is required to render echo tolerable. Figure 2.5, taken from C.C.I.T.T. Recommendation G. 131[34], shows overall reference equivalent versus round trip propagation time for three probabilities of encountering objectionable echo. Thus, this information is expressed in terms of the grade of service concept previously discussed.

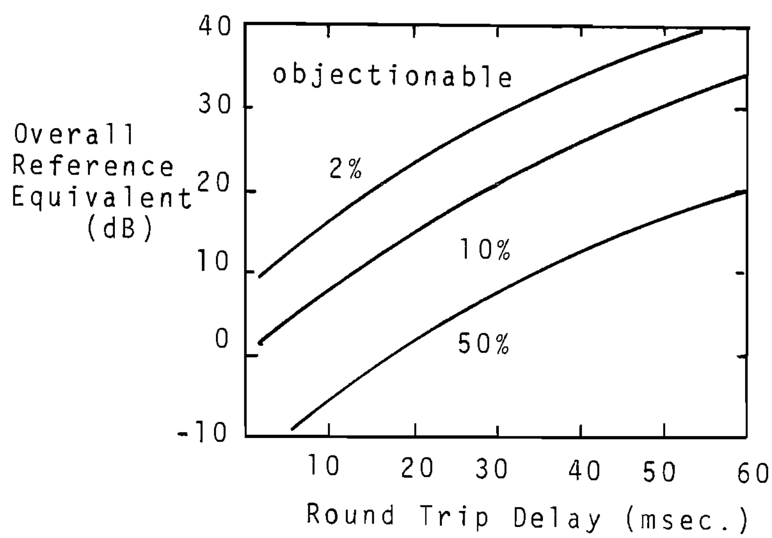

Figure 2.5 : C.C.I.T.T. Echo Tolerance.

CHAPTER III

ECHO SUPPRESSOR

An echo suppressor is a voice-operated device used on long circuits to prevent the echo of reflected speech from annoying the talker. The use of echo suppressors is much more important on long lines than on short lines because the annoyance increases as the time delay of the echo increases.

3.1 Functional Description

Figure 3.1 shows a telephone circuit set up between two speakers, A and B. Ideally, speech from A is channelled, by the hybrid A, through the upper path and then, by the hybrid B, to the telephone B. Speech from B to A is transmitted by the lower path.

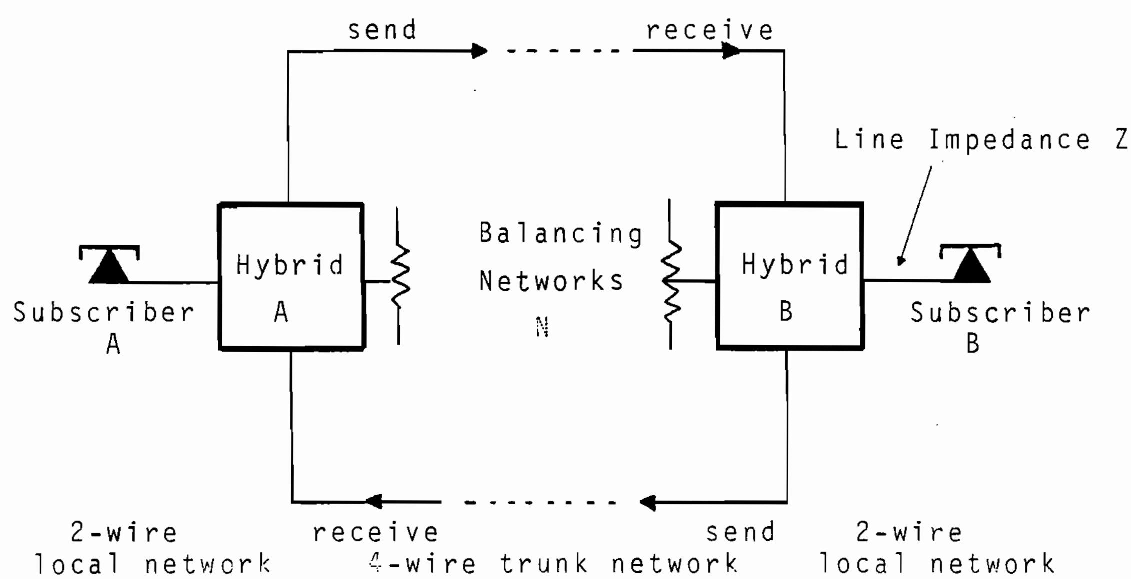

Figure 3.1 : Simplified 2-wire/4 wire long distance connection.

Since the match of network N to the impedance Z of the two-wire line at the hybrid is never perfect, echo is generated at the hybrids. To block the echoes, an echo suppressor, as shown in Figure 3.2, can be installed at each end of the four-wire portion of the circuit. Now if A is talking and B is not, the speech detector opens the switch S and a high loss is inserted in the outgoing path. When B is the only one talking the switch S remains close. However, if both talk at the same time and B's speech power becomes equal to or greater than that of A, the switched loss circuit is removed as a result of comparison of the two signals. Upon removal of the loss, A's echo is permitted to return and become mixed or interspersed with B's speech. Operation of the echo suppressor depends upon a fixed loss, which is switched into the suppressed direction of the four-wire circuit, while the dominant direction maintains control. Inserted loss is normally 50 dB.

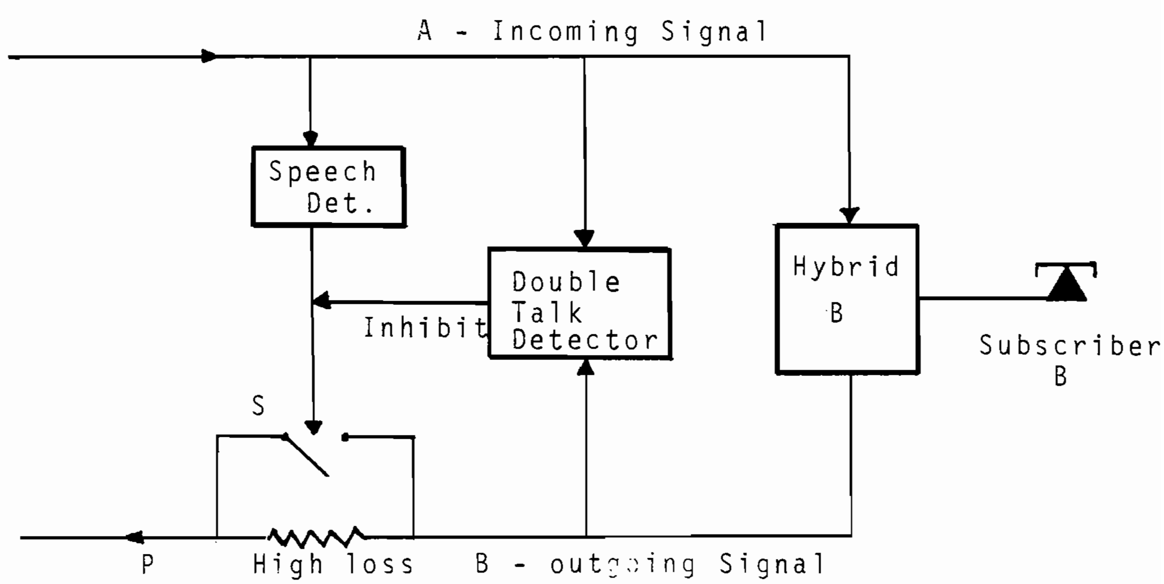

Figure 3.2: Echo Suppressor.

Switching between dominant directions sometimes results in speech mutilations called "chopping" of syllables, and also loud echo bursts can occur when the loss is removed. Echo suppressors are not very satisfactory in long-haul circuits, due to the delay. This extra delay causes confusion, and an increase in "double talking". As the echo suppressor relies on an absence of double talking, its effectiveness is significantly reduced.

3.2. Types of Echo Suppressors

There are two basic types, the full and the split echo suppressor. When one person starts talking, the suppressor shuts off the receiving path, that is, the line to which the talking party listens and which carries the echo. The suppressor turns on this receiving path when the other party begins talking, as explained before.

These operations are the same for both split and full echo suppression, but in split suppression a second suppressor acts on the second party's receiving path. Each suppressor reduces the echo created at the end at which it is located. In full suppression, the single suppressor at one end of the trunk operates, on both paths, suppressing the echo of the person who is talking.

When a full suppressor shuts off the receving path of the party nearer to it, it also cuts off the entire long-distance portion of the path. In so doing, the suppressor turns off the transmission noise normally heard on long-distance trunks. This does not occur on the distant party's path during full or split suppression because the suppressor shuts off the path at the other end of line, leaving the long-distance portion, and the transmission noise, intact. The switching on and off of the transmission noise can be annoying as the echo suppressor continually cuts in and out during a conversation. To remedy this, new suppressors generate noise and insert it into the near receiving path during suppression.

Under the Via Net Loss (VNL) Plan, losses up to 2.9 dB are built into trunks between toll offices. The amount of loss depends on the distance. Losses greater than 2.9 dB are required for trunks of 1850 miles and longer. Echo suppressors are used on all terrestrial trunks over 1850 miles and on all satellite trunks. In addition, if the echo-delay time is less than that for a 3800 miles trunk, a full echo suppressor is used. Full echo suppression is more economical than split echo suppression. The latter, is used on trunks with the longest delay, primarily overseas and satellite trunks. If performance in the full mode is accepted, certain other benefits may be obtained, e.g., easy of planning and maintenance. Echo cancellers can be operated only in the split mode. It will be seen in section 4.7 why they can not be operated in the full mode.

3.3. Tone Disabling Function

The normal operation of the echo suppressor must be inhibited for certain types of data transmission. This is accomplished by the "tone disabler" section of the echo suppressor circuit. A short burst of single-frequency energy, sent over the circuit immediately after the connection is set up, activates the disabler and prevents subsequent operation of the echo suppressor. The echo suppressor is disabled in the presence of a sine wave of approximately 2100 Hz transmitted from a data set.

3.4 Disadvantages of the Echo Suppressor

Since echo suppressors may be physically separated from the terminating sets by an appreciable distance, it is desirable to maintain the suppression loss after the cessation of speech in the receive path. This compensates for the propagation time around the echo path, between the receive-out and send-in ports of the echo suppressor, and is termed the "hangover time". If this time is insufficient, false break-in detection can occur after cessation of receive-path speech and a short burst of echo will be produced. It is also evident that if the ERL is low, the return-echo level can falsely activate the break-in detector and remove the suppression loss. Annoyance caused by these short bursts of echo is aggravated by large propagation delays. Brady and Helder[6] have suggested how the suppressors can disrupt a very interactive conversation. Because of the long delay in the circuit, a quick response by speaker B to something said by speaker A may cause suppression of something said by speaker A at a later time. The deletion is noticed by B, encouraging this speaker to stop and wait for A to get through. The resulting confusion, which may stop conversation entirely while each party waits for the other to say something, is only partially alleviated by restraining suppression during double talking.

Numerous techniques have been suggested to improve the performance of conventional echo suppressors; such as the use of continuously variable suppression loss[6] and adaptive speech-detector thresholds[41].

Bandpass filters have also been used to split the speech signal into several contiguous bands, then normal suppression techniques or center clipping operations[38] are applied separately to each band before reconstituting the signal. Although satisfactory performance under most conditions can usually be achieved, any method based on the suppression technique is inherently incapable of good echo control during double talk. Long propagation delay compounds the problem by causing a high incidence of double-talk. However, considerable effort continues to be put into improving these circuits[42].

CHAPTER IV

ECHO CANCELLER

4.1 Principles of Echo Cancellation

Echo cancellation is a method of echo suppression on telephone lines, which eliminates certain disadvantages of the differential echo suppressors used at present, i.e., the ocurrence of speech mutilations and unsuppressed echoes experienced when both parties speak simultaneously.

To remove the echo signal selectivety on the send path, it is necessary to have a reasonably close replica of the echo signal, which can then be subtracted from the send path signal without otherwise disturbing that path.

For each established connection, the echo path has an unknown transfer function H(f) with an impulse response.

$$h(t) = \int_{-\infty}^{\infty} H(f)\, e^{j2\pi ft}\, df \qquad (4.1)$$

Obtaining and storing an accurate estimate, $\hat{h}(t)$, of this impulse response in digital form makes it possible to construct a close echo signal replica, $\hat{y}(t)$, by digitally convolving the receive side signal, x(t), with $\hat{h}(t)$ over a finite window. To this end, four basic units are employed, as shown in Figure 4.1:

a) A recirculating shift register for storing and moving the receive-side signal samples;

b) A recirculating storage shift register for the model echo path impulse response samples;

c) A convolution processor and

d) An adaptive control loop for obtaining and updating the model echo path impulse response.

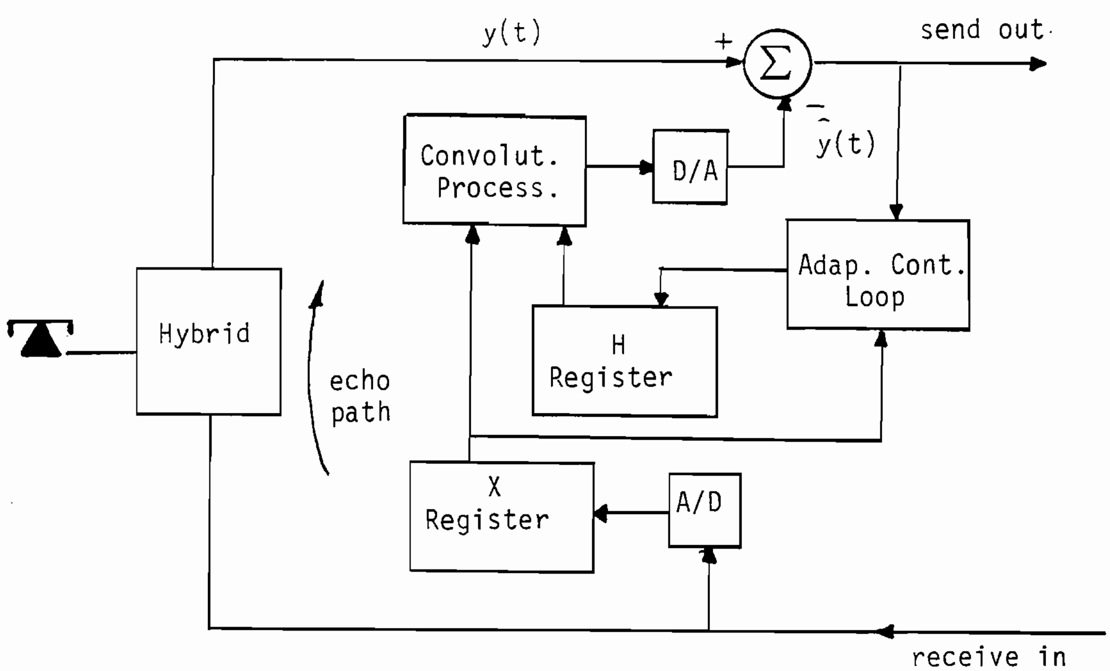

Figure 4.1 : Echo Canceller Block Diagram

The X register stores the most recent of the N receive signal samples, and recycles N + 1 positions every sampling interval so that the oldest sample is dumped and a new sample stored. The H register recyles at the same rate. During each recirculating cycle, each of its N contents is multiplied with the corresponding samples in the X register and the products are summed. The result is a single sample, $\hat{y}$, produced by the convolution processor. This sample is the most recent estimate of the real echo signal, y.

The values of the impulse response samples stored in the H register are obtained as follows. During each recirculating cycle, the correction sensor in the adaptive control loop accepts an error signal, $\varepsilon = y-\hat{y}$, and the $X[(j-i)T]$ samples from the X register to produce a positive, zero, or negative correction to the values stored in the H register. This correction is executed by the update unit. Figure 4.2 shows the movements in correction space for a mythical single tap weight; in the actual case of multiple tap weights, the convex function is a surface in a higher dimensional space.

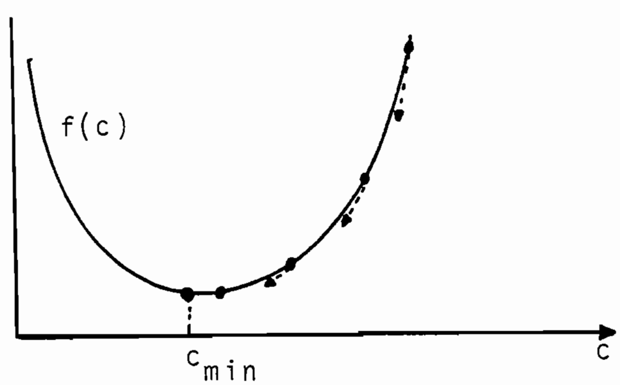

Figure 4.2 : Convex function of a single variable, showing how the minimum value of the function can be reached by a series of movements on the negative gradient direction.

The function performed by the correction sensor may be represented by the product of two functions $\phi_X(X)$ and $\phi_\varepsilon(\varepsilon)$, which are asymmetrical nondecreasing functions. The expression $\phi_X\phi_\varepsilon$ actually represents N products; each product results in a correction to one of the corresponding N values stored in the H register.

As mentionned before, values of ERL larger than 30 dB are required in circuits with delay in excess of 100 msec.; therefore, the echo canceller must attenuate the echo signal power by more than 26 dB under the worst-case condition; i.e., the echo return loss enhancement must be larger than 26 dB. The dynamic range of speech is about 40 dB, which requires 12 bits of linear quantization; and the possible amplitude range of the different impulse responses exceeds 24 dB[25,35] which requires at least 9 bits of linear quantization.

The end delay, t_ε, (time between the echo canceller and the hybrid) is usually less than 15 msec; consequently, the impulse response can be delayed by twice as much, or 30 msec. If the sampling rate is 8 KHz, and approximately 30 samples h_i are used to produce the echo response, at least 250 samples must be stored in the H and X registers for processing.

The echo estimate $\hat{y}_j$ is computed by performing 250 multiplications of 12 by 9 bits and 250 additions of these products during every sampling period.

4.2 Convergence Algorithm

In this section, the derivation of the correction method to obtain the optimum step size is discussed. Two developments are presented: the first is based on "piece wise" solution of the vector equation of the convolution process by using the concept of the pseudo-inverse of a matrix and the second is based on the method of steepest descend.

Representing the actual echo path memory function in discrete form by the row vector

$$\omega = [\omega_1, \omega_2, \ldots, \omega_N] \qquad (4.2)$$

where N is the number of sample points. If the samples are taken every T seconds, then the total duration of the sample window is NT. The simulated echo path memory function at the j^{th} iteration is represented by the row vector

$$\hat{h}_j = [\hat{h}_1, \hat{h}_2, \ldots, \hat{h}_N] \qquad (4.3)$$

The input sequence on the receive side is defined in general as

$$X_1, X_2, \ldots \qquad (4.4)$$

The input sequence involved in the j^{th} iteration in attempts to estimate $\hat{h}_j$ is the row vector

$$X_j = [X_{j-1}, X_{j-2}, \ldots, X_{j-N}] \qquad (4.5)$$

At the time of the j^{th} iteration, the value of the signal appearing at the output of the echo path is

$$y_j = X_j \omega^T \tag{4.6}$$

Where ω^T is the transpose of ω. At the same time, the value of the signal appearing at the output of the simulated echo path is

$$\hat{y}_j = X_j [\hat{h}_j]^T \tag{4.7}$$

The error between y_j and $\hat{y}_j$ at the j^{th} iteration is

$$\varepsilon_j = y_j - \hat{y}_j = X_j [\omega - \hat{h}_j]^T \tag{4.8}$$

A single error is thus expressed as one linear equation with N unknowns, which are the N values of the real impulse response. The information contained in this equation must be utilized to update the values $\hat{h}_j$ in order to minimize ε. Assuming that the updated values of $\hat{h}_j$ are

$$\hat{h}_{j+1} = \hat{h}_j + \Delta\hat{h}_j \tag{4.9}$$

It is wish to select values of Δh_j which minimize

$$y_j - X_j [\hat{h}_{j+1}]^T \tag{4.10}$$

This expression may be expanded as follows

$$y_j - X_j[\hat{h}_{j+1}]^T = y_j - X_j[\hat{h}_j]^T - X_j[\Delta\hat{h}_j]^T \tag{4.11}$$

$$= \varepsilon_j - X_j[\Delta\hat{h}_j]^T \tag{4.12}$$

The solution to this equation is, as shown by Greville[36].

$$\Delta\hat{h}_j = \underbrace{\left[X_j\ X_j^T\right]^{-1} \cdot X_j}_{A}\varepsilon_j. \tag{4.13}$$

The expression for A is

$$A = \frac{X_j}{\sum_i X^2_{j-i}} \tag{4.14}$$

Thus, the final solution for updating the $\hat{h}_{j+1}$ coefficients in the H register is

$$\Delta\hat{h}_j = \frac{X_j\ .\ \varepsilon_j}{\sum_i X^2_{j-i}} \tag{4.15}$$

It may be shown that this is also a special case of convergence by the steepest descend method.

In order to demonstrate this, the steepest descend method is derived as follows. Consider a positive even function, ϕ, of the error signal, ε, such that

$$\frac{d\ \phi(\varepsilon)}{d\varepsilon} = \phi_\varepsilon\ (\varepsilon) \tag{4.16}$$

is a monotonically nondecreasing odd function, and $d\phi_\varepsilon(\varepsilon)\ /\ d\varepsilon \geq 0$. These properties are required to ensure convergence of the process. For steepest descend, the increasing vector, Δh_j, must be proportional to the negative gradient of the function ϕ with respect to the values $\hat{h}_j$. Thus,

$$\Delta\hat{h}_j = -C\ \mathrm{grad}_h\ \phi(\varepsilon_j) \tag{4.17a}$$

$$= -C\ \mathrm{grad}_h\ \phi\left\{X_j W^T - X_j\ [\hat{h}_j]^T\right\} \tag{4.17b}$$

$$= C\ X_j\ \phi_\varepsilon\ \left\{X_j W^T - X_j\ [\hat{h}_j]^T\right\} \tag{4.17c}$$

$$= C\ X_j\ \phi_\varepsilon\ (\varepsilon_j) \tag{4.17d}$$

By letting $C = (X_j X_j^T)^{-1}$ and $\phi_\varepsilon(\varepsilon_j) = \varepsilon_j$, the same result as in equation (4.15) is obtained. The updating equation can alternately be written as

$$\hat{h}_{j+1} = \hat{h}_j + C\ X_j\ \phi_\varepsilon(\varepsilon_j) \tag{4.18}$$

Where C>0, and ϕ_ε fulfills equation (4.16). The choice of the functions $\phi(\varepsilon)$ and $\phi_\varepsilon(\varepsilon)$ critically affect the speed and accuracy of convergence and the complexity of the hardware implementation. The criterion commonly utilized in analogue implementation is the square-value criterion, $\phi(\varepsilon) = \varepsilon^2$. For digital implementation the absolute value criterion is used, often in the following comprise form.

$$\phi_\varepsilon(\varepsilon) = \phi_L(\varepsilon) \tag{4.19}$$

with

$$\phi_L(\varepsilon) = \begin{cases} \mathrm{sgn}(\varepsilon) & |\varepsilon| \geq L \\ 0 & |\varepsilon| < L \end{cases} \tag{4.20}$$

The small central "null zone" provides the adaptation process with immunity to low-level noise, but prevents reduction of the error sequence below the threshold L. A center clipper (set to operate at the level L) is sometimes inserted in the send-out path to remove this residual low-level noise. It is normally disabled while near-end speech is detected.

4.3 Convergence Time

The time for convergence is largely determined by the size of the correction step, $\Delta\hat{h}_j$, which is made to the coefficients in the H register. For a first approximation, it is assumed the following conditions at the start of convergence:

a) The X register is fully loaded with the N most recent samples of the received input signal,

b) The H register is cleared so that all N values are 0,

c) The true echo path is characterized by a unit pulse which occupies the full range of the H register (a most pessimistic assumption),

d) The H register values are linearly quantized using Q bits, and

e) The corrections are of magnitude $\Delta h = 2^u$ (using the u^{th} low-order bit).

Based on these conditions, it will take 2^{Q-u} iterations to reach convergence. Thus, the time to converge will be

$$T_c = T2^{Q-u} \tag{4.21}$$

This estimate is based on the assumption that no errors are made in the decisions for correcting the H register. Since some erroneous corrections will occur, increasing the time for convergence when the error, ε, approaches its smallest threshold, this assumption is true only for large values of error.

In the canceller described here, u (and thus Δh) is made a function of the magnitude of the error expressed by the updating equation; i.e., a large correction is made for a large error and correspondingly smaller corrections for smaller errors. This results in markedly less time to convergence than was predicted above.

4.4 Coding and Approximations in the Convolution Processor

4.4.1 Non-Linear Coding Format

A codec whose transfer characteristic has uniform step sizes is not usually the best choice for message signals for several reasons. First, the amplitude distribution of the message is not uniform. For a given talker volume, smaller amplitudes are more probable than larger amplitudes, and thus a better S/D ratio can be expected if the error characteristics is made smaller for the more probable amplitudes at the expense of larger errors for the less probable amplitudes. Second, as mentioned in section 4.1, speech signals have a dynamic range of up to 40 dB. For a uniform codec, weak signals will experience 40 dB poorer S/D ratio than strong signals. Third, in section 4.3 it was seen that the time of convergence depends on the number of bits; for less number of bits the convergence is faster. Further, non-linear coding are preferred because of the reduction in the hardware complexity and in the last years they have been used for input signal format so that the echo canceller is compatible with the TDM systems. The relationship between a uniform coder and an equivalent non-uniform coder, both having the same dynamic range is commonly known as the companding advantage.

Companding requirements are different for different signal distributions. For example, voice signals require constant S/D performance over a wide dynamic range, which means that the distortion must be proportional to signal amplitude for any signal level. To achieve this, a logarithmic compression law must be used. Of course a truly logarithmic assignment of code words is impossible because it implies both an infinite dynamic range and an infinite number of codes. Two methods for modifying the true logarithmic function have been used. In the first, called the u-law, for the normalized coding range of ±1,

$$F(x) = \operatorname{sgn}(x) \frac{\ln(1 + u|x|)}{\ln(1 + u)} \; ; \; -1 < x < 1 \qquad (4.22)$$

The second method of approximating the true logarithmic law is to substitute a linear segment to the logarithmic curve for small signals. It is called the A-law and is represented by:

$$F(x) = \operatorname{sgn}(x) . \begin{cases} \dfrac{A|x|}{1 + \ln A} \; ; & 0 \leq x \leq \dfrac{1}{A} \\[2ex] \dfrac{1 + \ln(A|x|)}{1 + \ln A} \; ; & \dfrac{1}{A} < x \leq 1 \end{cases} \qquad (4.23)$$

Both laws tend to produce flat S/D ratios over a wide dynamic range. Both are digitally linearizable and have a binary relationship between step size on adjacent segments, and both are segmented piece-wise linear approximations to the original laws. The segmented version of the u-law uses 15 segments with u = 255; a definition for positive values is given in Table I.

Table I

Segment	Input Level	Output	Interval Size
1	-31, 31	P111ABCD	2
2	31, 95	P110ABCD	4
3	95, 223	P101ABCD	8
4	223, 479	P100ABCD	16
5	479, 991	P011ABCD	32
6	991, 2015	P010ABCD	64
7	2015, 4063	P001ABCD	128
8	4063, 8159	P000ABCD	256

Encoding the samples in a pseudo-logarithmic format yields a reduction in the memory space. In addition, this kind of format converts the multiplications of the convolution process (4.7) to additions; with a consequent reduction in the number of components of the convolution processor. An additional advantage of this format is that the signal to quantization noise ratio (S/N) is nearly constant over a dynamic range of 30 dB, as shown in Figure 4.3. This type of encoding is less sensitive to the "phase roll" problem because of its faster convergence.

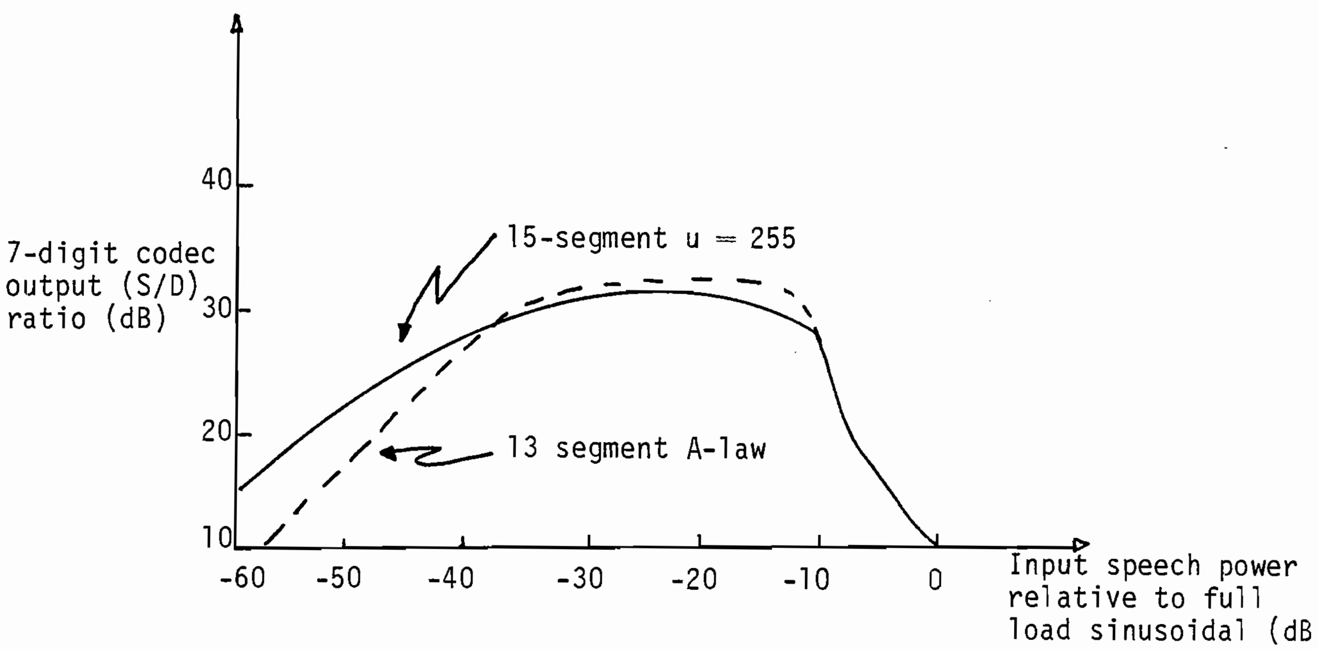

Figure 4.3: Signal-to-distortion performance of two digitally linearizable laws.

4.4.2 Approximations in the Convolution Processor

The function of the convolution processor is shown schematically in Figure 4.4. The n samples h_i of the echo response H (t) of the Dirac Impulse $\delta(t)$ are stored in the H register. The X-registers store samples x_{i+j} of signal waveform x(t). The number of stored samples, n, must be greater than

$$n > \frac{2t_\varepsilon}{T_s} \qquad (4.24)$$

where t_ε is the end delay and T_s is the sampling frequency. Because every sample x_{i+j} can be considered to be a Dirac pulse with amplitude X_{i+j} at time $(i+j)T_s$, the circuit response to every X_{i+j} is the product $X_{i+j} \cdot H(i)$, as shown in Figure 4.4. for samples X_0 to X_5. At every instant t_j, according to the superposition theorem, response r_j is the sum of the responses to individual sampes X_{i+j}

$$r_j = \sum_{i=o}^{m} h_i \cdot X_{i+j} = H(i) * X(i) \qquad (4.25)$$

For example, at time t_1,

$$r_1 = h_o X_1 + h_1 X_2 + h_2 X_3 \qquad (4.26)$$

The sequence of samples r_j represents the resulting response R(t),

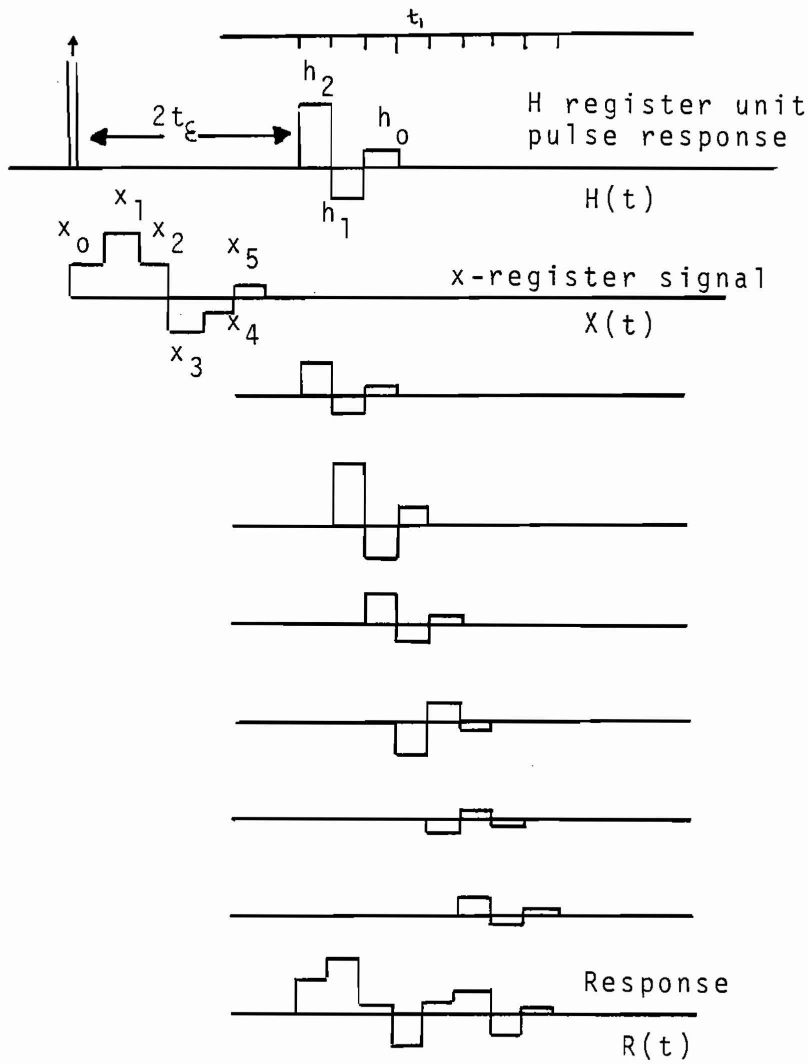

Figure 4.4 Schematic presentation of waveforms X(t) and H(t) entering the convolution processor and the formation of resulting response R(t).

i.e., the simulated echo of signal X(t) in the circuit with echo impulse response H(t).

Horna[23], in his echo canceller developed at Comsat Laboratories, uses analog-to-digital conversion performed directly into a seven bit A-law code by inexpensive hardware utilizing three stages of comparators for sign, exponent, and mantissa. This project follows the same procedure of Horna but using eight bit u-law format for the speech samples X_{i+j} and seven bit u-law format for the impulse response samples h_j. The encoding of the speech samples uses a 15 segment approximation which yields the resolution and dynamic range of an 11 bit linear code and the encoding of the impulse response samples uses a 13 segment approximation with a dynamic range equivalent to that of a 10 bit linear code. The 8 bit u-law code word consists of a polarity digit P followed by three segment-defining digits XYZ followed by four digits ABCD which specify the intrasegment value on a linear scale of 16 values. The convolution processor utilizes three stages of comparators for sign, exponent, and mantissa. Assuming unity digital-to-analog scaling factors the convolution processor functions as follows:

a) When the exponents e_x and e_h are zero, the samples are in the linear segment of the chosen code and their equivalent absolute values are

$$\begin{aligned} |x| &= 2 \cdot m_x \\ |h| &= 2 \cdot m_h \end{aligned} \qquad (4.27)$$

b) When the exponents are larger than zero the samples are encoded quasi-logarithmically. The equivalent values of the samples are

$$|x| = 2^{e_x} (1 + m_x)$$
$$|h| = 2^{e_h} (1 + m_h) \tag{4.28}$$

As there are four possible combinations of $|x|$ and $|h|$ formats entering the convolution processor's multiplier, four different algorithms must be used to perform the product P_i. If the exponents are zero the equivalent values of the samples are only three bits wide and the product is,

$$|P_i| = 2^2 \cdot m_x \cdot m_h \tag{4.29}$$

Product $|P_i|$ requires four bits resolution and can be founded in a 256 bit ROM.

If either e_x or e_h is larger or equal to one, then the product is

$$|P_i| = 2^{e_x}(1 + m_x) \times 2 \times m_h = 2^{e_x+1} (m_h + m_x \times m_h) \tag{4.30}$$

In this case, the 4-bit product $m_x \times m_h$ is read from the ROM and added to the mantissa with the zero exponent. The sum in brackets is less than 2 and less or equal to 5 bits wide; therefore, a simple 4-bit MSI adder can be used.

If both exponents are larger or equal to one an approximate expression for the binary logarithms is used in computing $|P_i|$. The following approximations are used

$$\begin{aligned} \log_2 (1 + m) &\doteq m \\ \text{antilog}_2\, m = 2^m &\doteq 1 + m \end{aligned} \qquad (4.31)$$

Using this approximation, the $\log_2$ of the product $|P_i|$ is therefore

$$\log_2 |P_i| = e_x + \log_2 (1 + m_x) + e_h + \log_2 (1 + m_h)$$

$$= e_x + e_h + m_x + m_h = E_i + M_i \qquad (4.32)$$

where E_i is the whole part of the sum and M_i is the remainder <1. Product $|P_i|$ is thus

$$|P_i| = \text{antilog}_2 (E_i + M_i) = 2^{E_i} (1 + M_i + \varepsilon_m) \quad (4.33)$$

Where ε_m is a correction which is a function of m_x, m_h, and M_i. Equation (4.32) can be implemented by a 8-bit adder. Multiplication by 2^{E_i} according to equation (4.33), where E_i is an integer

E_i $\varepsilon\{2, 3, \ldots 14\}$, is performed by shifting the partial products by E_i binary places before entering the acumulator.

In addition to circuit simplification, this coding also provides: a) a simple solution to the problem of zero values associated with logarithmic encoding and b) since the multiplication is performed by one read command from the ROM and by a single addition, the entire operation of forming product $|P_i|$ and adding it into the accumulator can be done in one clock period, which may be as short as 150 msec.

4.5 Cancellation at Convergence

There are three competing mechanisms which affect the magnitude of error when the system has attained convergence. The first is the minimum error threshold below which no corrections are allowed to be made to the impulse response model, i.e., the region $-\Delta\varepsilon < \varepsilon < \Delta\varepsilon$. This region will be referred to as an inert zone. Since the error within this region is equally likely to assume any value, the error distribution is uniform with zero mean. Hence, in this case the pdf for ε is

$$P_1(\varepsilon) = \begin{cases} \dfrac{1}{2\Delta\varepsilon} & ; \ -\Delta\varepsilon < \varepsilon < \Delta\varepsilon \\ 0 & ; \ \text{otherwise} \end{cases} \tag{4.34}$$

The error of variance is thus,

$$\bar{\varepsilon}^2 = \int_{-\infty}^{\infty} \varepsilon^2 P_1(\varepsilon)\, d\varepsilon = \int_{-\Delta\varepsilon}^{\Delta\varepsilon} \frac{\varepsilon^2}{2\Delta\varepsilon}\, d\varepsilon = \frac{\Delta\varepsilon^2}{3} \tag{4.35}$$

The second error mechanism is attributed to the process of impulse response correction. More specifically, there is a random component in the echo signal produced by the modeled impulse response. This component is caused by the effects of quantization of the $h(iT)$, $X\left[(j-i)T\right]$, and $\varepsilon(jT)$ values, and by truncation of the impulse response. In addition, a circuit noise component introduced in the real echo path contributes to the magnitude of the random component of the error signal. The net result of these errors is a random variation in the estimate of the actual send path signal which is superimposed on the error signal, ε . The distribution of this variation is designated as $P_m(\varepsilon)$ in Figure 4.5 where it is shown in relationship to the inert zone.

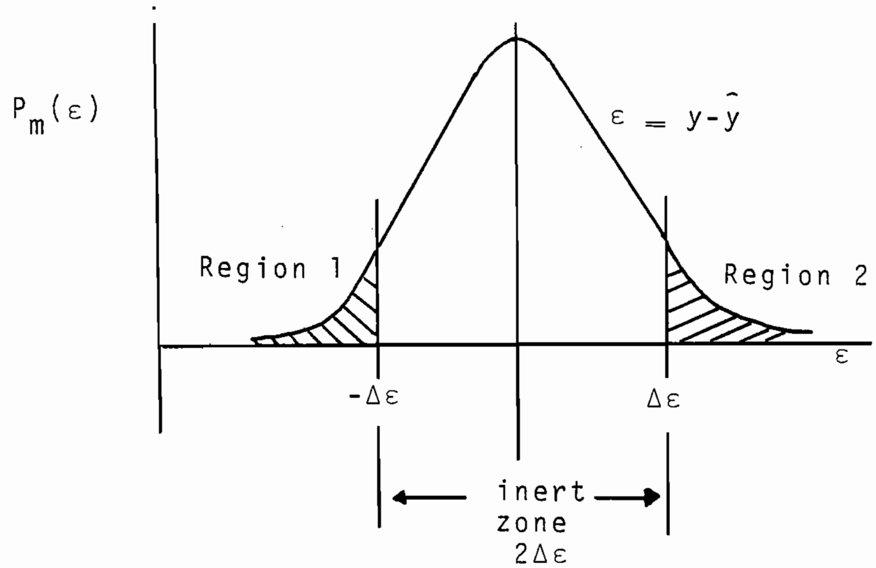

Figure 4.5: Distribution of errors.

If the error exceeds the upper bound, a correction of magnitude - ΔC will result from a change in the impulse response sample. Alternately, if the error is less than the lower bound, a correction of magnitude + ΔC is applied.

If the error lies inside the inert zone, no correction is applied. The probability of each of these types of correction is given by the area under the distribution $P_m(\varepsilon)$. Thus, the distribution of the corrections is

$$P_1 = P(C = \Delta C) = \int_{-\infty}^{-\Delta\varepsilon} P_m(\varepsilon)\, d\varepsilon \qquad (4.36a)$$

$$P_2 = P(C = 0) = \int_{-\Delta\varepsilon}^{\Delta\varepsilon} P_m(\varepsilon)\, d\varepsilon \qquad (4.36b)$$

$$P_3 = P(C = -\Delta C) = \int_{\Delta\varepsilon}^{\infty} P_m(\varepsilon)\, d\varepsilon \qquad (4.36c)$$

Where the error distribution $P_m(\varepsilon)$ is assumed to be Gaussian, i.e.,

$$P_m(\varepsilon) = \frac{1}{\sqrt{2\pi\,\sigma_m^2}}\, e^{-(\frac{1}{2})\,(\varepsilon^2/\sigma_m^2)} \qquad (4.37)$$

The process of correction also contributes noise; the expected value of this noise is

$$\overline{e^2} = \Sigma\; e^2\, P_i = \Delta C^2 \left[1 - 2 \int_0^{\Delta\varepsilon} P_m(\varepsilon)\, d\varepsilon\right] \qquad (4.38)$$

The third mechanism of error is the error due to the $\log_2$ approximation, ε_m.

The error, $\overline{\varepsilon^2}$, due to the width of the inert zone, increases with increasing $\Delta\varepsilon$, while the error, $\overline{e^2}$, due to the correction, increases with decreasing $\Delta\varepsilon$. Thus if all other factors are constant, there will be a value of $\Delta\varepsilon$ for which the error is minimum. If it is assumed that the three mechanisms of error are independent, then, the total error is given by

$$\bar{e}_T^2 = \bar{e}^2 - \bar{\varepsilon}^2 - \bar{\varepsilon}_m^2 \tag{4.39}$$

It is assumed that the principal contributor to e_T^2 is the quantizing error. This is a reasonable assumption, since, during normal operation, the echo canceller will experience a relatively low noise level on the send side (-50 dBm0), and during double talk, the feedback correction process will be disabled. Truncation of the impulse response is unlikely to cause a significant contribution to e_T^2 if sufficient storage time is allowed in the registers. Under these conditions, the degree of cancellation is determined mainly by the quantization errors Δ_h and Δ_x of samples h_i and X_{i+j}, respectively, entering the convolution processor[25]. As shown in equation (4.7), the j^{th} estimate of echo y_j is computed as a sum of products $P_i = h_i \cdot X_{i+j}$:

$$y_j = \sum_{i=1}^{m} P_i = \sum_{i=1}^{m} h_i \cdot X_{i+j} \tag{4.40}$$

If the process is stationary and samples h_i of stored response H_i converge to their optimal values, every product P_i will contain only a quantization error due to finite quantization steps Δ_h and Δ_x of samples h_i and X_{i+j} :

$$\begin{aligned} P_i + \Delta P_i &= (h_i + \Delta_h)(X_{i+j} + \Delta_x) \\ &= h_i X_{i+j} + \Delta_h X_{i+j} + \Delta_x h_i + \Delta_h \Delta_x \end{aligned} \tag{4.41}$$

Neglecting the last term in a first approximation leads to the

the quantization error

$$\Delta P_i = \Delta_h \; X_{i+j} + \Delta_x \; h_i \tag{4.42}$$

Because Δ_h and Δ_x are uncorrelated the variance σ_i^2 is

$$\sigma_i^2 = E\left[(\Delta P_i)^2\right] = \frac{\Delta_h}{12} \; X_{i+j}^2 + \frac{\Delta_x^2}{12} \; h_i^2 \tag{4.43}$$

Encoding both samples in a pseudo-logarithmic format except for the short linear range around zero the quantization steps Δ_h and Δ_x are approximately a constant fraction $\delta << 1$ of sample h_i or X_{i+j}. Therefore,

$$\Delta_h \doteq \delta \cdot |h_i| \tag{4.44a}$$

$$\Delta_x \doteq \delta \cdot |X_{i+j}| \tag{4.44b}$$

Which yields

$$\sigma_i^2 = \frac{\delta^2}{6} \; h_i^2 \; X_{i+j}^2 \tag{4.45}$$

and the expected value of the error ε_j^2 of all products P_i is

$$E\left[\varepsilon_j^2\right] = \sum_{i=1}^{m} \sigma_i^2 = \frac{\sigma^2}{6} \sum_{i=1}^{m} h_i^2 \; X_{i+j}^2 \tag{4.46}$$

which is proportional to the total power of the estimated echo signal $P(\hat{y}_j) = \hat{y}_j^2 / R$.

For small amplitudes $|h_i|$ and $|X_{i+j}|$, the equation (4.44) does not hold because the encoding is linear. That is, the quantization error increases with decreasing $|P_i|$. For this reason, the ERLE is decreased for low-level receive-in signal.

The relative errors of the $\log_2$ and antilog_2 approximations are

$$\varepsilon_e = \log_2 (1+m) - m = \phi(m) \qquad (4.47a)$$

$$\varepsilon_a = 2^m - (1+m) = \phi_a(m) \qquad (4.47b)$$

Where for given m the errors are approximately related, i.e.,

$$\varepsilon_a = -\varepsilon_e \longrightarrow \phi_a(m) \doteq -\phi(m) \qquad (4.48)$$

Figure 4.6 shows $\phi(m)$ as a function of m. It is advantageous in the convolution process to zero the average relative error, u, i.e., to add a constant term $\varepsilon_u = 0.0573$ to the mantissa m of $\log_2$ or to subtract it when computing 2^m.

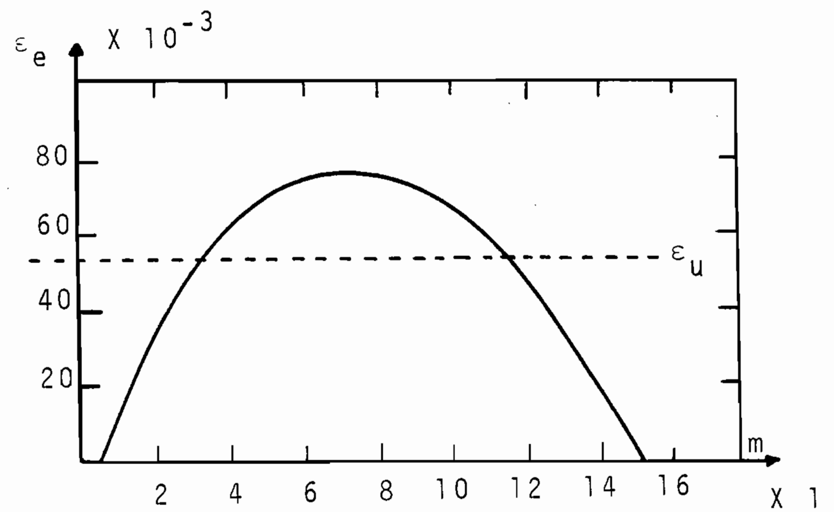

Figure 4.6 Error ε_e of the $\log_2$ approximat. as a function $\phi(m)$ of mantissa $0 < m < 1$.

For binary product $|P_i|$ according to equation (4.31) - (4.33), a correction ε_m, which is the binary representation of ε_u, is applied.

This rms value of the relative error of the $\log_2$ or antilog_2 approximation is

$$\varepsilon_{rms} = \left[\int_0^1 \left[\phi(m) - \varepsilon_u\right]^2 \, dm\right]^{\frac{1}{2}} = 0.027 \qquad (4.49)$$

which is also equal to the standard deviation, σ, because the average has been adjusted for zero.

4.6 The Control Algorithm

There are several approaches to adjust the tap weights of the digital filter to the optimum values. The equation (4.9) based on the least mean square is probably the most used. Its control algorithm is described as follows:

Step 1. During the i^{th} interval, a set of samples from X(t) and y(t) waveforms are taken.

Step 2. The coefficients stored at the end of the last interval $\hat{h}_j$, are convolved with the X data, then added to the stored Y data.

Step 3. Next, coefficients increments are calculated and added to the previously stored coefficients.

Step 4. The stored ε data is also updated using the same coefficients increments, to further reduce the echo component. This is equivalent to repeating step 2 using the most recent estimate of the optimum coefficients $\hat{h}_j$.

Step 5. Step 3 and step 4 are repeated many times, successively reducing the increment size until the end of the interval.

Step 6. Over the interval i+1 the samples ε_i of the interval i are sent at the Nyquist rate. These were calculated during interval i from samples collected during the interval i-1.

The correction to the tap weights in the H-register is as follows: If $e_h = 0$, i.e. h_j is in the linear segment of the code around zero, the correction formula is the same as that given by equation (4.9).

$$h_{j+1} = h_j + \Delta h_j = \mathrm{Sgn}(h_j) \times |2m_h| + \Delta h_j \qquad (4.50)$$

For exponents $e_h \geq 1$, the sample h_j is in non-linear format:

$$\log_2 h_j = e_h + m_h = \log_2 2^{e_h} (1 + m_h) \qquad (4.51)$$

The increment $\Delta_{i,j} \ll 1$ is added to the mantissa m_h of $\log_2 |h_j|$ to establish the new corrected value h_{j+1}. In linear format this is equivalent to the product

$$|h_{j+1}| = |h_j| \, (1 + \Delta_{i,j}) = |h_j| + |h_j| \Delta_{i,j} \qquad (4.52)$$

Thus, the sample h_j is multiplied by $(1 \pm \Delta_{i,j})$. In effect, the correction step Δh_j is directly proportional to the amplitude of the sample h_j :

$$|\Delta h_j| = |h_j| \, \Delta_{i,j} \qquad (4.53)$$

With this correction, the time of convergence is less dependent on the amplitude of the input response $H(i)$ [16,21]. For large values of h_j, the convergence is nearly eight times faster than with the correction method of equation (4.9), which adds the same correction to each value h_j.

4.7 Stability Criterion

The stability of the correction loop is ensured when absolute values of errors ε_j and ε_j' before and after the corrections of the response $H(j)$, respectively, fulfill the inequality

$$|\varepsilon_j'| < |\varepsilon_j| \tag{4.54}$$

In this section, the conditions necessary to satisfy equation (4.53) are derived.

In the echo canceller several hundred of samples X_{i+j} are stored and only 10-30 significantly participate in the computation of $\hat{y}_j$. Therefore, the square of the length of the vector $|X(j)|^2 = \sum_i (X_{i+j})^2$, which is proportional to the total signal power entering the adaptive filter, changes relatively slowly and can be approximated by the square of the input signal rms value $\overline{X}_j^2$. Substituting this in equation (4.15) and using equation (4.53) makes it possible to express the stability criterion as:

$$\Delta h_j = \Delta_{i,j}|h_j| = \frac{\varepsilon_j X_{i+j}}{\overline{X}_j^2} \leq \frac{\varepsilon_j}{\overline{X}_j} \cdot \frac{X_{i+j}}{\overline{X}_j} \tag{4.55}$$

$$= \alpha \beta \phi_{i+j} \quad F(\varepsilon_j)$$

where ϕ_{i+j} and $F(\varepsilon_j)$ are defined as follows

$$\phi_{i+j} = \begin{cases} 1 & ; \quad \left|\dfrac{X_{i+j}}{\alpha X_j}\right| > 1 \\ 0 & \quad \text{otherwise} \end{cases} \tag{4.56a}$$

$$F(\varepsilon_j) = \begin{cases} 1 & ; \quad \left|\dfrac{\varepsilon_j}{\beta X_j}\right| \geq 1 \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{4.56b}$$

α is a dimensionless constant, $1 \leq \alpha \leq 2$ and β determines the sensitivity of the error detector. The function ϕ_{i+j} selects the samples X_{i+j} with the highest amplitudes and its value is stored in the memory as a part of the digital signal X_{i+j} and fed into the correlation processor to compute the H-corrections according to equation (4.55). The function $F(\varepsilon_j)$ initiates the correction process only if the error $|\varepsilon_j|$ is greater than a fraction $|\beta|$ of the rms value X_j of the input signal. The dimensionless constant multiplier $\Delta_{i,j}$ has a fixed value of $\pm 2^{-4}$, and the optimum α is 1.4. The condition for correction is $F(\varepsilon_j)\ \phi_{i+j} = 1$; therefore, the coefficient β must be equal to

$$\beta \geq \frac{|h_i|}{1.4} \cdot 2^{-4} \qquad (4.57)$$

to satisfy the stability criterion.

4.8 Phase Roll and Nonlinear Distortions

To add to the problems of effective echo canceller design, it sometimes happens that part of the echo energy returns with distortions which cannot be replicated in an adaptive transversal filter. One of these forms of distortion is the frequency offset, which refers to a shift in all frequency components of a signal. Frequency offset, which is associated with the modulation and demodulation operations at the end of a carrier link, is perceived by an adaptive canceller as a constantly increasing phase, or "phase roll", in the echo signal, and the canceller can effectively cancel only if its adaptation can keep up with the phase roll.

CCITT recommends that the end-to end frequency shift does not exceed 2 Hz in either direction, which sets an upper bound of 4 Hz on the overall echo path offset. It has been found that phase-roll presents a real problem to the introduction of cancellers on a small, but significant, proportion of circuits examined.

Nonlinear distortions are much harder to compensate. An effort in this direction was made by Thomas[18], but the difficulty of identifying and characterizing low-level nonlinearities has constrained the development of practical designs.

Echo cancellers can not be operated in the full mode. The reason is that instead of modelling the impulse response of an echo path of 30 msec. round trip delay the echo path would be up to 600 msec. long for a single hop satellite circuit. Therefore, the speed required to perform the convolution processing will be drasticallly increased.

4.9 Double Talk Detector

When the incoming waveform from the hybrid contains an additive noise in addition to the echo signal, it will perturb the tap weights from their correct settings[37]. If the noise is uncorrelated with the reference waveform, as one hopes is the case, then the settings are correct on the average, but fluctuate about their true values. In reference to Figure 4.7, it is assumed that S,Y,X, and r have zero means and are statistically stationary, which is true at least during a period of several syllables, the send-out signal is

$$Z = S + Y - r \tag{4.58}$$

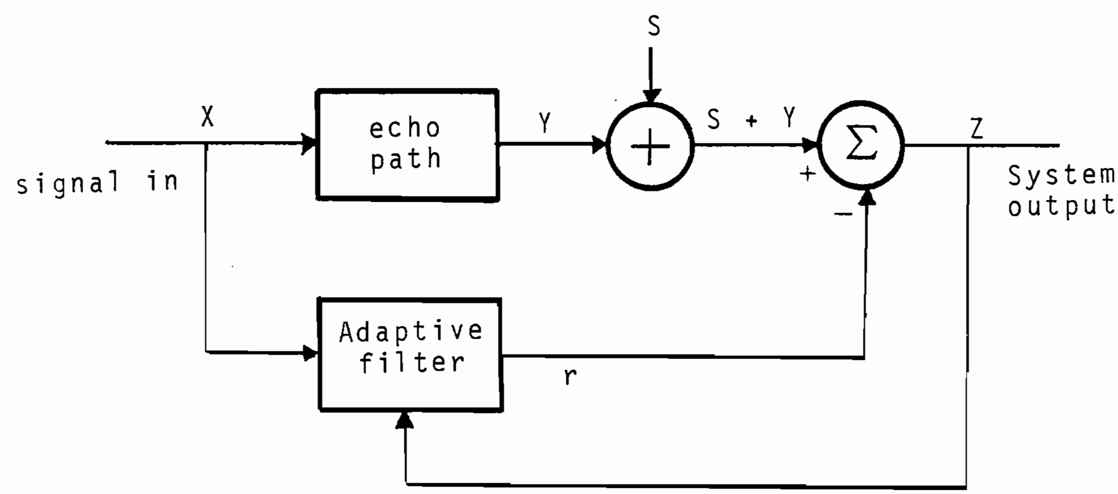

Figure 4.7 Equivalent diagram of the adaptive echo canceller.

Since S is uncorrelated with Y and r, the variance of Z is equal to the sum of average uncorrelated signal power levels

$$E = \left[Z^2\right] = E\left[S^2\right] + E\left[(Y-r)^2\right] \tag{4.59}$$

The adaptive algorithm minimizes the square of errors ε_j, i.e., the variance E (Z^2), since the uncorrelated signal S remains constant. Thus, minimizing $E(Z^2)$ minimizes the echo output power $E\left[(Y-r)^2\right] \rightarrow 0$, theoretically without regard to the presence of uncorrelated signal S. However, this is true only when echo signal Y is approximately equal to or larger than signal S, e.g., in circuits with very low ERL and low signal level S at the hybrid. However, under the double talk condition, when the send-out signal S is usually substancially larger than the echo Y, the correction algorithm does not work properly because the minimum value of E (Z^2) remains relatively constant as a function of h_i. Hence, it seems unavoidable that the tap weights will diverge from their proper settings and the echo will not be adequately cancelled. In order to avoid this possibility, the tap weights of currently proposed voice echo cancellers are "frozen" as soon as double talking is detected. A double-talking detector must therefore be provided in the echo canceller. Even with the assumption of a fast-acting double-talking detector, there is still the possibility of changes ocurring in the echo channel during the time that the canceller is frozen, which will lead to increased uncanceled echo. Fortunately, the duration of a period of double talk is usually very short. A relatively simple double talk detector as shown in Figure 4.8 can give satisfactory results even when the echo path is time variance or has some phase roll.

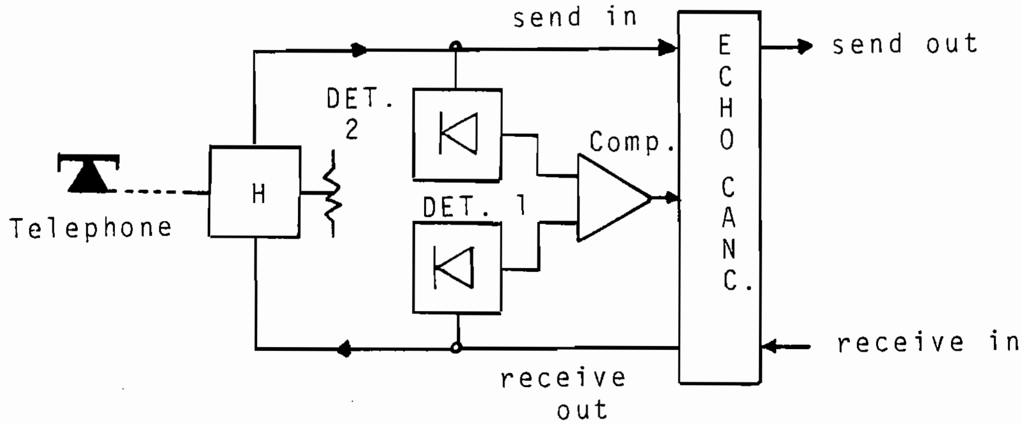

Figure 4.8 : Block diagram of the double talk detector.

Both receive-out and send-in rectified signals are compared in a voltage comparator. When the send-in signal is 0.5 dB stronger than the strongest possible echo signal at hybrid H, the output of the comparator inhibits the cross correlator's adaptation process.

The design of a good double-talking detector is difficult. Several design criteria must be met. The rise time of the inhibit signal must be substantially shorter than the convergence time. The rise time of the output of detector one (DET 1) must be longer than that of detector 2 (DET 2). The release time of the detector's output voltage must be between 50 and 200 msec, i.e., long enough to permit "rebuilding" of the impulse response during the pauses of near-end speech.

4.10 Adaptive Center Clipper

Even an echo canceller which is working well will leave some residual uncancelled error. Limitations on the achievable cancellation ratio are imposed by the presence of additive noise, by non-linear distortions, by echo dispersion beyond the length of the transversal filter, and by digital resolution constraints. To eliminate this residual echo signal, an adaptive center clipper may be included in the send-out path of the echo canceller, as shown in Figure 4.9.

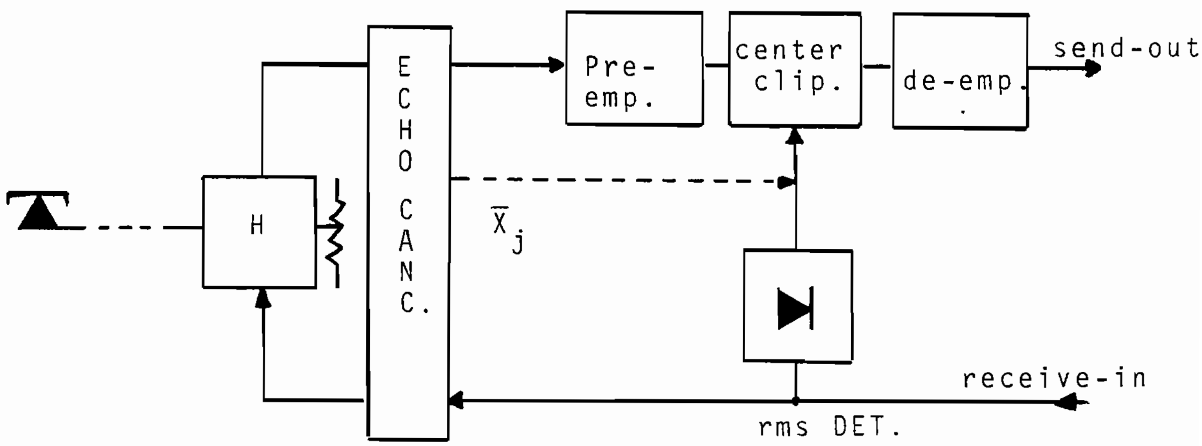

Figure 4.9 Center Clipper in the echo Canceller.

The nonlinear transfer characteristic of the center clipper suppresses signals with instantaneous amplitudes less than the threshold L, as shown in Figure 4.10. Tests have shown that best results can be obtained if the clipping level $|L|$ is held approximately proportional to the rms value $\overline{X}_j$ of the speech integrated over a window of 50 to 150 msec. When properly adjusted, this device can suppress the residual echo in the whole dynamic range of the telephone signal by another 8 to 50 dB without noticeably affecting the direct signal in the send-out path, which is usually much stronger than the residual echo.

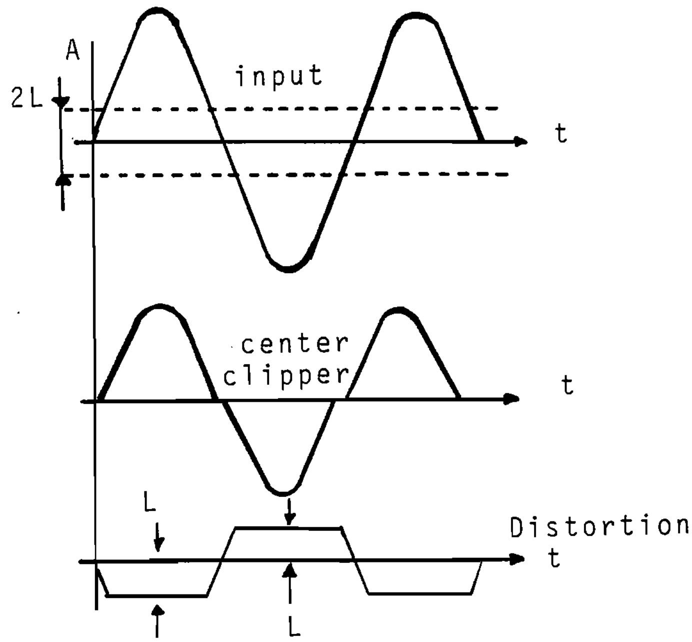

Figure 4.10 Effect of center clipping on a Sinusoidal waveform.

Only subjective tests can determine the effect of the center clipper distortion on speech, but an analysis based on the sine wave input can be used for design and evaluation. Figure 4.10 shows the distortion waveform which approaches a symmetric square wave. Theoretically, distortion is caused only by odd harmonics which are in the frequency band above 800 Hz, where the speech power is falling more than 12 dB/octave. The signal can therefore be pre-emphasized by 6 dB/octave before clipping and de-emphasized by

6 dB/octave after clipping to attenuate the harmonics. Berkley and Mitchell[38] have determined that a center clipper which is designed to operate separately on several subsets of the voice spectrum will not only effectively eliminate a substantial echo, but will also not noticeably degrade the desired signal which is present during activity of the speaker at the other end of the line, and thus does not have to be removed during double talking.

4.11 Interface with Digital Network

Figure 4.11 shows a satellite linked telephone circuit. Until this point, the signal between the telephone set and the satellite earth station has been assumed analog. Therefore, analog to digital conversion has been necessary.

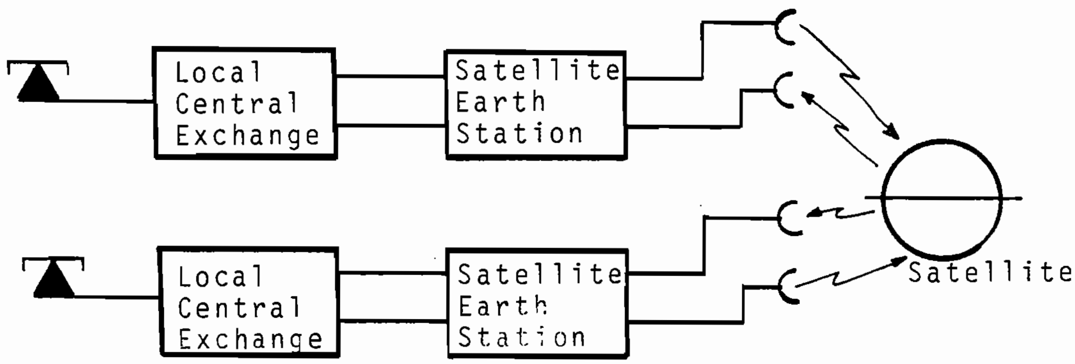

Figure 4.11: Satellite linked telephone circuit.

In this section it is assumed that the signal arriving at the satellite earth station from the local central exchange is digital; in particular, u-law PCM. The all-digital network in North America has adopted the 8-bit u =255 PCM format as an encoding standard for telephony. In European countries, the A-law is being used. The all-digital network operates on a Fixed Loss Plan[43] each toll connecting trunk operating at a loss of 3 dB and intertoll trunks at zero loss and since the noise performance is also essentially distance independent, the question arises as to whether the 1850 miles limit should be retained for the application of echo suppressors. At present it appears that the 1850 miles limit will be retained, however, one may conclude from Spang's Loss-Noise Echo study[44] that two options exist; firstly to extend the application mileage and still give loss-noise-echo grade of service comparable to the analog network, or, secondly to retain the 1850 miles limit and improve the loss-noise-echo grade of service. The latter approach has been adopted.

Two factors are of importance for planning echo control in the digital network. Firstly, the analog voice channel does not exist in the digital network and secondly, the single voice channel does not exist as a separate entity within the digital hierarchy. As a direct result echo control devices for the digital network must be digital and multi-channel unless demultiplexing and D/A conversion is to be performed both of which are undesirable for technical and economic reasons. The two candidate devices for active echo control in the future are the digital echo suppressor and the digital echo canceller.

4.11.1 Digital Echo Suppressors

Although operating on digital bit streams rather than analog voice signals, the digital echo suppressor is required to perform the same functions as an analog echo suppressor, i.e.,

a) to be transparent when voice channels are very low in level or present in the send path only;

b) to perform suppression (> 50 dB) when speech is present on the receive side only;

c) to permit break-in when both parties are talking (zero insertion loss on the send side; zero to several dB insertion loss on the receive side); and

d) to be disabled on command for data transmission

The subjective performance of digital echo suppressors is expected to be substantially better than that of analog suppressors partly due to the performance of the echo suppressor itself and partly due to characteristics of the digital network.

Table II shows a comparison of the digital echo suppressor[45,46] and the existing CCITT Rec. G. 161[47] for analog echo suppressors. From there it can be founded that the digital echo suppressor has a sharper suppression threshold, a more sensitive break-in threshold and that it reduced echo bursts and clipping effects.

Table II

Performance Comparison between DES and CCITT Rec. 161 for AES

Parameter	G.161	D.E.S.
Suppression threshold (dBmO)	$-33 \leq T \leq -29$	$-32 \leq T \leq -30$
Break-in threshold (dBmO)	$L_s = L_r \pm 1.5$ dB	$L_r - 4 \leq Ls \leq L_r$
Suppression Operate time (msec)	5	.125

Table II (Cont'd)

Parameter	G.161	D.E.S.
Break-in operate time (msec) (L_r constant)	30	.125
Suppression hangover time (msec)	40-75	24-32
Break-in hangover time (msec)	150-350	48-64

4.11.2 Digital Echo Canceller.

The digital echo canceller has developed based on the premise that it is a superior device than the echo suppressor. It follows the same process as that one explained in section 4.2, i.e., echo reduction is achieved by applying the receive-in signal to an estimate of the echo path transfer function to generate an estimate of the echo which is then substracted from the send-in signal path. The only difference with the echo canceller design for analog signals is that the A/D conversion is not necessary because the signal is already digital. Digital signals are converted to the echo canceller digital format by a digital interface. The digital interface obtains the uniform code equivalent of each signal and then follows the same procedure as mentioned in section 4.4.2. This method of processing signals gives the same performance as decoding the several digital signals, combining them by analog methods, and recoding, since digital combining is subject to round-off errors and analog combining is subject to quantizing errors.

In comparing the costs between the digital echo suppressor and the digital echo canceller the presently estimate are as follows:

2 Echo cancellers per circuit	\$3000
2 Echo suppressor per circuit (split)	\$ 400
1 Echo suppressor per circuit (full)	\$ 250

No universally accepted performance standards exist for digital echo suppressors or echo cancellers. However, CCITT is expected to produce a Recommendation for the former by 1980, and a preliminary form of recommendation for the latter by 1980.

CHAPTER V

SUMMARY

All transcontinental and intercontinental calls and calls between relatively close locations when satellites hops are used require echo control devices. Without them, a portion of long-distance calls would suffer from impairments ranging from hollow reverberations to severe echoes. The source of telephonic echoes is almost always a hybrid transformer, used for converting long-distance calls between two-wire and four-wire telephone transmission. Once an echo is produced, its effect on the quality of a telephone call depends on its intensity and delay time.

Until recently echo control has been accomplished by the application of two distinct techniques:

a) Insertion of loss as a function of trunk length on connections whose round trip delay does not exceed 45 msec.

b) Use of analog echo suppressors on trunks exceeding a certain distance (currently 1850 miles).

Two major network developments of the 1970's will have a significant effect on echo control techniques in the near future. These are, firstly, the emergence of the all digital toll network, which has encouraged the development of digital echo control devices, and secondly, the increasing use of satellite toll trunks, which has encouraged the development of an improved echo control device: the echo canceller.

An echo canceller is mainly composed of an adaptively controlled filter for simulating echo path transfer characteristics and synthesizing an echo replica, a subtractor and an adaptation processor. The subtractor is used for cancelling out a real echo with the echo replica. The processor is used for adaptively adjusting the filter characteristics to that of an echo path. The echo canceller operating characteristics of importance are speed of convergence and echo return loss enhancement. The expected enhancement of the echo canceller is about 23 dB and using the center clipping technique it can be increased over 30 dB. With a normal speech, the convergence time for a good canceller has to be expected shorter than 250 msec. In actual practice, a few speech syllables must be said over the circuit to obtain sufficient cancellation.

One area of concern when using the echo canceller is the reduction in ERLE due to "phase roll". Phase roll is due to small frequency translation offsets which may occur in the circuits analog tail. Horna[23] shows a 6 dB loss in ERLE for a frequency error of 1 Hz, but also states that this is not detectable for speech signals if the centre clipper is used.

During the next few years it is expected that the digital echo suppressor in either split or full mode will be used for terrestrial trunks, and echo cancellers will be used for satellite circuits. However, to win widespread acceptance echo cancellers must clearly drop substancially in price as present estimates project them several times as expensive as digital echo suppressors. In fact, in this year Bell Laboratories has developed an echo canceller integrated circuit which costs about one-tenth the price of present echo-canceller circuits, which makes it economically feasible for introduction into the telephone network.

LIST OF REFERENCES

(1) A.B. Clark and R.C. Mathes, "Echo suppressors for long telephone circuits", Proc. AIEE, Vol. 44, Apr. 1925.

(2) O.B. Blackwell, "The time factor in telephone transmission", BSTJ, Vol. 11, Jan. 1932.

(3) J.W. Emling and D. Mitchell, "The effects of time delay and echoes on telephone conversations", BSTJ, Vol. 42, Nov. 1963.

(4) R.R. Riesz and E.T. Klemmer, "Subjective evaluation of delay and echo suppressors in telephone communications", BSTJ, Vol. 42, Nov. 1963.

(5) P.T. Brady, "Effects of transmission delay on conversational behavior in echo-free telephone circuits", BSTJ, vol. 50, Jan. 1971.

(6) P.T. Brady and G.K. Helder, "Echo suppressor design in telephone communications", BSTJ, Vol 42, Nov. 1963.

(7) D.L. Richard and G.A. Buck, "Telephone echo tests", Proc. IEE, Vol. 107, Apr. 1960.

(8) E.T. Klemmer, "Subjective evaluation of transmission delay in telephone conversations", BSTJ, Vol. 46, July 1967.

(9) R.G. Gould and G.K. Helder, "Transmission delay and echo suppression", IEEE Spectrum, Vol. 7, Apr. 1970.

(10) D.L. Richards, "Conversational performance of speech links subject to long propagation times", Int. Conf. On Sat. Comm., IEE, Oct. 1962.

(11) R.M. Krauss and P.D. Bricker, "Effects of transmission delay and access delay on the efficiency of verbal communications", J.A.S.A., Vol 41, Feb. 1967.

(12) R.G. Gould, "Echo suppression and time delay, a subjective evaluation", IEEE Trans. Comm. Techn., Vol. Com-12, Sept. 1964.

(13) G.K. Helder, "Customer evaluation of telephone circuits with delay", BSTJ, Vol. 45, Apr. 1966.

(14) F.K. Becker and H.R. Rudin, "Application of automatic transversal filters to the problem of echo suppression", BSTJ, Vol. 45, Nov. 1966.

(15) A. Miura, R. Sato and K. Nagata, "A blockless echo suppressor", IEEE Trans. Comm. Techn. Vol. COM-17, Aug. 1969.

(16) M.M. Sondhi, "An adaptive echo canceller", BSTJ, Vol. 46, March, 1967.

(17) J.R. Rosenberger and E.J. Thomas, "Performance of an adaptive echo canceller operating in a noisy, linear, time-invariant environment", BSTJ, Vol. 50, March 1971.

(18) E.J. Thomas, "An adaptive echo canceller in a nonideal environment", BSTJ, Vol. 50, Oct. 1971.

(19) D. Mitra and M.M. Sondhi, "Adaptive filtering with nonideal multipliers: applications to echo cancellation", Proc. 1975, ICC, Vol. II.

(20) M.M. Sondhi and D. Mitra, "New results on the performance of a well-known class of adaptive filters", Proc. IEEE, Vol. 64, Nov. 1976.

(21) S.J. Campanella, H.G. Suyderhoud, and M. Onufry,"Analysis of an adaptive impulse response echo canceller", COMSAT Tech. Rev., Vol. 2, Spring 1972.

(22) H.G. Suyderhoud, S.J. Campanella and M. Onufry, "Results and analysis of a worldwide echo canceller field trial", COMSAT Tech. Rev., Vol. 5, Fall 1975.

(23) O.A. Horna, "Echo canceller with adaptive transversal filter", COMSAT Tech. Rev. Vol. 7, Fall 1977.

(24) H. Brueggemann, G. Kidd, L. Mackechnie, and A. Seyler, "Experiments on adaptive echo cancellation", ATR, Vol. 4, May 1970.

(25) N. Demytko, and L. MacKechnie, "A high speed digital adaptive echo canceller", ATR, Vol. 7, Jan. 1973.

(26) R. Wehrmann, "Suboptimum gradient methods for the adjustment of echo cancellers in a noisy environment", in Conf. Rec., 4th ICDSC., Oct. 78.

(27) A. Castanet, "Auto-adaptive echo canceller for long distance telephone circuits", L'Onde Electrique, Vol. 55, Jan. 1975.

(28) Y. Kato, S. Chiba, T. Ishiguro, Y. Sato, M. Tajima, T. Ogihara, S. Campanella, H. Suyderhoud, and M. Onufry, "A digital adaptive echo canceller", NEC Res. and Dev., No. 31, Oct. 1973.

(29) T. Araseki, Y. Itoh and K. Ochiai, "A microprocessor echo canceller", in Conf. Rec., 4th ICDSC, Oct. 78.

(30) G.K. Helder and P.C. Lopiparo "Improving transmission on domestic satellite circuits", BLR, Vol. 55, Oct. 77.

(31) "Communications system engineering handbook", D.H. Hamsher, Edit. McGraw-Hill, 1967.

(32) A.M. Noll, "Subjective effects of sidetone during telephone conversations", IEEE Comm. & Electronics, May 1964.

(33) "Telephone Transmission Quality", C.C.I.T.T., Green Book, Vol. V, 1973.

(34) "Line Transmission", C.C.I.T.T., Green Book, Vol. III-1, 1973.

(35) F.P. Duffy, G.K. McNees, I. Nasell, and T.W. Thatcher, "Echo performance of toll telephone connections in the United States, BSTJ, Vol. 54, Feb. 1975.

(36) T.N. Greville, "The pseudoinverse of a rectangular or singular matrix and its applications to the solution of system of linear equations", SIAM Rev., Vol. 1, Jan. 1959.

(37) B. Widrow, J. Glover, J. McCool, J. Kannitz, C. Williams, R. Hearn, J. Zeidler, E. Doug, and R. Goodlin, "Adaptive Noise Cancelling: Principles and Applications", Proc. of the IEEE, Vol. 63, Dec. 75.

(38) O. M. Mitchell and D.A. Berkley, "Full duplex echo suppressor using center clipping", BSTJ, Vol. 50, May 1971.

(39) Notes from course 304-7110, McGill University, Oct. 77.

(40) L. Schweizer and H. Hoss, "An experimental PCM encoder for 24 speech channels", IEEE Trans. On Comm. Tehn., Vol. COM-17 Oct. 1969.

(41) T. Areseki and K. Ochiai, "Adaptive echo suppressor", in Conf. Rec., 3rd Int. Conf. DSC, Kyoto

(42) "Session on digital echo suppressors", in Proc. 1976 ICC, Vol. III.

(43) "Telecommunications transmission engineering", Bell Laboratories Vol. III, Chapter 9, 1975.

(44) T.C. Spang, "Loss-noise-echo- study of the direct distance dialing network", BSTJ, Jan. 1976.

(45) R.C. Dreschler,"Echo suppressor terminal for No. 4 ESS, " ICC 1976, Conf. Rec. Vol. 3, June 1976.

(46) "Characteristics of Digital Echo Suppressors", CCITT Com XV-86-E Study Group XV, Contrib. from At&T, Oct. 1977.

(47) "Green Book, Vol. III", CCITT Fifth Plenary Assembly, Geneva, Dec. 72.