THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XLVI — MARCH 1967 — NUMBER 3

An Adaptive Echo Canceller

By M. M. SONDHI

(Manuscript received November 3, 1966)

A novel method is presented for echo-cancellation in long distance telephone connections. In contrast with conventional echo suppressors, the device described achieves echo-cancellation without interrupting the return path. A replica of the echo is synthesized and subtracted from the return signal. The replica is synthesized by means of a filter which, under the control of a feedback loop, adapts to the transmission characteristic of the echo path and tracks variations of the path that may occur during a conversation.

The adaptive control loop is described by a set of simultaneous, nonlinear, first-order differential equations. It is shown that under ideal conditions, the echo converges to zero. Estimates of the rate of convergence are obtained. Effects of noise are discussed. The results of computer simulations of various alternative configurations of the system are described.

I. INTRODUCTION

In telephone connections that involve both 4-wire and 2-wire links, an echo is generated at the hybrid that connects a 4- to a 2-wire link. The situation at one such hybrid is illustrated schematically in Fig. 1. Here S_1 and S_2 are the two speech signals and E is the echo of S_1 which is returned along with S_2. In practice, E is an the average about 15 dB lower than S_1, but in extreme cases may be only 6 dB lower.

This echo has a disturbing influence on the conversation, which appears to increase with increasing round-trip delay.[1] If no steps were taken to reduce this echo, conversation would be seriously impaired over satellite communication links with round-trip delays of hundreds of milliseconds. The devices used at the present time to combat echo are called echo suppressors. A number of different types of echo suppressors have been designed. They are all, basically, voice-operated switches (albeit ingenious and complex ones) which disconnect the return path or introduce a large attenuation in it whenever a decision mechanism indicates that the level of S_1 is large compared to that of $S_2 + E$. However, since E and S_2 both share the return path, the use of such echo suppressors introduces "chopping" or interruptions of S_2 during periods of double talking.* It has been shown that the

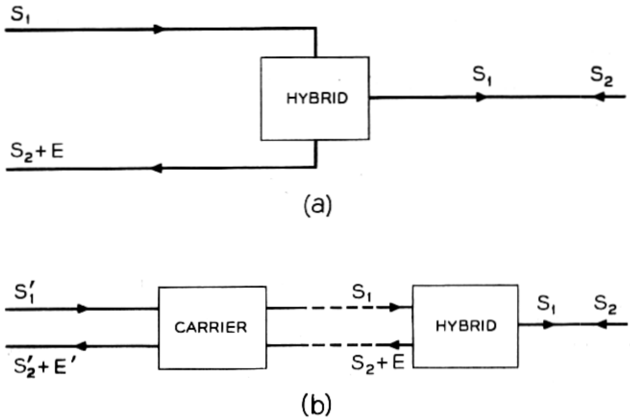

Fig. 1—Typical situations where an echo canceller could be used.

degrading effect of chopping also increases with increasing round-trip delay.[1] The characteristics, advantages and disadvantages of such echo suppressors have been described in a number of papers.[2,3,4]

It appears that improvements in such echo suppressors are not likely to solve the echo problem satisfactorily. Entirely different approaches are called for. One such approach was an open loop device suggested by J. L. Flanagan and D. W. Hagelberger and implemented by J. de Barbeyrac.[5] In this approach, E is regarded as a linearly filtered version of S_1. The impulse response of this filter is measured by means of a transmitted test pulse, and a transversal filter is synthesized to approximate this impulse response. With S_1 as an input to the transversal filter, the output approximates E, and may, therefore, be sub-

* In this paper, the term "double-talking" will be used for the simultaneous presence, at the echo suppressor, of speech signals of the two speakers.

tracted from the return signal to cancel the echo. In this manner, effective echo cancellation (as opposed to *suppression*) is achieved without interrupting S_2. J. de Barbeyrac demonstrated the feasibility and effectiveness of such a device on four- to two-wire junctions simulated in the laboratory.

An actual echo path is, however, not perfectly constant. Besides obvious step changes such as the connection or disconnection of extension phones during a conversation, or transfer of calls via key telephones or PBX's, there may be slow changes in gain and other fluctuations of the transfer function of the echo path. Also, economic consideration would force placement of echo cancellers at switching offices high in the hierarchical structure rather than at the hybrids where the echoes are generated. (The number of echo cancellers required in the latter case would be many thousand times the number required in the former.) In such a situation, one or more carrier links intervene between the echo canceller and the hybrid. A large percentage of these (e.g., the N carrier) use compandors which are nonlinear elements with memory. Thus, what are available to the echo canceller are not S_1 and $S_2 + E$ but the modified signals S_1' and $S_2' + E'$ (see Fig. 1(b)). E' is no longer a linearly filtered version of S_1', although a linear filter with an impulse response dependent upon the power level of S_1' could approximately transform S_1' to E'. Thus, for the open loop device to work in practice, it would seem necessary to intermittently adjust the transversal filter during a conversation. The transmission of test pulses required for such adjustments might prove quite intolerable to the customers.

A proposal made by John L. Kelly, Jr. avoids these difficulties. The speech signal itself is used in place of test pulses and a control loop continuously adapts the transversal filter to take care of fluctuations in the echo path.

In this paper, we will describe the system proposed by Kelly. We will describe various modifications of the system which simplify and improve it. Finally, we will report the results of tests of these systems by computer simulation, using artificially created echoes and also using two-track tape recordings made on an actual N-carrier link.

II. KELLY'S PROPOSAL

The adaptive control loop shown in Fig. 2 is the system proposed by Kelly, except for the introduction of the nonlinear function F. This function is chosen to be an odd nondecreasing function with $F(0) = 0$. (Kelly's proposal obtains as a special case when $F(e) = e$.)

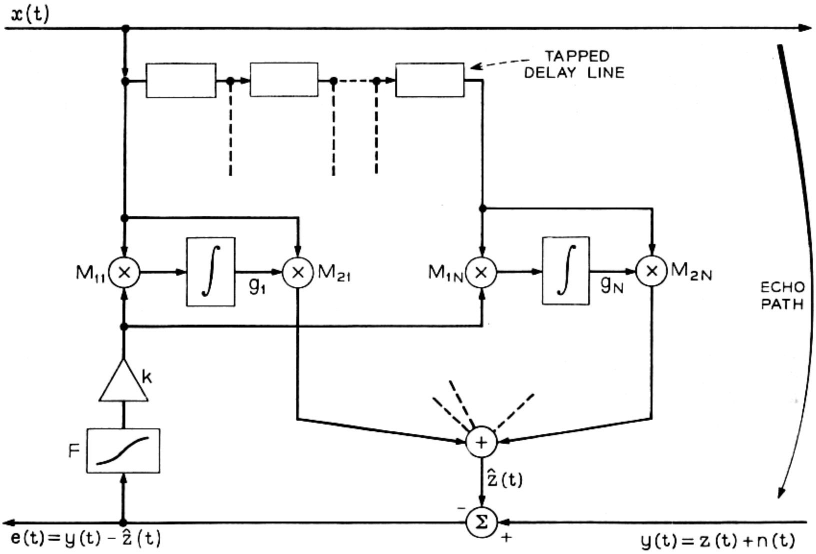

Fig. 2 — Schematic of the echo canceller using a transversal filter.

The signal $x(t)$ is the input speech signal (corresponding to S_1 or S_1' of Fig. 1). The signal $y(t)$ is the return signal (corresponding to $S_2 + E$ or $S_2' + E'$ of Fig. 1) and is given by

$$y(t) = n(t) + z(t),$$

where $z(t)$ is the echo of the input signal and $n(t)$ is a noise which is assumed statistically independent of $z(t)$. The noise may include a second speech signal besides circuit noise.

The N-tap transversal filter synthesizes an estimate of $z(t)$ given by

$$\hat{z}(t) = \sum_{1}^{N} g_k(t)x[t - (k - 1)T_d]$$

where T_d is the delay of each section of the transversal filter. The control loop uses the error $e(t) = y(t) - \hat{z}(t)$ to continuously improve the estimate $\hat{z}(t)$.

Ideally, the system should drive itself to the condition $e(t) = n(t)$ (not necessarily $e(t) = 0$, for as mentioned earlier $n(t)$ may contain a speech signal which must be left as undistorted as possible). Such ideal echo cancellation is possible with this system only if $n(t) \equiv 0$ and if $z(t)$ is exactly representable by passing $x(t)$ through an N-tap transversal filter with constant (or slowly varying) tap gains. In the next section we will exhibit a proof that under these conditions the system does indeed converge monotonically to this echo-less steady state. In the presence of noise the final state is one in which the tap gains fluctuate around their settings for perfect echo cancellation. The response of the system averaged over the noise ensemble will be shown to converge monotonically to this final state.

The system is a first-order system and is stable for any arbitrary input. As is the case with most control systems, speed of response can be traded for immunity to noise. The constant multiplier K of Fig. 2 allows adjustment of this trade-off. (Unfortunately, in the present case speed of response depends not only upon the feedback factor K but also upon the level and properties of the signal $x(t)$. Thus, K can be adjusted to give a certain speed of response only for some average level of $x(t)$.)

For immunity from noise, K should be made as small as possible; for fast convergence, it should be made as large as possible. We do not have any theory at present to calculate the optimum setting for K. This must be done by computer simulation and/or experiments with hardware implementations of the system.

III. CONVERGENCE

The proof of convergence given in this section is very similar to a proof given by Kelly. The introduction of the nonlinear function F necessitates only minor modifications. However, as we shall show in Section V, a judicious choice of this nonlinearity can considerably simplify and improve the performance and implementation of the system.

To simplify our discussion we introduce the following notation. We denote the output $x[t - (k - 1)T_d]$ of the kth tap of the transversal filter as $x_k(t)$. We will refer to N-tuples as vectors and consider them as column matrices. Thus, $\mathbf{X}(t)$ will be a column matrix with elements $x_1(t), x_2(t), \cdots, x_N(t)$, and $\mathbf{G}(t)$ the column matrix with elements $g_1(t), g_2(t), \cdots, g_N(t)$. The signal $\hat{z}(t)$ of Fig. 2 then becomes

$$\begin{aligned}\hat{z}(t) &= \sum_{k=1}^{N} g_k(t)x_k(t)\\ &= \mathbf{G}^T\mathbf{X}.\end{aligned}$$

Here the superscript T denotes the transpose of a matrix, and for brevity the dependence of $\mathbf{G}$ and $\mathbf{X}$ on t is not explicitly shown. The echo $z(t)$ will likewise be represented as

$$z(t) = \mathbf{H}^T\mathbf{X},$$

where $\mathbf{H}$ has elements $h_1, \cdots, h_N$ which are assumed fixed (or so slowly varying that their time derivatives may be neglected). Thus, we have

$$\begin{aligned} y(t) &= z(t) + n(t) \\ &= \mathbf{H}^T\mathbf{X} + n(t) \\ e(t) &= y(t) - \hat{z}(t) \\ &= \mathbf{R}^T\mathbf{X} + n(t), \end{aligned} \tag{1}$$

where $\mathbf{R} = \mathbf{H} - \mathbf{G}$.

Before proceeding to the proof of convergence let us give an interesting heuristic justification for the circuit of Fig. 2. Consider a function $C(e)$ such that $C(e) = C(-e)$ and $d^2C/de^2 \geqq 0$. $C(e)$ is then a monotonically nondecreasing function of the magnitude of e. Let us minimize $C(e)$ by varying the coefficients $g_k(t)$. If we choose to use the steepest descent method, we find the gradient of $C(e)$ with respect to the g_k and make the vector $\mathbf{G}$ change in the direction opposite to this gradient. Now

$$\begin{aligned} \text{grad } C(e) &= \text{grad } C(\mathbf{R}^T\mathbf{X} + n(t)) \\ &= -C'(\mathbf{R}^T\mathbf{X} + n(t))\mathbf{X} \\ &= -F(\mathbf{R}^T\mathbf{X} + n(t))\mathbf{X}, \end{aligned}$$

where $C'(\cdot) = F(\cdot)$ is the derivative of C with respect to its argument. To change $\mathbf{G}$ along the negative of the gradient we may set

$$\frac{d}{dt}\mathbf{G} = KF(\mathbf{R}^T\mathbf{X} + n(t))\mathbf{X}, \tag{2}$$

where K is a positive constant of proportionality.

By inspection, the matrix equation (2) is the equation that governs the dynamic behavior of the system of Fig. 2. Thus, the system is a steepest descent control system in the above sense.

The proof of convergence follows easily from (2) in the case when $n(t) \equiv 0$. Observe that since

$$\mathbf{R} = \mathbf{H} - \mathbf{G}, \qquad \frac{d}{dt}\mathbf{G} = -\frac{d}{dt}\mathbf{R},$$

premultiplying (2) with $-2\mathbf{R}^T$ gives

$$2\mathbf{R}^T \frac{d}{dt}\mathbf{R} = \frac{d}{dt}\mathbf{R}^T\mathbf{R}$$
$$= -2K\mathbf{R}^T\mathbf{X}F(\mathbf{R}^T\mathbf{X}). \tag{3}$$

By its definition F is a monotonic nondecreasing function and also an odd function. Thus, the right-hand side of (3) is always negative, hence $\mathbf{R}^T\mathbf{R}$ is nonincreasing. It is strictly decreasing whenever $\mathbf{R}^T\mathbf{X} \neq 0$, i.e., whenever there is an uncancelled echo. Now $\mathbf{R}^T\mathbf{R} = l_R^2$ is the square of the length of the vector $\mathbf{R} = \mathbf{H} - \mathbf{G}$. Thus, l_R is nonincreasing, and as long as there is an uncancelled echo the length keeps decreasing, i.e., $\mathbf{G}$ keeps approaching $\mathbf{H}$. To show that the echo goes to zero we integrate (3) between 0 and some time τ and obtain

$$l_R^2(0) - l_R^2(\tau) = K \int_0^\tau \mathbf{R}^T\mathbf{X}F(\mathbf{R}^T\mathbf{X})\, dt. \tag{4}$$

As l_R^2 is nonincreasing the left-hand side is bounded and $\leqq l_R^2(0)$. Thus, as $\tau \to \infty$ we note that the integrand on the right must approach zero. However, the integrand is a monotonic nondecreasing function of the magnitude of $\mathbf{R}^T\mathbf{X}$ (which is the uncancelled echo). Thus, the echo power must approach zero.*

If $y(t)$ contains a noise $n(t)$ besides the echo, then proceeding as before we find that

$$\frac{d}{dt}\mathbf{R}^T\mathbf{R} = -K\mathbf{R}^T\mathbf{X}F(\mathbf{R}^T\mathbf{X} + n(t)). \tag{5}$$

From our previous discussion it follows that the right-hand side of (5) is negative, if and only if $\mathbf{R}^T\mathbf{X} + n(t)$ has the same sign as $\mathbf{R}^T\mathbf{X}$. As long as the magnitude of the uncancelled echo is large compared to $n(t)$, this condition is met for a large percentage of time and convergence proceeds essentially monotonically as before. When the level of the uncancelled echo becomes of the same order as or lower than $n(t)$ the convergence clearly cannot be monotonic. There will be intervals when $n(t)$ is greater than R^TX in magnitude and of opposite sign. However, if the feedback gain constant K is small then $\mathbf{G}$ is a slowly varying function of time (typically K would be adjusted so that $\mathbf{R}$ has a "time constant" on the order of 0.5 sec or so). In such a quasi-stationary case it is justified to assume that $\mathbf{R}^T\mathbf{X}$ is independent of $n(t)$ provided

* I. W. Sandberg has pointed out that, strictly speaking, this argument does not prove that $|\mathbf{R}^T\mathbf{X}| \to 0$ for certain pathological cases. Although these cases are of no practical concern, it is interesting that weak conditions suffice to rule these out. For example, it is sufficient that: (*i*) $|\mathbf{X}|$ and $d/dt\,|\mathbf{X}|$ be bounded and (*ii*) the function F be such that $eF(e)$ and $d/de\,(eF(e))$ be bounded for all finite e.

$n(t)$ is a wideband signal. Then it is not hard to show that the average of $F(\mathbf{R}^T\mathbf{X} + n(t))$ over the noise ensemble has the same sign as the average of $\mathbf{R}^T\mathbf{X}$ provided only that noise has a symmetric distribution and thus, the system still converges in this average sense.

IV. RATE OF CONVERGENCE

While convergence can be proved by such relatively simple arguments, estimating the convergence rate is an extremely difficult problem. The convergence rate depends upon the properties of $\mathbf{X}$, upon the choice of F, and upon K and we do not have a solution to the problem in the general case. However, we will now derive an estimate of the mean convergence rate for the noiseless case under the assumption that (3) can be averaged over the $\mathbf{X}$ ensemble (which is assumed stationary) and the vector $\mathbf{R}$ assumed independent of $\mathbf{X}$ on the right-hand side. This assumption is justified if K is small, hence $\mathbf{R}$ slowly varying. Under this assumption, the expectation can be calculated for a variety of different functions F and random processes $\mathbf{X}$. We will give the result when $\mathbf{X}$ is a zero mean Gaussian process with (*i*) $F(x) = x$, and (*ii*) $F(x) = \text{sgn}\,(x)$ (here $\text{sgn}\,(x) = 1$ for $x \geqq 0$ and -1 for $x < 0$). In case (*i*)

$$\frac{d}{dt}\langle\mathbf{R}^T\mathbf{R}\rangle_{\text{av}} = -2K\sigma^2\langle\mathbf{R}^T\Phi\mathbf{R}\rangle_{\text{av}}, \tag{6}$$

where σ is the standard deviation of $x_i(t)$ (assumed identical for all the $x_i(t)$) and Φ is the normalized NxN correlation matrix of the $x_i(t)$. The angular bracket denotes ensemble averaging. In case (*ii*)

$$\frac{d}{dt}\langle\mathbf{R}^T\mathbf{R}\rangle_{\text{av}} = -2K\sigma\sqrt{\frac{2}{\pi}}\langle\mathbf{R}^T\Phi\mathbf{R}\rangle_{\text{av}}. \tag{7}$$

These equations follow easily since any linear combination of Gaussian variables is another Gaussian variable. Equations (6) and (7) give upper and lower bounds to the convergence rates for the two cases, when we observe that

$$\lambda_{\min}\mathbf{R}^T\mathbf{R} \leqq \mathbf{R}^T\Phi\mathbf{R} \leqq \lambda_{\max}\mathbf{R}^T\mathbf{R},$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and largest eigenvalues of Φ. If $x(t)$ is white noise, then Φ is an identity matrix with $\lambda_{\min} = \lambda_{\max} = 1$. Thus, for case (*i*)

$$\sqrt{\mathbf{R}^T\mathbf{R}}\,\Big|_{t=0}\exp\,(-K\sigma^2\lambda_{\max}t) \leqq \sqrt{\mathbf{R}^T\mathbf{R}} \leqq \sqrt{\mathbf{R}^T\mathbf{R}}\,\Big|_{t=0}\exp\,(-K\sigma^2\lambda_{\min}t) \tag{8}$$

and in case (*ii*)

$$\sqrt{\mathbf{R}^T\mathbf{R}\,|_{t=0}} - \sqrt{\frac{2}{\pi}}\,K\sigma\lambda_{\max}^{\frac{1}{2}}t \leqq \sqrt{\mathbf{R}^T\mathbf{R}} \leqq \sqrt{\mathbf{R}^T\mathbf{R}\,|_{t=0}} - \sqrt{\frac{2}{\pi}}\,K\sigma\lambda_{\min}^{\frac{1}{2}}t. \tag{9}$$

These upper and lower bounds are close to each other only if Φ is nearly diagonal (i.e., when $x(t)$ is broadband). More elaborate methods can be used to estimate the convergence rate (e.g., perturbation methods). However, although these methods would be extremely interesting from a theoretical point of view, they are not likely to yield much more insight into the convergence process. This is especially true in view of the fact that no satisfactory statistical description of a speech signal is available at present. For Gaussian noise (8) and (9) have been checked by computer simulation.

V. CHOICE OF NONLINEARITY

In the formulation of the echo suppressor problem discussed in Section III, the choice of the nonlinearity F depends upon the choice of the function C. This choice has a profound influence upon the behavior of the resulting system. One could set up the problem of determining the optimum F which would provide, according to some reasonable criterion, the fastest convergence and the maximum immunity from interfering noise. We do not know the solution to such a general optimization problem. In any case, since convergence rate depends upon the statistics of the signal and noise, any such optimization would be practically impossible for signals as difficult to characterize as speech signals. We have, however, found that a number of improvements over the linear case result upon making F an infinite clipper. The use of an infinite clipper has also been suggested independently by B. F. Logan.

One of the main drawbacks of making $F(e) = e$ (i.e., Kelly's proposal) is the dependence of the time constant of the control system on the signal level. Equation (6), although approximate, nevertheless indicates that the time constant (at least for a wide band signal) is proportional to the signal power. A 20-dB change in signal level thus changes the time constant by a factor of 100. If, however, an infinite clipper is used the same change in signal level changes the time constant by a factor of only 10, which is a considerable improvement.

Another important advantage of using an infinite clipper is the considerable simplification of the circuitry. The multipliers $M_{11} \cdots M_{1N}$ can all be replaced by switch modulators, which are far cheaper and simpler than broadband multipliers.

There is another advantage in using an infinite clipper which may be described as follows. Suppose the system is near equilibrium and the echo has been reduced to a very low value. If now there is a sudden burst of noise (say a spurt of double-talking) then in the linear system the rate at which the system moves away from equilibrium is proportional to the level of the noise, whereas it is more or less independent of noise level if an infinite clipper is used. (In Section VI we will give a detailed example of this effect.) We will argue in Section VII that during intervals of double-talking the control loop be opened and the vector **G** frozen at its last value. However, any decision mechanism that would make this possible would require a finite time to make the decision. It is therefore important that the system should not depart from equilibrium too rapidly upon the introduction of a large noise in the return signal.

VI. OTHER MODIFICATIONS

It may be noted that although we started out by taking the components of the vector $\mathbf{X}(t)$ to be delayed versions of the input $x(t)$, this fact was nowhere of any importance in the proof of convergence. All that is required is that the vector $\mathbf{X}(t)$ be derived from $x(t)$ in such a way that all transformations of $x(t)$, that may possibly be produced by the round trip transmission path, should be representable as $\mathbf{H}^T\mathbf{X}$ with a suitable choice of H. This immediately gives us the possibility of generalizing the circuit of Fig. 2 to that of Fig. 3. In Fig. 3 the $w_i(t)$ are a set of impulse responses such that linear combinations of them are good approximations to most practical echo path impulse responses.

Now there is an infinite variety of sequences of functions that are complete on the semi-infinite interval. One such set is the set of Laguerre functions which is of particular interest because it can be synthesized as a simple tapped RC ladder network. The impulse response of the nth Laguerre network is given by

$$l_n(t) = e^{\alpha t} \frac{d^n}{dt^n} \left(\frac{t^n}{n!} e^{-2\alpha t} \right) \tag{10}$$

with the corresponding transfer function

$$L_n(s) = \frac{\alpha}{s + \alpha} \left(\frac{\alpha - s}{\alpha + s} \right)^n. \tag{11}$$

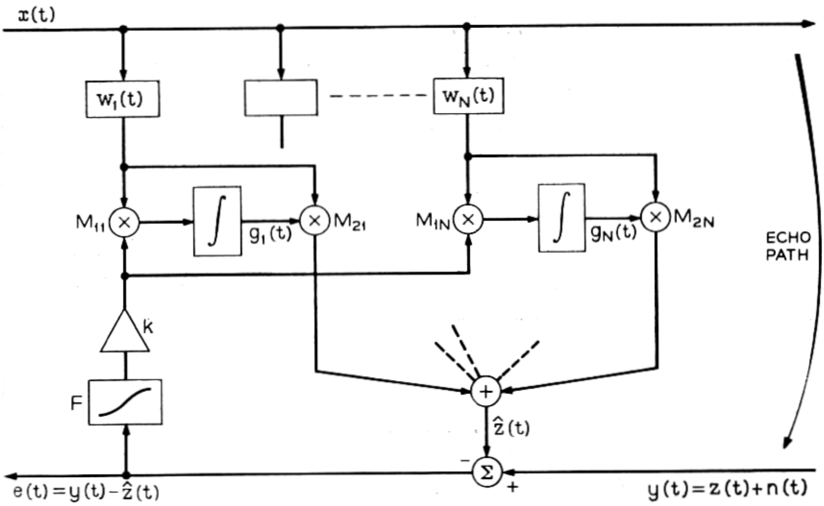

Fig. 3 — Generalization of the system of Fig. 2 using an orthonormal set of impulse responses.

As telephone speech is limited to about 3 kHz the choice of α appropriate in the present case is approximately $2\pi \times 2000$ radians per second, although it is not critical.

Tests by computer simulation, to be described in the next section, indicate that Laguerre networks are at least as satisfactory as a tapped delay line for the simulation of the echo path. However, a cascade of RC sections would be much cheaper than a delay line. The properties and synthesis of Laguerre networks is described in the literature.[6]

VII. COMPUTER SIMULATION

For a computer simulation of the system described by (2), it was converted to a difference equation. Thus, if a subscript n on a quantity is used to denote its value at the nth sampling instant, then the equation simulated on the computer is

$$\mathbf{G}_{n+1} = \mathbf{G}_n + KF(\mathbf{R}_n^T\mathbf{X}_n + n_n)\mathbf{X}_n.$$

In one class of simulations we used filtered Gaussian noise as the signal and computer-simulated echo paths. Equations (8) and (9) appear to be very good approximations if the time constant of the convergence (which we may define as the time taken for $l_R^2 = \mathbf{R}^T\mathbf{R}$ to become, say 30 percent of its initial value) is large compared to the reciprocal of the bandwidth of the input signal. If uncorrelated white noise is added to the echo before the system has converged then no convergence takes place if the noise level is about 15 dB *above* the echo. If the same noise is introduced after the system has converged, however, the balance is only slightly disturbed. A typical example will illustrate the orders of magnitude of these effects. The nonlinearity F was chosen to be an infinite clipper, and the level of the signal (which was a white gaussian noise) and the constant K were such that in the absence of noise $\mathbf{R}^T\mathbf{R}$ converged to about 55 dB below its initial value in 0.7 sec. The following two tests were performed:

(*i*) The same input signal, initial conditions, etc. were used as before, but an uncorrelated noise was added to the return signal at a level 18 dB above the echo.

(*ii*) Same as (*i*) except the noise was added after the system had converged for 0.6 sec so that l_R^2 (hence, the echo power) was about 23 dB below its initial value.

In case (*i*) no convergence took place and l_R^2 hovered around its initial value. In case (*ii*), after the onset of noise, l_R^2 increased slowly by about 3 dB in 1.5 seconds.

For comparison the same simulations were repeated with F replaced by a linear function and the constant K adjusted to give a time constant of the exponential decay of about 0.3 sec. The noiseless case and test (*i*) gave about the same results, except, of course, that the decay was exponential. In test (*ii*) the noise was introduced after 1.1 sec instead of 0.6, to allow l_R^2 to converge to about 30 dB. However, in this case, after the introduction of the noise l_R^2 rose by about 20 dB within 0.2 seconds. Thereafter it stayed at about this level.

The same kind of behavior was obtained when speech signals were used both as input and as interfering noise and when the echos were generated by the computer. Of course, it is much more difficult to estimate time constants of the convergence when speech signals are used.

We also used as inputs, digitized tape recordings of sentences spoken over an N2 carrier system. Two-track tapes were made of the input signal and the echo. Double-talking situations were also recorded and tested. For tests with these tapes, 40 to 50 delay line taps spaced 0.1 ms apart were used.

With very strong echos (0-dB return loss) the system provided a reduction of about 20 to 25 dB as measured on a VU meter. This reduction took place in 0.5 to 5 seconds depending on signal level as discussed in Section IV. The larger variation of convergence rate with signal level when the clipper is removed, was apparent in these tests also.

When a typical hybrid and a return loss of 6 dB were used the echo was reduced to the point of being unintelligible and almost inaudible even under quiet conditions. As in the case of tests with noise signals, the introduction of double-talking at a very high level after convergence had taken place produced little change in the balance.

The fact that in the case of these recordings over the N-carrier system the echo could not be reduced by more than about 20 dB or so is undoubtedly due to the compandors used in the N-system. The echo canceller can provide only a linear approximation to the transmission path of the echo, which in the case of the N-system has nonlinearities.

We have also simulated the echo canceller using the Laguerre expansion. In this case, the digital equivalent of the Laguerre impulse responses were used. In terms of the delay parameter, $z^{-1} = \exp(-j\omega T_s)$ where T_s is the sampling interval, these are

$$L_n(z^{-1}) = \frac{1}{1 - az^{-1}} \left(\frac{z^{-1} - a}{1 - az^{-1}} \right)^n,$$

where $a = (2 - \alpha T_s)/(2 + \alpha T_s)$, α being as defined in Section V. As mentioned earlier α is chosen so that the cut off frequency of the Laguerre function is about 2 or 2.3 kHz.

VIII. DISCUSSION AND CONCLUSIONS

We have described a new method of cancelling echos in telephone connections. From our theoretical discussion and simulations it appears that the method is feasible and can yield echo cancellation of about 20 dB or so with a convergence time of about 0.2 to 0.5 second for average speech levels. This convergence time increases to 10 times its value for a 20-dB decrease in signal level. Convergence much faster than this is not possible, as then the system becomes too sensitive to noise and behaves erratically with normal noise level to be expected on a telephone connection.

We have shown that the system would not appreciably depart from equilibrium even upon the incidence of double-talking. However, it needs, initially, a period of time in which only the echo is present in the return signal (of course low-level noise may be present also, but there should be no double-talking in this period). This initial interval can be as small as 0.5 seconds if the input signal is loud, but would have to be proportionately longer for weaker input signals.

It would be advisable to break the control loop during bursts of loud double-talking. This need not be done by very sophisticated means. The following simple method should be satisfactory. Assume that the maximum level of an echo is 6 dB below the input signal. Clearly if the return signal is much larger than this it indicates double-talking. Rectification and integration with a time constant of about 0.5 second gives a reasonably good estimate of the levels of the input and return signals. A switch could then be adjusted to open the feedback path in Fig. 3 immediately following the nonlinearity F, whenever the input level is less than, say, 3 dB above that of the return signal. It is important to note that this merely prevents the gain setting $\mathbf{G}$ from changing. It *does not* interrupt the return path.

We have tested our system only on an N2 carrier system (besides on artificially generated echos). This is a double-sideband modulation system in which compandors are used. There are also in use single-sideband carrier systems, in which no compandors are used but in which carrier frequency variation (during the round-trip transmission time) would introduce a time-varying nonlinearity. The degree to which this type of variation exists and its effect require further investigation.

The delay between the input signal and its echo must be compensated for. This delay may be as large as 60 ms. The problem of automatically determining this delay and compensating for it is a challenging problem and is being investigated by a systems group at Bell Telephone Laboratories. They are also collecting a large sample of impulse responses of various connections. This information would be very useful in the final design of the system. For example, this will enable us to decide upon the optimum number of taps. Also, a fixed weighting of the gain vector $\mathbf{G}$ depending upon the statistical distribution of the impulse responses would improve the average performance of the system.

The ultimate test of the system's performance and usefulness is its actual use during normal long distance telephone conversations. For this, actual hardware must be built. Two, rather different, instrumentations have been recently completed, one by A. J. Presti[7] and the other by F. K. Becker and H. R. Rudin,[8] and it should be soon possible to carry out such tests.

REFERENCES

1. Helder, G. K., Customer Evaluation of Telephone Circuits with Delay, B.S.T.J., *45*, September, 1966, pp. 1157–1191.
2. Emling, J. W. and Mitchell, D., Effects of Time Delay and Echoes on Telephone Conversations, B.S.T.J., *42*, November, 1963, pp. 2869–2891.
3. Brady, P. T. and Helder, G. K., Echo Suppressor Design in Telephone Communications, B.S.T.J., *42*, November, 1963, pp. 2893–2917.

4. Riesz, R. R. and Klemmer, E. T., Subjective Evaluation of Delay and Echo Suppressors in Telephone Communications, B.S.T.J., *42*, November, 1963, pp. 2919–2941.
5. Flanagan, J. L. and deBarbeyrac, J., unpublished memorandum.
6. Lee, Y. W., *Statistical Theory of Communication*, John Wiley & Sons, Inc., 1960.
7. Sondhi, M. M. and Presti, A. J., A Self-Adaptive Echo Canceller, B.S.T.J., *45*, December, 1966, pp. 1851–1854, also presented at the Acoustical Society of America, 72nd meeting, November 3, 1966.
8. Becker, F. K. and Rudin, H. R., Application of Automatic Transversal Filters to the Problem of Echo Suppression, B.S.T.J., *45*, December, 1966, pp. 1847–1850, also presented at the Acoustical Society of America, 72nd meeting, November 3, 1966.