UNCLASSIFIED

<table>
<thead>
<tr>
<th>AD NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD841587</td>
</tr>
</tbody>
</table>

LIMITATION CHANGES

TO:
Approved for public release; distribution is unlimited. Document partially illegible.

FROM:
Distribution authorized to U.S. Gov't. agencies only; Administrative/Operational Use; JUL 1968. Other requests shall be referred to Office of Naval Research, Code 462, Washington, DC 20360. Document partially illegible.

AUTHORITY

onr ltr, 9 nov 1973
GENERAL DYNAMICS CORP GROTON CONN ELECTRIC BOAT DIV

PROCESSING OF DATA FROM SONAR SYSTEMS. VOLUME V SUPPLEMENT 1. (U)

DESCRIPTIVE NOTE: ANNUAL REPT. 1 JUL 66-1 JUL 67, JUL 68 206P

CHANG, JOHN H.; McDONALD, VERNE H.; SCHULTHEISS, PETER M.; ITUTEUR, FRANZ

REPT. NO. U417-68-079

CONTRACT: NONR-2512(00)

UNCLASSIFIED REPORT DISTRIBUTION: USGO OTHERS TO CHIEF, OFFICE OF

NAVAL RESEARCH, ATTN: CODE 462, WASHINGTON, D.C. 20360

SUPPLEMENTARY NOTE: PREPARED IN COOPERATION WITH YALE

UNIV, NEW HAVEN, CONN DEPT OF ENGINEERING AND APPLIED SCIENCE. SEE ALSO AD-831 579.

DESCRIPTORS: (SONAR SIGNALS, DATA PROCESSING SYSTEMS), (UNDERWATER OBJECTS, DETECTION), SONAR ARRAYS, CLIPPER CIRCUITS, NOISE, INTERFERENCE, UNDERWATER TRACKING, HYDROPHONES, DELAY LINES, ELECTRIC FILTERS, CORRELATORS, SIGNAL-TO-NOISE RATIO, DOPPLER EFFECT, UNDERWATER SOUND, SOUND TRANSMISSION, SONAR TARGETS, REPORTS

IDENTIFIERS: SUBIC PROJECT, CLIPPING (ELECTRONICS), SUBIC (SUBMARINE INTEGRATED CONTROL), SUBMARINE INTEGRATED CONTROLS

THIS DOCUMENT PURSUES THE GENERAL SUBJECT OF PASSIVE SONARS OPERATING IN AN ANISOTROPIC NOISE ENVIRONMENT, AND HAS TAKEN TWO NEW DIRECTIONS. ENVIRONMENTS CONTAINING NOT ONE BUT SEVERAL POINT SOURCES OF INTERFERENCE OR A SPATIALLY DISTRIBUTED INTERFERENCE ARE STUDIED. CONVENTIONAL AS WELL AS OPTIMAL DETECTORS WERE ANALYZED. THE SECOND DIRECTION WAS AN EXAMINATION OF THE EFFECT OF SINGLE PLANE WAVE INTERFERENCE ON TRACKING ACCURACY. THIS VOLUME ALSO CONTINUES THE STUDY OF ACTIVE SONAR SYSTEMS INITIATED IN VOLUME IV, AND INITIATES AN EFFORT TO DEAL WITH THE SIGNAL DETECTION AND EXTRACTION PROBLEM IN A NOISE ENVIRONMENT WHOSE STATISTICAL PROPERTIES ARE LARGELY OR WHOLLY UNKNOWN. (AUTHOR)
Volume V, and its supplement, further pursues the general subject of passive sonars operating in an anisotropic noise environment, and has taken two new directions. Environments containing not one but several point sources of interference or a spatially distributed interference are studied. Conventional as well as optimal detectors were analyzed. The second direction was an examination of the effect of single plane wave interference on tracking accuracy. This volume also continues the study of active sonar systems initiated in Volume IV, and initiates an effort to deal with the signal detection and extraction problem in a noise environment whose statistical properties are largely or wholly unknown.
<table>
<thead>
<tr>
<th>KEY ABDUS</th>
<th>LINK A</th>
<th>ROLE</th>
<th>WT</th>
<th>LINK B</th>
<th>ROLE</th>
<th>WT</th>
<th>LINK C</th>
<th>ROLE</th>
<th>WT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.
GENERAL DYNAMICS CORPORATION
Electric Boat division
Groton, Connecticut

PROCESSING OF DATA FROM
SONAR SYSTEMS

VOLUME V
SUPPLEMENT 1

by

John H. Chang
Verne H. McDonald
Peter M. Schultheiss
Franz B. Tuteur
Yale University

Examined: J. W. Herring
SU3IC Project Manager

Approved: Dr. A. J. van Woerkom
Chief Scientist

This document has been approved for public release and sale — distribution is unlimited.

U417-68-079
July 31, 1968
<table>
<thead>
<tr>
<th>Report No.</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>The Effect of Clipping on the Performance of Replica Correlators</td>
<td>A-1</td>
</tr>
<tr>
<td>32</td>
<td>Some Comments on Optimum Bearing Extraction</td>
<td>B-1</td>
</tr>
<tr>
<td>33</td>
<td>The Effect of Noise Anisotropy on Detectability in an Optimum Array Processor</td>
<td>C-1</td>
</tr>
<tr>
<td>34</td>
<td>Methods of Stochastic Approximation Applied to the Analysis of Adaptive Tapped Delay Fine Filters</td>
<td>D-1</td>
</tr>
</tbody>
</table>
This is an unclassified supplement to Volume V of a series of reports describing work performed by Yale University under subcontract to Electric Boat division of General Dynamics Corporation. Volume V and this supplement cover the period 1 July 1966 to 1 July 1967. Electric Boat is prime contractor of the SUBIC (SUBmarine Integrated Control) Program under Office of Naval Research contract N00014-66-C-0012(00).

LCDR. E. W. Lull, USN, is Project Officer for ONR; J. W. Herring is Project Manager for Electric Boat division under the direction of Dr. A. J. van Woerkom.
THE EFFECT OF CLIPPING ON THE
PERFORMANCE OF REPLICA CORRELATORS

by

Peter M. Schultheiss

Progress Report No. 31
General Dynamics/Electric Boat Research
(8050-31-55001)

June 1967

DEPARTMENT OF ENGINEERING
AND APPLIED SCIENCE
YALE UNIVERSITY
Summary

This report deals with the effect of clipping on the performance of an active sonar system using conventional beamforming techniques followed by replica correlation. Detection as well as range and Doppler estimation are considered.

Two basic assumptions are made throughout the analysis:

a) The noise field (ambient or reverberation) is Gaussian and has the same power level at each hydrophone.

b) The input signal-to-noise ratio at each hydrophone is small.

In addition, much, though not all, of the work assumes a transmitted signal narrow in bandwidth compared with its center frequency. The model for reverberation noise postulates a series of stationary, Poisson distributed scattering centers. The array geometry is quite arbitrary, but certain computations require beam patterns narrow enough to permit approximations of the form \( \sin \theta = \theta \) over the effective dimensions of the pattern.

The quantity of primary interest is the clipping loss \( R \), defined as the output signal-to-noise ratio of the clipped instrumentation divided by the output signal-to-noise ratio of the unclipped (but otherwise identical) instrumentation.

The following results are obtained for detection:

1) If the noise (ambient or reverberation) is independent from hydrophone to hydrophone one can demonstrate with complete generality that \( R \leq 1 \). One can further demonstrate that, without additional restrictions on signal and noise, a lower bound of \( R = 0 \) can be approached arbitrarily closely. To set meaningful bounds on clipping loss one must therefore restrict the class of admissible signals and...
noise. A practically realistic and analytically fruitful restriction is the assumption of narrow-band signals, which underlies all remaining results.

2) If the noise (ambient or reverberation) is independent from hydrophone to hydrophone and if signal and noise are confined to the same frequency band (narrow compared with the center frequency) the clipping loss is bounded by \( 0.89 \leq R \leq 1 \).

3) If the noise does not fall into the same frequency band as the signal, large clipping losses can occur. This is practically important in a reverberation limited environment when the target return is subject to large Doppler shifts. In such situations \( R \) can approach arbitrarily close to zero if the Doppler shift is large enough.

4) If statistical dependences are allowed between the noise at different hydrophones, one requires further restrictions before useful lower bounds can be set on \( R \). An example is worked out to demonstrate that, even with purely isotropic ambient noise, values of \( R \) appreciably below 0.89 may be obtained. However, the example requires such careful matching of array geometry with carrier wavelength that it appears to fall more into the category of analytical pathologies than into that of practically important situations. A search (by no means exhaustive) for more realistic examples in which ambient noise would produce values of \( R \) below 0.89 led to negative results.

5) Probably the most important situation from a practical point of view is that of a reverberation limited environment. It was...
therefore studied in some detail. The most useful results were obtained under the assumption that the array dimensions are small compared with the wavelength of the highest modulating frequency, (i.e., the wavelength of the maximum frequency deviation from the carrier). The effect of this assumption is to permit complete separation of spatial and temporal effects in the reverberation. In the absence of target Doppler shifts one can then establish with considerable generality that \( R \geq 0.89 \). In the presence of target Doppler shifts one has, of course, the phenomenon discussed in 3). The key assumption concerning array dimensions can be weakened considerably if the beam pattern is narrow.

When one considers range and Doppler measurements one finds, not surprisingly, that the exact clipping loss depends to some extent on the specific instrumentation. It is therefore not possible to draw conclusions of quite the same generality as in the analysis of detection. Range (Doppler shift) is measured by cross-correlating the target return with a replica of the transmitted signal and "locating" the resulting correlation function in time (frequency). Different instrumentations result from different functional definitions of the term "location". It appears reasonable to speculate - and several sample computations tend to confirm this - that most reasonable definitions of "location" would lead to rather similar instrumentations, and in particular to instrumentations with very similar sensitivity to clipping. This report concerns itself primarily with range (Doppler) measurement in a reverberation limited environment. Range (Doppler shift) is measured by comparing
cross-correlations with two replicas of slightly different delay (frequency). Doppler shift is assumed known during the range measurement and range during the Doppler measurement. Clipping loss is defined as the ratio of the rms range (Doppler) errors with and without clipping. The results are

6) Under the conditions specified in 5), the clipping loss factor for range measurement has a lower bound not significantly different from the figure of 0.89 obtained for detection.

7) As might be anticipated from 3), the clipping loss in Doppler measurement depends heavily on the target Doppler shift. Separating out this effect by working with zero target Doppler shift, one can again show that the clipping loss factor has a lower bound close to 0.89.

Combining all of the above, one is lead to the following general conclusion: Serious clipping losses arise in a reverberation limited environment when the target Doppler shift is large enough to move the target return largely out of the reverberation band. In most other practically interesting situations the clipping loss in detection, ranging and Doppler estimation is limited to a factor of the order of 0.89, equivalent to about 1 db of input signal-to-noise ratio.
1. Introduction

This report is concerned with the effect of clipping on the performance of an active sonar array. Correlation with a replica (delayed and Doppler shifted) of the transmitted signal is used as the basic signal processing technique.

The general block diagram for detection is shown in Figure 1. The array geometry is entirely general.

![Figure 1]

The delays \( t_i \) bring the signal components of all hydrophone outputs into alignment. The clippers\(^1\) then generate a set of signals \( \{x_i(t)\} \) given

\[^1\]In practice the clippers would probably precede the delays (which would then be digital). It is clear from physical reasoning that this interchange would not alter the \( x_i(t) \).

A-5
by

\[ x_i(t) = \text{sgn}[s(t)+n_i(t-\tau_i)] = \begin{cases} 
1 & \text{if } s(t)+n_i(t-\tau_i) > 0 \\
0 & \text{if } s(t)-n_i(t-\tau_i) < 0
\end{cases} \] (1)

The \( x_i(t) \) are summed and multiplied by a suitable replica of \( s(t) \).

The resulting \( y(t) \) is finally smoothed over a period \( T \) comparable to the duration of the signal. If the noise has zero mean, the average output \( \bar{z} \) is zero in the absence of a signal component at the hydrophones.

When a signal component is present at the hydrophones, \( \bar{z} \) will differ from zero. Hence, if the output \( z(t) \) at time \( t \) exceeds a preset threshold (depending on signal-to-noise ratio and allowed false alarm rate) one concludes that a target is indeed present.

To measure range (or Doppler) one might employ two replicas with slightly different delays (or Doppler shifts), multiply each by the clipped and summed hydrophone outputs \( x \) and use the difference between the resulting \( y \)'s as a measure of range (or Doppler shift). A possible implementation is discussed in somewhat more detail at a later point.

The noise field is assumed to be Gaussian. This is reasonable even in a reverberation limited environment as long as no major portion of the noise power is contributed by large scatterers (false targets). [See Report No. 27].
11. General Relations for Detection

The output $y(t)$ of the multiplier can be written in the form

$$y(t) = \sum_{i=1}^{M} x_i(t) s(t) = \sum_{i=1}^{M} \text{sgn}[s(t)+n_i(t-T_i)] s(t)$$  \hspace{1cm} (2)

The delay of the replica is here assumed to be perfectly matched to the signal delay.

The mean value of the detector output is

$$\bar{z} = \sum_{i=1}^{M} \frac{1}{T} \int_{0}^{T} s(t) \text{sgn}[s(t)+n_i(t-T_i)] \, dt$$  \hspace{1cm} (3)

Now

$$\text{sgn}[s(t)+n_i(t-T_i)] = \Pr\{[s(t)+n_i(t-T_i)] > 0\} - \Pr\{[s(t)+n_i(t-T_i)] < 0\}$$  \hspace{1cm} (4)

For zero-mean Gaussian $n_i(t)$

$$\Pr\{[s(t)+n_i(t-T_i)] > 0\} = \frac{1}{\sqrt{2\pi N}} \int_{0}^{\infty} e^{-\frac{(n-s(t))^2}{2N}} \, dn = \frac{1}{2} \left[ 1 + \text{erf} \frac{s(t)}{\sqrt{2N}} \right]$$  \hspace{1cm} (5)

where

$$\text{erf} v = \frac{2}{\sqrt{\pi}} \int_{0}^{v} e^{-x^2} \, dx$$  \hspace{1cm} (6)

and $N$ is the average noise power.

Similarly

$$\Pr\{[s(t)+n_i(t-T_i)] < 0\} = \frac{1}{\sqrt{2\pi N}} \int_{-\infty}^{0} e^{-\frac{(n-s(t))^2}{2N}} \, dn = \frac{1}{2} \left[ 1 - \text{erf} \frac{s(t)}{\sqrt{2N}} \right]$$  \hspace{1cm} (7)
It follows that

\[
\text{sgn}[s(t)+n(t-T)] = \frac{\text{erf} \left( \frac{s(t)}{\sqrt{2N}} \right)}{2N}
\]

(8)

when \( \max s(t) \ll \sqrt{2N} \)

\[
\text{sgn}[s(t)+n(t-T)] = \frac{\sqrt{2}}{\pi} \frac{s(t)}{\sqrt{N}}
\]

(9)

Hence

\[
\frac{1}{T} \int_0^T \text{dt} \, s^2(t) \quad \text{for} \quad \max s(t) \ll \sqrt{2N}
\]

(10)

For a figure of merit of the detector we choose as usual the output signal-to-noise ratio, i.e., the average output due to the signal divided by the rms value of the output fluctuation. If the input signal-to-noise ratio is low, the output fluctuation is very largely due to the noise component of the input, so that

\[
z^2(t) = \sum_{i=1}^{M} \sum_{j=1}^{M} \frac{1}{T^2} \int_0^T \text{dt} \int_0^T \text{d}\lambda \, s(t) \, s(\lambda) \, \text{sgn}[s(t)+n_i(t-\tau_i)] \, \text{sgn}[s(\lambda)+n_j(\lambda-\tau_j)]
\]

(11)

From a well-known result in noise theory

\[
\text{sgn}[n_i(t)] \, \text{sgn}[n_j(\lambda)] = \frac{2}{\pi} \sin^{-1} \left[ \rho_{i,j}(t-\lambda) \right]
\]

(12)

where \( \rho_{i,j}(\tau) \) is the normalized cross-correlation function of the noise received at the \( i \)th and \( j \)th hydrophones. Hence, for low input signal-
The signal-to-noise ratio is

\[ r^2(t) = \frac{1}{T^2} \int_0^T \int \sum_{i=1}^M \sum_{j=1}^N d\lambda s(\lambda) \sin^{-1}\left[ \rho_{ij}(t-\lambda-\tau_i+\tau_j) \right] \] (13)

From Equations (10) and (13) the desired output signal-to-noise ratio is

\[
\frac{S}{N}_{\text{clipped}} = \frac{1}{T^2} \int_0^T s^2(t) dt
\]

(14)

In the absence of clipping

\[
\overline{z} = \frac{M}{1} \int_0^T s^2(t) dt
\]

(15)

and

\[
\overline{z}^2 = \frac{1}{T^2} \sum_{i=1}^M \sum_{j=1}^N \int dt \int d\lambda s(\lambda) s(t) N_{ij}(t-\lambda-\tau_i+\tau_j)
\]

(16)

where

\[ N_{ij}(t-\lambda) = \overline{n_i(t) n_j(\lambda)} \]

(17)

Hence, the output signal-to-noise ratio in the unclipped case is

\[
\frac{S}{N}_{\text{unclipped}} = \frac{1}{T^2} \int_0^T s^2(t) dt
\]

(18)
The present study is concerned with performance degradation due to clipping.

Hence the ratio \( R \) of Equations (14) and (18) serves as a convenient criterion.

\[
R = \frac{(S/N)_0 \text{ clipped}}{(S/N)_0 \text{ unclipped}} = \sqrt{\frac{\int_0^T dt \, s(t) \int_0^T ds(\lambda) \sum_{i=1}^{M} \sum_{j=1}^{M} N_{ij} \rho_{ij}(t-\tau_i+\tau_j)}{\int_0^T dt \, s(t) \int_0^T ds(\lambda) \sum_{i=1}^{M} \sum_{j=1}^{M} N \sin^{-1} \left[ \rho_{ij}(t-\lambda-\tau_i+\tau_j) \right]}}
\]

(19)

If the average noise power is the same at each hydrophone, then

\( N_{ij} = N \) and Equation (19) reduces to

\[
R = \sqrt{\frac{\int_0^T dt \, s(t) \int_0^T ds(\lambda) \sum_{i=1}^{M} \sum_{j=1}^{M} \rho_{ij}(t-\tau_i+\tau_j)}{\int_0^T dt \, s(t) \int_0^T ds(\lambda) \sum_{i=1}^{M} \sum_{j=1}^{M} \sin^{-1} \left[ \rho_{ij}(t-\lambda-\tau_i+\tau_j) \right]}}
\]

(20)
III. Noise Independent from Hydrophone to Hydrophone

An important special case is that of noise independent from hydrophone to hydrophone. In that situation, Equation (20) reduces to

\[
\begin{align*}
R &= \sqrt{\frac{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \rho(t-\lambda)}{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \sin^{-1}[\rho(t-\lambda)]}} \\
&= \sqrt{\frac{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \rho(t-\lambda)}{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \sin^{-1}[\rho(t-\lambda)]}}
\end{align*}
\]

where \( \rho(t) = \rho_{ii}(t) \), \( i = 1, 2 \ldots M \), is the normalized autocorrelation function of the noise at each hydrophone.

It is a simple matter to show that Equation (21) can never exceed unity. For greater ease in manipulation, consider the quantity \( 1/R^2 \) and expand the inverse sine into a power series (convergent for all values of \( t \) and \( \lambda \) since \( \rho(t-\lambda) \leq 1 \))

\[
\begin{align*}
\frac{1}{R^2} &= \frac{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) [\rho(t-\lambda) + 1/2 \times 1/3 \times 3/4 \times 1/5 \times 5/6 \times \ldots ]}{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \rho(t-\lambda)} \\
&= 1 + 1/6 \frac{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \rho^3(t-\lambda)}{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \rho(t-\lambda)} + \frac{3}{40} \frac{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \rho^5(t-\lambda)}{\int_{0}^{T} dt \int_{0}^{T} d\lambda \; s(t) s(\lambda) \rho(t-\lambda)} + \ldots
\end{align*}
\]

Since \( \rho(t) \), being a correlation function, is positive definite. Hence, the
denominator in each term is positive, regardless of the form of \( s(t) \).

Furthermore, \( \rho(t) \) is Fourier inverse of a non-negative spectral function \( G(w) \). It follows that \( \rho^n(t) \) is the Fourier inverse of the n-fold convolution of \( G(w) \) with itself. This convolution is clearly non-negative, so that \( \rho^n(t) \) is positive definite. Hence, all the numerators in equation (22) are non-negative. Thus all terms of the equation are non-negative and one concludes:

\[
\frac{1}{2} \geq 1
\]

or

\[
\gamma^2 \geq 1 \quad (23)
\]

In order to set a lower bound on \( R \) it is necessary to make subsidiary assumptions. In the absence of any restrictions on signal and noise properties one might postulate a signal confined to one frequency band and a noise confined to a different, disjoint frequency band. In such an environment the unclipped detector would be essentially perfect,\(^1\) while the clipping process could shift appreciable noise power into the signal band. Thus values of \( R \) arbitrarily close to zero could be obtained in principle by sufficiently artificial choices of signal and noise spectra. Reasoning more formally, one can proceed as follows to establish the impossibility of finding a general lower bound on \( R \) in excess of zero:

Replace the integrals in equation (22) by sums. The implied sampling in time can be very rapid, so that the approximation is arbitrarily good.

\(^1\)Limited only by the feasibility of maintaining perfectly disjoint spectra with a signal of finite duration.
A typical term in Equation (22) can then be written in the form

$$A_n = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} s_i s_j$$

where

$$s_i = \frac{1 \cdot 5 \cdot 9 \cdots (n - 2)}{2 \cdot 4 \cdot 6 \cdots (n - 1)} \cdot \frac{1}{n}$$

and

$$a_{ij} = \delta(t_i - t_j) : s_i = s(t_i)$$

Then

$$A_n = \sum_{i=1}^{n} \sum_{j=1}^{n} s_i s_j$$

so that

$$A_n = \left\{ \begin{array}{ll}
\sum_{i=1}^{n} \sum_{j=1}^{n} \left[ (a_{ij} - a_{ij}) s_i s_j \right] & \\
\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} s_i s_j & 
\end{array} \right\}$$

Now choose

$$s_i = \begin{cases} 
1 & \text{for } i = k \\
-1 & \text{for } i = k \\
0 & \text{all other } i
\end{cases}$$

then

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} s_i s_j = 0$$

and one obtains for the chosen $s_i$
\[ A = n \left( 1 + \frac{d_{11}}{d_{22}} \right) \]

(28)

It follows that

\[ R = n \left( 1 - \frac{d_{22}}{d_{11}} \right) \]

(29)

This upper bound on \( R \) can be diminished to zero by choosing the times \( t_k \) and \( t_{k-1} \) appropriately close together. With \( A_n \) at this upper bound, equation (2) reads

\[ 1 \leq \frac{1}{n} \sum_{j=1}^{n} \left( \frac{1}{x_{j} - 1} \right)^{1/2} \]

(30)

Thus, for any fixed \( m \), it is possible to find a signal \( s(t) \) which causes \( R \) to be arbitrarily close to zero.

Most practical cases are not necessarily same, are very different from the cases just described, one possible active linear signal would almost certainly be of the form

\[ s(t) = \sum_{j=1}^{m} s_{j}(t) \sqrt{ \left( x_{j} - 1 \right)^{1/2} } \]

(31)

where \( s_{j}(t) \) is the signal in the \( j \)-th channel, \( \sqrt{ \left( x_{j} - 1 \right)^{1/2} } \) determining the pulse shape and \( \sum_{j=1}^{m} \sqrt{ \left( x_{j} - 1 \right)^{1/2} } \) the total energy in the modulation. The bandwidth of \( s_{j}(t) \) and \( \sqrt{ \left( x_{j} - 1 \right)^{1/2} } \) can be considered as small compared with \( s_{j}(t) \), however, as small.
Given a narrowband signal, it appears reasonable to assume that the
receiving narrowband process also. For if the noise is primarily rever-
sion, its spectral properties are determined by the signal, while in
an ambient noise limited environment one would employ filters matched to
the bandwidth of the signal in order to improve signal-to-noise ratio. If
there is no Doppler shift in the target signal the center frequency of the
noise will be $\omega_0$. Then $s(t)$ would assume the form

$$s(t) = a(t) \cos \omega_0 t$$

(32)

The bandwidth of $a(t)$ is small compared with $\omega_0$. Clearly $\rho_1(0) = 1$
and $\rho_1(t) = 1$ for all $t$. Typical forms of $\rho_1(t)$ might be

$$\rho_1(t) = e^{-\alpha t^2}, \quad \phi_1(t) = e^{-\alpha t^2}$$

both of which have the computationally simple property $\rho_1(t) \geq 0$ for all $t$. This property will be
assumed in the immediately following computations.

Using Equations (31), (32) and (25) a typical term of Equation (22)
can now be written in the form

$$\sum_{n=1}^{T_N} \int \frac{dt}{t_0} \int \int \int d\tau d\sigma \int ds(t) s_1(t) \cos[\omega_0 t + \delta(t)] \cos[\omega_0 \lambda + \phi(t)] \rho_1^n(t) \cos n \omega_0 (t-\lambda)$$

$$n \leq n \leq T_N$$

(33)

This would remain approximately true for targets moving sufficiently
slowly relative to the receiver so that the Doppler shift is small compared
with the signal bandwidth.
Since \( n \) is an odd integer, \( \cos^n \omega_0(t-\lambda) \) can be expanded as follows

\[
\cos^n \omega_0(t-\lambda) = 
\sum_{k=0}^{n} \binom{n}{k} \cos^{n-k} \omega_0(t-\lambda) \cos^{k} \omega_0(t-\lambda)
\]

(34)

The coefficients \( \binom{n}{k} \) are related to the binomial coefficients. In particular, a straightforward expansion of \( \binom{n}{k} \) takes the form

\[
\binom{n}{k} = \frac{n!}{k!(n-k)!}
\]

(35)

Substituting Equation (35) into Equation (33) and invoking the narrowband assumption, one can use the Riemann-Lebesgue lemma to argue that only the term \( C_n \cos ^n \omega_0(t-\lambda) \) of Equation (34) contributes significantly to the importance. Therefore, Equation (33) becomes

\[
A_n = K \cos^n \omega_0 \int_0^1 d\lambda \int_0^1 dt \int_0^1 d\omega_1 \int_0^1 d\omega_2 
\cos[n(t-\lambda)] \cos[n(t+\lambda)]
\cos[(n-1)(t-\lambda)] \cos[(n-1)(t+\lambda)]
\]

(36)

Straightforward use of trigonometric identities yields

\[
\cos[n(t-\lambda)] \cos[n(t+\lambda)] = \frac{1}{2} \cos[2n\lambda(t-\lambda)] + \frac{1}{2} \cos[2n\lambda(t+\lambda)]
\]

(37)

Once more the \( K \) expansion rule is then employed to eliminate all but the term \( n \) in \( C_n \) becomes

\[
\]

(38)
\[ A_n = R_n \int_0^T \int_0^T \cos[\phi(t) - \phi(\lambda)] \sin_1(t) \sin_1(\lambda) \eta_1(t-\lambda) \cos[\phi(t) - \phi(\lambda)] dt \, d\lambda \]  

(38)

All factors in the integrand except for the cosine terms are non-negative. As for the cosine terms, note that \( \eta_1(t-\lambda) \) becomes small for \( (t-\lambda) \) larger than the correlation time of the noise process. For reverberation this correlation time is at most (stationary scatterers) equal to the correlation time of the signal, which in turn is the smaller of the following two quantities: 1) the signal duration, 2) the time within the signal noise during which the modulation \( \phi(t) \) changes by, roughly, one radian. Thus the cosine terms remain positive throughout the effective region of integration. In the ambient noise limited case the noise spectrum would generally be shaped by band limiting filters (or possibly transducer

---

1 The case of reverberation from stationary scatterers is treated with much greater rigor and care in Section IV.
characteristics, and hence have a bandwidth at least equal to that of the signal. Its correlation time would therefore be no longer than that of the signal, so that the above assumption remains valid.

Under these conditions, the integrands in both numerator and denominator of Equation (1) are positive over the effective range of integration. Since $A_k < A_{k+1}$ we have:

$$A_{k} = \frac{1}{n+1} \left( 1 + \frac{1}{n} \cdots \frac{n}{n+1} \right)^2 \frac{1}{n+1} \quad (39)$$

Hence, from Equation (1):

$$-\frac{1}{n^2} < 1 + 2 \sum_{k=1}^{n} \frac{1}{k(k+1)} \left( 1 + \frac{1}{k} \cdots \frac{k}{k+1} \right)^2 \frac{1}{n+1} \quad (40)$$

From the well-known inequality:

$$\left[ \frac{1}{2} \cdots \frac{1}{n+1} \right]^2 \leq \frac{1}{n+1} \quad (41)$$

one obtains, upon setting $n = n + 1$:

$$\left[ \frac{1}{2} \cdots \frac{1}{n+1} \right]^2 \leq \frac{1}{n+1} \quad (42)$$

It is now a simple matter to bound $1/n^2$ from above by evaluating the first few terms of Equation (40) computationally and using Equation (42) to set upper bounds on the remaining terms. Carrying the exact computation from Equation (40) to $n = 5$ we obtain:

\[ \text{[Continued on the next page]} \]

\[ \text{[Footnote]} \]
\[ \frac{1}{R^2} \leq 1.277 \]  

Combining equations (23) and (43) one can therefore rather generally bound the clipping loss for noise independent from hydrophone to hydrophone by

\[ 0.89 \leq R \leq 1 \]  

One situation in which the assumptions leading to Equation (44) are not satisfied is that of an ambient noise dominated environment containing a strong narrowband noise component. Here the effective range of integration is determined by the correlation time of the signal component of equation (38). It is clear from intuitive considerations that \( \cos[\cdot(t)-t(1)] \) does not change sign over this interval. More specifically, consider a signal of the form

\[ s(t) = e^{-\frac{t^2}{2}} \cos(w_0t+\frac{\pi}{2}t^2) \]  

Substitution into equation (38) yields, after an elementary computation,

\[ A_n = n \ln \int_{-\infty}^{\infty} e^{-\left(\frac{1}{2\alpha^2} + \frac{k^2}{8}\right)x^2} \frac{1}{\alpha^2} e_{1}(x) \frac{1}{\alpha^2} e_{1}(x) \]  

\[ = \int_{-\infty}^{\infty} e^{-\left(\frac{1}{2\alpha^2} + \frac{k^2}{8}\right)x^2} \frac{1}{\alpha^2} e_{1}(x) \frac{1}{\alpha^2} e_{1}(x) \]  

1The relation \( 1 + 1/9 + 1/25 + 1/49 + \cdots = \pi^2/8 \) has been used in summing the infinite series above \( n = 9 \).

2The limits were extended from \([0,\alpha]\) to \((-\infty,\alpha)\) because in practice the period of integration would undoubtedly cover the effective duration of the pulse.
Since both integrands of Equation (46) are non-negative and $\rho_1^n(x) \leq \rho_1(x)$, Equation (39) [and hence Equation (44)] remains valid.

Thus it appears that the only practically important exception to Equation (44) [for noise independent from hydrophone to hydrophone] occurs when a rapidly moving target is being detected in a reverberation limited environment. Here the returned signal might be of the form

$$s(t) = s_1(t) \cos[(\omega_D + \omega_0) t + \phi(t)]$$

(47)

where $\omega_D$ is the Doppler shift caused by the moving target. Following the same procedure as in Equations (33) to (38) one arrives at the following equivalent of Equation (38)

$$A_n = \frac{K_n L_n}{T} \int_0^T \int_0^T \rho_1^n(t-\lambda) \cos[(\omega_D t - \lambda) + \phi(t) - \phi(\lambda)]$$

$$\int_0^T \int_0^T \rho_1(t-\lambda) \cos[(\omega_D t - \lambda) + \phi(t) - \phi(\lambda)]$$

(48)

As a specific example, consider

$$s_1(t) = e^{-\frac{1}{2}t^2} , \quad \phi(t) = \frac{k}{2} t^2$$

(49)

a linearly frequency modulated pulse.

It is a simple matter to demonstrate [see Equation (86) with $i = j$, $d_{ji} = 0$] that $\rho_1(t)$ [for reverberation from stationary scatterers]

1The signal is assumed to be sufficiently narrowband so that the Doppler shift may be regarded as constant throughout the band.
assumes the form

\[
\phi_1(t) = e^{-\left(\frac{1}{2\sigma_T^2} + \frac{K^2}{1 - \epsilon^2}\right) t^2}
\]

(50)

\[
\left(\frac{1}{2\sigma_T^2} + \frac{K^2}{1 - \epsilon^2}\right)
\]

may be interpreted as the bandwidth of the signal and hence of the reverberation.

Substituting Equations (45) and (50) into Equation (48) and extending the ranges of integration to \((-\infty, \infty)\), one obtains, after some algebraic manipulation

\[
A_n = k C n^2 \ln \exp \left\{ \frac{1}{n+1} \omega D \sqrt{\frac{\omega_n^2}{\sigma_T^2} + K^2} \right\}
\]

(51)

If the signal bandwidth is determined primarily by the frequency modulation

\[
K^2 \sigma_T^2 \gg \frac{4}{\omega_n^2}
\]

and Equation (51) becomes

\[
A_n = k C n^2 \ln \exp \left\{ \frac{1}{n+1} K^2 \sigma_T^2 \right\}
\]

(52)

Conversely, if \(K^2 \sigma_T^2 \ll \frac{4}{\omega_n^2}\), [little or no frequency modulation]

\[
A_n = k C n^2 \ln \exp \left\{ -\frac{n-1}{4(n+1)} \omega_n^2 \omega_D^2 \right\}
\]

(53)

In either case the coefficient \(A_n\) becomes large when \(\omega_D\) greatly exceeds the bandwidth of the signal. Under the same conditions the factor \(K\) will therefore become very much smaller than unity. It may, in fact, come arbitrarily close to zero if the Doppler shift is large enough compared
with the signal bandwidth. This is quite reasonable from a physical point of view: Under the postulated conditions the signal and reverberation spectra are essentially disjoint so that the unclipped detector operates in an environment that is almost noise-free in the signal band. Clipping, on the other hand, shifts some of the reverberation power into the signal band and therefore sharply degrades detector performance.
IV. Noise Dependent from Hydrophone to Hydrophone

When the noise field exhibits significant coherence from hydrophone to hydrophone, it appears to be difficult to obtain results of the same generality as in the previous section. Thus even in the absence of Doppler shifts and with constraints on signal and noise, such as those specified by equations (31) and (32), one can readily specify realistic spatial covariances for the noise which result in values of $R$ lower than 0.89 [Equation (44)] by modest amounts. An example of this type is worked out in Appendix A, where an $R$ of 0.74 is shown to be attainable, even for isotropic ambient noise. With sufficient ingenuity, one suspects, one could devise noise models yielding even lower values of $R$. However, even the simple example of Appendix A requires fairly special assumptions and one feels that the search for more extreme cases would lead to more and more artificial assumptions. It appears more rewarding, therefore, to abandon the search for extremes and turn to the question whether serious clipping loss is likely to occur in cases commonly encountered in practice.

Some qualitative insight into this question may be gained by restating Equation (20) in the frequency domain. Defining

$$G_{ij}(w) = \frac{1}{\pi} \int_{-\infty}^{\infty} \rho_{ij}(\tau) e^{-j\omega \tau} d\tau$$  \hspace{1cm} (54)$$

and

$$P_{ij}(w) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2}{\pi} \sin^{-1} \left[ \frac{\rho_{ij}(\tau)}{\rho_{ij}(0)} \right] e^{j\omega \tau} d\tau$$  \hspace{1cm} (55)$$

one obtains the Fourier inverses.
\[ p_{ij}(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} G_{ij}(w) e^{j\omega \tau} \, dw \]  

(56)

and

\[ \frac{2}{\pi} \sin^{-1} \left[ \rho_{ij}(\tau) \right] = \frac{1}{2} \int_{-\infty}^{\infty} p_{ij}(w) e^{j\omega \tau} \, dw \]  

(57)

Substitution of Equations (56) and (57) into Equation (20) yields

(after a few steps of computation)

\[ R = \sqrt{\frac{2}{\pi}} \left\{ \int_{-\pi}^{\pi} \left| \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} G_{ij}(w) e^{j\omega (\tau_j - \tau_i)} \right| \int_{0}^{T} |S(w)| \sum_{i=1}^{M} \sum_{j=1}^{N} p_{ij}(w) e^{j\omega (\tau_j - \tau_i)} \right\} \]  

(58)

where

\[ S(w) = \int_{0}^{T} dt \, s(t) e^{-j\omega t} \]  

(59)

Since the observation interval \((0, T)\) would generally cover the entire signal pulse, \(S(w)\) is in effect the Fourier transform of the signal.

The expression \(\sum_{i=1}^{M} \sum_{j=1}^{N} G_{ij}(w) e^{j\omega (\tau_j - \tau_i)}\) in the numerator of Equation (58) is nothing other than the normalized noise power spectrum at point \(x\) in Figure 1 with the clippers removed. \(\sum_{i=1}^{M} \sum_{j=1}^{N} p_{ij}(w) e^{j\omega (\tau_j - \tau_i)}\) is the normalized noise spectrum at the same point in the presence of clipping. Thus the numerator (denominator) integral in Equation (58)
represents the power output of a filter matched to the signal whose inputs is \( x \) in the unclipped (clipped) instrumentation. Since clipping tends to spread the spectrum, one would expect a smaller percentage of the power to fall within the filter band in the clipped than in the unclipped case. Hence \( R \) should generally tend to exceed \( \sqrt{\frac{2}{\pi}} = 0.8 \). Important exceptions to this rule would be expected in two cases:

1) The unclipped noise spectrum is centered at a frequency quite different from the center frequency of \( |S(w)|^2 \). This would be the case when the target return is subject to a strong Doppler shift. Reverberation noise is not subject to this shift and it is only through the clipping operation that a significant amount of noise power is transferred into the signal band.

2) Strong negative correlation between closely adjacent \( x_i \) in Figure 1 causes the terms \( i \neq j \) in Equation (58) to subtract substantially from the power input into the filter matched to the signal. This effect can result in values of \( R \) smaller than \( \sqrt{\frac{2}{\pi}} \) only if clipping reduces the negative correlation so that the effect is less pronounced in the denominator than in the numerator of Equation (58). To see when this might be the case, consider the inverse sine in equation (20) expanded into a power series, as in III. The equivalent of Equation (22) is now

A-3.5
\[ \frac{1}{K^2} = 1 + \frac{1}{6} \int_0^T dt \int_0^T d\lambda \sum_{i=1}^M \sum_{j=1}^M \rho_{ij}^3(t-\lambda-\tau_i+\tau_j) \]

\[ \int_0^T dt \int_0^T d\lambda \sum_{i=1}^M \sum_{j=1}^M \rho_{ij}^3(t-\lambda-\tau_i+\tau_j) \]

Unless the peak negative value of \( \rho_{ij}(\tau) \) is fairly close to unity for a significant number of \( i \neq j \), the required subtraction in the denominator does not take place. On the other hand, if \( \rho_{ij}(\tau) \) comes very close to \((-1)\), \( \rho_{ij}^3(\tau) \) is not too far from \((-1)\) and a similar subtraction takes place in the numerator. Thus one looks for maximum clipping loss in cases where the negative correlation between closely adjacent phones is strong, but not strong enough so that \( \rho_{ij}(\tau)_{\text{max}}^3 \) has a magnitude comparable to unity. Since these conditions on the \( \rho_{ij} \) are quite restrictive one is not surprised to find only modest decreases in \( R \) for even rather carefully constructed examples, such as the situation analyzed in Appendix A. 1

1Note that the rather rapidly converging sequence of coefficients in equation (60) demands integral ratios of at least the order of 5 before \( R/K^2 \) begins to increase very substantially.
To complete the discussion the clipping loss will now be computed for a fairly general case of operation in a reverberation limited environment. The most important restriction is the assumption of negligible Doppler shift, both in the target return and in the reverberation.

The key step is the computation of the cross-correlation of the reverberation at the \( i \)th and \( j \)th hydrophone. If the signal assumes the narrow-band form of Equation (31) the reverberation observed at the \( i \)th hydrophone is

\[
V_i(t) = \sum_{k} \frac{a_k}{t_k^{1/2}} s_k(t-t_k) \cos[w_0(t-t_k) + \phi(t-t_k)]
\]

\( t_k \) is the travel time of sound from the origin of coordinates (nominal center of the source) to the \( k \)th scatterer and back to the \( i \)th hydrophone. \( a_k \) measures the amplitude of the signal reflected by the \( k \)th scatterer. It includes effects of the transmitter beam pattern as well as those of scatterer cross-section. Similarly, the reverberation at the \( j \)th hydrophone is

\[
V_j(t) = \sum_{k} \frac{a_k}{t_k^{1/2}} s_k(t-t_k) \cos[w_0(t-t_k) + \phi(t-t_k)]
\]

\( t_k \) is the sound travel time from the origin to the \( j \)th hydrophone via the \( k \)th scatterer. Hence the desired cross-correlation assumes the form:

---

1 The effect of strong target Doppler shifts has already been discussed in Section III.

2 The general nomenclature is that of Report No. 27.
\[ R_{ij}(\tau) = E \left\{ \sum_{\ell} \sum_{m} \frac{a_{\ell} a_{m}}{2\pi} s_{\ell}(t-t_{\ell}) s_{m}(t-T_{m}+\tau) \times \right. \\
\left. \times \cos[\omega_{0}(t-t_{\ell})+\phi(t-t_{\ell})] \cos[\omega_{0}(t-T_{m}+\tau)+\phi(t-T_{m}+\tau)] \right\} \]

Expressing the \( \cos() \cos() \) product in terms of sum and difference frequencies, one can invoke the Riemann-Lebesgue lemma to eliminate all but the difference terms of form \( \ell = m \). Hence

\[ R_{ij}(\tau) = \frac{1}{2} E \left\{ \sum_{\ell} \frac{a_{\ell}^{2}}{2\pi} s_{\ell}(t-t_{\ell}) s_{\ell}(t-T_{\ell}+\tau) \cos[\omega_{0}(t-t_{\ell}+\tau)+\phi(t-T_{\ell}+\tau)+\phi(t-t_{\ell})] \right\} \]

The next step is to express \( T_{\ell} \) in terms of \( t_{\ell} \). Consider the spherical coordinate system shown in Figure 2. By arbitrary convention

\[ \text{Figure 2} \]

\[ ^{1} \text{The symbol } E(\cdot) \text{ denotes the expectation of the bracketed quantity.} \]
the origin is placed at the \( i \)th hydrophone. \( d_{ij} \) is the distance between phones \( i \) and \( j \). A simple trigonometric computation now yields the distance \( R_i \) of the \( i \)th scatterer from phone \( j \) in terms of \( r_k \), the distance of the \( k \)th scatterer from phone \( i \).

\[
R_i = r_i \sqrt{1 - \frac{2d_{ij}}{r_i^2} \left[ \sin \phi \sin \theta \cos (\tau_k - \tau_j) + \cos \phi \cos \theta \right] + \frac{d_{ij}^2}{r_i^2}} \quad (64)
\]

In practice \( \frac{\tau_i}{r_i} \ll 1 \) for all scatterers sufficiently close to the target to receive some illumination simultaneously with the target. Using this approximation and dividing both sides of Equation (64) by the velocity of sound one obtains

\[
\frac{\tau_i}{r_i} = \frac{d_{ij}}{c} \left[ \sin \phi \sin \theta \cos (\tau_k - \tau_j) + \cos \phi \cos \theta \right] \quad (65)
\]

For greater ease in subsequent manipulation we introduce the notation

\[
o = \sin \phi \sin \theta \cos (\tau_k - \tau_j) + \cos \phi \cos \theta
\]

Then

\[
R_{ij}(t) = \frac{1}{2} \left\{ \sum_{k} - \frac{a_{ik}^2}{2} \frac{a_{jk}}{t_k^2 (t_k - \frac{d_{ij}}{c})^2} s_i(t-t_k) s_j(t-t_k + \frac{d_{ij}}{c} + \tau) \times
\]

\[
\cos \left[ \omega_0(t-t_k + \tau) + \omega(t-t_k + \frac{d_{ij}}{c} + \tau) - \phi(t-t_k) \right] \right\}
\]

(67)

The joint probability density of \( \tau_i \), \( \phi_k \), and \( \theta_k \) has been calculated.
in Report No. 27. For volume reverberation from scatterers independently and uniformly distributed over a large volume $V$ the result is

$$p(t, \theta, \phi) = \frac{c^3}{8V} t^2 \sin \theta$$  \hspace{1cm} (68)

The coefficient $a^2$ is now decomposed into its two primary components

$$a^2 = \frac{b^2}{\sin \theta} g(\theta - \theta_0, \phi - \phi_0)$$  \hspace{1cm} (69)

$b^2$ is proportional to the scattering cross-section while $g(\theta, \phi)$ is the transmitter pattern, centered at $(\theta_0, \phi_0)$.

With the introduction of equations (68) and (69) and the change of variable

$$x = t - t_0 + \frac{d_1 + 1}{2}$$  \hspace{1cm} (70)

Equation (67) becomes

$$R_{i1}(t) = \frac{c^3}{16V} \sum b^2 \int_0^{2\pi} \int_0^\pi \int_{-\infty}^{\infty} dx_s \left( x - \frac{1}{2} - \frac{d_1 + 1}{c} \right) s(x + \frac{1}{2})$$

$$\times \cos \left[ \omega_0 \left( \frac{d_1 + 1}{c} + 1 \right) + \phi(x + \frac{1}{2}) - \phi(x - \frac{1}{2} - \frac{d_1 + 1}{c}) \right]$$  \hspace{1cm} (71)^2

---

1 The use of a frequency independent pattern function implies that the signal is sufficiently narrow-band to have its directional properties described by a single frequency pattern function.

2 The bar indicates an averaging operation over the number of illuminated scatterers.
The slowly varying amplitude factor \( b^2_b / (t_c - \frac{d_{ij}}{c}) \) has here been replaced with \( b^2_b / t_0^2 \), where \( t_0 \) is the total travel time from the target to the origin of coordinates. For targets remote compared with the radial distance covered by the pulse this should be an excellent approximation.

We further approximations will now be made, both based on the observation that \( \frac{d_{ij}}{c} \) will be quite small. It is a simple matter to demonstrate that \( \frac{d_{ij}}{c} \ll 1 \). Hence, if the maximum distance between any pair of hydrophones in the array is 20 ft., \( \left| \frac{d_{ij}}{c} \right| \lesssim 0.004 \) sec. The typical pulse envelope function \( v_1(t) \) does not change significantly over an interval of 4 milliseconds. Hence,

1) \( s_1\left(x - \frac{x}{2} - \frac{d_{ij}}{2c}\right) \) is approximated by \( s_1(x - \frac{x}{2}) \) in Equation (71).

\( \dot{v}(t) \) is also a relatively slowly varying function. Using a Taylor series one can write

\[
\dot{v}(t) = \frac{d_{ij}}{c} \left( \frac{d_{ij}}{c} \right)^{\beta - 1} \dot{v}(x) \left( x - \frac{x}{2} \right) \frac{d_{ij}}{c}.
\]  

(72)

\( \dot{v}(t) \) is the instantaneous frequency modulation in radians/sec.

\( \frac{d_{ij}}{c} \) is the phase shift of the modulation frequency between hydrophones \( i \) and \( j \).

2) \( s_1\left(x - \frac{x}{2} - \frac{d_{ij}}{2c}\right) \) is approximated by \( s_1(x - \frac{x}{2}) \) on the assumption that the phase shift at the maximum frequency deviation (from \( v_0 \)) is small between any pair of hydrophones.

With a maximum deviation of 40 kHz, this bandwidth is both positive.
and negative deviations are allowed) and a maximum spacing between hydrophones of 20 ft., one finds that

\[
\max \left[ \int_{0}^{\max} \left| n'(x - \frac{\pi}{2}) - \frac{d_{ij}}{c} \right| \right] = 0.4n
\]  

Approximation 2) has certainly become questionable for these parameters. However, in any practical array a very small percentage of the total hydrophone pairs have spacings close to the maximum array dimension. Furthermore, one tends to use arrays in such a manner that few, if any, hydrophone pairs assume an endfire alignment relative to the target. Hence, \( a \) will generally be well below unity and \( \frac{d_{ij}}{c} \) in a 20 ft. array would be substantially below 0.004 sec for all, or almost all, pairs. Hence, approximation 2) appears not unreasonable for arrays of this general size and bandwidths up to the order of 100 cps.

With approximations 1) and 2) Equation (71) becomes

\[
R_{ij}(\tau) = \frac{C^3}{16V} \sum_{k=0}^{\infty} \frac{b_{k}^2}{c_{2}^2} \int_{0}^{2\pi} \int_{0}^{\pi} g(\theta, -\phi) \left[ \frac{d_{ij}}{c} \right] \int_{-\infty}^{\infty} \sin \left( \frac{x}{2} \right) \sin \left( \frac{x - \tau}{2} \right) dx
\]

Assuming, for the sake of computational simplicity, that \( s_{1}(\cdot) \) and \( \phi(\cdot) \) are both even, i.e.,

\[ s_{1}(x) = s_{1}(-x), \quad \phi(x) = \phi(-x) \]

By introducing an assumption of narrow beam patterns one can, in fact, considerably weaken the inequality \( \left| n'(x - \frac{\pi}{2}) - \frac{d_{ij}}{c} \right| \ll 1 \) implied by 2). The appropriate weaker inequality is worked out in Appendix B.
\[ s_1(x) = s_1(-x) \quad \text{pulse envelope symmetrical} \tag{75} \]

and

\[ \varphi(x) = \varphi(-x) \quad \text{even phase modulation or odd frequency} \tag{76} \]

modulation, such as linear FM

one can readily reduce Equation (74) to the form

\[
R_{ij}(\tau) = \frac{C^2}{16V} \int_0^\infty dx \int_0^\infty dy \int_0^\infty d\omega \int_0^\infty d\omega' \cos \left( \omega_0 \left( \frac{\omega}{c} + \tau \right) \right) \left( s_1(x - \frac{1}{2}) s_1(x + \frac{1}{2}) \cos \left( \varphi(x + \frac{1}{2}) - \varphi(x - \frac{1}{2}) \right) \right)
\]

\[
= \frac{C^2}{16V} p(\tau) \int_0^\infty dz \int_0^\infty d\omega \int_0^\infty d\omega' \cos \left( \omega_0 \left( \frac{\omega}{c} + \tau \right) \right) \left( s_1(x - \frac{1}{2}) s_1(x + \frac{1}{2}) \cos \left( \varphi(x + \frac{1}{2}) - \varphi(x - \frac{1}{2}) \right) \right)
\]

where

\[
p(\tau) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \left( s_1(x - \frac{1}{2}) s_1(x + \frac{1}{2}) \cos \left( \varphi(x + \frac{1}{2}) - \varphi(x - \frac{1}{2}) \right) \right)
\]

(77)

The double integral in Equation (77) depends only on transmitter and receiver geometry, while \( p(\tau) \) depends only on the waveshape of the transmitted signal. The complete separation of these two effects is the direct consequence of approximations 1) and 2).

The double integral can be further simplified if one considers symmetrical pattern functions narrowly concentrated near \((\varphi_0, \psi_0)\). In that case one can extend the \( \varphi \) and \( \psi \) limits to \((\infty, \infty)\) and represent \( R_{ij}(\tau) \) by the first three terms of a Taylor series.

[Page 1]
\[ a = a_{ij} + \beta_{ij}(\theta_i - \theta_0) + \gamma_{ij}(\phi_i - \phi_0) \]  

(79)

where

\[ a_{ij} = \sin \theta_0 \sin \theta_j \cos(\phi_i - \phi_j) + \cos \theta_0 \cos \phi_j \]  

(80)

\[ \beta_{ij} = \cos \theta_0 \sin \theta_j \cos(\phi_i - \phi_j) - \sin \theta_0 \cos \phi_j \]  

(81)

\[ \gamma_{ij} = \sin \theta_0 \sin \theta_j \sin(\phi_i - \phi_0) \]  

(82)

With the changes of variables \( \theta_0 - \theta_0 = u, \phi_i - \phi_0 = v \) the double integral becomes

\[
\int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv \, g(u, v) \cos \nu_0 \left[ \tau \pm \frac{d_{ij}}{c} \left( a_{ij} + \beta_{ij} u + \gamma_{ij} v \right) \right] 
\]  

(83)

According to Equation (20) we are ultimately interested in the value of the autocorrelation function not at \( \tau \), but at \( t - \lambda - \tau_i + \tau_j \).

However, with the array steered on target we obtain from Equations (65) and (66)

\[ \tau_i - \tau_j = \frac{d_{ij}}{c} a_{ij} \]  

(84)

Hence, from Equations (77), (83) and (84), using the postulated symmetry of the pattern function about \( (\theta_0, \phi_0) \)

\[
P_{ij}(\tau - \tau_i + \tau_j) = \frac{C^3}{16V} p(\tau) \cos \omega_0 \tau \sum_{\ell} \frac{b_\ell^2}{t_0^2} \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv \, g(u, v) \cos \nu_0 \cos \frac{d_{ij}}{c} (\beta_{ij} u + \gamma_{ij} v) \]

\[= \frac{C^3}{16V} \left[ \sum_{\ell} \frac{b_\ell^2}{t_0^2} \right] \left[ \left( \frac{d_{ij}}{c} \beta_{ij} \nu_0 \right), \left( \frac{d_{ij}}{c} \gamma_{ij} \nu_0 \right) \right] \]  

(85)
where

\[
G(\omega, z) = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv \, g(u, v) \cos(\omega u + \pi v)
\]

\[
= \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv \, g(u, v) \, e^{-j(\omega u + \pi v)}
\]  

(86)

Thus \(G(\omega, z)\) is the Fourier transform of the pattern function \(g(u, v)\).

To normalize the crosscorrelation function we need only recognize from Equation (85) that

\[
R_{11}(0) = R_{jj}(0) = \frac{C^3}{16V} \left( \sum_{i} b_i^2 \right) G(0, 0) \rho(0) \quad (87)
\]

Hence,

\[
\rho_{ij}(\tau - \tau_i + \tau_j) = \frac{G \left( \begin{array}{c} d_{ij} \\ c \end{array} , \begin{array}{c} \ell_i \nu_0 \\ Y_i \nu_0 \end{array} \right)}{G(0, 0)} \frac{\rho(t)}{\rho(0)} \cos \omega_0 \tau \quad (88)
\]

Equation (88) must now be substituted into Equation (60). Designating the general term of Equation (60) by \(A_n\) as before and using an obvious generalization of the steps leading to Equation (38) one finds

\[
A_n = \sum_{i=1}^{M} \sum_{j=1}^{M} \left( \frac{G \left( \begin{array}{c} d_{ij} \\ c \ell_i \nu_0 \\ \ell_j \nu_0 \end{array} \right)}{G(0, 0)} \right)^n \int_{-\infty}^{\infty} dt \, s_1(t) \int_{-\infty}^{\infty} \, \rho(t) \cos \left( \phi(t) - \phi(\lambda) \right) \frac{\rho(t)}{\rho(0)} \quad (89)
\]
g(u, v) describes the spatial distribution of radiated power and is therefore non-negative. It follows from Equation (86) that

\[ G(u, v) \leq G(0, 0) \text{ for all } u, v \]  

(90)

Furthermore, it is shown in Appendix C that for narrow beam patterns

\[ G(u, v) > 0 \]  

(91)

The ratio of the double sums Equation (89) is therefore no larger than unity and one obtains

\[ \frac{\int dt s_1(t) \int d\lambda s_1(\lambda) \cos[\phi(t)-\phi(\lambda)] \left[ \frac{p(t-\lambda)}{p(0)} \right]^n}{\int dt s_1(t) \int d\lambda s_1(\lambda) \cos[\phi(t)-\phi(\lambda)] \frac{p(t-\lambda)}{p(0)}} \leq K C \]  

(92)

Using Equations (75) and (76) it is a simple matter to show further that

\[ \frac{\int dt s_1(t) e^{j\phi(t)} \int d\lambda s_1(\lambda) e^{-j\phi(\lambda)} \left[ \frac{p(t-\lambda)}{p(0)} \right]^n}{\int dt s_1(t) e^{j\phi(t)} \int d\lambda s_1(\lambda) e^{-j\phi(\lambda)} \frac{p(t-\lambda)}{p(0)}} \leq K C \]  

(93)

The fact that both double integrals are in form of convolutions suggests the use of Fourier transforms. Define

\[ S_{\phi}(\omega) = \int_{-\infty}^{\infty} s_1(t) e^{j\omega t} dt \]  

(94)

and

\[ P(\omega) = \int_{-\infty}^{\infty} p(t) e^{-j\omega t} dt \]  

(95)
Thus $S_{\xi F}(\omega)$ is the Fourier transform of the low frequency signal, i.e., the signal after a downward shift by $\omega_0$. In terms of Equations (94) and (95) $A_n$ can now be written as follows

\[
A_n \leq K_n C_n \frac{\int_{-\infty}^{\infty} dw \ |S_{\xi F}(\omega)|^2 \ (P(\omega) * P(\omega) * ... * P(\omega))}{\int_{-\infty}^{\infty} dw \ |S_{\xi F}(\omega)|^2 P(\omega)}
\]

(96)

Here $A * B$ denotes the convolution of $A$ with $B$.

Consider next the relation between $P(\omega)$ and $S_{\xi F}(\omega)$. Using the postulated symmetry of $S_1$ and $\phi$ one obtains from Equation (78)

\[
p(\tau) = \int_{-\infty}^{\infty} dx \ s_1(x - \frac{\tau}{2}) s_1(x + \frac{\tau}{2}) e^{j\phi(x + \frac{\tau}{2})} e^{-j\phi(x - \frac{\tau}{2})}
\]

(97)

The change of variable $x + \frac{\tau}{2} = y$ leads to

\[
p(\tau) = \int_{-\infty}^{\infty} dy \ s_1(y) e^{j\phi(y)} s_1(y - \tau) e^{-j\phi(y-\tau)}
\]

(98)

which is a convolution of $S_1(y) e^{j\phi(y)}$ with $S_1(y) e^{-j\phi(y)}$. It follows that

\[
P(\omega) = |S_{\xi F}(\omega)|^2
\]

(99)

Thus, $P(\omega)$ is real. Now substituting Equation (99) into Equation (96) and using Parseval's theorem.
\[ A_n \leq \frac{K_n C_n [p(0)]^{-(n-1)}}{\int_{-\infty}^{\infty} dw \, P(w) \{P(w) \ast P(w) \ast \ldots \ast P(w)\}} \]

\[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\tau \left[ \frac{P(\tau)}{P(\phi)} \right]^{n+1} \]

\[ = \frac{K_n C_n}{\int_{-\infty}^{\infty} \left[ \frac{P(\tau)}{P(\phi)} \right]^2} \]

Equation (100)

Since \( n \) assumes only odd values, the integrands of both numerator and denominator are non-negative. Furthermore, from Equation (99)

\[ |P(\tau)| = \frac{1}{2\pi} \int_{-\infty}^{\infty} dw \, |S_{\phi f}(w)|^2 e^{j\omega \tau} \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} dw \, |S_{\phi f}(w)|^2 = p(0) \]

Equation (101)

Hence, the ratio of integrals in Equation (100) has an upper bound of unity and

\[ A_n \leq \frac{K_n C_n}{n \ln} \]

Equation (102)

But this is identical with Equation (39) so that one obtains immediately from Equation (44)

\[ R \geq 0.89 \]

Equation (103)

Thus, at least in the absence of Doppler shifts, clipping losses in a reverberation limited environment are quite small for a very general class of signals and arrays.
V. Range Estimation

An estimate of target range can be obtained by regarding the correlator output as a function of replica delay \((\tau)\) and establishing the "location" of this function on the \(\tau\) axis. Details of the required instrumentation depend on the precise definition of the term "location", but if the signal correlation function is sufficiently concentrated in \(\tau\) to permit useful range estimates one would expect any reasonable measure of "location" to lead to comparable results. One such measure is obtained by the arrangement shown in Figure 3. \(x(t)\) is the output of the beamformer as in Figure 1.

\[
s(t-t_1) \xrightarrow{\text{multiplier}} x(t) \xrightarrow{\text{integrator}} \frac{1}{T} \int_0^T dt y(t) = z(t)
\]

Figure 3

It is cross-correlated with two replicas of the signal, delayed by \(t_1\) and \(t_2\) seconds respectively. The resulting short time correlation functions are subtracted to yield the final output \(z(t)\). If the target delay \((t_0)\) is given by

\[
t_0 = \frac{t_1 + t_2}{2}
\]

the expected value of \(z\) is zero. Deviations of \(z\) from zero indicate
values of target delay other than that given in Equation (10A). The measurement error of such an instrumentation has been discussed in Report No. 29 (Section I). If the true target delay is \( t_0 \) then the rms measurement error is

\[
\sigma_{t_0} = \sqrt{\frac{\bar{z}^2}{t_0}} = \frac{t_{1}+t_{2}}{2} \tag{105} \]

This figure of merit must now be calculated for instrumentations with and without clipping. In the absence of clipping the equivalent of Equation (2) is

\[
x(t) = M s(t-t_0) + \sum_{i=1}^{M} n_i(t-\tau_i) \tag{106} \]

Hence,

\[
y_1(t) = M s(t-t_0) s(t-t_1) + \sum_{i=1}^{M} n_i(t-\tau_i) s(t-t_1) \tag{107} \]

\[\text{1For sufficiently high signal-to-noise ratio to make meaningful range measurement possible.}\]

\[\text{2The target delay } t_0 \text{ was omitted in Equation (1). Since range was assumed to be known in the detection study all delays could be measured relative to } t_0 .\]
\[ y_2(t) = M s(t-t_0) s(t-t_2) + \sum_{i=1}^{M} n_i(t-t_1) s(t-t_2) \]  

Since the average value of the noise is zero,

\[ \bar{z} = \frac{M}{T} \int_0^T dt \ s(t-t_0) \ [s(t-t_1) - s(t-t_2)] \]  

Similarly, in the presence of clipping one obtains from a computation parallel to Equations (2) - (10) [low input signal-to-noise ratio]

\[ \bar{z} = \sqrt{\frac{2}{\pi}} \frac{M}{N} \int_0^T dt \ s(t-t_0) \ [s(t-t_1) - s(t-t_2)] \]  

Hence, by a trivial computation

\[ \frac{\partial z}{\partial t} \bigg|_{\text{clipped}} = \frac{\partial z}{\partial t} \bigg|_{\text{unclipped}} = \sqrt{\frac{2}{\pi N}} \]  

In the absence of clipping the mean square value of \( z \) is

\[ \bar{z}^2 = \left[ \frac{1}{T} \int_0^T dt \ s(t-t_1) \sum_{i=1}^{M} n_i(t-t_1) - \frac{1}{T} \int_0^T dt \ s(t-t_2) \sum_{i=1}^{M} n_i(t-t_1) \right]^2 \]

\[ = \frac{1}{T^2} \int_0^T dt \int_0^T d\lambda [s(t-t_1) s(\lambda-t_1) - 2 s(t-t_1) s(\lambda-t_2) + s(t-t_2) s(\lambda-t_2)] \times \]

\[ \times N \sum_{i=1}^{M} \sum_{j=1}^{M} r_{ij} (t-t_1 + \lambda_j) \]  

\[ \text{(112)} \]

\[ ^1 \text{The contribution of the clipped fluctuations to } \bar{z}^2 \] has been ignored under the assumption of low input signal-to-noise ratio.
Similarly in the presence of clipping [see Equations (11) - (13)]

\[
\bar{z}^2 = \frac{1}{T^2} \int_0^T \int_0^T d\lambda \left[ s(t-t_1) s(\lambda-t_1) - 2s(t-t_1) s(\lambda-t_2) + s(t-t_2) s(\lambda-t_2) \right] \times
\]

\[
\times \frac{2}{\pi} \sum_{i=1}^{M} \sum_{j=1}^{M} \sin^{-1} \left[ \eta_{ij}(t-\lambda-t_i + t_j) \right]
\]

(113)

From Equations (105) and (111) - (113) the clipping loss is

\[
R = \frac{\sigma_{t_0}^2}{\sigma_{t_0}^2} \left( \frac{\text{clipped}}{\text{uncropped}} \right)
\]

\[
= \left\{ \begin{array}{c}
\frac{1}{T} \int_0^T \int_0^T d\lambda \left[ s(t-t_1) s(\lambda-t_1) - 2s(t-t_1) s(\lambda-t_2) + s(t-t_2) s(\lambda-t_2) \right]
\end{array} \right\}^2
\]

\[
= \left\{ \begin{array}{c}
\frac{1}{T} \int_0^T \int_0^T d\lambda \left[ s(t-t_1) s(\lambda-t_1) - 2s(t-t_1) s(\lambda-t_2) + s(t-t_2) s(\lambda-t_2) \right]
\end{array} \right\}^2
\]

\[
\times \left\{ \begin{array}{c}
\sum_{i=1}^{M} \sum_{j=1}^{M} \sin^{-1} \left[ \eta_{ij}(t-\lambda-t_i + t_j) \right]
\end{array} \right\}^2
\]

(114)

The similarity with Equation (20) is obvious. One can clearly carry through many of the general arguments of Section III and obtain similar results. Here we shall concern ourselves only with the case of reverberation in the absence of Doppler shift [treated in Section IV, Equations (61) - (93)]. Once again we work with the general narrowband signal

\[
s(t) = s_1(t) \left[ \cos \omega_0 t + \phi(t) \right]
\]

(115)
With the same restrictions on array dimensions as in Section IV

[array diameter small compared with wavelength of maximum frequency deviation from \( \omega_0 \)], one obtains from Equations (60) and (88)

\[
A_n = K_n \sum_{i=1}^{M} \sum_{j=1}^{M} \left\{ \frac{G \left[ \frac{d_{ij}}{c} \beta_{ij} \omega_0 \right]}{G(0, 0)} \right\}^{n} \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dt \, d\lambda \right]
\]

\[
[ s(t-t_1)s(\lambda-t_1)-2s(t-t_1)s(\lambda-t_2)+s(t-t_2)s(\lambda-t_2) ] \left[ \frac{p(t-\lambda)}{p(0)} \right]^{n} \cos n \omega_0 (t-\lambda)
\]

\[
[ s(t-t_1)s(\lambda-t_2)-2s(t-t_1)s(\lambda-t_2)+s(t-t_2)s(\lambda-t_2) ] \frac{p(t-\lambda)}{p(0)} \cos \omega_0 (t-\lambda)
\]

(116)

where \( K_n \) is defined by Equation (25), \( p(t) \) by Equation (78), and \( A_n \) is the \( n \)th term of the expansion of \( 1/R^2 \). The limits of integration have been extended to \((-\infty, \infty)\) on the assumption that the integration time \((0, T)\) at least covers the duration of the two replicas. As in Section IV, the ratio of the double sum has an upper bound of unity.

Hence,

\[
\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda \left[ s(t-t_1)s(\lambda-t_1)-2s(t-t_1)s(\lambda-t_2)+s(t-t_2)s(\lambda-t_2) \right]
\]

\[
\left[ \frac{p(t-\lambda)}{p(0)} \right]^{n} \cos n \omega_0 (t-\lambda)
\]

\[
\frac{p(t-\lambda)}{p(0)} \cos \omega_0 (t-\lambda)
\]

(117)
Only the term in $s(t-t_1) s(t-t_2)$ differs from the forms treated previously. One can express the constraint of Equation (104) by

$$t_2 = t_0 + \Delta$$  \hspace{1cm} (118)

$$t_1 = t_0 - \Delta$$  \hspace{1cm} (119)

Substituting Equations (115), (118) and (119) into (117), one obtains after some algebraic manipulation (invoking the Riemann-Lebesgue lemma)

$$A_n \leq K C n \ln n \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda \{ s_1(t) s_1(\lambda) \cos[\phi(t) - \phi(\lambda)] - s_1(t+\Delta) s_1(\lambda-\Delta) \}

\cos[2\omega_0 \Delta + \phi(t+\Delta) - \phi(\lambda-\Delta)] \left[ \frac{p(t-\lambda)}{p(0)} \right]^n \frac{p(t-\lambda)}{p(0)}$$  \hspace{1cm} (120)

The signal correlation function is monotone only over intervals of the order of half a carrier cycle. Hence, the separation $(2\Delta)$ between the two replica delays cannot exceed $\pi/\omega_0$ if the output from the device of Figure 3 is to have an unambiguous interpretation.\(^1\) For time increments of the order of $\pi/\omega_0$ the relatively slowly varying functions $s_1(t)$ and $\phi(t)$ do not change significantly. Hence,

\(^1\)The practically more interesting situation in which only the envelope of the correlation function is used in ranging is discussed on p. 42.
The change of variable

\[ t - \lambda = x \]

\[ t + \lambda = y \]

yields

\[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds_1(t) s_1(\lambda) \sin[\phi(t) - \phi(\lambda)] \left[ \frac{p(t - \lambda)}{p(0)} \right]^{n} \]

\[ = \frac{1}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx s_1(\frac{x+y}{2}) s_1(\frac{y-x}{2}) \sin[\phi(\frac{x+y}{2}) - \phi(\frac{y-x}{2})] \left[ \frac{p(x)}{p(0)} \right]^{n} \]

(123)

\( s_1 \) and \( \phi \) are even functions by assumption. Therefore \( \phi(\frac{x+y}{2}) - \phi(\frac{y-x}{2}) \) is odd in \( y \). It follows that the integrand of Equation (123) is odd in \( y \) so that the value of the integral is zero. Using this result in Equation (121) and substituting in Equation (120) one obtains finally

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds_1(t) s_1(\lambda) \cos[\phi(t) - \phi(\lambda)] \left[ \frac{p(t - \lambda)}{p(0)} \right]^{n} \]

\[
\Lambda_n \leq K C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds_1(t) s_1(\lambda) \cos[\phi(t) - \phi(\lambda)] \left[ \frac{p(t - \lambda)}{p(0)} \right]^{n} \]

\[ \Lambda = \Lambda \]

(124)
But this is identical with equation (92) so that Equation (103) remains true:

$$R \geq 0.89$$  \hspace{1cm} (125)

The discussion just concluded is unrealistic in one respect. The postulated instrumentation uses the carrier frequency in ranging, thus obtaining in effect extreme range accuracy (to a small fraction of a carrier wavelength) at the expense of ambiguity over multiples of the carrier wavelength. In practice one cannot tolerate such ambiguity. One would therefore almost certainly ignore carrier frequency effects and seek to locate the envelope of the signal correlation function on the $\phi$ axis. The formal analysis of an appropriate instrumentation is fairly cumbersome, but much of the desired insight into the question of clipping loss can be obtained from the following line of reasoning.

A typical signal autocorrelation function is sketched in Figure 4.

![Figure 4](image-url)

An instrumentation of the type of Figure 4 is designed to track the envelope would have to compare the amplitude of the quas-sinusoidal oscillations at a distance $\Delta$ from the origin. This clearly leads to lower sensitivities than in the previous computation, where $\Delta$ was of the order of a quarter wavelength of the carrier frequency. On the other hand, it is clear
from Equations such as (109) and (110) that the ratio of sensitivities for the clipped and unclipped instrumentations [Equation (111)] is unaffected by this change. Furthermore, the output fluctuation $\overline{z^2}$ for the envelope observation is simply the output fluctuation of Figure 3 with the delay adjusted for operation at the peak of the appropriate carrier cycle (as suggested in Figure 4). Thus, Equation (114) and hence, Equation (120) is still indicative of the clipping loss if $\Delta$ assumes the appropriate value.\(^1\)

One can now no longer make the approximation in the second line of Equation (121) and must write instead

$$\int_{-\infty}^{\infty} d\Delta \int_{-\infty}^{\infty} d\lambda \ s_1(t+\Delta) s_1(\lambda-\Delta) \cos[2\omega_0\Delta + \phi(t+\Delta) - \phi(\lambda-\Delta)] \left[ \frac{p(\lambda)}{p(0)} \right]^n$$

$$= \cos 2\omega_0\Delta \int_{-\infty}^{\infty} d\Delta \int_{-\infty}^{\infty} d\lambda \ s_1(t+\Delta) s_1(\lambda-\Delta) \cos[\phi(t+\Delta) - \phi(\lambda-\Delta)] \left[ \frac{p(\lambda)}{p(0)} \right]^n$$

$$- \sin 2\omega_0\Delta \int_{-\infty}^{\infty} d\Delta \int_{-\infty}^{\infty} d\lambda \ s_1(t+\Delta) s_1(\lambda-\Delta) \sin[\phi(t+\Delta) - \phi(\lambda-\Delta)] \left[ \frac{p(\lambda)}{p(0)} \right]^n$$

(126)

The change of variable (122) applied to the last term of Equation (126) still leads to an odd $y$ function so that this integral vanishes. The equivalent of Equation (124) is therefore

\(^1\)The zeros of $R_s(\tau)$ occur at values of $\tau$ such that $\omega_0\tau = \frac{\pi K}{2}$. $K$ odd. However, the peaks of $R_s(\tau)$ do not necessarily occur midway between the zeros, a fact which causes an apparent difficulty in the precise choice of $\Delta$. Fortunately, it turns out that changes of $\Delta$ by a small fraction of a carrier cycle do not affect the clipping loss computation.
\[
\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda \left| \frac{p(t-\lambda)}{p(0)} \right|^n s_1(t) s_1(\lambda) \cos [\phi(t) - \phi(\lambda)] - \\
A_n \leq K n \ln \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda \frac{p(t-\lambda)}{p(0)} \left| s_1(t) s_1(\lambda) \cos [\phi(t) - \phi(\lambda)] \right|
\]

\[
- \cos 2\omega_0 \Delta s_1(t+\Delta) s_1(\lambda-\Delta) \cos [\phi(t+\Delta) - \phi(\lambda-\Delta)]
\]

\[
- \cos 2\omega_0 \Delta s_1(t+\Delta) s_1(\lambda-\Delta) \cos [\phi(t+\Delta) - \phi(\lambda-\Delta)]
\]

The first terms in numerator and denominator are identical with the integrals in equation (92). The second terms can be reduced to an analogous form by the transformation

\[
x = t + \Delta \\
y = t - \Delta
\]

Following the same sequence of steps as in equations (92) - (100) one now obtains

\[
\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda \left| \frac{p(t)}{p(0)} \right|^n - \cos 2\omega_0 \Delta \left| \frac{p(t)}{p(0)} \right|^n \frac{p(t+2\Delta)}{p(0)} - \\
A_n \leq K n \ln \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda \left| \frac{p(t)}{p(0)} \right|^n - \cos 2\omega_0 \Delta \left| \frac{p(t+2\Delta)}{p(0)} \right|
\]

Because of the negative term it is difficult to draw conclusions of the same generality as in the case of detection [equations (100) - (102)].

The practical picture becomes clear, however, when one recalls that \( p(t) \) [Equation (97)] is in effect the autocorrelation function \( R_s(t) \) of the signal. Furthermore, \( A \) must be chosen in a range of \( s \) values where \( P_s(t) \) decays rapidly. If the envelope of \( R_s(t) \) is sufficiently

\( A \) is
concentrated on the $\tau$ axis to permit meaningful range measurements $p(2\Delta)$ should be very small. Hence, the product $\left|\frac{p(\tau)}{p(0)}\right|^n \frac{p(\tau-2\Delta)}{p(0)}$ should be small for all $\tau$ and the second terms in numerator and denominator of Equation (129) should have only a minor effect on the ratio of integrals. Thus, one expects Equation (102) to remain true with, at most, minor modifications.

As an example, consider a signal consisting of a Gaussian pulse with linear frequency modulation

$$s(t) = e^{-\frac{t^2}{2}} e^{\frac{\alpha T}{2}} \cos(\omega_0 t + \frac{\alpha}{2} t^2)$$

(130)

A straightforward computation yields

$$\frac{p(t)}{p(0)} = e^{-\Omega^2 \tau^2}$$

(131)

where

$$\Omega = \sqrt{-\frac{1}{2} + \frac{K^2 \sigma_T^2}{8}}$$

(132)

Thus, $\Omega$ is the effective bandwidth of the transmitted signal.

Substituting Equation (131) into Equation (129) one obtains

$$A_n \leq \frac{K C}{n} \ln \frac{2}{n+1} \frac{1 - \cos 2\omega_0 \cos - \frac{n}{n+1} \frac{4 \Omega^2 \Delta^2}{1 - \cos 2\omega_0 \cos - 2 \Omega^2 \Delta^2}}{1}$$

(133)

$p(t)/p(0)$ has maximum slope at $\tau = \frac{1}{\sqrt{2}}$ and $\Omega \Delta$. Hence, this value of delay is chosen for $\Delta$. Then

$A \approx 19$
\[ A_n \leq K \frac{C_n}{n} \ln \sqrt{\frac{2}{n+1}} \frac{1 - \cos \left( \sqrt{2} \frac{w_0}{n} \right) - 2 \frac{n}{n+1}}{1 - \cos \left( \sqrt{2} \frac{w_0}{n} \right)} - 1 \]  

(134)

The last fraction in Equation (134) has a maximum value when the cosine terms have arguments equal to multiples of \( 2\pi \) and when \( n \to \infty \). Therefore

\[ A_n \leq K \frac{C_n}{n} \ln \sqrt{\frac{2}{n+1}} \frac{1 - \epsilon}{1 - \epsilon} - 1 = 1.37 \frac{K C_n}{n} \ln \sqrt{\frac{2}{n+1}} \]  

(135)

\( n \) assumes only odd values. For \( n = 3 \)

\[ 1.37 \frac{K C_3}{3} \ln \sqrt{\frac{2}{4}} = 0.966 \frac{K C_3}{3} \]  

(136)

It follows that

\[ A_n \leq K \frac{C_n}{n} \ln \] for all \( n \geq 3 \)  

(137)

Since \( A_1 = 1 \), 0.89 remains a lower bound on \( R \). In this particular example the lower bound is definitely not reached, so that the clipping loss is actually even smaller.

VI. Doppler Estimation

Radial target velocity can be estimated by a procedure very similar to the one used in range estimation. Figure 5 shows a schematic Doppler estimator equivalent to the range estimator of Figure 3. The target delay is assumed to be known and the signal is
\[ x(t) = s_1(t) \cos[(\omega_0 + \omega_D)t + \phi(t)] \]

\[ y(t) = \int_0^T dt \; y(t) \]

\[ z(t) = s_1(t) \cos[\omega_2 t + \phi(t)] \]

Figure 5

Taken as sufficiently narrow-band so that Doppler shifts in the envelope and phase modulation may be ignored.

By analogy with Equation (105) the rms Doppler error is

\[ \sigma_{\omega_D} = \sqrt{2} \left| \frac{\omega_0 + \omega_D}{3\omega_0} \right| \frac{\omega_1 + \omega_2}{2} \]  

(138)

A computation entirely parallel to Equations (106) - (111) verifies that

\[ \frac{\partial^2}{\partial \omega_0^2} \left| \text{clipped} \right| = \sqrt{\frac{2}{\pi N}} \]

\[ \frac{\partial^2}{\partial \omega_0^2} \left| \text{unclipped} \right| \]

(139)

Proceeding as in Section IV one obtains in place of Equation (117)

\[ \hat{n} \leq K_n \frac{n + 1}{n} \left| 1 - \frac{1}{2} F \right| \]

(140)
where

\[ D_n = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda s_1(t) \cos \left[ \omega_1 t + \phi(t) \right] \cos \left[ \omega_1 \lambda + \phi(\lambda) \right] \left[ \frac{p(t-\lambda)}{p(0)} \right]^n \cos \omega_0 (t-\lambda) \]  

(141)

\[ E_n = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda s_1(t) \cos \left[ \omega_2 t + \phi(t) \right] \cos \left[ \omega_2 \lambda + \phi(\lambda) \right] \left[ \frac{p(t-\lambda)}{p(0)} \right]^n \cos \omega_0 (t-\lambda) \]  

(142)

\[ F_n = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda s_1(t) \cos \left[ \omega_1 t + \phi(t) \right] \cos \left[ \omega_2 \lambda + \phi(\lambda) \right] \left[ \frac{p(t-\lambda)}{p(0)} \right]^n \cos \omega_0 (t-\lambda) \]  

(143)

Set

\[ \omega_1 = \omega_0 + \omega_D - \Delta \omega \]

(144)

\[ \omega_2 = \omega_0 + \omega_D + \Delta \omega \]

Then a simple computation yields

\[ D + E = \frac{C_{1n}}{2} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\lambda s_1(t) \cos \omega_0 (t-\lambda) \cos \left[ \omega_0 (t-\lambda) + \phi(t) - \phi(\lambda) \right] \left[ \frac{p(t-\lambda)}{p(0)} \right]^n \]  

(145)

With the change of variable

\[ t - \lambda = x \]

\[ t + \lambda = y \]  

(146)

one obtains (using the even symmetry of \( \phi \))

\[ D + E = \frac{C_{1n}}{4} \int_{-\infty}^{\infty} dx \left[ \frac{p(x)}{p(0)} \right]^n \cos \omega_0 \cos \omega_D x \int_{-\infty}^{\infty} dy \frac{1}{2} s_1 \left( \frac{x+y}{2} \right) \cos \left[ \phi \left( \frac{x+y}{2} \right) - \phi \left( \frac{x-y}{2} \right) \right] \]

\[ = \frac{C_{1n}}{2} \int_{-\infty}^{\infty} dx \left[ \frac{p(x)}{p(0)} \right]^{n+1} \frac{p(0)}{p(0)} \cos \omega_0 \cos \omega_D x \]  

(147)
A similar sequence of steps yields

\[ F_n = \frac{C \ln n}{8} \int_{-\infty}^{\infty} dx \left[ \frac{p(x)}{p(0)} \right]^n \cos \omega x \int_{-\infty}^{\infty} dy \frac{1}{2} \left( \frac{x + y}{2} \right) s_n \left( \frac{y}{2} \right) \cos \left( \frac{x + y}{2} \right) + \frac{1}{2} \left( \frac{y}{2} \right) \cos \left( \frac{x + y}{2} \right) \]  

(148)

\[ p(x) < 1 \]. Hence \( \left[ \frac{p(x)}{p(0)} \right]^n \) decays more rapidly than \( p(x) \). It follows from Equation (147) that the ratio on the right side of Equation (140) can be made much larger than unity by choosing \( \omega_D \) sufficiently large.

This, of course, is simply the phenomenon discussed in Section III: If the target Doppler shift is large there will indeed be substantial clipping loss because signal and noise spectra in the unclipped instrumentation are almost disjoint. We are here concerned with any additional loss which may occur in Doppler measurement. Hence, we choose \( \omega_D = 0 \) and attempt to set bounds on \( A_n \) for that case.

From a qualitative point of view one can make the following observations: \( \Delta \omega \) would be chosen so as to place \( \omega_1 \) and \( \omega_2 \) near the points of maximum slope of the signal spectrum. Roughly speaking this identifies \( \Delta \omega \) with the half-bandwidth of the signal spectrum. The effective duration of \( \frac{p(x)}{p(0)} \) is given by the correlation time of the signal. Hence, the maximum value of \( \Delta \omega x \) in the effective range of integration is of the order of a radian for \( D_n + E_n \) and less for \( D_n + E_n \), \( n > 1 \). Therefore, the expression for \( D_n + E_n \) does not differ greatly from

\[ \frac{C \ln n}{8} \int_{-\infty}^{\infty} dx \left[ \frac{p(x)}{p(0)} \right]^{n+1} p(0) \]  

(149)

which is (except for a constant independent of \( n \)) the same as the numerator of Equation (100). In the expression for \( F_n \), on the other
hand, the effective range of the \( y \) integration covers the signal duration
and over that range \( d_wy \) could go through many complete periods (except in
the absence, or virtual absence, of frequency modulation). From the
Riemann-Lebesgue lemma one would then infer that \( 2F_n \) is small compared
with \( D_n + E_n \), except possibly in the absence of frequency modulation. One
would therefore generally expect the bound on \( A_n \) given by Equation (140)
not to differ drastically from Equation (137).

To give some quantitative support to this line of reasoning, consider
once more the Gaussian pulse with linear frequency modulation

\[
 s(t) = e^{-\frac{t^2}{\sigma_T^2}} \cos(\omega_0 t + \frac{K}{T} t^2)
\]  

(150)

Straightforward computations yield

\[
 D_n + E_n = \frac{C \ln}{4} \sqrt{2\pi \sigma_T} \sqrt{\frac{n}{n+1}} \frac{1}{\Omega} e^{-\frac{(\Delta \omega)^2}{4(n+1)\Omega^2}}
\]  

(151)

and

\[
 2F_n = \frac{C \ln}{4} \sqrt{2\pi \sigma_T} \sqrt{\frac{n}{n+1}} \frac{1}{\Omega} e^{-\frac{n}{n+1} \frac{\sigma_T^2}{2} \left[ 1 + \frac{1}{2\sigma_T^2 \Omega^2} \right]} (\Delta \omega)^2
\]  

(152)

where \( \Omega \) is the "signal bandwidth" defined by Equation (132).

Substituting into Equation (140) one obtains

\[
 A_n \leq K \frac{C}{n \ln} \left[ 1 - e^{-\frac{\sigma_T^2}{2} \left[ 1 + \frac{1}{2\sigma_T^2 \Omega^2} \right]} \right] (\Delta \omega)^2
\]  

(153)

A-54
The signal spectrum has maximum slope at \( w = \pm \sqrt{2} \Omega \). With the substitution \((\Delta \omega)^2 = 2\Omega^2\), Equation (153) becomes

\[
A_n = K_n c_n \ln \sqrt{\frac{2}{n+1}} \frac{1-e^{-\frac{n}{2(n+1)}}}{1-e^{-\frac{n+1}{n+1}} \frac{\sigma_T^2 \Omega^2}{e}}
\]

Equation (154)

One can readily demonstrate by direct computation that the right side of Equation (154) does not exceed \( K_c n \ln \) for any value of \( \sigma_T^2 \Omega^2 \) and any \( n \geq 3 \). As anticipated, the exponential terms \([\text{proportional to the ratio of } \frac{2F_n}{n+E_n}]\) are small unless \( \Omega = \frac{1}{\sigma_T^2} \). The lowest value of \( \Omega \) is reached when \( K = 0 \), in which case \( \Omega^2 = \frac{1}{2\sigma_1^2} \) and

\[
A_n = K_n c_n \ln \sqrt{\frac{2}{n+1}}
\]

Equation (155)

Thus the bound

\[
R \geq 0.89
\]

Equation (156)

remains valid for Doppler estimation,\(^1\) at least in this particular example and - from the preceding discussion - probably for most cases of practical interest.

---

\(^1\)Note that we are working with \( w_D = 0 \). Our conclusions, therefore, refer to clipping losses aside from those due to spectral separation of signal and reverberation.
Appendix A

Consider the problem of detection in a spherically isotropic ambient noise field. Here

\[ \rho_{ij}(\tau) = \frac{1}{2\tau_{ih}} \int_{-\tau_{ih}}^{\tau_{ih}} \rho(\lambda) \, d\lambda \]  

(A-1)

where \( \rho(\lambda) \) is the normalized autocorrelation of the noise and

\[ \tau_{ih} = \frac{d_{ih}}{c} \]  

(A-2)

\( d_{ih} \) is the distance between the \( i^{th} \) and \( h^{th} \) hydrophone. If the system processes only a narrow band near the nominal signal frequency, one can write as before

\[ \rho(\lambda) = \rho_1(\lambda) \cos \omega_0 \lambda \]  

(A-3)

Suppose, now, that \( \rho_1(\lambda) \) is essentially constant over the interval \( \tau - \tau_{ih} \leq \lambda \leq \tau + \tau_{ih} \) for every pair of hydrophones and every value of \( \tau \). This is the narrow-band assumption discussed in detail in Section IV.

Then Equation (A-1) becomes

\[ \rho_{ij}(\tau) = \frac{\rho_1(\tau)}{2\tau_{ih}} \int_{-\tau_{ih}}^{\tau_{ih}} \cos \omega_0 \sigma \, d\sigma \]

\[ = \frac{\sin \omega_0 \tau_{ih}}{\omega_0 \tau_{ih}} \rho_1(\tau) \cos \omega_0 \tau \]  

(A-4)

---

Returning to Equation (20) and expanding the $\sin^{-1}$ term as in Equation (22), one obtains the following equivalent of the latter equation

$$\frac{1}{R^2} = 1 + K_3$$

$$\frac{1}{R^2} = 1 + K_3 \left[ \int_0^T dt s(t) \int_0^T d\lambda s(\lambda) \sum_{i=1}^M \sum_{j=1}^M \rho_{ij}^3 \left( t-\lambda-\tau_{ij} \right) \right] +$$

$$\int_0^T dt s(t) \int_0^T d\lambda s(\lambda) \sum_{i=1}^M \sum_{j=1}^M c_{ij} \left( t-\lambda-\tau_{ij} \right)$$

$$+ K_5 \left[ \int_0^T dt s(t) \int_0^T d\lambda s(\lambda) \sum_{i=1}^M \sum_{j=1}^M \rho_{ij}^5 \left( t-\lambda-\tau_{ij} \right) \right] +$$

$$\int_0^T dt s(t) \int_0^T d\lambda s(\lambda) \sum_{i=1}^M \sum_{j=1}^M c_{ij} \left( t-\lambda-\tau_{ij} \right)$$

$$+ \ldots$$

(A-5)

where $K_n$ is given by Equation (25).

Equation (A-5) is of the same form as Equation (32). Hence, if one again postulates a signal of the form (31) one can pursue an argument entirely parallel to Equations (33) - (38). The typical term of Equation (A-5) now becomes

$$A_n = K_n C_{1n}$$

$$\sum_{i=1}^M \sum_{j=1}^M \left[ \sin \omega_0 \tau_{ij} \right]$$

$$\times \left[ \cos \left( \omega_0 (t_4 - \tau_{ij}) \right) \left( t_4 - \frac{\tau_{ij}}{\omega_0} \right) \right]$$

$$\times \left[ \cos \left( \omega_0 (t_5 - \tau_{ij}) \right) \left( t_5 - \frac{\tau_{ij}}{\omega_0} \right) \right]$$

(A-6)

In deriving this equation $\rho_{ij} (t-\lambda-\tau_{ij})$ has been replaced with $A_n$.\n
A-9
\( \rho_1(t-\lambda) \), invoking once more the narrow-band assumption, that the correlation time is large compared with sound travel time across the array. The definition of \( C_{1n} \) is given in Equation (35).

If we consider an array confined to a plane and a target producing a plane wave parallel to the plane of the array (remote broadside target), then \( \tau_i = \tau_j \) for all \( i \) and \( j \) and Equation (A-6) reduces to

\[
\hat{\gamma}_n = \frac{1}{n} \ln \frac{M}{M} \frac{\sin \omega_0 ij}{\sin \omega_0 ij} \sum_{i=1}^{M} \sum_{j=1}^{M} \frac{d t_s(t)}{d \lambda_s(\lambda) \rho_1(n(t-\lambda) \cos \theta_1(t) - \theta(\lambda))}
\]

The ratio of integrals is identical with that in Equation (38) and therefore has an upper bound of unity. This upper bound can be approached arbitrarily closely, for instance, by choosing \( \psi(t) = 0 \) (no frequency modulation) and taking \( \rho_1(t) = 1 \) for \( |t| \leq T \) (narrow-band noise). Thus

\[
\hat{\gamma}_n \leq \frac{1}{n} \ln \frac{M}{M} \frac{\sin \omega_0 ij}{\sin \omega_0 ij} \sum_{i=1}^{M} \sum_{j=1}^{M} \frac{d t_s(t)}{d \lambda_s(\lambda) \rho_1(n(t-\lambda) \cos \theta_1(t) - \theta(\lambda))}
\]

The upper bound can be approached as closely as desired by proper choice of \( \rho_1(t) \).

Consider now a planar array constructed from equilateral triangles. Figure A1 shows the most elementary version of such an array, with only
seven hydrophones. For it $A_n$ assumes a maximum value of $1.72 \times 10^{-1}$ at $\omega_0' = 5.0$.

![Figure A1](image)

For the next larger regular hexagon (Figure A2, nineteen hydrophones), the maximum $A_n$ is $2.53 \times 10^{-1}$ Going to a still larger regular hexagon (seven phones along a diameter for a total of thirty-seven phones), one arrives at a maximum $A_n$ of $3.02 \times 10^{-1}$. In each case $A_{5} , A_{7} \ldots$ are only slightly larger than $A_{3}$ because only the terms $1 = j$ contribute significantly to the numerator of Equation (A-9). Using $A_n = 3.02 \times 10^{-1}$ and following the procedure of Equations (A-9) one obtains

$$R > 0.74 \quad (A-9)$$

Thus, the clipping loss has increased to 0.6 db (from 1 db for uncorrelated phones). No attempt was made to evaluate more complex array structures, but it appears not unreasonable that still larger clipping losses might be encountered. The physical reasons for the observed phenomenon are clear: the system's inherent bandwidth is carefully

---

1In terms of equivalent input signal-to-noise ratio.
selected to yield high negative noise correlation between adjacent phones. This reduces the effective noise power in the unclipped case, but is ineffective in the clipped case because the generated harmonics do not retain the same phase relations as the fundamentals.

Figure A2
Appendix B

In order to exhibit the influence of the transmitter pattern on the narrow-bend approximation, return to equation (67) and evaluate $R_{ij}$ at the point $\tau - \tau_1 + \tau_j$. Using Equation (84)

$$R_{ij}(\tau - \tau_1 + \tau_j) = \frac{1}{2} \int \frac{2}{\pi} \frac{1}{\varphi_{ij}(x_0, \dot{r}_0, \varphi_0)} \left[ \cos \left( \varphi_{ij}(x_0, \dot{r}_0, \varphi_0) \right) - \sin \left( \varphi_{ij}(x_0, \dot{r}_0, \varphi_0) \right) \right] \right] \right]$$

Let

$$x = t - t_0 + \frac{d_{ij}}{c} (a - a_{ij}) + \frac{1}{2}$$

Then, using Equations (68) and (69) one obtains

$$R_{ij}(\tau - \tau_1 + \tau_j) = \frac{c}{16V} \int \frac{b}{t_0} \int \frac{d}{2} \left[ \omega(t) \right] \left[ \varphi_{ij}(x_0, \dot{r}_0, \varphi_0) \right] \right]$$

for narrow beams one obtains from Equation (81)

$$x = t_1 - t_1 + \frac{d_{ij}}{c} (a - a_{ij}) + \gamma_{ij}(x_0, \dot{r}_0, \varphi_0)$$

where $\gamma_{ij}$ and $\gamma_{ij}$ are defined by Equations (83) and (84). It is a simple matter to demonstrate that

$$\gamma_{ij} = 0$$

$\gamma_{ij} = 0$
Hence,

\[\max |a - a_{ij}| \leq \max_{\theta_k} |\theta_k - \theta_0| + \max_{\phi_k} |\phi_k - \phi_0| \quad \text{(B-6)}\]

Approximation 1) [p. 27] is at least as good as before. In place of approximation 2) one now has

\[\phi \left[ x - \frac{t}{2} - \frac{d_{ij}}{c}(a - a_{ij}) \right] = \phi(x - \frac{t}{2}) - \phi' (x - \frac{t}{2}) \frac{d_{ij}}{c}(a - a_{ij}) \]

\[\equiv \phi(x - \frac{t}{2}) \quad \text{(B-7)}\]

Thus, one makes the implicit assumption

\[\max \phi'(x - \frac{t}{2}) \frac{d_{ij}}{c} \left[ \max (a - a_{ij}) \right] \ll 1 \quad \text{(B-8)}\]

Using Equation (B-6) this becomes

\[\max \phi'(x - \frac{t}{2}) \frac{d_{ij}}{c} \left[ \max_{\theta_k} |\theta_k - \theta_0| + \max_{\phi_k} |\phi_k - \phi_0| \right] \ll 1 \quad \text{(B-9)}\]

The maximizations in \( \theta_k \) and \( \phi_k \) extend over the width of the transmitter pattern. For a maximum frequency deviation of 50 cps and a maximum distance between hydrometers of 20 ft., Equation (B-9) reads

\[0.4 \omega \left( \text{sum of } \omega \text{ and } \phi \text{ pattern half-widths} \right) \ll 1 \quad \text{(B-10)}\]

The pattern half-widths are measured in radians. Hence, the patterns need not be very narrow before Equation (B-9) becomes substantially less restrictive than its equivalent in Section IV.
Appendix C

Single Frequency Pattern Functions

Consider a transmitting array consisting of \( N \) omnidirectional point transducers in an arbitrary (but specified) geometrical arrangement. We wish to find the angular distribution of power in the far field. The basic geometry is shown in Figure C1 and is basically the same as that of Figure 2.

Ignoring spherical spreading losses [which are not part of the coefficients \( a_x \) in Equations (61) and (62)] the power at point \( P \) is

\[
\sum_{j=1}^{N} \frac{P_j}{d_j^2} \cos \theta_j \sin \phi_j
\]

\[(C-1)\]
where \( t_j \) is the delay of the signal applied to the \( j^{th} \) transducer (steering) and \( t_{ij} \) is the sound travel time from the \( j^{th} \) transducer to point \( P \). The pattern function \( g(\gamma) \) measures the angular dependence of power and is therefore directly proportional to (C-1).

From Equation (65) one obtains

\[
\tau_j = t_0 - \frac{d_j}{c} \left[ \sin \theta \sin \theta_j \cos(\psi - \phi_j) + \cos \theta \cos \theta_j \right] \tag{C-2}
\]

where \( t_0 \) is the sound travel time from the origin to point \( P \) and \( d_j \) is the distance of the \( j^{th} \) transducer from the origin. If the array is steered in the direction of \((\theta_0, \phi_0)\) then

\[
\tau_j = \frac{d_j}{c} \left[ \sin \theta_0 \sin \theta_j \cos(\phi_0 - \phi_j) + \cos \theta_0 \cos \theta_j \right] \tag{C-3}
\]

If the pattern is confined to small angular deviations from \((\theta_0, \phi_0)\) one can approximate (C-2) by the first terms of a Taylor series

\[
\tau_j = t_0 - \frac{d_j}{c} \left[ a_{0j} + \beta_{0j}(\theta - \theta_0) + \gamma_{0j}(\phi - \phi_0) \right] \tag{C-4}
\]

where \( a_{0j}, \beta_{0j}, \gamma_{0j} \) are analogous to the parameters \( a_{ij}, \beta_{ij}, \gamma_{ij} \) defined in Equations (80) - (82). From inspection of Equations (80) and (C-3)

\[
\tau_j = \frac{d_j}{c} a_{0j} \tag{C-5}
\]

Now, substituting Equations (C-4) and (C-5) into (C-1)

\[
g(\theta, \theta_0, \psi, \phi_0) = \left| \sum_{j=1}^{N} \sum_{c} c \right| \left| \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{c} c \right| \left( \beta_{0j} - \theta_0 \right) + \gamma_{0j}(\phi - \phi_0) \] \tag{C-6}

A-64
Hence, the Fourier transform of \( g( ) \) is

\[
G(u, v) = \sum_{j=1}^{N} \sum_{k=1}^{M} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \left[ \frac{e^{i \omega (x(\beta_{0j} - \beta_{0k})x + (\gamma_{0j} - \gamma_{0k})y)}}{1(ux+vy)} \right] e^{-i(\omega x + \omega y)}
\]

\[
= \sum_{i=1}^{N} \sum_{k=1}^{M} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \left[ \frac{e^{i \omega (x(\beta_{0j} - \beta_{0k})x + (\gamma_{0j} - \gamma_{0k})y)}}{1(u-i\omega x - i\omega y)} \right]
\]

\[
= 4\pi^{2} \sum_{i=1}^{N} \sum_{k=1}^{M} \delta[u-i\omega(\beta_{0j} - \beta_{0k})] \delta[y-i\omega(\gamma_{0j} - \gamma_{0k})]
\]

(C-7)

Thus, \( G(u, v) \) is a non-negative function. This result clearly remains true if one considers the pattern function defined by the total power over some frequency band, for the \( \omega \) integral of (C-7) is clearly non-negative.
SOME COMMENTS ON OPTIMUM BEARING ESTIMATION

by

Peter M. Schultheis

Progress Report No. 32

General Dynamics/Electric Boat Research

(8050-31-55001)

July 1967

DEPARTMENT OF ENGINEERING
AND APPLIED SCIENCE

YALE UNIVERSITY
Summary

The problem of optimum bearing estimation is discussed for the simplest possible case. The receiving array consists of two hydrophones. Signal and noise are stationary Gaussian random processes with zero means and spectra of the same form. The noise at one hydrophone is statistically independent from that at the other. The following results are obtained.

1) If the output signal-to-noise ratio is large enough to permit accurate bearing measurements, the rms bearing error of a simple cross-correlator (with pre-whitening) reaches the Cramer-Rao lower bound for all input signal-to-noise ratios. Hence the cross-correlator is an optimal bearing estimator.

2) In terms of the input signal-to-noise ratio (S/N) the rms bearing error of the cross-correlator varies as \( (S/N)^{-1} \) for \((S/N) < 1\) and as \( (S/N)^{-\frac{1}{2}} \) for \((S/N) > 1\). Thus the optimum bearing estimator has characteristics normally associated with a coherent (incoherent) device for high (low) input signal-to-noise ratios.

3) The rms error of the cross-correlator varies linearly with the correlation time (inverse bandwidth) of signal and noise. It varies inversely with the half power of the time-bandwidth product.

4) The rms bearing error of a 2 element split beam tracker is equal to that of the cross-correlator if the phase shift between beams is achieved by a pure differentiator. If a pure \(90^\circ\) phase shift is used in place of the differentiator there is a small loss in performance, equivalent to about 0.6 db of input signal-to-noise ratio.
I. Introduction

The purpose of this note is to correlate and extend certain previously reported results on optimum and suboptimum bearing estimation. The stimulus was provided by a recent paper by Middleton \textsuperscript{1} using a very different approach and obtaining, in part, different results.

In the following analysis the problem of optimality is approached by using the Cramér-Rao inequality to set a lower bound on the attainable rms bearing error. Realizability is then established by describing an instrumentation which actually reaches the lower bound. Only the simplest physical situation is considered. Signal and noise are assumed to be stationary Gaussian random processes with zero means and spectra of the same form. The noise is assumed to be statistically independent from hydrophone to hydrophone. The observation time is long compared with the correlation times of signal and noise. In most of the discussion the receiving array consists of a single pair of omnidirectional hydrophones.

II. Lower Bound on RMS Bearing Error

If the receiving array consists simply of a pair of omnidirectional hydrophones, the problem of bearing estimation is equivalent to the problem of estimating the signal delay between the two hydrophones. Let the true and estimated delay be $\delta_0$ and $\hat{\delta}$, respectively. For the case of signal and noise processes with the properties of Gaussian white noise limited in frequency to $0 \leq f \leq W$, McDonald\(^1\) has calculated the Cramèr-Rao lower bound on $(\delta_0 – \delta)^2$.

$$\frac{(\delta_0 – \delta)^2}{S^2} \geq \frac{1}{4W^2} \left( \frac{S^2}{S^2 + N^2} \right) \left( S^2 + N^2 \right) \left( \frac{S^2}{S^2 + N^2} \right) (\log(2WT) - 1) - 0.5772$$

S is the signal power, $N_1$, $N_2$ the noise power at phones 1 and 2 respectively and T the observation time. If $N_1 = N_2 = N$ and $WT \gg 1$, Equation (1) reduces to

$$\frac{(\delta_0 – \delta)^2}{S^2} \geq \frac{3}{8\pi^2W^2} \left( \frac{N}{S} \right)^2 + \left( \frac{N^2}{S^2} \right)$$

In order to translate Equation (2) into a lower bound on bearing accuracy, consider the geometry of Figure 1. If the target distance $r$ is very large compared with the spacing $d$ between hydrophones, one can write to an excellent approximation

$$\delta = \frac{d}{c} \sin \theta$$

\(^1\)A. L. Levesque, R. A. McDonald, R. N. Compton, J. Lamer, "Processing of Data from Sonar Systems," Vol. 11, 1966, Appendix E.
If the true bearing is \( \hat{\theta} \), then the estimation error \( \hat{\theta} - \theta \)
is related to \( \hat{\theta} - \theta \) by

\[
\hat{\theta} - \theta = \frac{d}{c} (\sin \hat{\theta} - \sin \theta) + \frac{d}{c} \sin \frac{\hat{\theta}}{2} \sin \frac{\theta}{2} \cos \frac{\hat{\theta} + \theta}{2}
\]

In practice, one is concerned with the numerical value of the rms tracking
error only if it is quite small, i.e., if \( \hat{\theta} - \theta \) does not exceed a
small fraction of a radian with any significant probability. In that
case, one can approximate

\[
\hat{\theta} - \theta \approx \frac{d}{c} (\sin \hat{\theta} - \sin \theta)
\]
and, except for $\theta_0$ extremely close to $\pi/2$,

$$\cos \frac{\delta}{4}(\theta^* + \theta_0) = \cos \theta_0$$ (4)

Substituting these approximations in Equation (4) one obtains

$$\theta^* - \theta_0 = (\theta^* - \theta_0) \frac{d}{c} \cos \theta_0$$ (7)

Hence, the standard deviation of $(\theta^* - \theta_0)$ has the lower bound

$$B(\theta^*) = \sqrt{(\theta^* - \theta_0)^2} \geq \frac{\sqrt{2/3} \sqrt{S}}{\sqrt{2/3} \sqrt{W} \sqrt{\pi}} \frac{\sqrt{x^2 c}}{\sqrt{2/3} \sqrt{S}}$$ (8)

This result generalizes in trivial fashion to non-white signal and noise spectra, as long as the ratio of the two spectral functions is a constant over the processed frequency band $0 < f < W$. For then one can regard the output of each hydrophone as prewhitened by an appropriate linear filter, prior to further processing. The (invertible) linear filtering operation can clearly not have any effect on the minimum attainable mean square error. Thus, Equation (5) remains valid.

---

1In practice this condition is often at least approximately true, because signal and noise spectra are shaped by the same hydrophone and receiver characteristics.

2This argument does not prove formally that the Cramér-Rao lower bound might not be lower for non-white spectra. However, it does prove that any such lower value cannot be realizable. The optimum estimator after prewhitening cannot achieve a performance better than the right side of Equation (8). But the optimum estimator would include linear filters to shape the spectra to the most desirable form. Therefore, no realizable estimator working with the initial spectra can improve on the lower bound of Equation (8).
III. An Optimal Instrumentation: The Cross-correlator

When only two hydrophones are available in obvious instrumentation for delay measurement (and hence bearing measurement) is the simple cross-correlator shown in Figure 2.

\[
x_1(t) = s(t) + n_1(t)
\]

Phone 1

\[
x_2(t) = s(t) + n_2(t)
\]

Phone 2

\[
x_2(t) \mid \text{delay } t_1
\]

\[
y(t) \rightarrow \text{low pass filter} \rightarrow z(t; t_1)
\]

\[
x \rightarrow y(t) \rightarrow \text{multiplier}
\]

\[
x \rightarrow y(t)
\]

\[
= s(t) + n_2(t)
\]

**Figure 2**

The system output is the short-time cross-correlation (averaged over the smoothing time of the filter) of the two hydrophone outputs, delayed relative to each other by \( t_1 \). At any given time \( t \), \( z(t; t_1) \) is computed for all possible values of \( t_1 \). The value \( t_1^k \) of \( t_1 \) which maximizes \( z(t; t_1) \) is designated as the instantaneous estimate of signal delay. The bearing estimate is then obtained from an obvious analogue of Equation (8):

\[
t_1^k = \frac{\theta}{2 \pi}
\]

(9)
The maximum of $z(t; t_1)$ occurs at a point where

$$\frac{\partial z}{\partial t_1} = 0$$  \hfill (10)

Thus the rms error in $t_1$ is equal to the rms fluctuation of the null of this derivative. This fluctuation is given by

$$t_1 = \sqrt{\frac{\partial^2 z}{\partial t_1^2}}$$  \hfill (11)

$D(\ )$ denotes the standard deviation of the bracketed quantity.

The output $y(t)$ of the multiplier in Figure 2 is given by:

$$y(t) = [s(t-\delta_0) + n_1(t)] [s(t-t_1) + n_2(t-t_1)]$$  \hfill (12)

If the noises at the two receivers are statistically independent of each other and of the signal (and have zero means)

$$\bar{z} = \bar{y} = \bar{s} \rho_s(t_1-\delta_0)$$  \hfill (13)

$\rho_s(t)$ is the normalized autocorrelation function of the signal.

---

1 We are assuming that $\partial z/\partial t_1$ has only one zero in the neighborhood of the correct value, $t_1 = \delta_0$. For signal-to-noise ratios high enough to permit tracking with reasonable accuracy this should be a good assumption.

In evaluating slope and standard deviation at $t_1 = \delta_0$, we imply further that the null of $\partial z/\partial t_1$ fluctuates by less than the correlation time of the signal. Examination of Equation (16) reveals that this is true when the output signal-to-noise ratio is large compared with unity. This is the only condition under which the bearing accuracy problem has much practical interest.
Differentiating equation (13) twice with respect to \( t_1 \) and evaluating at \( t_1 = 0 \), one obtains

\[
\frac{d^2}{dt_1^2} y(t_1) = x(t_1) \tag{14}
\]

If the lowpass filter shown in Fig. 2 has a weighting function \( h(\sigma) \), the output \( z(t) \) is given by

\[
z(t) = \int_0^t dt\ h(t-\sigma) \ v(t-\sigma) \tag{15}
\]

hence,

\[
\frac{dz}{dt_1} = \int_0^t dt\ h(t-\sigma) \ \frac{dV(t-\sigma)}{dt_1} \tag{16}
\]

We must consider a filter which smooths the output over the past \( T \) seconds. Such a filter has the weighting function

\[
h() = \begin{cases} 
1/2^n & \text{if } n < 0 \\
0 & \text{elsewhere}
\end{cases} \tag{17}
\]

If \( T \) is much larger than the correlation time of signal and noise, it is a simple matter to demonstrate [Report No. 10, Equation (9)] that

\[
\frac{d^2}{dt_1^2} \left[ \frac{dy}{dt_1} \right] = \frac{1}{T} \int_{-\infty}^\infty \left[ \frac{d^2}{dt_1^2} \left[ R_{yy}(\tau) - R_{yy}(\infty) \right] \right] d\tau \tag{18}
\]

\( R_{yy}(\tau) \) is the auto-correlation function of \( \frac{dy}{dt_1} \).
From Equation (12)

$$\frac{\partial y}{\partial t_1} = - [s(t-\delta_o) + n_1(t)] [s'(t-t_1) + n_2'(t-t_1)]$$ (19)

Hence,

$$R_{xy}(t) = \frac{\partial y}{\partial t_1}$$

$$E\{[s(t-\delta_o)+n_1(t)] [s(t+\tau-\delta_o)+n_1(t+\tau)] [s'(t-t_1)+n_2'(t-t_1)] [s'(t+\tau-t_1)+n_2'(t+\tau-t_1)]\}$$ (20)

Twelve of the sixteen terms in this average vanish immediately, because they contain one of the noise components only once. The remaining terms are

$$R_{xy}(t) = E\{s(t-\delta_o)s(t+\tau-\delta_o)s'(t-t_1)s'(t+\tau-t_1)\}$$

$$+ SN_s\rho_n'(t) + SN_s'\rho_n(t) + SN_s'\rho_n'(t)$$ (21)

$$\rho_x(t)$$ designates the normalized autocorrelation function of the random process $$x(t)$$. Both noise processes are assumed to have the same autocorrelation function $$\rho_n(t)$$.

Since the signal process is Gaussian, the first term of Equation (21) can be expressed in terms of second order moments. Thus

$$R_{xy}(t) = 5\rho_s(t)\rho_s'(t) + 5\rho_s'(t)\rho_s(t) + R_{ss'}(\delta_o-t_1) + R_{ss'}(\delta_o-t_1+\tau)$$

$$+ SN[\rho_s(t)\rho_n'(t) + \rho_s'(t)\rho_n(t)] + N^2\rho_n(t)\rho_n'(t)$$ (22)

$${ }^1E( )$$ stands for the expectation of the bracketed quantity.
Here $R_{xx}(t)$ is the (unnormalized) cross-correlation of $x(t)$ and $y(t)$.

There remains the computation of the various correlation functions

$$
\rho_S'(t) = \frac{E[s(t)s'(t+t)]]}{S} \cdot \lim_{\Delta t \to 0} \left\{ \frac{[s(t+\Delta t) - s(t)][s(t+\Delta t - \Delta t) - s(t+\Delta t)]}{S(\Delta t)^2} \right\}
= \lim_{\Delta t \to 0} \frac{2\rho_S(t) - \rho_S'(t) - \rho_S(t - \Delta t)}{(\Delta t)^2}
$$

(23)

Now expanding the last two terms into Taylor series about the point $t$,

$$
\rho_S'(t) = \lim_{\Delta t \to 0} \frac{1}{(\Delta t)^2} \left\{ 2\rho_S(t) - \rho_S'(t)\Delta t + \rho_S''(t) \frac{(\Delta t)^2}{2} + o(\Delta t)^2 \right\}
- \left[ \rho_S(t) - \rho_S'(t)\Delta t + \rho_S''(t) \frac{(\Delta t)^2}{2} + o(\Delta t)^2 \right]
$$

(24)

The notation $o(\Delta t)^2$ indicates that the remainder term in question approaches zero faster than $(\Delta t)^2$ as $\Delta t \to 0$.

Now carrying out the limiting operation

$$
\rho_S'(t) = -\rho_S''(t)
$$

(25)

By completely analogous computations one obtains

$$
\rho_p'(t) = -\rho_p''(t)
$$

(26)

and

$$
R_{bb}'(t) = S_b'(-1) = -R_{bb}(-1)
$$

(27)

In the absence of DC and other deterministic components of signal...
and noise, all of the correlation functions and their derivatives tend to zero as $t \to \infty$. Hence from Equation (22)

$$R_{3y}(\omega) = \frac{R_{gs}^2(\delta-o-t_1)}{\hat{S}_o^2} = \hat{S}_o'(\delta-o-t_1)$$

(28)

Now substituting Equations (22) and (25) - (28) into Equation (18)

$$p^2 \left( \frac{\partial \varphi}{\partial t_1} \right) = -\frac{1}{T} \int_0^\infty d\tau \left\{ \varepsilon^2 \left[ \rho_s(\tau) \rho_o''(\tau) + \rho_s'(\tau+\delta-o-t_1) \rho_o'(\tau+\delta-o+t_1) \right] 
+ SN \left[ \rho_s(\tau) \rho_o''(\tau) + \rho_o(\tau) \rho_o''(\tau) \right] + N^2 \rho_o(\tau) \rho_o''(\tau) \right\}$$

(29)

Using Parseval's theorem and the real translation theorem

$$D^2 \left( \frac{\partial \varphi}{\partial t_1} \right) = \frac{2\pi}{T} \int_0^\infty d\omega \omega^2 \left\{ S^2 g_s(\omega) \left[ 1-e^{2j\omega(\delta-o-t_1)} \right] + 2 SN g_s(\omega) g_n(\omega) + N^2 g_n^2(\omega) \right\}$$

(30)

where $g_s(\omega)$ and $g_n(\omega)$ are the normalized spectral functions of signal and noise respectively. Note that the first term

$$\frac{2\pi}{T} \int_0^\infty d\omega \omega^2 S^2 g_s(\omega) \left[ 1-e^{2j\omega(\delta-o-t_1)} \right] = \frac{2\pi}{T} S^2 \int_0^\infty d\omega \omega^2 g_s(\omega) \left[ 1-\cos^2\omega(\delta-o-t_1) \right]$$

(31)

exhibits the effect on the output variance of a misadjustment in $t_1$.

For $t_1 = \delta_o$ this term vanishes and one obtains

$$p^2 \left( \frac{\partial \varphi}{\partial t_1} \right) \bigg|_{t_1 = \delta_o} = \frac{2\pi}{T} \int_0^\infty d\omega \omega^2 \left[ 2 SN g_s(\omega) g_n(\omega) + N^2 g_n^2(\omega) \right]$$

(32)

B-11
Noting that

\[ s_t^{-2}(\omega) = -\int_{-\infty}^{\infty} dw \, w^2 \, s_t(w) \tag{33} \]

we can now substitute Equations (14) and (32) into Equation (11) to obtain

\[ s_t \left( \frac{\sqrt{\frac{1}{\pi} \int_{-\infty}^{\infty} dw \, w^2 \left| 2 \, 2N \, \kappa_n(\omega) \, \nu_n(\omega) \right|}}{N \int_{-\infty}^{\infty} dw \, w^2 \, \kappa_n(\omega)} \right) \]

Consider now the white spectra

\[ g_s(\omega) = g_r(\omega) = \begin{cases} \frac{1}{\pi \omega} & \text{if } |\omega| \leq 2\pi \omega \\ 0 & \text{if } |\omega| > 2\pi \omega \end{cases} \tag{35} \]

Straightforward evaluation of the integrals in Equation (34) yields

\[ s_t \left( \frac{1}{2} - s_t(\omega) \right) = \left( \frac{\sqrt{\frac{1}{N} + \frac{\pi^2}{\omega^2}}}{\sqrt{\frac{1}{N} + \frac{\pi^2}{\omega^2}}} \right) \]

which is identical with the square root of the right side of Equation (2). Translation of delay error into bearing error proceeds as in Equations (3) - (3), yielding the result

\[ D(\delta) = \frac{1}{2} \frac{\sqrt{\frac{1}{2N} + \frac{\pi^2}{\omega^2}}}{\sqrt{\frac{1}{2N} + \frac{\pi^2}{\omega^2}}} \cdot \frac{2 \, \delta \, \cos \theta}{2 \, \delta \, \cos \theta} \]

\[ = \frac{1}{\sqrt{1 + \frac{2 \theta}{\pi}}} \cdot \frac{1}{\sqrt{1 + \frac{2 \theta}{\pi}}} \tag{37} \]
Thus the instrumentation of Figure 2 actually attains the Cramér-Rao lower bound. The value of $t_1$ which maximizes $z(t, t_1)$ is therefore the optimum bearing estimate. It is obvious by inspection that it is an unbiased estimate, as demanded by the formulation of the Cramér-Rao inequality used in deriving Equation (2).

While the last part of the argument has been carried out for white signal and noise, a trivial modification of Figure 2 generalizes the result to Gaussian signals and noises with arbitrary spectra, as long as the ratio of signal and noise spectra is a constant over the processed frequency range. In such cases one simply inserts a prewhitening filter after each hydrophone, thus reducing the problem to the one just treated.

It is interesting to observe that the rms error of the optimum tracker varies as the inverse first power of $S/N$ for small input signal-to-noise ratios (the typical behavior for incoherent processing) whereas, for large input signal-to-noise ratios it varies as the inverse half power of $S/N$ (the typical behavior for coherent processing). For large input signal-to-noise ratios one can reason, of course, that the waveshape at each hydrophone is a good approximation of the signal waveshape. Thus, one has something akin to knowledge of the signal waveshape, the distinguishing characteristic of coherent operation.
IV. Conventional Split Beam Trackers

An instrumentation frequently used in bearing estimation is the split beam tracker, an elementary version of which is shown in Figure 3.

This configuration has been analyzed in detail in Report No. 29. Omitting the interference (the quantity of primary interest in Report No. 29), but retaining the signal dependent terms in the fluctuation, one obtains from Equation (55), [Report No. 29], after a few steps of computation.
\[
\sigma_z^2 = \frac{2\pi}{T} \int_{-\infty}^{\infty} dw \left( M^2 S g_B(w) + MN g_n(w) \right)^2 - \left( M^2 S g_B(w) \right)^2
\]  

(38)

\(\sigma_z^2\) is the output variance and \(M\) is the number of hydrophones in each half of the array. All other symbols retain their earlier meaning.

Using again the white spectra of Equation (35), one obtains, after evaluation of the integrals

\[
\sigma_z^2 = \frac{M^2}{2TW} \left( 1 + 2M \frac{S}{N} \right)
\]

(39)

From Report No. 29, Equation (65)

\[
\frac{\partial z}{\partial \theta} = \pi SW M^{-1} \frac{d}{c} \cos \theta_0
\]

(40)

Hence,

\[
D(\theta) = \left. \frac{\sigma_z}{\frac{\partial z}{\partial \theta}} \right|_{\text{on target}} = \frac{c}{\sqrt{2} \pi W \sqrt{TW} M^2 \cos \theta_0} \cdot \frac{1}{\sqrt{1 + 2M \frac{S}{N}}}
\]

(41)

For direct comparison with the optimum bearing estimator, consider \(M = 1\) (array consisting of two hydrophones) and take the ratio of Equations (37) and (41)

\[
\frac{D(\theta^*)}{D(\theta)} = \frac{\sqrt{3}}{2}
\]

(42)
Thus the split beam tracker performs almost optionally: The ratio of rms errors is 0.87, equivalent to about 0.6 db of input signal-to-noise ratio.

If in Figure 3 one replaces the pure phase shift

$$H(j\omega) = \frac{j\omega}{\omega}$$

with pure differentiator

$$H(j\omega) = j\omega$$

Equations (39) and (40) assume the form

$$E_z^2 = \frac{2\pi N^2}{M^2} |2SNM^2 + N^2M^2|$$

on target

and

$$\frac{\partial^2}{\partial \theta^2} = \frac{4n^2N^2}{3} \frac{S M^3}{c} \cos \theta_o$$

on target

hence, for $N = 1$

$$D(\theta) = \frac{\sqrt{S}}{2\sqrt{2} NW \sqrt{M^2 \cos \theta_o}}$$

This is identical with Equation (37) and with the Cramer-Rao lower bound.

Thus, at least in the simple case under discussion, the conventional split beam tracker actually achieves the optimum possible performance.
THE EFFECT OF NOISE ANISOTROPY
ON DETECTABILITY IN AN OPTIMUM
ARRAY PROCESSOR

By
Franz B. Tuteur

Progress Report No. 33
General Dynamics/Electric Boat Research

September 1967

DEPARTMENT OF ENGINEERING
AND APPLIED SCIENCE
YALE UNIVERSITY
Introduction

The object of this report is to investigate the effect of localized noise sources on the performance of the likelihood-ratio detector.

In a previous investigation[1] expressions for the performance loss of likelihood-ratio detection were derived for the central part and a component for the side lobes of the performance obtained. In the present report these results are extended to the detection of noise in an attempt to also get some estimate of the performance loss caused by analogously sources that are not strongly localized. Such sources can, presumably, be represented by a large number of closely spaced point sources.

The notation used follows that of Rabiner et al. [2], since this permits important simplification over the notation used previously. For the sake of completeness the expressions for the likelihood-ratio processor are re-derived here and a brief comparison between the old and new notation is given.

Unfortunately, in spite of the simplified notation the final result is still difficult to evaluate. An approximate result showing that the performance loss due to point noise sources is equivalent to the loss of one hydrophone per point source can be obtained by elimination of all of the terms causing computational difficulties. However, since these terms also contain all the information about the central location of the noise sources this approximation is fairly crude. It is made only if the number of localized noise sources is small and if they are weak and well separated from each other and from the target direction. Another approximate result can be obtained for small angular separations between the sources. It is found that noise sources distributed around a circle with the same true separation in
azimuth has approximately the same effect on detectability as a single point source of the same strength located at the center of the distributed source. For large anisotropic-to-isotropic noise ratio a similar result can be demonstrated only if the anisotropic noise source extends only over a much smaller angle.

In order to get some idea of performance loss in cases not covered by these analytical results computations were also performed on the digital computer. Although these computations are not conclusive they indicate that the analytical results for distributed noise sources at low anisotropic-to-isotropic noise ratio can probably be extended to large anisotropic-to-isotropic noise ratio.
II. Basic Analysis

Assume that the array consists of \( M \) hydrophones, and that the received signal at the \( i^{th} \) hydrophone is \( x_i(t) \). Then if the spectrum of \( x_i(t) \) is limited to frequencies below \( W \) cps, and the \( x(t) \) are observed over an interval \( T \), such that \( WT \gg 1 \), \( x_i(t) \) can be expanded in a Fourier series:

\[
x_i(t) = \sum_{n=-\infty}^{\infty} x_i(n)e^{j2\pi nt/T}
\]

where the \( x_i(n) \) are complex Fourier coefficients satisfying \( x_i(-n) = \overline{x_i(n)} \) and where the asterisk stands for complex conjugate. Then all the available information about the signals received by the entire array is contained in the set of vectors

\[
X(n) = \begin{bmatrix}
x_1(n) \\
\vdots \\
x_M(n)
\end{bmatrix}
\]

Following Bryn [3] or Edelblute, Fisk, and Kinnison [2], we assume that \( X(n) \) and \( X(m) \) are statistically independent for \( n \neq m \). Suppose that the signal \( x_i(t) \) received at the \( i^{th} \) hydrophone consists of signal and noise, then we let the signal be given by

\[
s_i(t) = \sum_{n=-\infty}^{\infty} y_i(n)e^{j2\pi nt/T}
\]

so that the signal at all hydrophones is represented by

\[
Y(n) = \begin{bmatrix}
y_1(n) \\
\vdots \\
y_M(n)
\end{bmatrix}
\]

Here again we assume the \( Y(n) \) to be independent from \( Y(m) \) for \( n \neq m \).

The optimum detector is known to be the likelihood ratio detector, which
detected presence or absence of a signal by comparing the likelihood ratio

\[ R = \frac{f_s(x)}{f_n(x)} \]  

(5)
to a fixed threshold. Here \( f_s(x) \) is the conditional probability density of the received samples (over all time, over all signal and over all frequencies) when signal is assumed to be present; similarly \( f_n(x) \) is the conditional probability density when signal is assumed to be absent. Since \( x(n) = s^x(n) \), and since \( x(n) \) and \( n(n) \) are independent for \( i \neq i \):

\[ L^2 = \prod_{i=1}^{T} \frac{f_s[x(n_i)]}{f_n[x(n_i)]} \]  

(6)

We assume now that whether signal is present or not \( x(t) \) is a stationary Gaussian random process with zero mean value. The normalized covariance matrix for noise only is

\[ Q(n) = \frac{1}{N(\Omega)} \langle \dot{x}_n^*(n) \dot{x}_n(n) \rangle_H \]  

(7)

where the superscript \( T \) refers to transposition and the symbol \( \langle \rangle_H \) means ensemble average subject to the noise only hypothesis, \( N(\Omega) \) to the average noise power at frequency \( 2\pi \Omega \) rad/sec. Then

\[ f_n[x(n_i)] = C_H(n_i) \exp \left[ \frac{-1}{T(\sigma^2)} \left( \dot{x}_n^*(n) \dot{x}_n(n) - \dot{x}_n^*(n) \dot{x}_n(n) \right) \right] \]  

(8)

where \( C_H(n_i) \) is the normalization constant to be determined.

Assume that signal and noise are independent, and that the normalized covariance matrix for signal alone is given by

\[ F(n) = \frac{1}{S(n)} \dot{x}_s^*(n) \dot{x}_s(n) \]  

(9)

where \( S(n) \) is the average signal power.
If the signal is a plane wave, the elements \( y_i(n) \) of \( Y(n) \) are all delayed replicas of each other; thus:

\[
y_i(n) = c_i s(n) e^{j \frac{2\pi n}{T} r_i}
\]

(10)

where \( s(n) \) is the \( n^{th} \) Fourier coefficient of the signal waveform; the \( c_i \) are weighting factors to take into account that the signal strength or gain at different hydrophones may be different, and \( r_i \) is the delay at the \( i^{th} \) hydrophone. The \( c_i \)'s are conveniently defined in such a way that:

\[
\langle s^*(n) s(n) \rangle = s(n)
\]

(11)

for all \( n \). Hence

\[
Y(n) = s(n) \begin{bmatrix} c_1 e^{j \frac{2\pi n}{T} r_1} \\ \vdots \\ c_N e^{j \frac{2\pi n}{T} r_N} \end{bmatrix} = \mathcal{S}(n) \mathcal{W}(n),
\]

(12)

and therefore

\[
P(n) = \frac{\langle s^*(n) s(n) \rangle}{s(n)} \frac{V(n) V^T(n)}{s(n)} = V^*(n) V^T(n)
\]

(13)

\( P(n) \) is seen to be of rank 1. Because of the independence of signal and noise the covariance matrix of signal and noise together is \( N(n) Q(n) + S(n) P(n) \).

Therefore

\[
f_s(x(n)) = C_{S+N}(n) \exp \left\{ -\frac{1}{2} x^T(n) \left[ N(n) Q(n) + S(n) P(n) \right]^{-1} x(n) \right\}
\]

(15)

and therefore the likelihood ratio is

\[
LR = \prod_{n=1}^{\infty} \frac{C_{S+N}(n)}{C(n)} \exp \left\{ -\frac{1}{2} x^T(n) \left[ \frac{Q^{-1}(n)}{N(n)} - \left( N(n) Q(n) + S(n) P(n) \right)^{-1} x(n) \right] \right\}
\]

(16)

Since \( P(n) \) is of rank 1 the inversion of the second term in the brackets is easily accomplished by using the following identity [4]: For an arbitrary nonsingular
matrix $A$ and two vectors $U$ and $V$

$$\left(A + UV^T\right)^{-1} = A^{-1} - \frac{A^{-1} U V^T A^{-1}}{1 + V^T A^{-1} U} \tag{17}$$

Using this identity

$$\left[ N(n) Q(n) + S(n) P(n) \right]^{-1} \frac{Q^{-1}(n)}{N(n)} = \frac{S(n) Q^{-1}(n) P(n) Q^{-1}(n)}{N^2(n) \left[ 1 + S(n) G_o(n)/N(n) \right]} \tag{18}$$

where $G_o(n) = V^T(n) Q^{-1}(n) V(n)$

As is shown by Edelblute et. al. [3], $G_o(n)$ is the maximum value of the array gain. Using this result one finds that the logarithm of the likelihood ratio is given by

$$\log LR = C + \sum_{n=1}^{\infty} \frac{S(n) X^T(n) Q^{-1}(n) P(n) Q^{-1}(n) X^*(n)}{N^2(n) \left[ 1 + S(n) G_o(n)/N(n) \right]} \tag{20}$$

where $C = \sum_{n=1}^{\infty} C_{N(n)} / C_{N(n)}$

Since $Q^T(n) = Q^*(n)$ the quadratic form appearing in this expression can be written in the form:

$$X^T(n) Q^{-1}(n) P(n) Q^{-1}(n) X^*(n) = X^T(n) Q^{-1}(n) V^T(n) Q^{-1}(n) X^*(n)$$

$$= \left\| V^T(n) Q^{-1}(n) X(n) \right\|^2 = \left\| Z^T(n) X(n) \right\|^2 \tag{21}$$

and therefore

$$\log LR = C + \sum_{n=1}^{\infty} \frac{\left\| H(n) X(n) \right\|^2}{N(n) \left[ 1 + S(n) G_o(n)/N(n) \right]} \tag{22}$$

where $H(n) = \frac{S(n) Q^{-1}(n) V^T(n)}{N(n) \left[ 1 + S(n) G_o(n)/N(n) \right]} \tag{23}$
This implies that log LR can be obtained from a circuit of the form shown in Fig. 1 where the elements of the filter bank for $f_n$ are the elements of the vector $H(n)$.

![Diagram of Likelihood Ratio Detector](image)

Figure 1  Likelihood Ratio Detector

Except for the scalar multiplier this result is the same as that obtained by Knapp [5]. The difference between the result given here (which is that of Edelblute, et. al.) and Knapp's stems from the fact that Knapp's filter maximizes the output signal-to-noise ratio defined in terms of a signal-plus-noise covariance matrix $Q_x$, rather than producing log LR.

III. Detection Index

In this section it is assumed that the signal-to-noise ratio is small; i.e. $S(n)G_o(n) \ll N(n)$ for all frequencies. In this case one can assume that log LR is approximately a Gaussian random variable and that the performance of the detector is specified by

$$d = \frac{\mu_{LR}E_{LR} - \mu_{N}}{\sqrt{\text{Var}(LR)}}$$

(24)
where \( u = \log LR \), and where \( \sigma_N(u) \) is the standard deviation of \( u \) under the hypothesis that there is only noise at the input. The computation for \( d \) is straightforward; for details see Appendix A. The result is

\[
d = \frac{\sum_{n=1}^{WT} K(n) S(n) G_\alpha^2(n)}{\sqrt{\sum_{n=1}^{WT} K^2(n) N^+(n) \sigma_\alpha^2(n)}}
\]

where \( K(n) = \frac{S(n)/N^2(n)}{1 + S(n) G_\alpha^2(n)/N(n)} \) if \( S(n) G_\alpha^2(n) \ll N(n) \) (26)

Using the small signal approximation to \( K(n) \)

\[
d = \sqrt{\sum_{n=1}^{WT} \frac{S^2(n) G_\alpha^2(n)}{N^2(n)}} - \frac{\sum_{n=1}^{WT} S(n) G_\alpha^2(n)}{N^+(n)} \sum_{n=1}^{WT} \frac{S^2(n)}{N^2(n)} [ \frac{1}{\sigma^T(n) \sigma^{-1}(n) Y^b(n)} ]^2
\]

Note that in earlier reports we used \( d = \frac{1}{\sqrt{2}} \)

\[
d = \sqrt{\sum_{n=1}^{WT} - \frac{1}{2} \text{tr} \left[ \sigma(n) \sigma^{-1}(n) \right]}
\]

This is equivalent to Eq. (27) as can be seen by using the identity

\[
U^T A U = \text{tr} A \text{tr} U^T U = \text{tr} U^T A
\]

where \( U \) is a vector and \( A \) is a square matrix. Hence

\[
\text{tr} \left[ \sigma(n) \sigma^{-1}(n) \right] = \text{tr} \left[ \sigma^2(n) \sigma^{-1}(n) Y^b(n) \sigma^T(n) \sigma^{-1}(n) Y^b(n) \sigma^{-1}(n) \right]
\]

The bracketed term \( \text{tr} \left[ \sigma(n) \sigma^{-1}(n) \right] \) can be factored out; in fact, this
scalar is \( G_r(n) \). The remaining terms are also equivalent to \( G_r(n) \) by Eqs. (29) and (19); thus \( \int P(n) G_r^{-1}(n) \overline{a}^2 = G_r^2(n) \). The factor of 1/2 appearing in Eq (28) does not appear here because of the complex notation that is used. Thus Eqs. (27) and (28) are identical.

As \( T \to \infty \) the equation appearing in Eq (27) can be converted to an integral as follows:

\[
\overline{d} = \sqrt{\frac{n}{N^2(n)}} \int \overline{T} \frac{S^2(f) G_r^2(f)}{N^2(f)} \, df \tag{31}
\]

IV. Effect of Directional Interference

Suppose that the noise component of \( x(t) \) consists of two parts, an isotropic part and an interference part. It is assumed that the interference is generated by \( R \) point sources. The \( r \)th point source is located at an azimuth angle \( \Theta_r \), and its spectral density is \( I_r(\omega) \); hence the interference power from the \( r \)th interference source at the frequency \( \omega_n \) is \( I_r(n) \). The desired target is at the azimuth angle \( \Theta_0 = 0 \), and it is assumed that the array is steered in the target direction. The isotropic noise power at the frequency \( \omega_n \) is \( N_0(n) \). The isotropic noise component, the interference sources, and the target signal are all assumed to be mutually independent Gaussian processes with zero mean. Then the total noise power density is given by

\[
N(n) = N_0(n) + \sum_{r=1}^{R} I_r(n) \tag{32}
\]

and the normalized noise covariance matrix has the form:

\[
\begin{pmatrix}
\delta_{rr} & \overline{a} & 0 & \cdots & 0 \\
\overline{a} & \delta_{rr} & \overline{a} & \cdots & 0 \\
0 & \overline{a} & \delta_{rr} & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & \overline{a} & \delta_{rr}
\end{pmatrix}
\tag{33}
\]
where \((n')\) is the normalized covariance matrix of the isotropic noise component and where each element of the matrix results from one of the interference point sources. By direct analogy to Eqs. (12), (13), and (14)

\[
\Sigma_p(r) = \begin{bmatrix}
\sigma_r^2 & \sigma_{r1} & \cdots & \sigma_{rn} \\
\sigma_{r1} & \sigma_{11} & & \\
\vdots & & \ddots & \\
\sigma_{rn} & & & \sigma_{nn}
\end{bmatrix}
\]

where \(\sigma_{ij}\) is the delay of the plane wave from the \(i\)th interference source at the \(j\)th hydrophone.

The matrix \(Q(n)\) can be inverted by using the following generalization of Eq. (17): if \(A\) is a non-square matrix of dimension \(M\) and \(\bar{B}\) is a matrix of \(M\) rows and \(R\) columns then

\[
(A + \bar{B})^{-1} = \bar{B}^{-1} - \bar{B}^{-1} A (I + \bar{B}^{-1} A)^{-1} \bar{B}^{-1}
\]  

(35)

This identity is easily proved by multiplication:

\[
(A + \bar{B})^{-1} = \bar{B}^{-1} - \bar{B}^{-1} A (I + \bar{B}^{-1} A)^{-1} \bar{B}^{-1} A
\]

\[
= \bar{B}^{-1} - \bar{B}^{-1} A (I + \bar{B}^{-1} A)^{-1} \bar{B}^{-1} A
\]

\[
= \bar{B}^{-1} - \bar{B}^{-1} A (I + \bar{B}^{-1} A)^{-1} \bar{B}^{-1} A
\]

\[
= \bar{B}^{-1} + \bar{B}^{-1} A - \bar{B}^{-1} A (I + \bar{B}^{-1} A)^{-1} \bar{B}^{-1} A
\]

\[
= \bar{B}^{-1} + \bar{B}^{-1} A - \bar{B}^{-1} A (I + \bar{B}^{-1} A)^{-1} \bar{B}^{-1} A
\]

\[
= \bar{B}^{-1} + \bar{B}^{-1} A - \bar{B}^{-1} A (I + \bar{B}^{-1} A)^{-1} \bar{B}^{-1} A
\]

In the present application, set

\[
A = \frac{\mathbf{\Sigma}_p(n)}{\mathbf{\Sigma}_r(n)}
\]

(37)

\[
\mathbf{\Sigma}_p(n) = \begin{bmatrix}
\mathbf{\Sigma}_p(1) \\
\vdots \\
\mathbf{\Sigma}_p(R)
\end{bmatrix}
\]

\[
\mathbf{\Sigma}_r(n) = \begin{bmatrix}
\mathbf{\Sigma}_r(1) \\
\vdots \\
\mathbf{\Sigma}_r(R)
\end{bmatrix}
\]

(38)
where $K_r = \frac{u_r(n)}{N_r(n)}$ for $r = 1, \ldots, R$

Also, to simplify the notation let

$$G_{rs} = G_{rs}^0(n) = \sqrt{T_r(n) C_{rr}} \sum_{k=0}^{N_r} |v_s(n)|^2$$

Note that $G_{rs}^0(n)$ has the general form of an array and hence Eq. (17'), we can call it a "cross-array". It is clear that

$$G_{rs}^0(n) = G_{rs}^0$$

In terms of this notation the matrix $\sqrt{\kappa^{-1} G}$ of Eq. (45) becomes

$$I + \begin{bmatrix} \sqrt{K_1 V_1^T} & -1 & \sqrt{K_2 V_2^T} & \cdots & \sqrt{K_R V_R^T} \\ -1 & \sqrt{K_1 V_1^T} & \sqrt{K_2 V_2^T} & \cdots & \sqrt{K_R V_R^T} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{K_1 V_1^T} & \sqrt{K_2 V_2^T} & \sqrt{K_3 V_3^T} & \cdots & \sqrt{K_R V_R^T} \end{bmatrix}$$

by an obvious definition of the "cross-array" matrix $G$. Note that $G$ is square, of dimensionality $R$, and invertible.

The detection index is given by Eq. (72) except that for the sake of consistency we rewrite $V(n)$ as $\tilde{V}(n)$. Then

$$d = \sqrt{\frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{r=1}^{R} |G_{rr}^0 \tilde{V}_r(n)|^2}$$

$$= \sqrt{\frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{r=1}^{R} |u_r(n)|^2}$$
where

\[
F(n) = \mathbf{V}(n) \mathbf{G}(n) \mathbf{V}(n) = \mathbf{v}^T(n) \mathbf{v}(n) \mathbf{R}^{-1}
\]

This equation can be written in a more compact form by defining a vector \( \mathbf{g} \) such that

\[
\mathbf{g}^T = \begin{bmatrix} \sqrt{K_{11}G_{11}} & \sqrt{K_{22}G_{22}} & \ldots & \sqrt{K_{nn}G_{nn}} \end{bmatrix}
\]

Then, since \( \mathbf{G}_{ss} = \mathbf{G}_{rs} \)

\[
F(n) = \mathbf{G}_{ss} - \mathbf{g}^T \left[ \mathbf{I} + \mathbf{g} \mathbf{g}^T \right]^{-1} \mathbf{g}^T
\]

It is clear that \( \mathbf{G}_{ss} \) is the optimum array gain with isotropic noise and that \( \mathbf{g}^T \left[ \mathbf{I} + \mathbf{g} \mathbf{g}^T \right]^{-1} \mathbf{g}^T \) represents the effect of the interference. In general the evaluation of the interference term in specific instances is difficult for two reasons:

a) The cross-array gain \( \mathbf{G}_{ij} \) are quadratic forms involving the inverted

\[
\mathbf{G}(\mathbf{R}) \nonumber
\]

b) Even if the \( \mathbf{G}_{ij} \) are known the \( \mathbf{I} + \mathbf{g} \mathbf{g}^T \) matrix must be inverted.

Thus it is necessary either to solve the (4.11) by computer or to make approximations permitting analytic results.

The standard approximations used in most previous analyses and which eliminates part of the difficulty involved in evaluating the \( \mathbf{G}_{ij} \) is to assume that there is no cross talk between different hydrophones due to the isotropic noise - i.e., the noise in hydrophones have been considered before and is assumed of an isotropic nature, i.e., \( \mathbf{G}_{ij} = 0 \) in any significant way.
A further small simplification results from the fact that the array is steered on target; this implies that \( \gamma_1 = \gamma \), and therefore

\[
G_{or}(n) = \frac{M}{i=1} e^{-i\omega_\gamma (r_i - r_i)}
\]

Eq. (47) is now explicitly evaluated for a number of simple special cases.

V. Single Point Interference

If there is only a single interference, the Eq. (47) becomes

\[
F(n) = G_{oo} - \frac{K_1 |G_{ol}|^2}{1 + K_1 G_{ll}}
\]

and with the simplification \( G_{oo}(n) = 1 \); this becomes

\[
F(n) = M - \frac{K_1 |G_{ol}|^2}{1 + K_1 M}
\]

By Eq. (50)

\[
|G_{ol}|^2 = \frac{M}{i=1} \sum_{k=1}^{M-1} e^{i\omega_\gamma (r_i - r_k)} = \frac{M}{i=1} \sum_{k=1}^{M-1} \cos \omega_\gamma (r_i - r_k)
\]

where, for simplicity \( r_1 (1) = r_i \).

Hence the term representing the loss of detectability due to interference in Eq. (47) becomes

\[
\frac{K_1 M}{1 + K_1 M} \left[ 1 + \frac{2}{M} \sum_{i=1}^{M-1} \sum_{k=i+1}^{M} \cos \omega_\gamma (r_i - r_k) \right]
\]

This result is essentially identical to Eq. (72) of Ref. 1.
The value of the series sum in Eq.(56) can be further evaluated for the special case of $Q = 1$, or $\phi = 0$, in which case all the array elements, $u$, the velocity of sound, and the reference distance, $r$, are from the interference source to the center of the array which is perpendicular to the array axis. If $Q = 1$ in Eq.(56), the sum in $\sum_{k=1}^{M-k}$ can be replaced by a single sum as shown:

$$
\frac{K \sum_{j=k}^{M-k}}{j + k} \quad \text{for} \quad j = -k, -k + 1, \ldots, k - 1, k
$$

(56)

Substituting this expression into Eq.(56) results in

$$
d = \sqrt{\frac{\pi}{n \sigma_{\Delta n}^{2}}} \left[ 1 - \frac{1}{1 + \sqrt{\frac{\pi}{n \sigma_{\Delta n}^{2}}}} \right] \sum_{k=1}^{M-k} (H - k \cos k \theta_0) \left( \frac{M-k}{\sum_{j=k}^{M-k} j + k} \right)^{1/2}
$$

(57)

where $\sigma_{\Delta n} = \sigma_{\Delta n}(n)$.

Since our interest is the effect of the oscillatory term, assume that $\Delta(n) = 0$ and $\sigma_{\Delta n}(n)$ are independent of $n$ and are equal, interference, and electronic noise spectra. Then Eq.(57) can be written as a function of $\theta_0$ and can be performed by converting it into an integral, as in Eq.(1): i.e.

$$
d = \sqrt{\frac{\pi}{n \sigma_{\Delta n}^{2}}} \left[ 1 - \frac{1}{1 + \sqrt{\frac{\pi}{n \sigma_{\Delta n}^{2}}}} \right] (M-k \cos k \theta_0) \left( \frac{M-k}{\sum_{j=k}^{M-k} j + k} \right)^{1/2}
$$

(58)
\[ d^2 = \frac{1}{N_0^2} \left( \left( 1 - \frac{K_1 M}{1 + K_1 M} \right)^2 - \frac{L K_1 M}{1 + K_1 M} \frac{|1 + K_1 (H-1)|}{(1 + K_1 M)^2} \right) \sum_{k=1}^{M} \frac{\sin 2\pi(k-k)M}{2\pi(k-k)M} \]

\[ + \frac{2K_1^2 M^3}{(1 + K_1 M)^2} \sum_{k=1}^{M} \sum_{i=1}^{M} \frac{(H-k)(H-i)}{2\pi(k-i)M} \left( \frac{\sin 2\pi(k-i)M}{2\pi(k-i)M} + \frac{\sin 2\pi(k+i)M}{2\pi(k+i)M} \right) \]

(59)

In most practical cases the maximum frequency processed is such that \(2\pi M_{\text{io}} \gg 1\); this is also consistent with the assumption that \(Q_{\text{o}}(n) = 1\). Then the summand of the single summation in the second term is small and oscillatory; the last term except when \(k = j\); thus the last term is approximately equal to \(\frac{2K_1^2 M^3}{3(1 + K_1 M)^2}\) and

\[ d^2 \approx T_{o} \frac{S^2}{N_0^2} \left[ R - \frac{K_1 M}{1 + K_1 M} \left( 1 - \frac{1}{3} \frac{K_1 M}{1 + K_1 M} \right) \right] + O\left( \frac{K_1 M}{1 + K_1 M} \right) \]

(60)

where \(O(\cdot)\) is a remainder term which is of order \(1/M\) relative to the first term.

Hence if \(K_1 M \gg 1\),

\[ d^2 \approx T_{o} \frac{S^2}{N_0^2} \left[ M - \frac{2}{3} \right]^2 \]

(61)

which indicates that the effect of the interference is equivalent to the loss of \(2/3\) of a hydrophone from the array. This is substantially in agreement with the conclusion of Ref. [11] where it was concluded that the loss under these conditions was equivalent to one hydrophone; this was arrived at by completely neglecting the summation term in Eq. (54). Since the figure of \(2/3\) has been derived only for a linear array, and since it probably depends on the array configuration, the figure of one hydrophone is probably a reasonable estimate in general.

VI. Two Point Interferences

with two interference, Eq. (61) becomes.
\[ F(n) = G_{oo} - \left[ G_{oo}^* \frac{K_1G_{11} + K_2G_{12}}{G_{oo}^*} \right]^{-1} \left[ \frac{G_{11}K_2^* + G_{12}K_1^*}{G_{oo}^*} \right] \]  

\[ G_{oo} = G_{11} - G_{12} - K_1 K_2 \frac{F(n)}{2} \]  

where \( \text{Re}(\ ) \) means \( \text{real part} \) 

As before the term \( G_{oo} \) represents the effect of the interference. If, as before, 

\[ G_{oo}(n) = 1 \]

then 

\[ |G_{oo}|^2 = \sum_{i=1}^{M-1} \sum_{k=1}^{M} \cos \omega_n \left( \tau^{(i)}_k - \tau^{(i)}_{k+1} \right) + \sum_{i=1}^{M} \sum_{k=1}^{M} \cos \omega_n \left( \tau^{(i)}_k - \tau^{(i+1)}_k \right) \]

and 

\[ \text{Re}(G_{oo}^* G_{11}) = \sum_{i=1}^{M} \sum_{k=1}^{M} \cos \omega_n \left( \tau^{(i)}_k - \tau^{(i)}_{k-1} \right) \]

If the two interference sources are widely separated in angle from each other and from the target, then \( \tau^{(i)}_k \) and \( \tau^{(i+1)}_k \) differ substantially for all \( k \) and are not close to zero. From these conditions the coefficient of \( K_1 K_2 \) in both the numerator and denominator is substantially increased with \( n \) while the other terms increase with \( n^2 \). Hence, for large \( n \) these coefficients become negligible with the result.

\[ F(n) \rightarrow \frac{K_1G_{11} + K_2G_{12}}{G_{oo}^*} \]  

\[ \text{Re}(\ ) \]  

(62)  

(63)  

(64)  

(65)  

(65a)
where the second approximation involves neglect of the oscillating terms $|G_{01}|^2$ and $|G_{02}|^2$. Thus the effect of interferences is seen to be additive under these conditions. For small interference-to-ambient-noise ratio, where $K_1 M$ and $K_2 M$ are very much less than unity, $F(n)$ is reduced roughly by $(K_1 M)^2$ while for very large interference-to-ambient-noise ratio, the reduction is no greater than 2. Thus for small interference-to-noise ratio the detection index $J$ decreases roughly with the first power of interference-to-noise ratio (see Eq. (44)), but the maximum effect is no greater than the loss of two hydranges.

Suppose next that the two interference sources are sufficiently close together so that for all frequencies of interest, and for all $t$

$$\omega_n |\tau_1^{(1)} - \tau_1^{(2)}| \approx 0$$

then $|G_{02}|^2 \approx |G_{01}|^2$ 

$|G_{12}|^2 \approx n^2$

and $\text{Re}(G_{01} G_{12}^* G_{02}) \approx n |G_{01}|^2$

Then $F(n) \approx G_{00} - \frac{K_1 (1 + K_2 k) |G_{01}|^2 + K_2 (1 + K_1 M)|G_{01}|^2}{1 + (K_1 + K_2) M} - 2nK_1 K_2 |G_{01}|^2$

$$= G_{00} - \frac{(K_1 + K_2) |G_{01}|^2}{1 + (K_1 + K_2) M}$$

(68)

Thus the result converges to the case of a single plane-wave interference of strength $K_1 + K_2$ in this case.

For a linear array of $N$ elements spaced $d$ ft. apart we can set

$$|\tau_1^{(1)} - \tau_1^{(2)}| = |d/2 - 1| |d/2| \sin \theta_1 - \sin \theta_2$$

where the delay at the relative peak effect is determined over the $N$ elements.

For small $\theta_0 - \theta_1$,
\[
|r_1^{(1)} - r_1^{(2)}| = 2 |K/2 - i \frac{d}{c} \sin \left( \frac{\theta_1 - \theta_2}{2} \right) \cos \left( \frac{\theta_1 + \theta_2}{2} \right) |
\]

where \( \theta_m = \frac{\theta_1 + \theta_2}{2} \) is the angle half-way between the two interferences. Then, since \( \omega_n \) is 2\( \pi \)W

\[
\omega \left| r_1^{(1)} - r_1^{(2)} \right|_{\text{max}} = 2 \omega_n \frac{d}{c} \left| \theta_1 - \theta_2 \right| \cos \theta_m
\]

and the two interferences are close enough together so that Eq. (68) holds if

\[
\omega_n \frac{d}{c} \left| \theta_1 - \theta_2 \right| \cos \theta_m \ll 1
\]

or

\[
\left| \theta_1 - \theta_2 \right| \ll \frac{1}{\omega_n \frac{d}{c} \cos \theta_m}
\]

As an example let \( W = 5000 \) cps, \( d = 2 \) ft, \( c = 5000/\)ft and \( N = 10 \), then if \( \left| \theta_1 - \theta_2 \right| \ll 0.06/\cos \theta_m \) radians the two interference points have the same effect as a single one with a higher power level and therefore the maximum detectability loss can be no greater than 2/3 of a hydrophone as shown by Eq. (61).

Note that for \( \theta_m = \pi/2 \) the two interferences are located symmetrically to the end-fire axis of the array; therefore their effect is always that of a single interference. This, however, is due to the symmetry of the linear array and does not hold in other cases.

Note further that Eq. (72) is a rather conservative limit since neither the effect of integration over frequency or over hydrophone spacing has been considered. Depending on the exact form of the power spectrum these integrations should result in increasing the value of \( \left| \theta_1 - \theta_2 \right| \) by a factor of 4 or 5 over that given in Eq. (72).

If the interference-to-noise ratio is small enough so that \( (K_1 + K_2)M \ll 1 \), then Eq. (68) and Eq. (65a) are approximately the same; thus under this condition
the effect of two interferences on the detectability is proportional to the interference power, and independent of the spacing of the two interference sources from each other. Note, however, that approximating $|G_{o1}|^2$ and $|G_{o2}|^2$ by $M$ still implies that the interference direction is substantially different from the target direction.

VII. More than Two Interferences

Extension of the results obtained so far to more interferences is difficult and involves further approximations. Consider first the large interference-to-noise ratio case, with all interferences widely separated. We assume as before that $Q_o = 1$ and that therefore $G_{rr} = M$ for $r = 0, 1, 2, \ldots, R$. It can be seen from Eq. (64) that the off-diagonal elements of $G$ are of order $\sqrt{M}$. It can be shown in general (see Appendix B) that for the purpose of approximate inversion an $n$-dimensional matrix whose diagonal elements are of order $k$ relative to the off-diagonal elements can be approximated by a diagonal matrix if $k \gg n$. Thus if $\sqrt{M} \gg R$ the off-diagonal elements of the matrix $[I + G]$ can be neglected in forming the inverse, with the result that

$$F(n) \approx G_{oo} - \sum_{r=1}^{R} \frac{K_r |G_{ro}|^2}{1 + K_r G_{rr}}$$

$$= M - \sum_{r=1}^{R} \frac{K_r |G_{ro}|^2}{1 + K_r M}$$

$$= M - \sum_{r=1}^{R} \frac{K_r M}{1 + K_r M} \left[ 1 + \frac{r-1}{2} \sum_{i=1}^{r} \cos \theta_{i} + \sum_{i=1}^{r} k_{i+1} \right]$$

where $\theta_{i}$ is the interference delay from the $i$th interference source at the $i$th hydrophone. Eq. (73) is the general expression of Eq. (64) and indicates that for
widely spaced point interference sources the detectability loss is approximately equivalent to the loss of one hydrophone per interference source. The approximation is good only for \( R \ll \sqrt{M} \).

VIII. Effect of Distributed Interference Source

A distributed interference source can be represented by a large number of closely spaced point sources. Suppose that the interference source has a spectral density \( I(n) \) and that the interference power is uniform for angular inside the interval \( \theta_1 \leq \theta \leq \theta_2 \) and zero outside. Then the interference can be represented by \( R \) points of spectral density \( I(n)/R \) equally spaced in the interval, where \( R \) is a large number. Initially it will be assumed that the interference-to-ambient noise ratio is small. Although the result obtained under this assumption is somewhat academic (since the interference effect is very small in any case) it is possible to obtain an analytic result which is usually applicable with some modifications to larger interference-to-ambient-noise ratios as well. Under this assumption, the elements of the matrix \( G \) are all very small, and it is approximately true that

\[
I(n) = G_{rs} I(n)
\]

Then the matrix inversion is, of course, trivial. To precise conditions for Eq.(76) to be a good approximation may be deduced if \( |G_{rs}| \): a simple sufficient condition is that

\[
\frac{R}{s=1} \left| K_r K_s \right| |G_{rs}| \ll 1 \quad \text{for all } r, s = 1, \ldots, R
\]

In the present discussion \( K_r = K_I/k_r \) for all \( r = 1, \ldots, \), where \( K_I = I(n)/N_o(n) \) is the total interference-to-ambient-noise ratio. A conservative upper bound on \( K_I \) such that Eqs.(76) and (77) are good approximations is obtained by letting

\[
|G_{rs}| = K_I k_r < 1 \quad \text{for all } r, s (\text{see Eq.(40)}). \]

Hence, if

\[
K_I k_r < 1
\]
Eq. (76) is a good approximation, and under these conditions Eq. (77) becomes

\[ F(n) = G_{oo} - \sum_{r=1}^{R} \frac{R}{r} |G_{rr}|^2 \]

and by use of Eq. (64) this becomes:

\[ F(n) = G_{oo} - \frac{K_i}{R} \sum_{r=1}^{R} \left( \sum_{i=1}^{\infty} \frac{1}{k_{i+1}} \sum_{k=1}^{\infty} \cos \omega_n \left( \frac{r}{r+1} \right) \right) \]

We assume now that the azimuth angle subtended by the interference is small enough so that the \( \tau_i \) do not differ very greatly as \( r \) goes from 1 to \( R \). Then, it is possible to expand \( \tau_i \) in a Taylor series in \( r \) as follows:

\[ \tau_i(r) = \tau_i(m) + (r - m) \Delta \tau_i \]

where \( m = \frac{R}{2} \) is used as the point about which the expansion is performed; \( \tau_i(r) \) is effectively the mean delay of the interference wavefront.

As \( R \) is allowed to go to infinity the summation in \( r \) can be converted into an integral and evaluated as follows:

\[ R \sum_{r=1}^{\infty} \cos \omega_n \left( \tau_i - \tau_i \right) \]

\[ R \sum_{r=1}^{\infty} \cos \omega_n \left( \tau_i - \tau_i \right) \]

\[ R \int_{0}^{\infty} \cos \omega_n \left( \tau_i - \tau_i \right) \sin \frac{n}{2} \Delta \tau_i \Delta \tau_i \]

\[ = R \frac{n}{2} \frac{\cos \omega_n \left( \tau_i - \tau_i \right)}{\sin \frac{n}{2} \Delta \tau_i \Delta \tau_i} \]

(82)
Hence Eq. (80) becomes, after some reduction:

\[
P(n) = G_0 - K_1 \left[ H + 2 \sum_{i=1}^{N-2} \sum_{k=1}^{M} \frac{\sin \frac{\omega R}{2} (\Delta \tau_i - \Delta \tau_k)}{\sin \Delta \tau_1 (m) - \tau_1 (m)} \right] (83)
\]

\[
= G_0 - K_1 \left[ H + 2 \sum_{i=1}^{N-2} \sum_{k=1}^{M} \frac{\sin \omega R}{2} (\Delta \tau_i - \Delta \tau_k) \cos \omega (\tau_1 (m) - \tau_1 (m)) \right] (84)
\]

where, in going from Eq (83) to (84) we have used the fact that

\[
\frac{\omega R}{2} (\Delta \tau_i - \Delta \tau_k) = m(\Delta \tau_i - \Delta \tau_k) = \tau_i (m) - \tau_k (m)
\]

As before, the term representing the loss of detectability is the bracketed term in Eq. (84). Except for the \(\frac{\sin x}{x}\) term the form of the double summation is the same as that which would be obtained for a single interference, (Eq. 5h) with a mean delay \(\tau_i (1)\) at the \(i^{th}\) hydrophone, under the condition \(K_1 \ll 1\). In fact, the argument that the summation of oscillating terms tends to become negligible applies here with ever greater force, because of the \(\frac{\sin x}{x}\) term. One can conclude, therefore, that for interference-to-signal-to-noise ratio small enough to satisfy Eq. (78), and if the angle subtended by the interference is relatively small, a distributed interference source affects the performance in essentially the same way as a single point interference.

In order to obtain an estimate of the magnitude of azimuth angle that can be considered "small", consider a linear array with \(M\) hydrophones spaced \(d\) feet apart. For such an array

\[
\tau_i (r) = i \frac{d}{c} \sin \theta_i
\]

where \(\theta_i\) is the azimuth angle of the \(i^{th}\) interference point. Assume that the interference power is confined to the range \(\theta_1 \leq \theta \leq \theta_2\) and is zero outside this range.
range. The center of the interference is at the angle

$$\eta_n = \frac{1}{2} (\eta_1 + \eta_2)$$  \hspace{1cm} \text{(87)}$$

Then, by analogy with Eq. (81), we expand \( \sin \eta_n \) about \( \eta_n \); i.e.,

$$\sin \eta_n \approx \sin \eta_n + \eta_n - \frac{\eta_n^3}{6} \cos \eta_n \hspace{1cm} \text{(88)}$$

All the other steps leading to Eq. (88) can then be performed in exactly the same way, with the summation over \( \Lambda \) replaced by an integration over \( \eta_n \). The final result can be put into the form:

$$F(n) = G_{co} - K_I \left| \eta_n - \frac{1}{2} \sum_{k=1}^{\infty} \frac{k\eta_n}{n} \cos \left( \frac{k \eta_n}{n} \right) \sin \left( \frac{k \eta_n^3}{6n \cos \eta_n} \right) \cos \left( \frac{\eta_2 - \eta_1}{m} \right) \right|^2$$ \hspace{1cm} \text{(89)}$$

Except for the \( \frac{\sin x}{x} \) term in the summation, this is again the expression that we would have obtained for a single point interference located at the angle \( \eta_n \). It is clear that the accuracy of this expression depends on the accuracy of Eq. (88), which, in turn, is a fairly good approximation for \( \eta_2 - \eta_1 \) less than about one radian. Thus we conclude that an interference source spread over no more than one radian affects the detectability essentially like a single point interference, provided that \( K_I \ll 1 \).

Since the effect of interference for \( K_I \ll 1 \) is very small, the result just obtained is not very interesting and it would be desirable to extend it somehow to the case of \( K_I \gg 1 \). Unfortunately this is quite difficult; in fact, the only simple result that has been obtained is an extension of Eq. (88) to more than 2 interference sources. If, in the case of the interferences, it is assumed that the interference points are close enough together so that for all frequencies of interest:

$$\frac{K_I}{1}$$ \hspace{1cm} \text{(90)}$$
then $|G_{or}|^2 \geq |G_{el}|^2$ for all $r=1 \ldots R$ (91)

and $G_{rs}(n) \geq M$ for all $r,s,n.$ (92)

Then the matrix $I - G$ becomes

$$
I - G = \begin{bmatrix}
1 + \frac{K_1^M}{1 + R_1 M} & \frac{\sqrt{K_1^M}}{1 + R_1 M} & \ldots & \frac{\sqrt{K_1 M}}{1 + R_1 M} \\
\frac{\sqrt{K_2^M}}{1 + R_2 M} & 1 + \frac{K_2^M}{1 + R_2 M} & \ldots & \frac{\sqrt{K_2 M}}{1 + R_2 M} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\sqrt{K_R^M}}{1 + R_R M} & \frac{\sqrt{K_R M}}{1 + R_R M} & \ldots & 1 + \frac{K_R^M}{1 + R_R M}
\end{bmatrix}
$$

This is easily inverted by use of Eq. (17)

$$
|I - G|^{-1} = I - \frac{G}{1 + M \cdot \sum_{r=1}^{R} K_r} \tag{95}
$$

Also the vector $g$ defined in Eq. (16) becomes:

$$
g^T = G_{el} \left[ \frac{\sqrt{K_1}}{1 + R_1} \frac{\sqrt{K_2}}{1 + R_2} \ldots \frac{\sqrt{K_R}}{1 + R_R} \right]^T \tag{96}
$$

and therefore $F(n)$ of Eq. (17) becomes

$$
F(n) = M - \frac{G_{el}^2}{1 + M \cdot \sum_{r=1}^{R} \frac{K_r}{1 + R_r} \cdot \frac{1}{1 + M \cdot \sum_{r=1}^{R} \frac{K_r}{1 + R_r}}} \tag{97}
$$

Obviously, if all the $K_r$ are equal to $K_1/R$, $F(n)$ becomes

$$
F(n) = M - \frac{G_{el}^2}{1 + R_1 M} \tag{98}
$$
Thus the detectability loss is again equivalent to that of a single interference source of strength $K_f$, as one could have expected. It is clear from the result for the case of two point sources that for a linear array having $M$ equally spaced hydrophones the maximum value of $\frac{\alpha}{\phi} - 1$ for which Eqs. (97) or (98) hold is given by Eq. (72).

IX. Computational Results

Since it has not been possible to obtain meaningful analytic results for cases in which the approximations made in the above work are not applicable, $F(n)$ has been evaluated on a digital computer for a number of different array and interference patterns, and for specific frequencies. The results of some of these computations are presented in Figures 1 through 6. In all computations it is assumed that the array is steered on target at an angle $\theta = 0$ and that an interference exists at some angle $\theta_I$. The curves are then plots of $F(n)$ as $\theta_I$ is varied. Thus, if $\theta_I$ is near 0 the interference is near the target in azimuth, and $F(n)$ is small. Also, the assumption that correlation of ambient noise waveforms between different hydrophones is zero has not been used; instead the exact form of the $Q_0(n)$ matrix as given by Bryn [3] was used. As a result $F(n) \neq M$ in the absence of interference as would be inferred from equations such as (52), (65), (75) or (98). In fact, $F(n) < M$ in all cases; however, this is a coincidence; it is possible for $F(n) > M$ as is shown by Bryn [3]. In all cases the interference-to-ambient-noise ratio is large.

Figure 1 shows the effect of a single point interference with a small circular array. It shows that if the interference directions differ by more than about $40^\circ$ from the target direction the effect on $F(n)$ is essentially negligible. It must be borne in mind, however, that this has been demonstrated for a single frequency ($F(n)$, frequency). It may not be the same at other frequencies, and
the integrated effect of all frequencies therefore, has the effect of the loss of one hydrophone as is predicted by the analysis of Section V.

Figures (2) and (3) are similar to Figure 1 except that the interference pairs are respectively of two or four near points, separated by .1 radian. Since the interference covers a larger area with segment, the effect on F(n) covers a larger angle; however, it is still true that for interference sources at angles for removed from zero the effect on F(n) is small. A similar result is shown in Fig. (4) which shows the effect of two interferences separated by a large angle (90°). The last two figures show the effect of a strong interference (K_1=200) distributed over a relatively large angle (17°). Again the effect at angles far removed from the target angle is small, but, as is shown in Fig. (6), the relative effect is quite different at different frequencies, as has already been pointed out.

The computations leading to the results shown in Fig. 1 through 6 were quite time consuming, with computing times on the order of several minutes on the IBM 7094 for the cases with large numbers of interference points. For this reason no attempt was made to compute the complete detection index, since this would have required summation of F(n) over a large number of frequencies. The computer results therefore still don't conclusively answer the question of how serious the effect of large distributed interferences is. The indications are, however, that the results of Sections VII and VIII are valid under considerably wider conditions than those assumed there to produce analytical approximations. In fact, it appears that loss of detectability for rather widely distributed interference is equivalent to an appropriately increased
X. Conclusions

The major difficulty in obtaining general estimates of the effect of directional noise on the detectability in an array processor is that the mathematical manipulations required to derive these are quite complex. Results have therefore been obtained only in a restricted number of simple cases.

The general tenor of these results is that if the anisotropic-to-isotropic-noise ratio is small the effect of a number of local noise sources is additive; that is the loss of detectability resulting from two noise sources of equal strength is twice that resulting from a single source. For large anisotropic-to-ambient-noise ratio the effect depends on whether the directional noise sources are close together or not. For a single point source it has been shown previously and corroborated here, that the loss in detectability is approximately equivalent to the loss of one hydrophone from the array. If there are R noise sources, widely separated from each other and from the target direction the loss is approximately equivalent to the loss of R hydrophones, provided that $R \ll M$, where M is the number of hydrophones.

Point noise sources that are close together affect the system like a single distributed noise source, and the indications are that if such an anisotropy is spread over a relatively small angle its effect is essentially that of a single point noise. Unfortunately this has not been conclusively demonstrated, even by use of a digital computer, and only a rather conservative estimate of azimuth angle that can be considered to be "small" has been obtained.
Appendix A Derivation of the Detection Index

The detection index \( d \) is given by

\[
d = \frac{\langle u \rangle_{SN} - \langle u \rangle_N}{\sigma_N(u)} \tag{A-1}
\]

where \( u = \sum_{n=1}^{WT} \left| H_T(n) X(n) \right|^2 \) \tag{A-2}

and where

\[
H(n) = \sqrt{K(n)} \, S^{-1}(n) \, \psi_n(n) \cdots K(n) \, \psi_T(n) \, S^{T^{-1}}(n) \tag{A-3}
\]

with

\[
K(n) = \frac{S(n)/N^2(n)}{1 + S(n)G_0(n)/N(n)} \tag{A-4}
\]

Then

\[
\langle u \rangle_N = \sum_{n=1}^{WT} H_T(n) \langle X(n) \, \psi_T(n) \rangle_N \, H^*(n) \tag{A-5}
\]

But from Eq. (A-3)

\[
\langle X(n) \, \psi_T(n) \rangle_N = H(n) \, \psi_k(n) \sim N(n) \, \psi_T(n) \tag{A-6}
\]

Therefore

\[
\langle u \rangle_N = \sum_{n=1}^{WT} H_T(n) \, H^*(n) \, H^T(n) = \sum_{n=1}^{WT} N(n) \, K(n) \, \psi_T(n) \, \psi_T^{-1}(n) \, \psi(n) \tag{A-7}
\]

Similarly, and using the fact that

\[
\langle X(n) \, \psi_T(n) \rangle_{SN} = N(n) \, \psi_T(n) + S(n) \, \psi_T(n) \tag{A-8}
\]

\[
\langle u \rangle_{SN} = \sum_{n=1}^{WT} N(n) \left[ S(n) \, G^2_T(n) + F(n) \, G^2_0(n) \right] \tag{A-9}
\]

Hence

\[
\langle u \rangle_{SN} - \langle u \rangle_N = \sum_{n=1}^{WT} K(n) \, S(n) \, G^2_0(n) \tag{A-10}
\]

To find \( \sigma_N(u) \) it is necessary to find \( \langle u^2 \rangle_N \), given by

\[
\langle u^2 \rangle_N = \sum_{n=1}^{WT} \left| H_T(n) \, X(n) \right|^2 \left| H_T(n) \, X(n) \right|^2 \tag{A-11}
\]
Let \( w(n) = H(n) x(n) \), then \( w(n) \) is a Gaussian random variable since \( x(n) \) is. In terms of \( w(n) \)

\[
\langle u^2 \rangle_N = \sum_{n=1}^{WT} m=1 \langle w(n) w^*(n) w(m) w^*(m) \rangle_N
\]

\[
= \sum_{n=1}^{WT} m=1 \left[ \langle w(n) w^*(n) \rangle_N \times \langle w(m) w^*(m) \rangle_N \times \langle w(n) w(m) \rangle_N \langle w^*(n) w^*(m) \rangle_N \right]
\]

\[
+ \langle w(n) w^*(m) \rangle_N \langle w(n) w(m) \rangle_N
\]

\[(A-12)\]

This expansion is permissible because \( w(n) \) is Gaussian \( \| h \| \). The first term in the expansion is simply the square of the mean \( \langle u^2 \rangle_N \), the second term vanishes \( \langle u^2 \rangle_N \), and in the last term all terms for which \( n \neq m \) vanish because components at different frequencies are assumed to be independent. Hence

\[
\sigma^2_N(u) = \langle u^2 \rangle_N - \langle u^2 \rangle_N = \sum_{n=1}^{WT} \left\{ \langle w(n) x^*(n) x(n) \rangle_N \right\}^2
\]

\[
= \sum_{n=1}^{WT} K^2(n) S(n) \sigma^2_N(u)
\]

\[(A-13)\]

as in Eq. (A-7).

Thus, finally

\[
d = \sqrt{\sum_{n=1}^{WT} K(n) S(n) \sigma^2_N(u)}
\]

\[(A-14)\]

For small signal-to-noise ratio, such that \( \langle u(n) x(n) S(n) \rangle N \approx 1 \)

\[
K(n) \approx S(n) / K^2(n)
\]

Then

\[
\sigma^2_N(u) \approx \sqrt{\sum_{n=1}^{WT} \left( \frac{S(n)}{K^2(n)} \right) \sigma^2_N(u)}
\]

\[(A-15)\]
Appendix B

Approximate Inversion of a Matrix whose Diagonal Terms are Large Relative to the Off-Diagonal Terms.

Let the \( n \times n \) nonsingular matrix \( A \) be given by

\[
A = D : B
\]

where \( D \) is diagonal and \( B \) is a matrix with zero diagonal elements. It is assumed that all the non-diagonal elements of \( B \) are of about the same order of magnitude, and that the elements of \( D \) are of about \( K \) times that magnitude, with \( K \gg 1 \).

The inverse of \( A \) is given by

\[
A^{-1} = (D + B)^{-1} = D^{-1}(I + BD^{-1})^{-1} = D^{-1} \left[ I - BD^{-1} + (BD^{-1})^2 \ldots \right]
\]

Since the elements of \( D \) are of order \( K \) relative to \( B \) the elements of \( BD^{-1} \) are of order \( 1/K \) relative to unity.

It can be shown [8] that a sufficient condition for convergence of Eq.(B-2)

\[
\sum_{j=1}^{n} |b_{ij}| < 1 \quad i = 1, \ldots, n
\]

where \( b_{ij} \) are the elements of \( BD^{-1} \). Assuming all of these elements to be of about the same order of magnitude, condition (B-3) can be expressed in the approximate form

\[
nb_o < 1
\]

where \( b_o \) is a representative element of \( BD^{-1} \). This element is of order \( 1/K \); therefore convergence requires

\[
n/K < 1
\]

The convergence will clearly be more rapid if this inequality is sharper; hence one can approximately neglect the matrix \( B \) in the inversion of \( A \) if \( n/K \ll 1 \).
L. t of references:


4. Edelblute, et al., op cit., Eq. (141)


METHODS OF STOCHASTIC APPROXIMATION APPLIED TO THE ANALYSIS OF ADAPTIVE TAPPED DELAY LINE FILTERS

by

John H. Chang
Franz B. Tutundjian

Progress Report No. 34

General Dynamic/Electric Boat Research

November 1967

DEPARTMENT OF ENGINEERING AND APPLIED SCIENCE

YALE UNIVERSITY
ABSTRACT

In this report some studies have been carried out for a class of adaptive filters consisting of tapped delay lines and adjustable gains. The method of stochastic approximation and mean square error criterion are employed to adjust the gains automatically. It is shown that it is not necessary that the desired signal generally used to obtain the error function is available. Either signal or noise correlation functions will suffice to generate the error gradient. Problems basic to all adaptive processes such as the conditions for convergence, rate of convergence, effect of misadjustment, effect of time-varying parameters, and the relationship between mean square error and the number of delay elements are answered with explicit expressions.
Contents

I. Introduction
   1. Optimum linear filters
   2. Adaptive systems and techniques
   3. Adaptive filters and state of the art
   4. Summary of the report
   5. Research work in progress

II. Method of Stochastic Approximation
   1. Historical development
   2. Basic definitions
   3. The system
   4. Convergence properties
   5. Geometrical significance of the conditions for convergence

III. Adaptive Delay Line Filters
   1. Convergent tapped delay line filters
   2. Minimum mean square error and the effect of non-optimum settings
   3. Adaptive tapped delay line filters

IV. Adaptive Filters and Rate of Convergence
   1. Adaptive schemes
      a. Statistically variable
      b. Stationary
      c. I-sided
   2. Convergence
      a. Converges
      b. Converges in probability
      c. Convergence is constant

V. Adaptive Tapped Delay Line Filters with Time-Varying Parameters
   1. Statistical properties of delay line filters
      a. Stationary case
      b. Non-stationary case
      c. Time-varying case
   2. Adaptive schemes for delay line filters with slowly time-varying parameters
      a. For the variable method
      b. Applications on adaptive filters
      c. Effects of time variation on minimum mean square error

References

D-8
I. INTRODUCTION

1.1 Optimum Linear Filters

The problem of designing a device to eliminate noise or to predict the future behavior of an incoming signal has been considered by Norbert Wiener more than twenty years ago. This kind of device has been termed as "filter" in general. Consider a linear filter shown in Figure 1, where the input $x(t)$ is a combination of the useful signal $s(t)$ and noise $n(t)$. Assuming that $n(t)$ is additive to and statistically independent of $s(t)$, we have

$$x(t) = s(t) + n(t) \quad (1.1)$$

and

$$s(t) n(t) = 0 \text{ if } s(t) = 0, n(t) = 0 \quad (1.2)$$

Figure 1. A linear filter

The output $y(t)$ of the filter is to approximate a desired function $d(t)$ which is related to the signal $s(t)$. The performance criterion to be minimized is the mean square error

$$e^2(t) = (d(t) - y(t))^2 \quad (1.3)$$

This is the classical problem of Wiener, the analytic solution (for the impulse response of a realizable filter) is known to be the solution of the Wiener-Hopf integral equation.2
\[ R_{xy}(T) = \int_{0}^{\infty} h(t) R_{xx}(T-t) \, dt \]  
(1.4)

with the solution

\[ H_0(\omega) = \frac{\Phi_{xd}(\omega)}{\Phi_{xx}(\omega)} \]  
(1.5)

and the minimum mean square error

\[ e_{\text{min}}^2 = \int_{-\infty}^{\infty} \left( \Phi_{yd}(\omega) - H_0(\omega)^2 \Phi_x(\omega) \right) \, d\omega \]

\[ = d^2(t) - \gamma^2_0(t) \]  
(1.6)

In the above equations \( R_{xy}(T) \) and \( \Phi_{xy}(\omega) \) are the cross correlation function and spectral density functions between \( x(t) \) and \( y(t) \). 

\( y_0(t) \) is the output of the optimum filter. The results are valid for stationary signals.

Kalman and Bucy \(^3\) have presented a new method to design optimal linear filters for nonstationary signals. They considered the model for the signal process as

\[ \frac{dx}{dt} = Fx + Gu \]  
(1.7)

where \( u(t) \) is white noise, i.e.,

\[ \overline{u(t)u(t)} = \delta(t-\tau) \]  
(1.8)

The observed signal is assumed as

\[ z = Hx + v \]  
(1.9)

where \( H = (1, 0) \) and \( v \) is white noise of spectral density \( \phi \).

*This type of filters may not be physically realizable. Further discussions can be found in Wiener.\(^1\)  

D-2
The optimum filtering problem consists in determining
\[ \hat{x}(t) = E\left[ x(t) \mid z(T) \right] \]  
(1.10)
where \( \hat{x}(t) \), the conditional expectation of \( x(t) \) given the observation \( z(T) \) in the interval \( (0,t) \), is the minimum variance unbiased estimator of \( x(t) \).

The optimum system is described succinctly by the following four equations:
\[ \frac{d\hat{x}}{dt} = F\hat{x} + PH\Phi^{-1} (z - H\hat{x}) \]
with
\[ \hat{x}(0) = 0 \]  
(1.11)
and
\[ \frac{dT}{dt} = FP + F^T \Phi^{-1} - PH\Phi^{-1} HP + \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix} \]
\[ \Gamma'(0) = E\left[ x(0) x^T(0) \right] \]

For time varying systems, Kalman-Bucy filters can provide much better performance than Wiener filters. From (1.5) and (1.11) it is seen that the statistical properties of both signal and noise should be known in order to design either type of the two filters.

Thus, if the a priori information is not known completely, optimal performance of the filters cannot be expected. In an attempt to recover some of the missed a priori information by evaluation of the actual performance of the operating system, the concept of adaptation has been developed and accepted as one of the possible solutions to many such problems.

1.2 Adaptive systems and techniques

A system is adaptive if there is a learning mechanism box in the
feedback loop where the observable output is compared with the desired output so as to adjust the system for better performance. In other words, an adaptive system is one which is provided with a means of continuously monitoring its own performance according to a given performance index, and also a means of adjusting its own parameters by closed loop action so as to optimize its own operation. Relaxation method and method of steepest descent (or ascent) are two of the most commonly used adaptive techniques.

The relaxation method involves making a change in the value of only one of the controller parameters and then re-evaluating the performance measure. If the performance has been improved, a second change in the same direction is made; otherwise, the first change is retracted and a change in the opposite direction is made. This process is continued until no further improvement in the performance measure can be accomplished by adjusting that particular parameter; whereupon the same process is repeated for each of the remaining controller parameters. After several iterations through the entire procedure, the controller parameters tend toward that set of values which yields the optimum performance measure.

The methods of steepest descent (or ascent), referred as gradient techniques are operated in a manner similar to the relaxation method, with the notable exception that all parameters are adjusted simultaneously rather than sequentially. This is done by measuring the partial derivative of the performance measure with respect to each of the controller parameters and then adjusting all the parameters in such a way that the net effect is the largest possible improvement in the performance measure.
A number of techniques have been developed for determining the partial derivatives.

The most straightforward method is to perturb each of the parameters sequentially and measure the derivatives directly. This procedure, however, offers little advantage over the relaxation method. A second technique is to perturb the parameters simultaneously in such a manner that the effect of the perturbation of each parameter on the performance measure will be distinguishable from the effects of the perturbations of all the other parameters. Ways in which this may be done include perturbation by independent random noise, distinguishing the individual effects by correlation detection; or perturbation by frequency-separated sinusoids, distinguishing the effects by narrow-band detection. Gradient techniques can be considered as the special case of the more general method of stochastic approximation, by which either deterministic or random problems can be solved with ease.

1.3 Adaptive Filters and State of the Art

Adaptive filters have been investigated by a number of researchers. Their methods differ chiefly in the ways of implementation, but all are designed with the same purpose in mind — to extremize the performance index by gradually adjusting the system parameters. One of the simplest implementations in this area is the use of tapped delay lines which can be constructed easily with shift registers in digital computers. Weaver considered tapped delay line filters, but his adaptive scheme was rather ineffective and required the solution of many simultaneous equations. Narendra and the author used delay lines to identify the characteristics of some unknown nonlinear linear systems. Although the
feasibility of the above-mentioned methods was indicated by computer simulations, some figures of merit to judge these schemes, such as the rate of convergence of the parameters to the optimum, remain untackled. Widrow\textsuperscript{12, 13} has attempted to solve these problems by defining some adaptive constants and misadjustment formulas. His system consists of delay lines and adjustable gains, and has been acclaimed to perform nearly as the Kalman-Bucy filter when perfect a priori information is available. Under circumstances in which the a priori information is not perfectly known, it is quite possible that the performance of such an adaptive filter could exceed that of either a Wiener or a Kalman-Bucy filter. However, Widrow's system requires the availability of a desired signal to generate the real time error function. Convergence proof of his LMS adaptation algorithm (least-mean-square-error algorithm) was not given, and the development of the rate of adaptation was not mathematically rigorous. Moreover, how to make a time-varying system adaptive has not been considered.

1.4 Outlines of the Report

A. Consider a random function $Q(x|c)$ where $x = \{x_1, x_2, \ldots, x_n\}$ is a vector of stationary random process with distribution $P(x)$. In an attempt to minimize the criterion

$$I(c) = E_x \left[ Q(x|c) \right]$$  \hfill (1.12)

it is natural to set the gradient of $I(c)$ to zero,

$$\nabla I(c) = E_x \left[ \nabla_c Q(x|c) \right] = 0$$  \hfill (1.13)

Since $P(x)$ is generally unknown, an algorithm derived from the method of stochastic approximation to obtain $c^*$, the optimum value of $c$, is
Algorithm (1.14) converges with probability one

\[ P \left\{ \lim_{j \to \infty} (e_j - e^*) = 0 \right\} = 1 \]  

as well as in mean square

\[ \lim_{j \to \infty} E \left\{ \|e_j - e^*\|^2 \right\} = 0 \]  

under the following conditions

A. \( \sum_{j=1}^{\infty} \gamma_j = \infty, \sum_{j=1}^{\infty} \gamma_j^2 < \infty, \gamma_j > 0 \)  

B. \( \inf E \left\{ (e - e^*)^T \nabla_c Q(x|c) \right\} > 0 \)  

\[ \epsilon < \|e - e^*\| < \frac{1}{\epsilon} \quad \epsilon > 0 \]  

C. \( E \left\{ \nabla_c^T 0(x|c) \nabla_c Q(x|c) \right\} < d(e^T c^* + e^T c) \)  

for all \( c \) and \( d > 0 \)

For the tapped delay line filter under study, the transfer function of the filter is represented by

\[ H(\omega) = \sum_{k=0}^{N} c_k e^{-j\omega k} \]  

and its impulse response is

\[ h(t) = \sum_{k=0}^{N} c_k (1 - \gamma_k t) \]  

If we attempt to minimize the mean square error
\[
E^2 = E \left \{ \left \{ d(t) - z(t) \right \}^2 \right \} = E \left \{ (d(t) - \sum_{k=0}^{N} c_k x(t-T_k))^2 \right \} \tag{1.22}
\]

the optimum values of \( c \) is obtained as

\[
\hat{c} = P_{\eta}^{-1} R_{d\eta} \tag{1.23}
\]

where

\[
R_{\eta} = E \begin{pmatrix}
\eta_{\omega} \eta_{\omega} & \cdots & \eta_{\omega} \eta_{\omega} \\
\eta_{\omega} \eta_{\omega} & \eta_{\omega} \eta_{\omega} & \cdots & \eta_{\omega} \eta_{\omega}
\end{pmatrix}, \quad R_{d\eta} = E \begin{pmatrix}
d\eta_{\omega} \\
d\eta_{\omega}
\end{pmatrix}
\]

with

\[
\hat{\eta}_k(t) = x(t-T_k)
\]

it is seen that in \( (1.23) \) the values of \( \hat{c} \) cannot be determined unless we have full knowledge about both signal and noise correlations.

Substituting \( (1.23) \) into \( (1.22) \) we can obtain the expression for the minimum mean square error as

\[
e_{\text{min}}^2 = d^2(t) - R_{d\eta} R_{\eta}^{-1} R_{d\eta} \tag{1.24}
\]

\[
e_{\text{min}}^2 = d^2(t) - (t-T_k) \eta \eta_{\omega} \eta_{\omega} \eta_{\omega}
\]

where \( z_{\omega}(t) \) is the output of the optimum filter.

The mean square error at any time for arbitrary values of \( c \) is

\[
c^2(t) = e_{\text{min}}^2 + (\hat{c} - c) R_{\eta}^{-1} (\hat{c} - c) \tag{1.25}
\]

and the effect of non-optimum setting is bounded by

\[
e^2(t) - e_{\text{min}}^2 \leq (N+1) \max_{i,\eta,\omega} |c_i - \hat{c}_i| \max_{i,j} |\hat{\eta}_i\hat{\eta}_j| \tag{1.26}
\]
A relationship between the minimum mean square error and the number of delay elements is
\[ e_{\text{min}}^2 = r_n(\alpha) - B_s^T R_q^{-1} B_s \] (1.27)

The last term in the right-hand side of (1.27) is a functional of \( N \), the number of delay elements and the correlation functions. For any known forms of \( R_n(T) \) and \( R_s(T) \), a plot of \( e_{\text{min}}^2 \) versus \( N \) can be constructed. It is anticipated that the larger \( N \) is, the smaller \( e_{\text{min}}^2 \) will be.

C. Under various practical situations we may not have perfect information about both \( R_n(T) \) and \( R_s(T) \). Techniques of adaptation can be employed to estimate the incomplete a priori information and to make the filter adaptive to changing operating conditions.

From (1.22) the error gradient without averaging operation is obtained as
\[ \nabla_c Q(x | c) = -2 ( d(t) - z(t) ) \eta(t) \]
so that the desired adaptive algorithm is
\[ c_{j+1} = c_j + 2 \gamma_j e_j \eta_j \] (1.28)

The above algorithm converges if 1. \( Q(c) \) is strictly convex, 2. \( \frac{\partial^2 Q}{\partial c^2} \) exists and is uniformly bounded, 3. \( \eta(t) \) and \( n(t) \) are uniformly bounded.

In (1.29) the desired signal \( d(t) \) is used to generate the error function. This assumption is not so practical when dealing with
detection problems. If only the noise correlation function is known, we can change Eq. (1.28) to

\[ x_{j+1} = x_j + 2v_j y_j \eta_j (x_j - z_j) - 2v_j R_n \tag{1.29} \]

where

\[ R_n^T = \left( R_n(o), \ldots, R_n(T_n) \right) \]

On the other hand, if only the signal correlation function is known, we have

\[ c_{j+1} = c_j + 2v_j R_s - 2v_j z_j \eta_j \tag{1.30} \]

where

\[ R_s^T = \left( h_s(o), \ldots, R_s(T_N) \right) \]

D. The other problem that we have to consider is the rate of adaptation, or the rate of convergence. We want to estimate how fast the gains approach their optimum values. Defining

\[ E\{\eta \} = \frac{1}{\sum H} \eta \tag{1.31} \]

\[ \omega^p = \frac{1}{2(j+1)\lambda_p} \text{ for the } p\text{th component of } \omega, \tag{1.32} \]

we can express the component of \( \omega \) at any time during the adaptation period as a function of the initial choice and the optimum values of \( \omega \), i.e.,

\[ \omega_j = \frac{1}{j+1} (\omega_j - \omega^*) + \omega^* \tag{1.33} \]

Using \( (1.24) \) we obtain the difference between \( e^2(t) \) at any time and \( e^2_{\text{min}} \) as

\[ e_j^2 - e^2_{\text{min}} = \sum_{p=0}^{N} \lambda_p \left( \frac{\omega_j^p}{j} - \frac{\omega_j^p}{j} \right)^2 \tag{1.34} \]

\[ \frac{1}{j+1} e^2_{\text{min}} \sum j^2 \left\{ (\epsilon_j - c^*)^T R_n (\epsilon_j - c^*) \right\} \]

D-10
\[
\hat{t} = \beta \tilde{t}
\]

then the method of two time scales can be used to modify the algorithms obtained previously. The mean square error is changed to

\[
e^2(t) = E \left\{ \left[ d(t) - \sum_{i=0}^{N} c_i \hat{a}(t) \right]^2 \right\}
\]

\[
+ 2\beta T_{av} \tilde{\sum} \left\{ \left( \hat{a}(t) - \sum_{i=0}^{N} c_i \hat{a}(t) \right) \sum_{i=0}^{N} \eta_i(t) \frac{\partial \hat{a}(\tilde{t})}{\partial \tilde{t}} \right\}.
\]

(1.39)

where \( T_{av} \) is the average delay time. Algorithm (1.28) is modified to the form

\[
\hat{c}_{j+1} = \hat{c}_j + 2\gamma_j \eta_j \hat{c}_j + 2\gamma_j \beta T_a \eta_j \hat{c}_j \eta_j^T \delta
\]

(1.40)

where

\[
\delta_j = \frac{\partial \hat{c}_j(\tilde{t})}{\partial t}
\]

The minimum mean square error for this case is

\[
e^2_{\min} \leq e^2_{\min} + \beta T_{av} \tilde{\sum} \left\{ R_{\delta q} \right\} + 2\beta^2 T_{av} \tilde{\sum} \left\{ \delta \right\} (1.41)
\]

where \( e^2_{\min} \) is the minimum mean square error of the time-invariant filters.

1.5 Research Work in Progress

So far preliminary results have been obtained for a single input single output filter under very general situations. Practical examples for different types of signal, noise and ways of parameter variation are being worked out. In order to verify the results, digital computer simulation will be conducted.
where \( \lambda \) is the \( p \)th component of \( A \) and \( w_1^p \) is the \( p \)th component of \( w_1 \).

E. So far we have assumed that the whole system is time-invariant and the output signals are stationary at least in the wide sense. However, sometimes for one reason or another, system parameters or signal properties may change slowly. One instance of this situation is the fluctuation of power levels. Some schemes to adjust the gains under this case would be highly desirable. First consider the quasi-stationary case where only slow time variable is involved. If a time function \( x(t) \) is delayed by an amount of \( T \) and multiplied by \( c \) such that

\[
y(t) = cx(t-T)
\]

or

\[
y(t+T) = cx(t)
\]

which can be written as

\[
Ly(t) = cx(t) \quad (1.35)
\]

\( L \) is a linear differentiation operator

\[
L = e^{tpT} = \sum_{i=0}^{\infty} \frac{T^i p^i}{i!} \quad \text{and} \quad p = \frac{d}{dt} \quad (1.36)
\]

Let \( x(t) = e^{\lambda t} \) and \( y(t) = z(t,\lambda)e^{\lambda t} \), then \( z(t,\lambda) \) can be obtained as \( z_0(t,\lambda) = ce^{-\lambda t} \)

if \( c \) is a constant, and

\[
z(t,\lambda) = \frac{1}{e^{\lambda t}} \left( c - \sum_{i=1}^{n} \frac{1}{i!} \frac{d^i z(t)}{dt^i} \right) \quad (1.37)
\]

if \( c \) is time-varying. The effect of time-varying parameter is observed.

During the training period the system is operated in real (fast) time \( \hat{t} \). If at the same time some parameters are changing slowly in slow time \( \hat{t} \) such that

D-11
A very interesting and practical application of adaptive filters is the sonar detecting system consisting of many hydrophones steered or not steered on target. The output of each hydrophone passes through an adaptive filter, and the sum of the filter outputs is squared and averaged to indicate the presence or absence of a target. When the input signal to noise ratio is small and the levels of signal and noise are the same at all hydrophones, the whole system can be adjusted to form a likelihood ratio detector. Some features of this detector will be studied with emphasis on adaptive schemes, rate of convergence and detectability for stationary and time-varying processes. Performance analysis will be carried out with analytic and numerical examples.
II. METHODS OF STOCHASTIC APPROXIMATION

1. Historical Developments

The methods of stochastic approximation were originally developed by Robbins and Monro in 1951. Their purpose was to find the root of a noisy function. The term "stochastic" refers to the random character of the experimental errors, while the term "approximation" refers to the continued use of past measurements to estimate the approximate position of the goal. Kiefer and Wolfowitz adapted the idea of stochastic approximation to the problem of finding the maximum of a unimodal function obscured by noise. Blum used the gradient method to extend the above techniques to multi-dimensional case. Later on Dvoretzky greatly generalized and unified the whole theory and Kesten derived some formulas to speed up the rate of convergence in terms of the number of changes in sign before a certain step.

2. Basic Considerations

Stochastic approximation, much like ordinary successive approximation in the absence of experimental error, involves two basic considerations - first choosing a promising direction in which to search and selecting the distance to travel in that direction. Picking a search direction is no more difficult for stochastic than for deterministic approximations for one simply behaves as if he believed the experimental results, ignoring entirely the possibility of error. This means of course that the experimenter will move away from his goal whenever he is misled by the vagaries of chance error. It will be seen that such temporary setbacks do not prevent
ultimate convergence if the step sizes are chosen properly.

In both stochastic and deterministic schemes, the corrections are made progressively small as the search proceeds so that the process will eventually converge. To make this convergence rapid, one would like to shrink the step size as speedily as possible. The main difference between stochastic and deterministic procedures is in fact the speed with which the steps can be shortened. When noise is totally absent one can reduce the steps very rapidly, but when there is danger of an occasional jump in the wrong direction, shortening the steps too rapidly could make it impossible to erase the long-run effects of a mistake. In the latter case the process would still converge, but to the wrong value.

3. The Methods

Many problems in modern cybernetical systems design can be reduced to that of finding the extrema of functions of several variables

\[ I = Q(c_1, c_2, \ldots, c_n) = Q(c) \]

where \( c = \{c_1, c_2, \ldots, c_n\} \) (2.1)

Denoting the optimal values of \( c \) by \( c^* \) and assuming that the extremum of interest to us is a minimum, we can obtain the solution of \( c = c^* \) by setting the gradient of \( Q(c) \) equal to zero; i.e.,

\[ \nabla Q(c^*) = 0 \] (2.2)

where \( \nabla Q(c) = \left\{ \frac{\partial Q(c)}{\partial c_1}, \ldots, \frac{\partial Q(c)}{\partial c_n} \right\} \)
Generally a closed-form solution cannot be obtained for \((2.2)\), so iteration methods are required, especially the gradient method.

The gradient method relates the coordinates of a given point with the coordinates of the preceding point and the gradient \(\nabla Q(c)\). The algorithm for determining \(c^H\) can be written in the form

\[
\begin{align*}
C_{j+1} &= C_0 - \gamma \nabla Q(C_j) \\
&= C_j - \gamma \nabla Q(C_j)
\end{align*}
\tag{2.3}
\]

Here \(\gamma_j\) determines the pitch of the algorithm and generally depends on the index of the step and the function itself.

When \(Q(c)\) is not given analytically or is not differentiable, the gradient \(\nabla Q(c)\) can be approximately determined with the formula

\[
Q(c, a) = \frac{Q(c, a) - Q(c, a)}{2a}
\]

where

\[
Q(c, a) = \{Q(c + ae_1), \ldots, Q(c + ae_n)\}
\tag{2.4}
\]

and \(e_i\) denotes the base vectors

\[
e_1 = \{1, 0, \ldots, 0\}, \quad e_n = \{0, 0, \ldots, 1\}
\tag{2.5}
\]

The corresponding algorithm is then

\[
C_{j+1} = C_j - \gamma \frac{Q_+(C_j, a) - Q_-(C_j, a)}{2a_j}
\tag{2.6}
\]

In the above we assumed that \(Q(c)\) is a deterministic function. If we consider a random function \(Q(x|c)\), where \(x = \{x_1, x_2, \ldots, x_n\}\) is a vector of stationary random processes with distribution \(p(x)\), it is natural to attempt to find the extrema of the mathematical expectation:
\[
I(c) = \int_{\mathcal{X}} Q(x|c) p(x) \, dx = E_{x} \left\{ Q(x|c) \right\}.
\]  

(2.7)

The condition for determining the optimal value \( c = c^* \) is of the form

\[
J I(c) = E \left\{ \nabla_{c} Q(x|c) \right\} = 0
\]

(2.8)

We can apply the algorithms (2.3) and (2.6) to (2.8) and functional (2.7) only when the a priori distribution \( p(x) \) is known and, consequently, the mathematical expectation (2.7) can be determined beforehand. Frequently, however, the probability density function \( p(x) \) is unknown. Nonetheless, the optimal vector \( c = c^* \) can still be determined by applying the gradient method using \( \nabla_{c} Q(x|c) \) instead of \( E \left\{ \nabla_{c} Q(x|c) \right\} \).

This is the advantage of using the method of stochastic approximation. With this method the algorithms for determining \( c = c^* \) can be written in the form

\[
c_{j+1} = c_{j} - \gamma_{j} \nabla_{c} Q(x|c)_{j}
\]

(2.9)

if \( Q(x|c) \) is analytic and differentiable, and

\[
c_{j+1} = c_{j} - \frac{\gamma_{j}}{2a_{j}} \left\{ Q_{+}(x|c_{j}, a_{j}) - Q_{-}(x|c_{j}, a_{j}) \right\}
\]

(2.10)

if \( \nabla_{c} Q(x|c) \) does not exist.

Algorithm (2.9) is a multivariate form of the Robbins-Monro procedure, while algorithm (2.10) is a multivariate form of the Kiefer-Wolfowitz scheme. The analogy between deterministic and stochastic algorithms is apparent. It should be emphasized however, that stochastic algorithms deal with stationary random variables which may contain random noise in addition to the useful signal. The convergence properties of the above algorithms will be considered in the next section.
4. Convergence Properties

In this section the conditions under which the above-mentioned algorithms converge will be described. Since mean square error is used in this report as the only performance criterion, $Q(x|c)$ is analytic and differentiable, and we therefore need to consider only algorithm (7.9).

Let $c^*$ satisfy the equation

$$E\left\{ \nabla_c Q(x|c) \right\} = 0 \quad (2.11)$$

$E\left\{ \nabla_c Q(x|c) \right\}$ is a set of real measurable functions of real variables $c$ such that $\nabla_c Q(x|c) \quad (2.12)$

where $c \geq c^*$ means $c_i \geq c_i^*$ for all $i$.

**Theorem:** Let $\gamma_1, \gamma_2, \ldots$ be a sequence of positive numbers such that

(A1) $\lim_{j \to \infty} \gamma_j = 0 \quad (2.13a)$

(A2) $\sum_{j=1}^{\infty} \gamma_j = \infty \quad (2.13b)$

(A3) $\sum_{j=1}^{\infty} \gamma_j^2 < \infty \quad (2.13c)$

(A4) $\gamma_j > 0 \quad (2.13d)$

Let the following conditions be satisfied

(B) $\inf_{\epsilon} E\left\{ (c - c^*)^T \nabla_c Q(x|c) \right\} > 0 \quad (2.14)$

$$\epsilon < \|c - c^*\| < \frac{1}{\sqrt{2}} \quad \epsilon > 0$$
Then the sequence $c_j$ defined by (2.9) converges with probability one to $c^*$. 

Proof: Subtracting both sides of Eq. (2.9) by $c^*$ we have

$$c_{j+1} - c^* = c_j - c^* - \gamma_j \nabla Q$$

(2.16)

where, for simplicity, $0 = Q(x|c)$

Squaring Eq. (2.16)

$$(c_{j+1} - c^*)^T (c_{j+1} - c^*) = (c_j - c^*)^T (c_j - c^*)$$

$$- 2 \gamma_j (c_j - c^*)^T \nabla Q$$

$$+ \gamma_j^2 \nabla^T Q \nabla Q$$

and taking the conditional mathematical expectation for given $c_1, c_2, \ldots,$ $c_j$, we obtain

$$E \left\{ \|c_{j+1} - c^*\|^2 \left| c_1, c_2, \ldots, c_j \right. \right\}$$

$$= \|c_j - c^*\|^2 - 2 \gamma_j E \left\{ (c_j - c^*)^T (\nabla Q) \right\}$$

$$+ \gamma_j^2 E \left\{ \nabla^T Q \nabla Q \right\}$$

(2.17)

From condition (C), (2.17) becomes

$$E \left\{ \|c_{j+1} - c^*\|^2 \left| c_1, c_j \right. \right\} \leq \|c_j - c^*\|^2 - 2 \gamma_j E \left\{ (c_j - c^*)^T \nabla Q \right\}$$

$$+ \gamma_j^2 d (c^{*T} c^* + c_j^T c_j)$$

(2.18)
Using condition (B), (2.18) is reduced to

\[
E \left\{ \left\| c_{j+1} - c^* \right\|^2 \left| c_1, \ldots, c_j \right\} \right.
\leq \left\| c_j - c^* \right\|^2 \left( 1 + \gamma_j^2 d \right) + 2 \gamma_j^2 d c^T c^*
\]

(2.18a)

Let \( z_j = \left\| z_j - c^* \right\|^2 \sum_{k=j}^{\infty} \gamma_k d (1 + \gamma_k^2 d) \)
\( + \sum_{k=j}^{\infty} 2d \gamma_k c^T c^* \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d) \)
(2.19)

Then \( z_{j+1} = \left\| z_{j+1} - c^* \right\|^2 \sum_{k=j+1}^{\infty} \gamma_k d (1 + \gamma_k^2 d) \)
\( + \sum_{k=j+1}^{\infty} 2d \gamma_k c^T c^* \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d) \)
(2.20)

Taking the conditional mathematical expectation for given \( c_1, c_2, \ldots, c_j \), we have

\[
E \left\{ z_{j+1} \mid c_1, \ldots, c_j \right\} = E \left\{ \left\| z_{j+1} - c^* \right\|^2 \left| c_1, \ldots, c_j \right\} \right.
\sum_{k=j+1}^{\infty} \gamma_k d (1 + \gamma_k^2 d) \)
\( + \sum_{k=j+1}^{\infty} 2d \gamma_k c^T c^* \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d) \)
\leq \left\{ \left\| z_j - c^* \right\|^2 \left( 1 + \gamma_j^2 d \right) + 2 \gamma_j^2 d c^T c^* \right. \sum_{k=j+1}^{\infty} \gamma_k d (1 + \gamma_k^2 d) \)
\( + \sum_{k=j+1}^{\infty} 2d \gamma_k c^T c^* \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d) \)
\( = z_j \)

or \( E \left\{ z_{j+1} \mid c_1, \ldots, c_j \right\} \leq z_j \)
(2.21)
Next taking the conditional mathematical expectation for given $Z_1$, ..., $Z_j$ on both sides of (2.21), we have

$$E\left\{Z_{j+1} \mid Z_1, \ldots, Z_j\right\} \leq Z_j$$

(2.22)

Since $Z_j = f(c_1, c_2, \ldots, c_j)$

Inequality (2.22) shows that $Z_j$ is a semimartingale, where

$$E\frac{Z_{j+1}}{Z_j} \leq E\frac{Z_j}{Z_j} \leq \cdots \leq E\frac{Z_1}{Z_1} < \infty$$

(2.23)

so that, according to the theory of semimartingales 2 the sequence $Z_1$ converges with probability one, and hence by virtue of Eq. (2.19) and (2.13c) the sequence $(c_j - c^k)$ also converges with probability one to some random number $\xi_j$. It remains to show that $P(\xi_j = 0) = 1$. It is seen that from (2.23), (2.19) and (2.13c) the sequence $E(c_j - c^k)$ is bounded. Now taking the mathematical expectation on both sides of the inequality (2.18),

$$E\left\{\|c_{j+1} - c^k\|^2\right\} \leq E\left\{\|c_j - c^k\|^2\right\} - 2\gamma_j E\left\{(c_j - c^k)^T \nabla q\right\}$$

and adding the first $j$ inequalities together, we have by deduction

$$E\left\{\|c_{j+1} - c^k\|^2\right\} \leq E\left\{\|c_1 - c^k\|^2\right\} - \sum_{k=1}^{j} \left(d c^T c_{-k} c - \delta \gamma_k^2 \sum_{k=1}^{j} (c_{-k} c_{-k})\right)$$

(2.24)

Since $E\left\{\|c_j - c^k\|^2\right\}$ is bounded and condition (2.13c) is fulfilled, from this (2.24) it follows that
\[
\sum_{k=1}^{\infty} \gamma_k E \left\{ (c_j - c^*)^T \nabla Q \right\} < \infty \tag{2.25}
\]

Using condition (2.13b), i.e., \( \sum_{j=1}^{\infty} \gamma_j = \infty \) and noting Eq. (2.14)

\[
\inf \epsilon < ||c - c^*|| < \frac{1}{\epsilon}
\]

We deduce from (2.25) that

\[
E \left\{ (c_j - c^*)^T \nabla Q \right\} \rightarrow 0 \text{ with probability one for some sequence } N.
\tag{2.26}
\]

Now taking \( E \left\{ ||c_j - c^*||^2 \right\} \rightarrow \frac{1}{2} \) with probability 1, and comparing (2.26) with (2.14) we obtain

\[
\zeta = 0 \text{ with probability } 1
\tag{2.27}
\]

Therefore, algorithm (2.9) converges with probability one

\[
P \left\{ \lim_{j \rightarrow \infty} (c_j - c^*) = 0 \right\} = 1
\tag{2.28}
\]

as well as in mean square sense, i.e.,

\[
\lim_{j \rightarrow \infty} E \left\{ ||c_j - c^*||^2 \right\} = 0
\tag{2.29}
\]

5. Geometrical Significances of the Conditions for Convergence

In the last section we mentioned several restrictions imposed on the properties of the sequence \( \{ \gamma_1, \ldots, \gamma_j \} \) as well as on the behavior of the function \( \nabla cQ(x|g) \). These conditions not only guarantee the convergence of the algorithms but also possess certain geometrical meanings. 23
A. \( \gamma_j > 0 \). This is to assure that the corrections, on the average, are to be made in the right directions.

B. \( \gamma_j \to 0 \) as \( j \to \infty \). This is to assure that \( e_j \) calculated from algorithm (2.9) will converge on some specific value. Suppose we let the measured error gradient be \( \nabla_c Q(x|c) \) and the real gradient be \( \mathbb{E}\{ \nabla_c Q(x|c) \} \).

Normally there is random noise in measurement

\[
\nabla_c Q(x|c) = \mathbb{E}\{ \nabla_c Q(x|c) \} + \xi_j
\]

\( j = 1, 2, \ldots \)

Thus \( \nabla_c Q(x|c) \to 0 \) even if \( e_j = 0 \). For \( e_j \) to converge on any value at all, the condition \( \gamma_j \to 0 \) as \( j \to \infty \) must be satisfied.

It is seen that the method of stochastic approximation is extremely noise resistant. Random independent additive noise \( \xi_j \) is eliminated and does not affect the final results.

C. \( \sum_{j=1}^{\infty} \gamma_j^2 < \infty \) or \( \sum_{j=1}^{\infty} \gamma_j^2 \to 0 \) as \( j \to \infty \). This condition is to account for the accumulative effect of \( \xi_j \). One application of this condition has been seen in the last section. When random noise \( \xi_j \) is added at each iteration step, algorithm (2.9) becomes

\[
e_{j+1} - e_j = \gamma_j \nabla_c Q(x|c) + \gamma_j \xi_j \quad (2.30)
\]

Summing the above equation from \( j = J \) upward gives

\[
e_n - e_J = -\sum_{j=J}^{\infty} \gamma_j \nabla_c Q(x|c) + \sum_{j=J}^{n} \gamma_j \xi_j \quad (2.31)
\]
(2.31) expresses the total variation in $c$ from the $j$th step onward.

Since

$$
\left( \sum_{j=1}^{\infty} \gamma_j \xi_j \right)^2 = \frac{2}{J} \sum_{j=1}^{\infty} \gamma_j^2
$$

and

$$\sum_{j=1}^{\infty} \gamma_j^2 \to 0$$

assures that the total random variation of $\left( \sum_{j=1}^{\infty} \gamma_j \xi_j \right)^2$ approaches zero as $J$ becomes very large.

D. $\sum_{j=1}^{\infty} \gamma_j \to \infty$. The above conditions assure that $c_j$ converges on some value $c_\infty$. $\sum_{j=1}^{\infty} \gamma_j \to \infty$ assures that $c_\infty = c^*$. Since this condition also implies $\sum_{j=1}^{\infty} \gamma_j \to \infty$, if $c_j$ approaches any value other than $c^*$, the total correction effect $\sum_{j=1}^{\infty} \gamma_j \nabla_c Q(x|c)$ is infinite. On the other hand, we have no fear of overshoot because each step is very small as $\gamma_j \to 0$ when $j \to \infty$. Conditions A-D state that the rate with which $\gamma_j$ decreases must be such that, on the one hand, the variance of performance index vanishes, and on the other hand, the variation in $\gamma_j$ over the variation period is large enough for the law of large numbers to hold.

E. \[ \inf_{c} \left\{ \frac{1}{\varepsilon} \left\| \left( c - c^* \right)^T \nabla_c Q(x|c) \right\| \right\} \geq 0 \text{ for } \varepsilon > 0. \]

This condition determines the behavior of the surface $E_x \left\{ \nabla_c Q(x|c) \right\} = 0$ close to the root and, consequently, the sign of the increments of $c_j$.

Actually, if the error criterion does have a unique minimum, the above condition is generally satisfied.

F. \[ P \left\{ \nabla_c^T Q(x|c) \nabla_c Q(x|c) \right\} \leq d(c^* T c^* + c^T c) \text{ for } d > 0. \]

This condition requires that the mathematical expectation of the quadratic forms

$$E_x \left\{ \nabla_c^T Q(x|c) \nabla_c Q(x|c) \right\}$$

increase, as $c$ increases, no faster than a quadratic paraboloid.
III. TAPPED DELAY LINE FILTERS

1. Optimum Tapped Delay Line Filters

Linear filters can be continuous or discrete. The optimum linear filter developed by Wiener has the form of Eq. (1.5)

\[ H_0(\omega) = \frac{\phi_{\text{xd}}(\omega)}{\phi_{\text{xx}}(\omega)} \]  \hspace{1cm} (3.1)

\( H_0(\omega) \) may be either physically realizable or not. If it is physically realizable, then standard techniques in network theories can be applied to obtain \( H_0(\omega) \) consisting of RLG elements with or without transformers.\(^2\)

Another method of synthesizing a continuous linear filter of arbitrary transfer function (and, hence, impulse response) is to represent it as an infinite linear combination of filters

\[ H(\omega) = \sum_{i=1}^{\infty} c_i F_i(\omega) \]  \hspace{1cm} (3.2)

where the functions \( F_i(\omega) \) are independent and together form a complete set. While an infinite sum is necessary to reproduce exactly the optimum filter response \( h(t) \), in practice it might be more useful to find the best filter which can be constructed from a finite number \( K \) of such independent components.

One particular type of (3.2) but discrete in nature is the tapped delay line filter. This filter consists of a tapped delay line, or equivalent, with adjustable weights at each tap. In this case

\[ H(z) = \sum_{k=0}^{N} c_k e^{-j\omega k} \]  \hspace{1cm} (3.3)

and the impulse response is

\[ h(t) = \sum_{k=0}^{N} c_k \delta(t-kT_k) \]  \hspace{1cm} (3.4)

where \( T_k = kT \), \( T \) is the delay increment between delay line taps, \( c_k \) is
the weight at the \( k \)th tap on the filter, and \( \delta \) is the Dirac delta function.

The configuration of such a filter is shown in Fig. 2. \( D_i \) denotes a delay of 
\( T_i \) in time.

![Fig. 2 Tapped Delay Line Filter](image)

The signal \( n_i \) obtained at a point after the delay element \( D_i \) is of course
\[
\eta_i(t) = \int_0^t x(\tau) \delta(t-\tau) d\tau = \int_0^t x(\tau) \delta(t-T_i) d\tau = x(t-T_i).
\] (3.5)

The weights \( C_i^*(i = 1, 2, \ldots, N) \) which optimize any performance criterion may
be found by using standard techniques such as calculus of variation or by setting
the partial derivatives of the performance criterion with respect to the adjustable gain to zero. Mean squared error criterion \( E\{[d(t) - z(t)]^2\} \) is used here
because it is simple to use as any and most configurations are not very sensitive
to error criterion.25

A. Frequency domain optimization using calculus of variations.

We are interested in determining \( H(z) \) minimizing
\[
F = E\{[d(t) - z(t)]^2\} = d^2(t) + z^2(t) - 2R_dz(c)
\] (3.6)

Each term of (3.6) can be related to \( H(z) \) by means of frequency integral.
Since the filter output is
\[ z(t) = \sum_{k=0}^{N} c_k \eta_k(t) \tag{3.7} \]
and its spectral density function is
\[ \phi_z(\omega) = \sum_{k=0}^{N} \sum_{m=0}^{N} \phi_z(\omega) = \sum_{k=0}^{N} \sum_{m=0}^{N} H_k(\omega) H_k^*(\omega) \phi_x(\omega) \]
\[ = \phi_x(\omega) \sum_{i=0}^{N} \sum_{k=0}^{N} C_k C_k e^{-j\omega(\ell-k)T} \tag{3.8} \]
the second term in (3.6) \( z^2(t) \), is given by
\[ \overline{z^2(t)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_z(\omega) d\omega = \frac{1}{2\pi} \sum_{i=0}^{N} \sum_{k=0}^{N} \frac{1}{2\pi} C_k C_k e^{-j\omega(\ell-k)T} \int_{-\infty}^{\infty} \phi_x(\omega) e^{-j\omega(\ell-k)T} d\omega \tag{3.9} \]

The third term is
\[ R_{d_z}(\tau) = E \left[ d(t) Z(t) \right] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_z(\omega) d\omega \]
\[ = \sum_{k=0}^{N} \frac{1}{2\pi} \int_{-\infty}^{\infty} d\varphi_z(\omega) d\omega \tag{3.10} \]
and finally
\[ \overline{d^2(z)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi^2(\omega) d\omega \tag{3.11} \]

Combining (3.9), (3.10) and (3.11), we obtain
\[ F = \overline{z^2(t)} = \sum_{k=0}^{N} \sum_{i=0}^{N} \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_z(\omega) e^{j\omega(i-k)T} d\omega \]
\[ -2 \sum_{i=0}^{N} C_k \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_d(\omega) e^{j\omega(i-k)T} d\omega \right] + \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_d(\omega) d\omega \tag{3.12} \]
The frequency integrals are simply correlation functions; that is,

\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_x(\omega) e^{i\omega(k-k')} d\omega = R_x(k'T-kT) \]  \hspace{1cm} (3.13)

and

\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_d(\omega) e^{i\omega k'} d\omega = R_d(k'T) \]  \hspace{1cm} (3.14)

Now let \( c_k = c_k^0 + \delta_k \); \( k = 0, 1, \ldots, N \)

where \( c_k^0 \) is the optimum and \( \delta_k \) is an arbitrary constant. For \( c_k^0 \) optimum, \( F(\epsilon) \) must now have a minimum at \( \epsilon = 0 \). Or

\[
\frac{dF}{d\epsilon} \bigg|_{\epsilon = 0} = \sum_{k=0}^{N} \sum_{\ell = 0}^{N} R_x(\ell-k') (\delta_k (c_k + \epsilon \ell_k) + \delta_\ell (c_k + \epsilon \delta_k))
\]

\[
-\sum_{\ell = 0}^{N} R_d(\ell) \delta_\ell
\]

\[
= 2 \sum_{\ell = 0}^{N} \left( \sum_{k=0}^{N} c_k R_x(\ell-k) - R_d(\ell) \right) \delta_\ell = 0
\]

for any and all \( \delta_\ell \). Therefore

\[
\sum_{k=0}^{N} c_k R_x(\ell-k) = R_d(\ell) \quad \text{for } \ell = 0, 1, \ldots, N. \hspace{1cm} (3.15)
\]

In matrix form

\[ R^T \eta = \begin{pmatrix} R_d \end{pmatrix} \]

or

\[ \eta^* = (R^T \eta)^{-1} R_d \eta \]  \hspace{1cm} (3.16)

where

\[ R^T \eta = \mathbb{E} \left\{ \begin{pmatrix} \eta_0 & \cdots & \eta_0 \\ \cdots & \cdots & \cdots \\ \eta_N & \cdots & \eta_N \end{pmatrix} \right\} \]  \hspace{1cm} (3.17)
and
\[ R_{\eta} = E \left\{ \left( \frac{d\eta}{\eta} \right) \right\} \]  

(3.18)

B. Direct Differentiation

Since
\[
e(t) = d(t) - \sum_{i=0}^{N} c_i \eta_i(t) = d(t) - \eta^T c
\]
\[ e^2(t) = d^2(t) - 2d^T \eta c - \eta^T c \eta c \]

Taking the mathematical expectation, we have
\[
e^2(t) = d^2(t) - 2d^T \eta c + \eta^T c R \eta c
\]

(3.19)

Let
\[
\nabla C^Q = \begin{pmatrix}
\frac{\partial}{\partial \eta} \\
\frac{\partial}{\partial c}
\end{pmatrix}^Q
\]

Since
\[
\nabla_c \{ c^T R \eta c \} = R \eta c + R \eta^T c
\]

together with \[ R = R^T \]

The gradient of \( e^2 \) is
\[
\nabla_c e^2 = -2 d \eta c + 2 R \eta^T c
\]

(3.20)

For \( e^2 \) to be minimum, we set \( \nabla_c e^2 = 0 \).
Then \( c^* = (R_{\eta}^T)^{-1} R_{d\eta} \) same as obtained above.

C. Derivation from Wiener Filter.

Another derivation of \( c^* \) can be obtained directly from the Wiener filter. From (1.5), the optimum linear filter for additive noise is

\[
H_0(\omega) = \frac{\phi_{dx}(\omega)}{\phi_{xx}(\omega)}
\]

(3.21)

Setting \( c = c^* \) in Eq. (3.3) and combining with Eq. (3.2) give

\[
H_0(\omega) = \sum_{k=0}^{N} c_k^e e^{-j\omega t} e^{-j\omega t} = \left[ \begin{array}{c} e^{-j\omega t} \\
\end{array} \right]^T \left( c^* \right) = \frac{\phi_{dx}(\omega)}{\phi_{xx}(\omega)}
\]

or

\[
\phi_{xx}(\omega) \left[ \begin{array}{c} e^{-j\omega t} \\
\end{array} \right]^T \left( c^* \right) = \phi_{dx}(\omega)
\]

(3.22)

Multiplying both sides of (3.22) by \( e^{j\omega t} = (e^{j\omega t}, \ldots, e^{j\omega t})^T \) and integrating from \(-\infty\) to \(+\infty\), we obtain

\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{xx}(\omega) \left[ e^{-j\omega t} \right]^T \left( e^{j\omega t} \right) d\omega \left( c^* \right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{dx}(\omega) \left( e^{-j\omega t} \right) d\omega
\]

(3.23)

Comparing (3.23) with (3.15), we see that they are identical.

2. Minimum Mean Square Error and Effect of Non-optimum Settings.

A. Expressions of mean square error.

The minimum mean square error of the tapped delay line filter is
obtained by substituting the expression of the optimum filter into $e^2$.

Using

$$e^{c^2} = \begin{bmatrix} \hat{c}_o^* \\ \hat{c}_n^* \end{bmatrix} R \eta^T \eta^{-1} R \eta$$  \hspace{1cm} (3.16)

we have

$$e^{2 \min} = \overline{d^2(t)} - 2 \overline{\eta^T \eta} \hat{c} + \hat{c}^T R \eta \hat{c}$$

$$= \overline{d^2(t)} - 2 \overline{\eta^T \eta} R \eta^T \eta^{-1} \overline{d \eta}$$

$$+ \overline{\eta^T \eta} R \eta^T \eta^{-1} \overline{d \eta}$$

$$= \overline{d^2(t)} - \overline{R \eta^T \eta R \eta^T \eta^{-1} \overline{d \eta}}$$  \hspace{1cm} (3.24a)

or

$$= \overline{d^2(t)} - \overline{R \eta^T \eta \hat{c}^*}$$  \hspace{1cm} (3.24b)

or

$$= \overline{d^2(t)} - \overline{\overline{e^2_o(t)}}$$  \hspace{1cm} (3.24c)

where $z_0(t)$ is the output of the optimum filter.

In terms of $e^{2 \min}$, $e^2(t)$ can be expressed as follows:

From $e^2(t) = \overline{d^2(t)} - 2 \overline{\eta^T \eta} \hat{c} + \hat{c}^T R \eta \hat{c}$  \hspace{1cm} (3.19)

Using Eq. (3.16) and (3.24b), the mean square error is then expressed as

$$e^2(t) = \overline{d^2(t)} - 2 \overline{\eta^T \eta} \hat{c} + \hat{c}^T R \eta \hat{c}$$

$$= e^{2 \min} + \overline{\eta^T \eta} \hat{c}^* - 2 \overline{\eta^T \eta} \hat{c} + \hat{c}^T R \eta \hat{c}$$

$$= e^{2 \min} + \hat{c}^T R \eta \hat{c} - 2 \hat{c}^T \overline{\eta^T \eta} \hat{c} + \hat{c}^T R \eta \hat{c}$$

$$= e^{2 \min} + \hat{c}^T R \eta \hat{c} - \overline{\hat{c}^T R \eta \hat{c}} (\hat{c} - \hat{c}^*)$$  \hspace{1cm} (3.25)
B. Effect on Minimum Mean Square Error due to Non-optimum Settings.

From (3.25) the difference in mean square error due to non-optimum values of \( c \) is

\[
\Delta F = e^2(t) - e_{\text{min}}^2 = (c^T - c^\ast)^T R \eta (c - c^\ast)
\]

Let \( c_i = c_i^\ast + \delta_i \)

then

\[
\Delta F = \delta^T \sum_i \delta_i \eta_i \eta_j = \sum_i \sum_j \delta_i \delta_j \eta_i \eta_j \leq (N+1)^2 \max_{i,j} \delta_i \eta_i \eta_j
\]

Thus, the error due to non-optimum settings is bounded if the deviations of the weights and the input correlation functions are bounded. Note that for delay line filters \( \max_{i,j} |\eta_i \eta_j| = R_x(0) = R_x(0) + R_n(0) \).

C. Relationship between the minimum mean square error and the number of time delay elements used in the filter.

From (3.24a) it is seen that

\[
e_{\text{min}}^2 = d^2(t) - \sum \eta_i \eta_j \sum \eta_i \eta_j
\]

Taking the case \( d(t) = s(t) \) and noting that

\[
d(t) \eta_i(t) = s(t) x(t - T_i)
\]

\[
= s(t) \left[s(t - T_i) + n(t - T_i)\right]
\]

\[
= \frac{s(t)}{s(\cdot - T_i)} \frac{s(\cdot - T_i)}{R_x(T_i)}
\]

D32
we have

$$e_{\text{min}}^2 = R_s(o) - \left[ R_s(o) R_s(T_1) \cdots R_s(T_n) \right] \begin{bmatrix} R_x(o) & \ldots & R_x(T_n) \\ R_x(T_1) R_x(o) R_x(T_{n-1}) & \cdots & R_s(T_1) \\ \vdots & \ddots & \vdots \\ R_x(T_n) R_x(o) & \cdots & R_s(T_n) \end{bmatrix} \begin{bmatrix} R_s(o) \\ R_s(T_1) \\ \vdots \\ R_s(T_n) \end{bmatrix}$$

where $R_x(T_i) = R_s(T_i) + R_n(T_i)$

(3.24d)

The last term in the right hand side of (3.24d) is a functional of $N$, the number of delay line elements, and the correlation functions. For any given forms or values of $R_x(T_i)$ a plot of $e_{\text{min}}^2$ versus $N$ can be constructed. It is anticipated that the larger $N$ is the smaller $e_{\text{min}}^2$ will be.

3. Adaptive Tapped Delay Line Filters

The above discussion presented a means of determining the optimum values of the gains provided that the statistical properties of both the desired signal and the noise are known. Unfortunately, in practice, it is not always possible to know all this information very accurately. If only the filter input and output are available and nothing else, no systematic procedures can be found to adjust the gains. However, if we know something about the system, then we can develop some algorithms to make the filter optimum. It will be shown that if a desired signal is available, or correlation functions of the desired signal, or (not and) correlation functions of the noise can be estimated within acceptable accuracy, the methods of stochastic approximation can be employed to make the filter adaptive to changing operating conditions. These changes may be due to variation in the input signal or the internal structure of the filter. Adaptation is accomplished by observation of the reaction of the
filter to an external signal or to an internal variation with subsequent
goal-directed variation of the filter parameters so as to minimize some
quality criterion.

The quality criterion may be represented in the form of the mathematical
expectation of some strictly convex (not necessarily quadratic) function
of the deviation of the output variation from the desired function.

For simplicity we shall use the mean squared criterion. Thus,

$$I(c) = E\left\{ Q(d(t) - z(t)) \right\} \text{ with } Q(e) = e^2$$  \hspace{1cm} (3.28)

For the tapped delay line filter shown schematically in Fig. 2, we know,

$$x(t) = s(t) + n(t)$$  \hspace{1cm} (1.1)

It is assumed here that these functions are stationary random processes.
The desired function is the function obtained by applying an arbitrary
operation on $s(t)$. This operator may be a differential operator, in-
tegral operator, predictor, etc. It can even be a unity operator such
that $d(t) = s(t)$. We shall first of all consider the case where $d(t)$
is available. Those cases for which signal or noise correlation functions
are known will be treated in a later section. They will turn out to be
slight modification of the first case. Nonstationary or time varying systems
will be considered subsequently.

For the first case

$$I(c) = E \left\{ Q(d(t) - z(t)) \right\}$$  \hspace{1cm} (3.29)

since

$$z(t) = \sum_{k=0}^{N} c_k \eta_k(t) = \sum_{k=0}^{N} c_k x(t-kT)$$

D-34
we have
\[ I(c) = E \left\{ Q(\alpha(t)) - \sum_{k=0}^{N} c_k \eta_k(t) \right\} \]

\[ = \int_{x} Q(d(t) - \sum_{k=0}^{N} c_k \eta_k(x)) P(x) \, dx \]  

(3.30)

Since \( P(x) \) is generally unknown, algorithm (2.9) will be used. For 
\( Q(e) = e^2(t) \), we see that
\[ \nabla_c Q(x|c) = 2e \nabla_c e \]

But
\[ \nabla_c e = \nabla_c (d(t) - \sum_{k=0}^{N} c_k \eta_k(t)) \]

therefore
\[ \begin{bmatrix} \frac{\partial e}{\partial c_0} \\ \frac{\partial e}{\partial c_k} \end{bmatrix} = 2(d(t) - \sum_{k=0}^{N} c_k \eta_k(t)) \begin{bmatrix} -\eta_0(t) \\ -\eta_n(t) \end{bmatrix} \]

and the desired algorithm is
\[ c_{j+1} = c_j + 2\gamma_j e_{j} \eta_{j/\dot{e}_j} \]  

(3.31)

This is precisely the LMS algorithm used by Widrow\textsuperscript{13} derived from intuitive reasoning rather than from rigorous mathematical proofs.

It would be desirable and instructive to give some physical interpretations of the conditions under which algorithm (3.31) converges.

D-35
Algorithm (2.9)

\[ c_{j+1} = c_j - \gamma_j \nabla_c Q(x_j | c_j) \]  

(2.9)

converges if the following conditions are satisfied:

\begin{align*}
  \text{(a)} & \quad \lim_{j \to \infty} \gamma_j = 0, \quad \sum_{j=1}^{\infty} \gamma_j = \infty, \quad \sum_{j=1}^{\infty} \gamma_j^2 < \infty \\
  \text{(b)} & \quad \inf_{\epsilon < \|c - c^*\|} E \left\{ (c - c^*)^T \nabla_c Q(x|c) \right\} > 0 \\
  \text{in the neighborhood of} \ c^* \quad \epsilon > 0 \\
  \text{(c)} & \quad \left\{ \nabla_c^T Q(x|c) \nabla_c Q(x|c) \right\} \leq d(c^* c + c c^*). \quad (2.15)
\end{align*}

The choice of \( \gamma_j \) which satisfies (a) is rather at our own disposal.

For example, \( \gamma_j = \frac{a}{3^j} \) with \( a, b > 0 \) will definitely fulfill the requirement of (a). The remaining conditions depend on the surface of the error gradient, which in turn depends on the choice of error criterion and the physical system under consideration.

Condition (b) is satisfied as long as the function \( Q(e) \) is strictly convex. Since \( Q(e) \) has a minimum at \( c = c^* \), it is evident that

\[ \frac{\partial Q}{\partial c_i} > 0 \text{ for } c_i > c_i^* \]

\[ = 0 \text{ for } c_i = c_i^* \]

\[ \leq 0 \text{ for } c_i = c_i^* \]

\[ i = 0, 1, 2, \ldots, N \]

Consequently

\[ (c_i - c_i^*) \frac{\partial Q}{\partial c_i} \geq 0 \text{ for all } i \]
and
\[ \varepsilon < \inf \left\| \mathbf{c} - \mathbf{c}^* \right\| \leq \frac{1}{\varepsilon} \mathbb{E} \left\{ (\mathbf{c} - \mathbf{c}^*)^T \nabla_{\mathbf{c}} Q(\mathbf{x}|c) \right\} > 0 \]

\[ \varepsilon > 0 \]

Condition (c) is satisfied if

(a) \( \frac{\partial^2 Q}{\partial \mathbf{c}^2} \) exists and is uniformly bounded.

(b) \( s(t) \) and \( n(t) \) are uniformly bounded.

Using a Taylor expansion about \( \mathbf{c} = \mathbf{c}^* \), we have

\[
 \frac{\partial Q(c)}{\partial c} = 0 + \sum_{i=0}^{N} (c_i - c_i^*) \left. \frac{\partial^2 Q(c)}{\partial c_i \partial c_j} \right|_{c=c^*} \]

for arbitrary \( j \), with \( j = 0, 1, 2, \ldots, N. \)

For the tapped delay line filter with mean square error criterion

Therefore,

\[
 Q(c) = Q \left( d(t) - \sum_{i=0}^{N} c_i \gamma_i(t) \right) \]

\[ (3.34) \]

\[ \frac{\partial Q(c)}{\partial c_i} = \frac{\partial Q}{\partial \varepsilon} \left( - \gamma_i(t) \right) \]

\[ (3.35) \]

\[ \frac{\partial^2 Q(c)}{\partial c_i \partial c_j} = \frac{\partial^2 Q}{\partial \varepsilon^2} \gamma_i(t) \gamma_j(t) \]

By definition

\[ \gamma_i(t) = x(t-T_i) = s(t-T_i) + n(t-T_i) \]

\[ (3.36) \]

It is evident that \( \frac{\partial^2 Q}{\partial c_i \partial c_j} \) is bounded if conditions (a) and (b) are satisfied.
Therefore,
\[
\nabla_c Q(x|c) \leq k_1 \sum_{i=0}^{N} (c_i^*-c_i^0)
\]

(3.37)

where \( k_1 = k \sup_{all \, i} \left| \frac{\partial^2 Q(c)}{\partial c_i \partial c_j} \right| \)

Taking the inner product and mathematical expectation on each side of (3.37) gives
\[
E\left\{ \nabla_c^T Q(x|c) \nabla_c Q(x|c) \right\} \\
\leq k_1^2 \sum_{i=0}^{N} (c_i^*-c_i^0)^2 \\
\leq k_1^2 \sum_{i=0}^{N} (c_i^2 + c_i^{02})
\]

(3.38)

In practice conditions (a) and (b) are easily satisfied. Thus the methods of stochastic approximation can be employed in a variety of adaptive processes.
IV. Adaptive Schemes and Rate of Convergence

1. Adaptive Schemes

A. An algorithm has been presented to adjust the gains in the tapped delay line filter. If the desired signal is available to generate the error gradient, the adaptive scheme is given by

\[ c_{j+1} = c_j + 2 \gamma_j \frac{e_j}{\tau_j} \]

with

\[ e(t) = d(t) - \frac{N}{\sum_{k=0}^{N} c_k \eta_k(t)} \]

The scheme is shown below.

The complete adaptive system is shown in Fig. 3.

Fig. 3. Adaptive system with \( d(t) \) available
B. When \( d(t) \) is not available but the statistical properties of the noise are known, the algorithm is modified as follows.

Using \( s(t) = x(t) - n(t) \) in the expression of error criterion we have

\[
I(c) = E \left\{ Q(e) \right\} = E \left\{ \left( s(t) - z(t) \right)^2 \right\} \\
= E \left\{ \left( x(t) - n(t) - z(t) \right)^2 \right\} \\
= E \left\{ \left( x(t) - z(t) \right)^2 \right\} + E \left\{ n^2(t) \right\} \\
- 2 E \left\{ n(t) \left( x(t) - z(t) \right) \right\} \\
\]  

(4.3)

Since

\[
x(t) = s(t) + n(t) \\
z(t) = \sum_{k=0}^{N} c_k y_k(t) = \sum_{k=0}^{N} c_k x(t - T_k) \\
= \sum_{k=0}^{N} c_k \left( s(t - T_k) + n(t - T_k) \right) \\
\]

and \( E \left\{ s(t) n(t) \right\} = 0 \)

Eq. (4.3) becomes

\[
I(c) = E \left\{ \left( x(t) - z(t) \right)^2 \right\} \]

\[
- 2 E \left\{ n^2(t) \right\} + 2 E \left\{ n(t) \sum_{k=0}^{N} c_k n(t - T_k) \right\} \\
= E \left\{ \left( x(t) - z(t) \right)^2 \right\} - E \left\{ n^2(t) \right\} \\
+ 2 E \left\{ \sum_{k=0}^{N} c_k n(t) n(t-T_k) \right\} \\
= E \left\{ \left( x(t) - z(t) \right)^2 \right\} - R_n(o) + 2 \sum_{k=0}^{N} c_k R_n(T_k) \\
\]  

(4.4)

where \( R_n(T_k) \) is the noise correlation function.

In comparing the algorithm used for the case when \( d(t) \) is available, if \( s(t) \) is replaced by \( x(t) \), we would adjust the gains \( c \) to minimize the
first term on the right hand side of (4.4), i.e., to solve the equation
\[ \text{grad } E \left( x(t) - z(t) \right)^2 = 0 \]
Now consider the rest of the right-hand side of (4.4). The second term
is independent of \( C \), and the entries \( R_n(T_k) \) appearing in the third term
are known. We wish to minimize the sum of the three terms, i.e., find the
solution
\[ \text{grad } E \left( x(t) - z(t) \right)^2 + \text{grad} 2 \sum_{k=0}^{N} C_k R_n(T_k) = 0 \]
At this point we shall use a modified algorithm whose convergence properties
and proofs are found at Appendix A. It is shown that if we let \( Q = Q_1 + Q_2 \),
the algorithm
\[ C_{j+1} = C_j - \gamma_j \left( v_c Q_1 + \overline{v_c Q_2} \right) \]  
also converges in the same sense and under the same physical conditions as
algorithm (2.9) for tapped delay line filters.
In (4.3) we can set
\[ Q_1 = \left[ x(t) - z(t) \right]^2 \]
\[ Q_2 = u(t) - \zeta_n(t) [x(t) - z(t)] \]
But
\[ \overline{Q_2} = R_n(0) + 2 \sum_{k=0}^{N} C_k R_n(T_k) \]
and
\[ \overline{Q_2} = 2R_n = 2 \left[ R_n(0) R_n(T) ... R_n(NT) \right]^T \]
Thus, algorithm (4.1) is modified to
\[ C_{j+1} = C_j - \gamma_j \left( v_c x(t) - z(t) \right)^2 + 2R_n \]
\[ = C_j + 2\gamma_j \eta_j (x(t) - z(t)) - 2\gamma_j R_n \]  
The adaptive scheme is shown below and the whole system is drawn in Fig. 4.
Fig. 4. Adaptive system with known noise statistics

C. Now we shall consider the case when the statistical properties of the signal are known.

Since

$$I(g) = E \left\{ \left( s(t) - Z(t) \right)^2 \right\}$$

$$= E \left\{ s^2(t) \right\} - 2 E \left\{ s(t) Z(t) \right\} + E \left\{ Z^2(t) \right\}$$

$$= E \left\{ s^2(t) \right\} - 2 E \left\{ s(t) \sum_{k=0}^{N} c_k \left( s(t - T_k) + n(t - T_k) \right) \right\}$$

$$+ E \left\{ c_i c_j \eta_i(t) \eta_j(t) \right\}$$

$$= R_s(0) - \sum_{k=0}^{N} 2 c_k R_n(T_k) + \sum_{i=0}^{N} \sum_{j=0}^{N} c_i c_j \eta_i(t) \eta_j(t)$$

We cond set

$$Q_1 = \eta_1(t)$$

$$Q_2 = s^2(t) - 2s(t) z(t)$$

D-42
Following the same procedure as before, we have in this case an algorithm

$$S_{j+1} = S_j + 2 \gamma_j R_S - 2 \gamma_j \eta_j z_j$$  \hspace{1cm} (4.11)

the scheme for Eq. (4.11) is then

\[ Z(t) \xrightarrow{\otimes} \xrightarrow{+} \xrightarrow{\gamma} \Delta C_c \]

\[ \eta_j(t) \xrightarrow{R_j(T_i)} \]

while the whole system is shown in Fig. 5.

**Fig. 5.** Adaptive system with known signal statistics.

In the above schemes no distinction between continuous and discrete processes has been made because their connection is obvious. In the discrete case we can set \( \gamma_j = \frac{1}{j} \) while its continuous counterpart is \( \gamma(t) = \frac{1}{t} \). Theorems concerning the choice of \( \gamma(t) \) already exist\(^{26} \) and are not discussed here.
2. Rate of convergence

Having found an algorithm which converges, we shall investigate how fast it converges. In other words, we would like to know the mean square error at each stage during the adaptation period.

From (4.1)

\[ \mathbf{C}_{j+1} = \mathbf{C}_j + 2 \gamma_j \mathbf{e}_j \]

and using the expression

\[ e_j = d_j - z_j = d_j - \sum_{i=0}^{N} c_i^T \eta_i \]

\[ = d_j - \eta_j^T \mathbf{c}_j \]

we get the corresponding matrix form

\[ \mathbf{C}_{j+1} = \mathbf{C}_j - 2 \gamma_j \mathbf{e}_j^T \mathbf{e}_j \mathbf{C}_j + 2 \gamma_j \mathbf{d}_j \]

\[ = (1 - 2 \gamma_j \mathbf{e}_j^T \mathbf{e}_j) \mathbf{C}_j + 2 \gamma_j \mathbf{d}_j \]

Taking the mathematical expectation of Eq. (4.12) and diagonalizing the matrix \( \mathbf{E} \left\{ \mathbf{e}_j^T \mathbf{e}_j \right\} \) such that

\[ \mathbf{E} \left\{ \mathbf{e}_j^T \mathbf{e}_j \right\} = \mathbf{R} \] \[ = \mathbf{P}^{-1} \lambda \mathbf{P} \]

where \( \mathbf{P} \) is an orthonormal matrix, and \( \lambda = \left( \begin{array}{c} \lambda_1 \\ \vdots \\ \lambda_N \end{array} \right) \) is the eigenvalue matrix, we obtain

\[ \mathbf{C}_{j+1} = (1 - 2 \gamma_j \mathbf{R} \mathbf{e}_j) \mathbf{C}_j + 2 \gamma_j \mathbf{d}_j \]

\[ = (1 - 2 \gamma_j \mathbf{P}^{-1} \lambda \mathbf{P}) \mathbf{C}_j + 2 \gamma_j \mathbf{d}_j \]

\[ = \mathbf{C}_j (1 - 2 \gamma_j \mathbf{P}^{-1} \lambda \mathbf{P}) + 2 \gamma_j \mathbf{d}_j \]

\[ \mathbf{C}_{j+1} = \mathbf{C}_j + 2 \gamma_j \mathbf{d}_j \]

In the above we assumed that \( \mathbf{c} \) is statistically independent of \( \eta \).

Although \( \mathbf{c} \) cannot affect \( \eta \) in any manner, the increment of \( \mathbf{c} \) at each
stage is, however, related to \( n \) by (4.1). Since the increment is generally very small and the total effect involves addition of a large number of small increments, we can assume \( \frac{\Delta n}{\Delta t} = \frac{\Delta \bar{y}}{\Delta t} \) in a manner similar to that used in the analysis of phase-locked loops.

Let us define

\[
\overline{W} = P \overline{C}, \quad \overline{r}' = P \overline{r}
\]  

(4.14)

then (4.13) becomes

\[
\overline{W}_{j+1} = (1 - 2\gamma_j \lambda) \overline{W}_j + 2\gamma_j \overline{d\eta}_j
\]  

(4.15)

Since \( \overline{d\eta}_j = R_{\eta} C^* \) as seen from (3.16), we have

\[
\overline{W}_{j+1} - \overline{W}^* = (1 - 2\gamma_j \lambda) (\overline{W}_j - \overline{W}^*)
\]  

(4.16)

Now consider any particular component \( w \) of \( \overline{W} \) and for clarity no subscript or superscript indicating the component is used. Then

\[
\overline{W}_{j+1} - \overline{W}^* = (1 - 2\gamma_j \lambda) (\overline{W}_j - \overline{W}^*)
\]  

(4.17)

Using Eq. (4.17), recursively gives

\[
\overline{W}_j = (\overline{W}_1 - \overline{W}^*) \sum_{k=1}^{j-1} (1 - 2\gamma_j \lambda) + \overline{W}^*
\]  

(4.18)

We shall now find $\omega_j^2$.

From (4.1) and taking the product of $c_j$ and $c_j^T$, we obtain

$$c_{j+1} = (c_j + 2\gamma_j e_j n_j) (c_j^T + 2\gamma_j e_j n_j^T)$$

$$= c_j c_j^T + 2\gamma_j e_j (c_j n_j^T + n_j c_j^T) + 4\gamma_j^2 e_j^2 n_j n_j^T$$

(4.19)

Since

$$e_j (c_j n_j^T + n_j c_j^T)$$

$$= (d_j - n_j c_j)(c_j n_j^T + n_j c_j^T)$$

$$= c_j d_j n_j^T + d_j n_j c_j^T - c_j c_j^T n_j - n_j c_j^T c_j$$

Note $A B^T + B A^T = 2 (A B^T)^s$

where $s$ denotes the symmetrical part of a matrix. For example, if

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

then

$$A^s = \begin{bmatrix} \frac{1}{2} (a_{11} + a_{11}) & \frac{1}{2} (a_{12} + a_{21}) \\ \frac{1}{2} (a_{21} + a_{12}) & \frac{1}{2} (a_{22} + a_{22}) \end{bmatrix}$$

we have

$$e_j (c_j n_j^T + n_j c_j^T)$$

$$= 2 \{ R_{\eta} c^* c_j^T \}^s - 2 \{ R_{\eta} c_j c_j^T \}^s$$

$$= 2 \{ R_{\eta} (c^* - c_j) c_j^T \}^s$$

(4.20)
Taking mathematical expectation on both sides of (4.19) and using (4.20) yield

\[ c_{j+1}^T c_{j+1} = c_j^T c_j + 4\gamma_j \left( R_\eta (c^* - c_j) c_j^T \right)^s + 4\gamma_j^2 e_j^2 n_j n_j^T \]  

(4.21)

For large \( j \), the following approximation can be made

\[ e_j^2 n_j n_j^T = e_{\text{min}}^2 n_j n_j^T = e_{\text{min}}^2 R_\eta \]  

(4.22)

(4.22) can be viewed as a Taylor series expansion around the optimum point and with higher order terms neglected for large \( j \).

Therefore, (4.21) becomes

\[ c_{j+1}^T c_{j+1} = c_j^T c_j + 4\gamma_j \left( R_\eta (c^* - c_j) c_j^T \right)^s + 4\gamma_j^2 e_{\text{min}}^2 R_\eta \]  

(4.23)

Using the transformation \( c = P^{-1} w \) as defined in (4.14), we can change (4.23) to the form

\[ \tilde{w}_{j+1}^T \tilde{w}_{j+1} = \tilde{w}_j^T \tilde{w}_j + 4\gamma_j P^{-1} - P^{-1} (\tilde{w}^* - \tilde{w}) P^{-1} \left( P^T \tilde{w}^* - \tilde{w}^T \right) \] 

\[ + 4\gamma_j^2 P e_{\text{min}}^2 P^{-1} - P P^{-1} \]  

And

\[ \left( \tilde{w}_{j+1}^T \tilde{w}_{j+1}^T \right)^D = \left( \tilde{w}_j^T \tilde{w}_j^T \right)^D + 4\gamma_j \left( (\tilde{w}^* - \tilde{w})^T \tilde{w}^T \right)^D + 4\gamma_j^2 e_{\text{min}}^2 R_\eta \]  

(4.24)

In the above \( D \) denotes the diagonal elements of a matrix. These elements have
the desired form \( \bar{w}^2 \), which can be expressed as

\[
\bar{w}_{j+1}^2 = \bar{w}_j^2 + 4\gamma_j \lambda (w^* - w_j) \bar{w}_j + 4\gamma_j^2 \lambda \varepsilon_{\min}^2
\]

\[
= (1 - 4\gamma_j \lambda) \bar{w}_j^2 + 4\gamma_j \lambda w^* \bar{w}_j + 4\gamma_j^2 \lambda \varepsilon_{\min}^2
\]  

(4.25)

From (4.16) we have

\[
\bar{w}_{j+1} \bar{w}_{j+1}^T = (1 - 4\gamma_j \lambda) (\bar{w}_j \bar{w}_j^T) + 4\gamma_j \lambda \bar{w}^* \bar{w}_j^T
\]

(4.26)

Let \( \theta_j = (\bar{w}_j - \bar{w}_j) (\bar{w}_j - \bar{w}_j)^T = (\bar{w} \bar{w}^T)_{j} - \bar{w}_j \bar{w}_j^T \)

Subtracting the diagonal terms of (4.24) from those of (4.26), we obtain

\[
\theta_{j+1}^D = (1 - 4\gamma_j \lambda) \theta_j^D + 4\gamma_j \lambda \varepsilon_{\min}^2
\]

(4.27)

Since \( \theta_j^D \) has the elements \( \bar{w}_j^2 - \bar{w}_j^2 \), we see that for any particular component of \( \theta_j^D \),

\[
\bar{w}_{j+1}^2 - \bar{w}_j^2 = \theta_{j+1}^D (1 - 4\gamma_j \lambda) \theta_j + 4\gamma_j \lambda \varepsilon_{\min}^2
\]

(4.28)

Iterating backward,

\[
\theta_{j+1}^D = \theta_j^D \sum_{k=1}^{j} (1 - 4\gamma_k \lambda) + 4\gamma_k \lambda \varepsilon_{\min}^2 \sum_{k=1}^{j} \gamma_k^2 \sum_{l=k+1}^{j} (1 - 4\gamma_l \lambda)
\]

(4.29)

But \( \theta = 0 \) because \( \bar{w}_1 = w_1 \),

\[
\theta_{j+1}^D = 4\gamma \varepsilon_{\min}^2 \sum_{k=1}^{j} \gamma_k^2 \sum_{l=k+1}^{j} (1 - 4\gamma_l \lambda)
\]

(4.30)

or

\[
(w_{j+1} - \bar{w}_{j+1})^2 = 4\gamma \varepsilon_{\min}^2 \sum_{k=1}^{j} \gamma_k^2 \sum_{l=k+1}^{j} (1 - 4\gamma_l \lambda)
\]

D-18
Several special cases will be considered.

(1) Setting \( \gamma_j = \frac{1}{2(j+1)\lambda} \) \hspace{1cm} (4.31)

This is a legitimate expression as \( \gamma_j \) defined by (4.31) satisfies all the required conditions for convergence.

Note \( \frac{1}{\lambda} \sum_{k=1}^{j} \left( 1 - \frac{1}{k+1} \right) = \frac{1}{\lambda} \sum_{k=1}^{j} \frac{k}{k+1} = \frac{1}{j+1} \) \hspace{1cm} (4.32)

(4.18) gives us

\[
\hat{w}_{j+1} = \frac{1}{j+1} (w_1 - w^*) + w^* \]

(4.33)

Note also that [see Eq. (B.8) of Appendix B],

\[
\sum_{k=1}^{j} \gamma_k \left( 1 - 4\gamma_k^2 \right) = \sum_{k=1}^{j} \left( 1 - \frac{2}{k+1} \right) = \frac{(k+1)^2}{(j+1)^2} \]

\[
\sum_{k=1}^{j} \frac{1}{4\lambda^2} (1 - 4\gamma_k^2) = \frac{1}{4\lambda^2} \left( \frac{(k+1)^2}{(j+1)^2} \right) \]

\[
= \frac{1}{4\lambda^2} \sum_{k=1}^{j} \frac{1}{(j+1)^2} = \frac{1}{4\lambda^2} \left[ \frac{1}{(j+1)^2} \right] \]

Thus (4.30) gives us

\[
\frac{(\hat{w}_{j+1} - \hat{w}_{j+1})^2}{(w_1 - w^*)^2} = \frac{e^2_{\text{min}}}{\lambda} \frac{1}{(j+1)^2} \]

(4.36)

As derived in (3.25), the mean squared error at any time is given by

\[
e_j^2 = e_{\text{min}}^2 + (c_j - c^*)^T R_n (c_j - c^*)
\]

\[
= e_{\text{min}}^2 + (c_j - c^*)^T F^{-1} \Lambda \ P (c_j - c^*)
\]

\[
= e_{\text{min}}^2 + (\bar{w}_j - \bar{w}^*)^T \Lambda (\bar{w}_j - \bar{w}^*) \]

(4.37)

The expected difference between the mean squared error at each stage during the adaptation period and the minimum mean squared error is then
\[
E \left\{ \frac{\mathbf{e}_{j+1}^2}{ \mathbf{e}_{\text{min}}^2} \right\} = E \left\{ \begin{pmatrix} \mathbf{w}_{j+1} \end{pmatrix}^T \begin{pmatrix} \mathbf{w}_{j+1}^* \end{pmatrix} \right\} \\
= E \left\{ \sum_{i=0}^{N} \lambda_i \left( \mathbf{w}_{j+1,i} - \mathbf{w}_{i}^* \right)^2 \right\} \\
= \sum_{i=0}^{N} \lambda_i \left( \mathbf{w}_{j+1,i} - \mathbf{w}_{i}^* \right)^2 \\
(4.38)
\]

But

\[
\left( \mathbf{w}_{j+1} - \mathbf{w}^* \right)^2 = \left( \mathbf{w}_{j+1} - \mathbf{w}_{j+1}^* \right)^2 + \mathbf{w}_{j+1}^2 - 2 \mathbf{w}^* \mathbf{w}_{j+1}^* + \mathbf{w}^2
\]

Using (4.33) and (4.36), we have

\[
\left( \mathbf{w}_{j+1} - \mathbf{w}^* \right)^2 = \frac{1}{\lambda_{\text{min}}} \left( \mathbf{w}_{1} - \mathbf{w}_{1}^* \right)^2 + \frac{1}{\lambda_{1}} \left( \mathbf{w}_{1,j} - \mathbf{w}_{1,j}^* \right)^2 \\
(4.39)
\]

(4.38) becomes

\[
E \left\{ \frac{\mathbf{e}_{j+1}^2}{ \mathbf{e}_{\text{min}}^2} \right\} = \frac{1}{\lambda_{j+1}} \sum_{k=0}^{N} \lambda_k \left( \mathbf{w}_{k,j} - \mathbf{w}_{k,j}^* \right)^2 \\
= \frac{1}{\lambda_{j+1}} \left( \mathbf{c}_{j+1} \right)^2 \mathbf{e}_{\text{min}}^2 + \frac{1}{\lambda_{j+1}} \left( \mathbf{c}_{j+1} - \mathbf{c}_{1} \right)^T \mathbf{R}_{\mathbf{w}} \left( \mathbf{c}_{j+1} - \mathbf{c}_{1} \right) \\
(4.40)
\]

The last step is obtained from

\[
\sum_{k=0}^{N} \lambda_k \mathbf{w}_{k,j}^2 = \mathbf{w}_{j+1}^T \mathbf{w}_{j+1} = \mathbf{c}_{j+1}^T \mathbf{R}_{\mathbf{w}} \mathbf{c}_{j+1}
\]

Thus for large \( j \),

\[
E \left\{ \frac{\mathbf{e}_{j+1}^2}{ \mathbf{e}_{\text{min}}^2} \right\} \approx \frac{(N+1) \mathbf{e}_{\text{min}}^2}{j+1} \\
(4.41)
\]

(4.40) is the desired expression for the rate of convergence. For large \( j \), the mean squared error decreases approximately as the first power of time.

(2) Setting \( \gamma_j = \frac{1}{2(j+1)} \)

\[
(4.42)
\]
The choice of \( \gamma_j \) defined by (4.31) requires some a priori knowledge about the signal and noise properties. Otherwise, if the correlation matrix \( R_n \) is not known, the eigenvalues \( \lambda_j \) cannot be determined. The arbitrary choice of \( \lambda_j \) defined by (4.42) will be studied.

From (4.18) with \( \gamma_j = \frac{1}{\gamma(j+1)} \) we have

\[
\omega_{j+1} = (\omega_1 - \omega^*) \sum_{k=i}^{j} \left( 1 - \frac{\lambda}{j+1} \right) + \omega^*
\]

(4.43)

But

\[
\sum_{k=1}^{j} \left( 1 - \frac{\lambda}{j+1} \right) = \frac{\Gamma(j+1-\lambda)}{(j+1)! \Gamma(2-\lambda)}
\]

\[
= \frac{1}{\Gamma(2-\lambda)(j+1)^{\lambda}} \quad \text{for } j \gg 1 \text{ and } j \gg \lambda
\]

(4.44)*

Thus

\[
\bar{\omega}_{j+1} = \frac{(\omega_1 - \omega^*)}{\Gamma(2-\lambda)(j+1)^{\lambda}} + \omega^*
\]

(4.45)

Note also that

\[
\sum_{k=1}^{j} \frac{2 \lambda}{j+1} = \sum_{k=1}^{j} \frac{(k+1)^{2\lambda}}{(j+1)^{2\lambda}}
\]

(4.46)

Therefore,

\[
\gamma_k \sum_{k=1}^{j} \gamma_k \sum_{k=1}^{j} \frac{1}{4(k+1)^{2\lambda}} = \sum_{k=1}^{j} \frac{1}{4(k+1)^{2\lambda} - 2\lambda} - \frac{1}{4(k+1)^{2\lambda} - 2\lambda}
\]

Using the formula (No. 29.9, Tables of Integrals by Dwight)

\[
\sum_{u=1}^{n} u^{p-1} = \frac{n^{p+1}}{p+1} + n^{p-1} \left( \frac{1}{12} n^{p-1} - \frac{1}{30} n^{p-1} (p-1) (p-2)n^{p-3} + \cdots \right)
\]

* Derivation appears in Appendix B.

D-51
\[ \frac{j}{k+1} = (k+1)^{2\lambda - 2} = \frac{j + 1}{u} = \frac{j + 1}{u} = \frac{p}{u - 1} \]

\[ = -1 + \frac{(j+1)^{2\lambda - 1}}{2 \lambda - 2} + \frac{(j+1)^{2\lambda - 3}}{2} + \frac{1}{12} (2\lambda - 2) (j+1)^{2\lambda - 3} + \ldots \]  

(4.47) then becomes

\[ \frac{j}{k+1} \gamma_k = \gamma_{k+1}^2 \gamma_{k+1}^4 (1 - \gamma_k \lambda) = \frac{1}{4(j+1)^{2\lambda}} [-1 + \frac{(j+1)^{2\lambda - 1}}{2 \lambda - 2} - \frac{(j+1)^{2\lambda - 2}}{2} + \ldots] \]

\[ = \frac{1}{4} \left[ \frac{1}{(j+1)^{2\lambda}} + \frac{1}{(2\lambda - 2)} \frac{1}{j+1} \right] \text{ for large } j \]  

(4.49)

Substituting (4.49) into (4.30) and combining with (4.43) yield

\[ \frac{\lambda}{j+1 - w}^2 \gamma_{\min}^2 \frac{1}{j+1} + \frac{1}{(j+1)^{2\lambda}} \left\{ \frac{(w_1 - w)^2}{\sigma^2 (2\lambda)} - \frac{\gamma_{\min}^2}{j+1} \right\} \]  

(4.50)

and

\[ E \{ \gamma_{j+1}^2 - \gamma_{\min}^2 \} = \sum_{k=0}^{N} \lambda_k \left( \frac{(w_{j+1,k} - w)^2}{\sigma^2 (2\lambda)} \right) \]

\[ = \sum_{k=0}^{N} \frac{\lambda_k^2}{(j+1)^{2\lambda - 2} + \frac{1}{(j+1)^{2\lambda}} \left[ \frac{(w_{1,k} - w)^2}{\sigma^2 (2\lambda)} k \right]} \gamma_{\min}^2 \]  

(4.51)

(3) \( \gamma_j = \gamma = \text{constant} \)

The expressions for \( \gamma_j \) defined by (4.31) and (4.42) satisfy the conditions for convergence as stated in (2.13). In these cases the \( \gamma_j \) and thus the gain increment \( \Delta c_j \) become smaller and smaller as time \( j \) proceeds during the adaptation period. It is anticipated that the rate of convergence will be increased if a small constant value is set for \( \gamma \). As shown by Coner, the algorithm with constant \( \gamma_0 \) has comparatively little noise resistance. Furthermore, in the presence of measuring error with variance \( \sigma^2 \), convergence in the usual sense
does not occur, but

$$\lim_{j \to \infty} E \left\{ \left\| C_j - C^* \right\|^2 \right\} < F(\gamma_0, \sigma^2)$$

(4.53)

and

$$F(\gamma_0, \sigma^2) \to 0 \text{ as } \gamma_0 \to 0$$

Now we shall study the rate of convergence when \( \gamma \) is a constant.

From Eq. (4.18), we see that with \( \gamma_j = \gamma = \text{const} \),

$$\begin{align*}
\frac{\bar{w}_{j+1}}{w_j} &= \frac{(w_j - w^*)}{w_j} \sum_{k=1}^{j} (1 - 2\gamma \lambda)^k + w^* \\
&= (1 - 2\gamma \lambda)^j (w_1 - w^*) + w^*
\end{align*}$$

(4.54)

Since

$$a + a\gamma + a\gamma^2 + \ldots + a\gamma^{n-1} = \frac{a(1 - \gamma^n)}{1 - \gamma}$$

We can obtain

$$\sum_{k=1}^{j} (1 - 4\gamma \lambda)^{-k} = \frac{1}{4\gamma \lambda} \left[ (1 - 4\gamma \lambda)^{-j-1} \right]$$

Thus

$$\sum_{k=1}^{j} \gamma_k (1 - 4\gamma \lambda)^{-k} = \frac{1}{4\gamma \lambda} \frac{1}{(1 - 4\gamma \lambda)^{j-1}}$$

and (4.30) becomes

$$\left( w_{j+1} - \frac{\bar{w}_{j+1}}{w_j} \right)^2 = e^{-2 \gamma (1 - 4\gamma \lambda)^{j-1}}$$
The mean squared error is then

\[ \bar{E} \left\{ \frac{e^2\min_{j+1}}{e^2\min_j} \right\} = \frac{e^2\min_j}{e^2\min_j} \sum_{k=0}^{N} \lambda_k \left(1-(1-4\gamma \lambda)^{j-1}\right) \]

\[ + \sum_{k=0}^{N} \lambda_k (\omega_1^k - \omega_1^k)^2 (1 - 2\gamma_k)^2 \]  

(4.55)

It is seen from (4.35) that if the error is to decrease at all, one basic requirement should be met, i.e.,

\[ 0 < 1 - 4\gamma \lambda < 1 \quad \text{with} \quad \gamma > 0 \]  

(4.56)

which implies

\[ 0 < \gamma < \frac{1}{4\lambda_{\max}} \]  

(4.57)

\( \lambda_{\max} \) is the largest eigenvalue of the correlation matrix \( R_{\eta} \). Thus \( \gamma = \) constant cannot be set at will if stability of the adaptive loop is to be maintained.

The rate of convergence has been obtained so far only for the algorithm with the availability of a desired signal to generate the real time error function \( \omega(t) \). Now we shall compare the algorithm

\[ \xi_{j+1} = \xi_j + 2\gamma_j \xi_j \eta_j \]

\[ = \xi_j + 2\gamma_j s_j \eta_j - 2\gamma_j z_j \eta_j \]  

(4.58)

where \( s_j \) replaces \( d_j \) for the desired signal with the other two

\[ \xi_{j+1} = \xi_j + 2\gamma_j \xi_j \eta_j - 2\gamma_j z_j \eta_j \]  

(4.59)

\[ \xi_{j+1} = \xi_j + 2\gamma_j \xi_j \eta_j (x_j - z_j) - 2\gamma_j \xi_j \eta_n \]  

(4.60)

when signal or noise correlation functions are used.
Taking mathematical expectation on both sides of (4.58) gives

\[ \bar{c}_{j+1} = \bar{c}_j + 2\gamma_j \bar{s}_j \bar{\eta}_j - 2\gamma_j \bar{z}_j \bar{\Delta}_j \]  

(4.61)

But

\[ \bar{s}_n = E \left\{ \begin{array}{c} s(t) \\ s(t-T) \\ s(t-NT) \end{array} \right\} \]

(4.62)

and

\[ \bar{z}_j \bar{\eta}_j = \bar{n}_j \bar{\Delta}_j = R_n \bar{c}_j \]

(4.63)

we thus have

\[ \bar{c}_{j+1} = (1-2\gamma_j R_n) \bar{c}_j + 2\gamma_j \bar{x}_j \]

(4.64)

Taking the average on both sides of (4.59) gives

\[ \bar{c}_{j+1} = \bar{c}_j + 2\gamma_j \bar{n}_j - 2\gamma_j \bar{z}_j \bar{\Delta}_j \]

which is identical to (4.64) by virtue of (4.62).

Taking the average on both sides of (4.60) gives

\[ \bar{c}_{j+1} = \bar{c}_j + 2\gamma_j \bar{n}_j (\bar{x}_j - \bar{z}_j) - 2\gamma_j R_n \]

(4.65)

But

\[ \bar{n}_j \bar{x}_j - R_n = E \left\{ \begin{array}{c} s(t) + n(t) \\ s(t-T) + n(t-T) \\ s(t-NT) + n(t-NT) \end{array} \right\} - R_n \]

(4.66)

\[ \begin{pmatrix} R_s (o) + R_n (o) \\ R_s (T) + R_n (T) \\ \vdots \\ R_s (NT) + R_n (NT) \end{pmatrix} = \begin{pmatrix} R_s (o) \\ R_s (T) \\ \vdots \\ R_s (NT) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \]

(4.67)
(4.65) can then be reduced to (4.64).

However, (4.64) is just (4.13) if \( d(t) \) is replaced by \( s(t) \). We therefore can conclude that for filtering problem where \( d(t) = s(t) \), the expected values for the gains at any stage are given by the same formula, i.e.,

\[
\omega_{j+1}^{\ast} = (\omega_{j} - \omega^{\ast}) \sum_{k=1}^{j} \left( (1 - 2\gamma) \lambda \right) + \omega^{\ast}
\]  

(4.67)

which is valid for the transformed gain components.

Let us now consider the variation of \( c_{j}^{T} c_{j} \) for the other two cases.

Taking the product of each side with its transpose in (4.59) gives

\[

c_{j+1} = c_{j}^{T} + 2\gamma \sum_{j=1}^{j} R_{s}^{T} - 2\gamma z_{j} c_{j}^{T} n_{j}^{T} \\
+ 2\gamma z_{j} R_{s} c_{j}^{T} + 4\gamma R_{s} z_{j} c_{j}^{T} - 4\gamma z_{j} \left( R_{s} \right)^{T} n_{j}^{T} + 4\gamma z_{j} n_{j}^{T} R_{s}^{T} + 4\gamma z_{j} n_{j}^{T} n_{j}^{T} \\
- 2\gamma z_{j} c_{j}^{T} - 2\gamma z_{j} R_{s} c_{j}^{T} + 4\gamma z_{j} n_{j}^{T} R_{s}^{T} + 4\gamma z_{j} n_{j}^{T} n_{j}^{T}
\]

which is valid for the transformed gain components.

Let us now consider the variation of \( c_{j}^{T} c_{j} \) for the other two cases.

Taking the product of each side with its transpose in (4.59) gives

\[

c_{j+1} = c_{j}^{T} c_{j} + 2\gamma (R_{s} c_{j}^{T} c_{j}) + (R_{s} c_{j}^{T} c_{j} T)
\]

\[
- 2\gamma z_{j} (c_{j}^{T} c_{j} + n_{j}^{T} c_{j}^{T})
\]

\[
- 4\gamma z_{j} (R_{s} c_{j}^{T} c_{j} + n_{j}^{T} R_{s}^{T})
\]

\[
+ 4\gamma z_{j} R_{s} R_{s}^{T} + 4\gamma z_{j} n_{j}^{T} n_{j}^{T}
\]

\[
= c_{j}^{T} c_{j} + 4\gamma (R_{s} c_{j}^{T} c_{j} T) - 4\gamma (R_{s} c_{j}^{T} c_{j} T) s
\]

\[
- 8\gamma (R_{s} c_{j}^{T} c_{j} T) b + 4\gamma z_{j} R_{s} R_{s}^{T} + 4\gamma z_{j} n_{j}^{T} n_{j}^{T}
\]

When average is taken on both sides, we have

\[

c_{j+1}^{T} c_{j+1} = c_{j}^{T} c_{j} + 4\gamma (R_{s} c_{j}^{T} c_{j} T) - 8\gamma (R_{s} c_{j}^{T} c_{j} T) b + 4\gamma z_{j} R_{s} R_{s}^{T} + 4\gamma z_{j} n_{j}^{T} n_{j}^{T} (4.68)
\]
Similarly, if we take the product of each sides with its transpose in (4.58), we have

\[
\begin{align*}
C_{j+1}^T C_{j+1} &= C_j C_j^T + 2\gamma_j (C_j R_j + R_j C_j) \\
&- 2\gamma_j z_j (C_j R_j + R_j C_j) \\
&- 8\gamma_j^2 s_j z_j R_j + 4\gamma_j^2 s_j^2 R_j R_j \\
&+ 4\gamma_j^2 z_j^2 R_j R_j
\end{align*}
\]

(4.69)

We shall see that (4.68) and the average of (4.69) are equivalent by virtue of the following terms.

(1)\[
C_j R_j + R_j C_j \approx C_j R_j + R_j C_j
\]

(2)\[
z_j (R_j R_j + R_j R_j) \approx z_j s_j R_j + z_j R_j s_j
\]

(3)\[
R_j R_j = s_j R_j + R_j s_j
\]

The last expressions are approximately correct if the number of taps is large.

Thus \(C_{j+1}^T C_{j+1}\) either derived from (4.58) or derived from (4.59) are equivalent. Similar steps can be applied to (4.60). In conclusion we can state that the rates of convergence are the same regardless of the choice of algorithms.
V. Adaptive Tapped Delay Line Filters with Time-varying Parameters

The adaptive schemes using the methods of stochastic approximations have been studied for tapped delay line filters. An implicit assumption made so far is that the system under study is time-invariant and all the signals and noise are generated from stationary sources. Although ergodicity of the process has not been required, wide-sense stationarity is implied. If the system itself or the input signals are nonstationary or time-varying, the adjustments made for minimizing certain error criteria may not produce the desired effects. Suppose that the rate of parameter variation is faster than that of convergence, we can never expect to have the algorithm converge at any time. However, if the rate of parameter variation is slow, we can estimate its effects in a qualitative fashion. Let us say that \( \phi(t) \) is a slowly varying time function if the relative change in its value in any interval of length \( \Delta t \) is small; here \( \omega_0 \) is the minimum frequency of the natural oscillation of the system. If the transient behavior is aperiodic for any initial conditions, the function \( \phi(t) \) is said to be slowly varying when its change is small in comparison with the relative change of the output. The term "slowly varying" used throughout this report is defined in the above sense. The statistical properties of the delay line filter will be studied. For stationary and nonstationary input signals the results seem trivial as a delay element does not change any statistical properties at all, but for the time-varying case the method developed gives us some insights about the system.

1. Statistical properties of delay line filters

The statistical properties studied here refer only to the auto correlation functions and variance of the output as a measure of the accuracy of the system.
Some other properties like output distributions, probability density functions, etc. are not considered.

A. Stationary case.

Let us first of all consider a single delay element. The input and output are related by

\[ y(t) = \frac{\sin c}{\omega_0 T} - b \cdot \sin j_0 \]

and the transfer function is given by

\[ H(j \omega) = \frac{\sin c}{\omega_0 T} e^{-j \omega T} \]  \( \text{(5.1)} \)

If \( x(t) \) is a stationary random function with its covariance function given by

\[ R_x(T) = D_x e^{-\alpha |T|} \]  \( \text{(5.2)} \)

The above expression corresponds to a Markov process and its spectral density function is

\[ \phi_x(\omega) = D_x \left( \int_{-\infty}^{\infty} e^{-\alpha |\tau|} e^{-j \omega \tau} d\tau \right) \]

\[ = D_x \left( \int_{0}^{\infty} e^{-\alpha \tau} e^{-j \omega \tau} d\tau + \int_{0}^{\infty} e^{-\alpha \tau} e^{j \omega \tau} d\tau \right) \]

\[ = \frac{2 D_x}{\alpha^2 + \omega^2} \]  \( \text{(5.3)} \)

The output spectral density function is accordingly

\[ \phi_y(\omega) = \phi_x(\omega) \mid H(j \omega) \mid^2 = \frac{2 \alpha D_x}{\alpha^2 + \omega^2} \cdot 2 \frac{c^2}{\alpha^2 + \omega^2} \]  \( \text{(5.4)} \)

D-59
The output variance is then

\[ D_y = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_y(\omega) \, d\omega = \frac{2\alpha c^2}{\pi} D_x \int_{-\infty}^{\infty} \frac{d\omega}{\alpha^2 + \omega^2} \]

\[ = D_x \frac{\alpha c^2}{\pi} \cdot 2\pi \cdot \frac{1}{2\alpha} = \frac{1}{2} c^2 D_x \]  

(5.5)

as we expected.

A tapped delay line filter consists of several delay elements, gain constants, and a summer so that the transfer function of the filter is

\[ H(j\omega) = \sum_{k=0}^{N} c_k e^{-j\omega T_k} \]  

(5.6)

Direct combination of \((5.6)\), \((5.5)\), and \((5.4)\) yields the output spectrum

\[ \Phi_y(\omega) = D_x \sum_{k=0}^{N} c_k e^{-j\omega T_k} e^{-\alpha |T_k - k|} \]  

(5.7)

and the output variance is the Fourier transform of \((5.7)\)

\[ D_y = D_x \sum_{k=0}^{N} |c_k|^2 e^{-\alpha |T_k - k|} \]  

(5.8)

B. Nonstationary case

Suppose that the input \(y\) is a non-stationary time function with the correlation function

\[ R_y(t, t') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_y(\tau) \rho_y(\tau') e^{-\alpha |\tau - \tau'|} \]  

(5.9)

The random function \(y\) may be expressed by

\[ y(t) = \int_{-\infty}^{\infty} \rho_y(\tau) x(t - \tau) \]  

(5.10)

where \(x(t)\) is a stationary random time function with correlation function given by \((5.2)\)

\[ R_x(T) = e^{-\alpha T} \]  

(5.11)
and spectral density function given by \( \phi_{x1}(\omega) = \frac{2 \alpha}{\alpha^2 \omega^2} \) \(^{(5.11)}\)

Assuming that the standard deviation \( \sigma_x(t) \) of \( x(t) \) can be approximated closely enough by the exponential function

\[
\sigma_x(t) = \sigma_0 e^{\mu t} \quad (5.12)
\]

According to Ref. 27, a function \( x(t) \) can be put into an integral expansion of the type

\[
x(t) = m_x(t) + \sum_{r=1}^{N} \int_{-\infty}^{\infty} v_r(\omega) x_r(t, \omega) \, d\omega, \quad (5.13)
\]

where \( m_x(t) \) is the mean value of \( x(t) \),

\( v_r(\omega) \) are uncorrelated white noise

and

\( x_r(t, \omega) \) are coordinate functions defined by

\[
x_r(t, \omega) = \sum_{k=1}^{S} C_{rk}(\omega) e^{(\mu_k + j\omega)t} \quad (5.14)
\]

\( r = 1, 2, \ldots, N \), \( C_{rk}(\omega) \) are coefficients.

While the output can also be represented by an integral canonical expansion like \( x(t) \) with the coordinate functions

\[
Y_r(t, \omega) = \sum_{k=1}^{S} C_{rk}(\omega) H(\mu_k + j\omega) e^{(\mu_k + j\omega)t} \quad (5.15)
\]

\( r = 1, 2, \ldots, N \)

The general formulae for the covariance function and dispersion of the output have also been provided by (p. 209, Ref. 27)

\[
R_y(t, t') = \sum_{r=1}^{N} \frac{1}{2\pi} \int_{-\infty}^{\infty} C_r(\omega) \sum_{k, l=1}^{S} C_{rk} C_{rl}^*(\omega) \cdot H(\mu_k + j\omega) H^*(\mu_l + j\omega) e^{(\mu_k + j\omega)(t-t')} \, d\omega
\]

\( D-61 \)
\[ D_y(t) = \sum_{r=1}^{N} \frac{1}{2\pi} \int_{-\infty}^{\infty} G_r(\omega) \left| \sum_{k=1}^{s} C_{rk}(\omega) \phi(\varphi_k t + j\omega) e^{\mu_k t} \right|^2 d\omega \]  

(5.17)

where \( G_r(\omega) \) are white noise intensities.

In our present case \( N = 1 \), and the coordinate function is
\[ x(t, \omega) = \sigma_0 e^{(\mu + j\omega)t} \]  

(5.18)

and the filter transfer function is modified as
\[ H(\mu + j\omega) = \sum_{i=0}^{N} C_i e^{-j(\mu + \omega)T_i} \]  

(5.19)

Therefore, using (5.16) and (5.17), we have the output correlation function
\[ R_x(t, t') = \sum_{i=0}^{N} \frac{\varphi_0^2 C_i^2}{\pi} e^{(t + t')} \int_{-\infty}^{\infty} e^{\mu_t (t-t')/\sigma^2} d\omega \]

\[ = \sum_{i=0}^{N} C_i^2 \]  

(5.20)

and the variance
\[ D_y(t) = \sum_{i=0}^{N} \frac{\varphi_0^2 C_i^2}{\pi} \int_{-\infty}^{\infty} e^{\mu_t /\sigma^2} d\omega \]

\[ = D_x(t) \sum_{i=0}^{N} C_i^2 \]  

(5.21)

where \( D_x(t) = \sum_{i=0}^{N} \varphi_0^2 = \int_0^\infty e^{\mu_t} \), the variance of \( x(t) \).

C. Time-varying case.

It is a very important case when a linear system is not stationary throughout its total operating time but its behavior is close to it for a comparatively short period. If such a system receives an input which is near to an exponential function, then the output is also near to an exponential function at the end of transient behavior.
The technique used here is confined to the situations where only slow time variation is involved.

Consider a single delay element

\[ Y(t) \xrightarrow{\text{Delay} T} y(t) \]

Since \( y(t) = C x(t-T) \)
therefore,
\[
Y(j\omega) = C X(j\omega) e^{-j\omega T} = e^{+j\omega T} Y(j\omega) = C x(j\omega)
\]

Or in time domain
\[
e^{\mathcal{H}T} y(t) = c x(t)
\]  \( (5.22) \)

where
\[
e^{\mathcal{H}T} = \sum_{j=0}^{\infty} \frac{(-T)^j}{j!} \frac{d^j}{dt^j} \quad \text{with} \quad \mathcal{H} = \frac{d}{dt}
\]  \( (5.23) \)

At a first glance \( e^{\mathcal{H}T} \) appears to be a strange-looking operator.

Actually \( e^{\mathcal{H}T} y(t) = \sum_{i=0}^{\infty} \frac{(-T)^i}{i!} \frac{d^i y(t)}{dt^i} \) is just the Taylor series expansion of \( y(t + T) \) around \( T = 0 \). The role of \( \mathcal{H} \) played here is clear.

Let \( x(t) = e^{\lambda t} \) with \( \lambda = \mu + j\omega \) and \( y(t) = z(t, \lambda) e^{\lambda t} \),
then Eq. (5.22) becomes
\[
\sum_{j=0}^{\infty} \frac{T^j}{j!} \frac{d^j}{dt^j} \left( z(t, \lambda) e^{\lambda t} \right) = c e^{\lambda t}
\]

or
\[
z(t, \lambda) \left( e^{\mathcal{H}T} \right) e^{\lambda t} + c e^{\lambda t} \sum_{j=0}^{\infty} \frac{T^j}{j!} \frac{d^j}{dt^j} \left( z(t, \lambda) \right) = c e^{\lambda t}
\]

or
\[
z(t, \lambda) e^{\lambda T} e^{\lambda t} + e^{\lambda t} \sum_{j=1}^{\infty} \frac{T^j}{j!} \frac{d^j}{dt^j} \left( z(t, \lambda) \right) = c e^{\lambda t}
\]  \( (5.24) \)
We shall develop a procedure to approximate \( z(t, \lambda) \). In the first approximation, we neglect the derivative of the slowly varying function \( z(t, \lambda) \) and obtain

\[
z_1(t, \lambda) e^{\lambda T} = c
\]

thus,

\[
z_1(t, \lambda) = c e^{-\lambda T}
\]

and

\[
y_1(t) = z_1(t, \lambda) e^{\lambda t} = c e^{\lambda (t-T)}
\]  (5.25)

as we expected since \( y(t) = c x(t-T) \).

In the second approximation, the first derivative of the slowly varying function \( z(t, \lambda) \) is taken to be equal to the derivative of \( z_1(t, \lambda) \) from the first approximation

\[
\frac{\partial z(t, \lambda)}{\partial t} = \frac{\partial z_1(t, \lambda)}{\partial t} = \frac{c}{e^{\lambda T}}
\]  (5.26)

where 

\[
\cdot = \frac{dc(t)}{dt}
\]

We then get the following equation for the second approximation

\[
z_2(t, \lambda) e^{\lambda T} + T \frac{\partial z_1(t, \lambda)}{\partial t} = c
\]

or

\[
z_2(t, \lambda) e^{\lambda T} = c - T \cdot c e^{-\lambda T}
\]

\[
z_2(t, \lambda) = \frac{1}{c e^{\lambda T}} (c - T \cdot c e^{-\lambda T})
\]  (5.27)

the \( n \)th approximation of \( z(t, \lambda) \) is

\[
z_n(t, \lambda) = \frac{1}{c e^{\lambda T}} \left( c - \sum_{i=1}^{n} \frac{T^i}{i!} \frac{d^i c(t)}{dt^i} \right)
\]

depending upon the order \( n \) to which the time derivatives of \( c(t) \) exist.
For the actual delay line filter, the output is represented by

\[ y(t) = \frac{1}{e^{-\lambda T}} \sum_{j=0}^{N} \left( c_j - \frac{1}{e^{-\lambda T}} \sum_{i=1}^{n} \frac{d}{dt} c_j(t) \right) \]  

(5.28)

the variance of \( y(t) \) is given respectively by

\[ D_y(t) = \frac{\alpha D_x(t)}{\pi} \int_{-\infty}^{\infty} \frac{|z(t,\mu + j\omega)|^2}{\alpha^2 + \omega^2} \, d\omega \]  

(5.29)

for the nonstationary input signal used for part B.

If \( c \) is time-invariant, then \( z(t,\lambda) = c e^{-\lambda T} \), \( (5.29) \) will be reduced to \( (5.21) \) as it should be. The above formulations are only applicable to asymptotically stable systems, and instants which are sufficiently remote from the initial time \( t \).

2. Adaptive schemes for delay line filters with slowly time-varying parameters.

When the system or input characteristics vary slowly with time, it is found convenient to think in terms of two time scales by using a "fast" time variable \( \hat{t} \) and a "slow" time variable \( \hat{t} \). The ratio of the two scales is a small number so that

\[ \hat{t} = \beta \tilde{t} \]  

(5.30)

The "fast" time variable \( \tilde{t} \) refers to the time variable in which the adaptive system operates while the "slow" time variable \( \hat{t} \) refers to the time variable in which some parameters in the system vary. One example of the latter case is the fluctuation of signal or noise power levels. The levels change but very slowly so that nearly all the techniques developed for time-invariant cases can be applied if additional modifications are made to account for the effect of slow variation. For the adaptive delay line filters under study, if we know the forms of the signal or noise correlation functions (even they are changing very slowly),
the schemes described previously can readily be used. However, if we do not
know exactly how the correlation functions change (but we know the cause of
variation, for example, sinusoidal or exponential amplitude modulation, frequency
modulation, etc.) then we can assume that the weight parameters are functions
of slow time variables and leave the unknown fluctuations untouched. In what
follows the method of two time variables is described and then applied to
the tapped delay line filters with slowly time-varying parameters.
A. Two time variable method

Consider a linear system whose input \( x(t) \) and output \( e(t) \) are
related by the differential equation
\[
Le(\hat{t}, \hat{t}) = Mx(\hat{t}, \hat{t}) \tag{5.31}
\]
In the above equation \( L \) and \( M \) are linear differential operators and can be
written in the form
\[
L = L(c, \hat{t}, \hat{t}) = \sum_{i=0}^{n} a_i(c, \hat{t}) \hat{P}^i \tag{5.32}
\]
\[
M = M(c, \hat{t}, \hat{t}) = \sum_{j=0}^{m} b_j(c, \hat{t}) \hat{P}^j \tag{5.33}
\]
where the symbol \( \hat{P} \) denotes the total derivative with respect to time and
is defined by
\[
\hat{P} = \frac{d}{dt} = \frac{\partial}{\partial \hat{t}} + \beta \frac{\partial}{\partial \hat{t}} = \hat{\beta} + \beta \hat{\beta} \tag{5.34}
\]
Noting that the adaptive parameter is a slowly varying function of the
slow time variable,
\[
c = c(\hat{t}) \tag{5.35}
\]
We can now expand the operations \( L \) and \( M \) in Taylor series about \( \beta = 0 \).
\[ L(c, \hat{t}, \hat{p}) = L(c, \hat{t}, \hat{p} + \beta \hat{p}) \]
\[ = L(c, \hat{t}, \hat{p}) + \frac{\partial L}{\partial p} \bigg|_{\hat{p}=0} (\beta \hat{p}) + \frac{1}{2} \frac{\partial^2 L}{\partial p^2} \bigg|_{\hat{p}=0} (\beta \hat{p})^2 + \ldots \]  
\( (5.36) \)

\[ M(c, \hat{t}, \hat{p}) = L(c, \hat{t}, \hat{p} + \beta \hat{p}) \]
\[ = M(c, \hat{t}, \hat{p}) + \frac{\partial M}{\partial p} \bigg|_{\hat{p}=0} (\beta \hat{p}) + \frac{1}{2} \frac{\partial^2 M}{\partial p^2} \bigg|_{\hat{p}=0} (\beta \hat{p})^2 + \ldots \]  
\( (5.37) \)

Since \( L \) terminates at the power \( n \) and \( M \) terminates at the power \( m \), \( (5.36) \) and \( (5.37) \) can be rewritten as

\[ L = \sum_{j=0}^{n} L_j (\beta \hat{p})^j \]  
\( (5.38) \)

\[ M = \sum_{i=0}^{m} M_i (\beta \hat{p})^i \]  
\( (5.39) \)

where

\[ L_j = \frac{1}{j!} \frac{\partial^j L}{\partial \hat{p}^j} \bigg|_{\hat{p}=0} = \frac{1}{j!} \frac{\partial}{\partial \hat{p}} \frac{\partial^{j-1} L}{\partial \hat{p}^{j-1}} = \frac{1}{j!} \frac{\partial^{j+1} \hat{L}_0}{\partial \hat{p}^{j+1}} \]  
\( (5.40) \)

\[ M_i = \frac{1}{i!} \frac{\partial^i M}{\partial \hat{p}^i} \bigg|_{\hat{p}=0} = \frac{1}{i!} \frac{\partial}{\partial \hat{p}} \frac{\partial^{i-1} M}{\partial \hat{p}^{i-1}} = \frac{1}{i!} \frac{\partial^{i+1} \hat{M}_0}{\partial \hat{p}^{i+1}} \]  
\( (5.41) \)

The solution can be expanded in the form

\[ \varepsilon(\hat{t}, \hat{\gamma}) = \sum_{j=0}^{N} \beta^j \varepsilon_j \]  
\( (5.42) \)

Substituting the expression of \( \varepsilon(\hat{t}, \hat{\gamma}) \) into \( (5.31) \) and equating the terms with the same power of \( \beta \), we have

\[ \sum_{j=0}^{n} L_j (\beta \hat{p})^j \sum_{i=0}^{N} \beta^i \varepsilon_i = \sum_{j=0}^{m} M_j (\beta \hat{p})^j \sum_{i=0}^{N} \beta^i \varepsilon_i \]

\( D=6' \)  
\( (5.43) \)
thus, for \( j = 0 \),
\[
I_0 \, e_0 = M_0 \, x \quad (5.44)
\]
\[
j = 1 \quad L_0 \, e_1 = -L_1 \, \hat{p} \, e_0 + M_1 \, \hat{p} \, x \quad (5.45)
\]
Consequently,
\[
e_o = \frac{M_o}{L_o} \, x \quad (5.46)
\]
\[
e_1 = -\frac{1}{L_o} \left( L_1 \, \hat{p} \left( \frac{M_o}{L_o} \, x \right) - M_1 \, \hat{p} \, x \right) \quad (5.47)
\]
\( e_o \) is just the solution for time-invariant case while \( e_1 \) with \( i \geq 1 \) are additional terms as a result of slow time variation.

The differentiation with respect to \( \hat{t} \) implied by \( \hat{p} \) can be carried out explicitly.
\[
\hat{p} \left( \frac{M_o}{L_o} \, x \right) = \frac{1}{L_o} \left( \frac{\partial M_o}{\partial \hat{t}} \right) c = \text{const} + \frac{\partial M_o}{\partial c} \left( \frac{\partial c(t)}{\partial \hat{t}} \right) x
\]
\[
- \frac{M_o}{L_o^2} \left( \frac{\partial L_o}{\partial \hat{t}} \right) c = \text{const} + \frac{\partial L_o}{\partial c} \left( \frac{\partial c(t)}{\partial \hat{t}} \right) x
\]
\[
+ \frac{M_o}{L_o} \frac{\partial x}{\partial \hat{t}} \quad (5.48)
\]
If the variation in \( c \) is slowly enough such that \( \frac{\partial^2 c(t)}{\partial \hat{t}^2} \approx 0 \),
then the solution has the form
\[
e(\hat{t}, \tilde{t}) \approx e_o + \beta \, e_1 \quad (5.49)
\]
where \( e_1 \) is obtained by combining \( (5.47) \) and \( (5.48) \)
\[
e_1 = -\frac{1}{L_o} \left( L_1 \, \hat{p} \left( \frac{M_o}{L_o} \, x \right) - M_1 \, \hat{p} \, x \right)
\]
\[
= -\frac{1}{L_o} \left( L_1 \, \hat{p} \, e_0 - M_1 \, \hat{p} \, x \right)
\]
\[
= - \frac{\partial c(t)}{\partial \hat{t}} \frac{L_1}{L_o} \frac{\partial e_0}{\partial c} + \frac{\partial c(t)}{\partial \hat{t}} \frac{M_1}{L_c} \frac{\partial x}{\partial c} \quad (5.50)
\]
From Eq. (5.49), the mean square of \( e \) is approximately

\[
\overline{e^2} = \overline{e_0^2} + 2 \beta \overline{e_0 e_1} + \cdots
\]  
\( (5.51) \)

B. Applications on adaptive filters

Let us now turn our attention to the tapped delay line filter. As has been shown previously, the filter output is

\[
z(t) = \sum_{i=0}^{N} c_i x(t - T_i)
\]  
\( (5.52) \)

Using Eq. (5.22), we can express Eq. (5.52) as

\[
\mathcal{Z}(\omega) = \sum_{i=0}^{N} c_i X(\omega) e^{-j\omega T_i} = X(\omega) \sum_{i=0}^{N} c_i e^{-j\omega T_i}
\]

or in time domain

\[
\frac{1}{\sum c_i e^{-\beta T_i}} z(t) = x(t)
\]  
\( (5.53) \)

Comparing with (5.31), we see that

\[
L = \frac{1}{\sum_{i=0}^{N} c_i e^{-\beta T_i}}
\]  
\( (5.54) \)

\[M = 1, \text{ unity operator}\]

the operator \( L_1 \) is

\[
L_1 = \frac{\partial L}{\partial \beta} = \frac{\partial}{\partial \beta} \left( \frac{1}{\sum_{i=0}^{N} c_i e^{-\beta T_i}} \right)
\]  
\( (5.55) \)
and
\[
\frac{L_1}{L} = \frac{\sum_{i=1}^{N} T_i c_i e^{-\frac{\hat{p}}{\dot{p} T_i}}}{\sum_{i=0}^{N} c_i e^{-\frac{\hat{p}}{\dot{p} T_i}}} \triangleq T_{av}
\] (5.55)

\(T_{av}\) defined by (5.56) can be thought of as the average delay time of the filter.

Following (5.49) the filter output is
\[
z(t) \approx z_o(\hat{t}) + \beta z_1(\tilde{t}, \hat{t})
\]
with
\[
z_1(\tilde{t}, \hat{t}) = -\sum_{i=0}^{N} \frac{\partial c_i(\hat{t})}{\partial \hat{t}} \left( \frac{\partial L/\partial p}{L} \right) \frac{\partial z_o(\tilde{t})}{\partial c_i}
\]
\[
= -\sum_{i=0}^{N} \frac{\partial c_i(\hat{t})}{\partial \hat{t}} (T_{av}) \frac{\partial}{\partial c_i} \left( \sum_{i=0}^{N} c_i x(t - T_i) \right)
\]
\[
= -T_{av} \sum_{i=0}^{N} \frac{\partial c_i(\hat{t})}{\partial \hat{t}} x(t - T_i) = -T_{av} \eta(\tilde{t})^T \delta
\] (5.57)

where
\[
\eta(\tilde{t})^T = \begin{pmatrix}
    x(t) \\
    x(t - T) \\
    x(t - T_N)
\end{pmatrix}
\]
and
\[
\delta = \begin{pmatrix}
    \delta_0 \\
    \delta_1 \\
    \delta_N
\end{pmatrix} = \begin{pmatrix}
    \frac{\partial c_o(\hat{t})}{\partial \hat{t}} \\
    \frac{\partial c_1(\hat{t})}{\partial \hat{t}} \\
    \frac{\partial c_N(\hat{t})}{\partial \hat{t}}
\end{pmatrix}
\] (5.58)

If the desired signal \(d(t)\) is not slowly time-varying, the error function is then
\[
e(t) = d(t) - z(t)
\]
\[
= d(t) - z_o(t) - \beta z_1(\tilde{t}, \hat{t})
\]
\[
= e_o + \beta e_1
\] (5.59)
The mean square error is approximately
\[
\frac{e^2(t)}{e_0} \cong \frac{e_1^2}{e_0^2} + 2 \beta \frac{e_1}{e_0} e_1 .
\] (5.60)

Or
\[
e^2(t) = E \left\{ \left( d(t) - \sum_{i=0}^{N} c_i \eta_i(t) \right)^2 \right\} + 2 \beta E \left\{ \left( d - \sum_{i=0}^{N} c_i \eta_i(t) \right) (T \eta(t) \delta) \right\}
\]
\[
\quad = E \left\{ \left( d(t) - \sum_{i=0}^{N} c_i \eta_i(t) \right)^2 \right\} + 2 \beta T_{av} E \left\{ \left( d(t) - \sum_{i=0}^{N} c_i \eta_i(t) \right) (T \eta(t) \delta) \right\} \] (5.61)

Taking the partial derivative of \( e^2 \) with respect to \( c_1 \), we have
\[
\frac{\partial e^2}{\partial c_1} = 2 E \left\{ \left( d(t) - \sum_{i=0}^{N} c_i \eta_i(t) \right) \left( -\eta_1(t) \right) \right\} - 2 \beta T_{av} E \left\{ \eta_1(t) \eta(t)^T \delta \right\}
\]

Thus the gradient of \( e^2 \) is
\[
E \left\{ \nabla_c Q(\mathbf{x} | \mathbf{z}) \right\} = -2 E \left\{ \eta(t)^T \left( d(t) - \eta(t)^T \right) c \right\} - 2 \beta T_{av} E \left\{ \eta(t) \eta(t)^T \delta \right\}
\]
or
\[
\nabla_c Q(\mathbf{z} | \mathbf{e}) = -2 \eta(t)^T \left( d(t) - \eta(t)^T \right) c
\]
\[
\quad - 2 \beta T_{av} \eta(t) \eta(t)^T \delta \] (5.62)
Using the adaptive scheme

\[ \mathcal{C}_{j+1} = \mathcal{C}_j - \mathcal{Y}_j \nabla_{\mathcal{C}} Q(\mathcal{C} | \mathcal{C}) \]

we obtain a new scheme for the time-varying case

\[ \mathcal{C}_{j+1} = \mathcal{C}_j + 2 \gamma_j \mathbf{d}_j \eta_j - 2 \gamma_j \mathbf{v}_j \eta_j \mathbf{v}_j^T - 1
\]

\[ + 2 \gamma_j \beta_j \mathbf{T}_{av} \eta_j \mathbf{v}_j^T \delta \]

\[ = \mathcal{C}_j + 2 \gamma_j \mathbf{d}_j \eta_j + 2 \gamma_j \beta_j \mathbf{T}_{av} \eta_j \mathbf{v}_j^T \delta \] (5.63)

Algorithm (5.63) can readily be implemented by

Thus the increment for the time-varying parameter \( \mathcal{C}_j \) is obtained.

Similar schemes can be obtained through proper transformations, Eqs. (4.42) and (4.46), for the cases where only the signal or the noise correlation functions are assumed to be known. The results are

\[ \mathcal{C}_{j+1} = \mathcal{C}_j + 2 \gamma_j \beta_j \mathbf{T}_{av} \eta_j \eta_j^T \delta
\]

\[ - 2 \gamma_j \eta_j \mathbf{z}_j + 2 \gamma_j \eta_j \mathbf{x}_j - \mathbf{R}_n \] (5.64)

when \( \mathbf{R}_n(\mathbf{T}) \) is known, and

\[ \mathcal{C}_{j+1} = \mathcal{C}_j + 2 \gamma_j \beta_j \mathbf{T}_{av} \eta_j \eta_j^T \delta
\]

\[ - 2 \gamma_j \eta_j \mathbf{z}_j + 2 \gamma_j \mathbf{R}_s \] (5.65)

when \( \mathbf{R}_s(\eta) \) is known.
3. Effects of slow time variation on minimum mean square error.

If we set \( V e^2 = 0 \), we see from Eq. (5.61) that

\[
\begin{align*}
2 \eta^T c^* &= 2 \eta^T d + 2 \beta T_{av} \eta^T c^* \\
R_{\eta} c^* &= d \eta + \beta T_{av} R_{\eta} c^* \\
or c^* &= \left( R_{\eta} \right)^{-1} \left( d \eta + \beta T_{av} c^* \right)
\end{align*}
\]  
(5.66)

Substituting the expression of \( c^* \) for \( c \) into (5.61) and using (3.24 b), we get the minimum mean square error

\[
\begin{align*}
e_{min}^2 &= d^2 - d \eta^T c^* \\
\bar{e}_{min}^2 &= d^2 - d \eta^T R_{\eta}^{-1} d \eta - \beta T_{av} d \eta^T c^* \\
\bar{e}_{min}^2 &= d^2 - d \eta^T R_{\eta}^{-1} d \eta - \beta T_{av} d \eta^T \left( R_{\eta}^{-1} d \eta + \beta T_{av} c^* \right) \\
&+ 2 \beta T_{av} c^* d \eta \\
\bar{e}_{min}^2 &= d^2 - d \eta^T R_{\eta}^{-1} d \eta - \beta T_{av} d \eta^T c^* \\
&+ 2 \beta T_{av} c^* d \eta \\
&\leq \bar{e}_{o_{min}}^2 + \beta T_{av} d \eta^T c^* \\
&+ 2 \beta T_{av} c^* d \eta
\end{align*}
\]  
(5.67)

where \( \bar{e}_{o_{min}}^2 = d^2 - d \eta^T R_{\eta}^{-1} d \eta \) is the minimum mean square error of the time-invariant filter as derived in (3.24 a).

The expressions \( \bar{e}_{min}^2 \) for other cases (known \( R_{\eta} (\tau) \)) or \( R_S (\tau) \) as well as the effect of slow time variation on the rate of convergence can be obtained in a straightforward manner.
Appendix A

Proof of the modified algorithm

It was mentioned in chapter IV that algorithm (4.7) and (4.11) were derived from the formula

\[ C_{j+1} = C_j - \gamma_j (Q_1 + Q_2) \]  

(A.1)

rather than from

\[ C_{j+1} = C_j - \gamma_j (\overline{Q}_1 + \overline{Q}_2) \]

(A.2)

where \( Q_1 + Q_2 = Q \) is a function of error, and the average of it is the performance criterion to be minimized.

Comparing (A.1) and (A.2) with the regular gradient method with constant \( \gamma \)

\[ C_{j+1} = C_j - \gamma (\overline{Q}_1 + \overline{Q}_2) \]

(A.3)

we see that in (A.2) no average is taken while in (A.1) partial average is taken. The error \( \xi \) caused by measurement

\[ \overline{Q} = Q + \xi \]

is eliminated by the properly chosen sequence \( \{\gamma_j\} \). Intuitively speaking, the same \( \{\gamma_j\} \) which eliminates the error caused by \( (\overline{Q}_1 + \overline{Q}_2) \) minus \( (Q_1 + Q_2) \) can definitely eliminate that caused by \( (\overline{Q}_1 + \overline{Q}_2) \) minus \( (Q_1 + Q_2) \). This stems from the fact that the measuring noise in the second case is smaller on the average than in the first case. Although intuition does not generally warrant mathematical correctness, we can state with mathematical rigour that either signal or noise statistical properties will suffice to generate the error gradient used in the adaptive schemes. The physical conditions under which these algorithms converge remain unchanged. Two lemmas and one theorem will be proved in sequence.
Lemma 1,

For the tapped delay line filters considered in Chapter 4, if

(a) \( Q = Q_1 + Q_2 \)
(b) \( Q(c) = e^2 \)
(c) \( \forall Q_2 \) is independent of \( c \),

then at the neighborhood of \( c^* \), which minimizes \( T(c) = \cdots \), the following statement is true:

\[ \inf_{\varepsilon < |c - c^*| < \frac{1}{\varepsilon}} E \left\{ (c - c^*)^T (VQ_1 + VQ_2) \right\} > 0 \]  \( (A.4) \)

\[ \varepsilon > 0 \]

Proof: If \( Q = Q_1 + Q_2 \) has a minimum at \( c = c^* \), then

\[ \frac{\delta (Q_1 + Q_2)}{\delta c_1} > 0 \quad \text{for} \quad c_1 > c_1^* \]
\[ = 0 \quad \text{for} \quad c_1 = c_1^* \]
\[ < 0 \quad \text{for} \quad c_1 < c_1^* \]  \( (A.5) \)

thus

\[ \frac{\delta (Q_1 + Q_2)}{\delta c_2} \]  \( c_1 - c_1^* \) \( (c_1 - c_1^*) \frac{\delta (Q_1 + Q_2)}{\delta c_2} \) \( > 0 \quad \text{for all} \quad i \]  \( (A.6) \)

and

\[ \inf_{\varepsilon < |c - c^*| < \frac{1}{\varepsilon}} E \left\{ (c - c^*)^T (VQ_1 + VQ_2) \right\} > 0 \]  \( (A.7) \)

\[ \varepsilon > 0 \]

Since \( VQ_2 \) is independent of \( c \), we have

\[ E \left\{ (c - c^*)^T (VQ_1 + VQ_2) \right\} \]
\[ = E \left\{ (c - c^*)^T VQ_1 \right\} + E \left\{ (c - c^*)^T VQ_2 \right\} \]
\[ = E \left\{ (c - c^*)^T WQ_1 \right\} + E \left\{ (c - c^*)^T E VQ_2 \right\} \]
\[ = E \left\{ (c - c^*)^T (VQ_1 + VQ_2) \right\} \]  \( (A.8) \)

Therefore, by virtue of \( (A.7) \)

\[ \inf_{\varepsilon < |c - c^*| < \frac{1}{\varepsilon}} E \left\{ (c - c^*)^T (VQ_1 + VQ_2) \right\} > 0 \]  \( (A.9) \)
Lemma 2

If \( Q = Q_1 + Q_2 \), \( Q(e) = e^2 \), \( \forall Q \) is independent of \( e \), and

(a) \( \frac{\partial^2 Q_1}{\partial e^2} \) exists and is uniformly bounded

(b) \( s(t) \) and \( n(t) \) are uniformly bounded,

then for the tapped delay line filter

\[
E \left\{ (\bar{Q}_1 + \bar{Q}_2)^T (\bar{Q}_1 + \bar{Q}_2) \right\} < d(c^* + c^T c)
\]

\( d > 0 \) (A.10)

Proof:

Using a Taylor expansion about \( e = e^* \), we have

\[
\frac{\partial Q_1}{\partial c_j} = \frac{\partial Q_1}{\partial c_j} \bigg|_{c = c^*} + \sum_{i=0}^{N} (c_i - c_i^*) \frac{\partial^2 Q_1}{\partial c_i \partial c_j} \bigg|_{c = c^*}
\]

for arbitrary \( j \), and \( j = 0, 1, 2, \ldots, N \).

Since \( e(t) = s(t) - z(t) \)

\[
= \sum_{i=0}^{N} c_i n_i(t)
\]

we see that

\[
\frac{\partial Q_1(e)}{\partial e} = \frac{\partial Q_1}{\partial e} = \frac{\partial Q_1}{\partial e} \cdot [-n_1(t)]
\]

(A.13)

and

\[
\frac{\partial^2 Q_1(e)}{\partial c_i \partial c_j} = \frac{\partial^2 Q_1}{\partial e^2} \cdot n_i(t) \cdot n_j(t)
\]

\[
= \frac{\partial^2 Q_1}{\partial e^2} \cdot [s(t - T_1) + n(t - T_1)] \cdot [s(t - T_j) + n(t - T_j)]
\]

(A.14)
Thus
\[
\frac{\Delta^2 Q_1}{\Delta c_i \Delta c_j} \quad \text{is bounded if (a) and (b) are satisfied.}
\]

Therefore, from (A.11)
\[
V_c Q_1 \leq V_c Q_1 \bigg|_{c = c^*} + k_1 \sum_{i=0}^{N} (c_i - c_i^*)
\]

where
\[
k_1 = \sup_{\text{all } i} \left| \frac{\Delta^2 Q_1}{\Delta c_i \Delta c_j} \right|
\]

As \( Q_2 \) contains \( c \) only in the first order and \( V_c Q_2 \) is independent of \( c \), we can write
\[
\overline{V_c Q_2} = \overline{V_c Q_2} \bigg|_{c = c^*}
\]

and
\[
V Q_1 + \overline{V Q_2} \leq V Q_1 \bigg|_{c = c^*} + \overline{V Q_2} \bigg|_{c = c^*} + k_1 \sum_{i=0}^{N} (c_i - c_i^*)
\]

Note
\[
E \left( V Q_1 \bigg|_{c = c^*} + \overline{V Q_2} \right) = E \left( \overline{V Q_1} + \overline{V Q_2} \bigg|_{c = c^*} \right)
\]

= \( E \left( \overline{V Q_1} + \overline{V Q_2} \bigg|_{c = c^*} \right) = 0 \)

(A.19)

Taking mathematical expectation on both sides of (A.17) gives
\[
E \left( \overline{V Q_1} + \overline{V Q_2} \bigg|_{c = c^*} \right) \leq k_1 \sum_{i=0}^{N} (c_i - c_i^*)
\]

(A.20)

Lemma 2 is obtained by taking the inner product of (A.20)
\[
E \left( \overline{V Q_1} + \overline{V Q_2} \bigg|_{c = c^*} \right) \overline{V Q_1} + \overline{V Q_2}
\]

\[
\leq k_1^2 N \sum_{i=0}^{N} (c_i - c_i^*) (c_j - c_j^*)
\]

\[
= d(c^T c^* + c^T c)
\]

(A.21)
Theorem

Let \( \gamma_1, \gamma_2, \ldots \) be a sequence of positive numbers such that

\[
\begin{align*}
(A1) & \quad \lim_{j \to \infty} \gamma_j = 0 \\
(A2) & \quad \sum_{j=0}^{\infty} \gamma_j = \infty \\
(A3) & \quad \sum_{j=1}^{\infty} \gamma_j < \infty
\end{align*}
\]

Let the following conditions be satisfied

\[
\begin{align*}
(B) & \quad \inf \left\{ \epsilon : |c - c^*| < \frac{1}{\epsilon}, \epsilon > 0 \right\} > 0 \\
(C) & \quad E \left\{ \left( V_{c+1} + \bar{V}_{c+2} \right)^T \left( V_{c+1} + \bar{V}_{c+2} \right) \right\} < d \left( c^* + c^* \right),
\end{align*}
\]

\( d > 0 \), for all \( c \) in a bounded set.

Then the algorithm

\[
\begin{align*}
\epsilon_{j+1} = \epsilon_j - \gamma_j \left( V_{c+1} + \bar{V}_{c+2} \right)
\end{align*}
\]

which minimizes the performance criterion

\[
I(c) = E \left\{ Q(e) \right\} = E \left\{ Q_1(c) + Q_2(c) \right\}
\]

converges with probability and to \( c^* \).

Proof:

Substratting both sides of (A.26) by \( c^* \), we have

\[
\begin{align*}
\epsilon_{j+1} - c^* = \epsilon_j - c^* - \gamma_j \left( V_{c+1} + \bar{V}_{c+2} \right)
\end{align*}
\]

taking the inner product on both sides of (A.23)

\[
\begin{align*}
\left( \epsilon_{j+1} - c^\ast \right)^T \left( \epsilon_{j+1} - c^\ast \right) & = \left( \epsilon_j - c^\ast \right)^T \left( \epsilon_j - c^\ast \right) - 2 \gamma_j \left( \epsilon_j - c^\ast \right)^T \left( V_{c+1} + \bar{V}_{c+2} \right) \\
& \quad + \gamma_j^2 \left( V_{c+1} + \bar{V}_{c+2} \right)^T \left( V_{c+1} + \bar{V}_{c+2} \right)
\end{align*}
\]
and taking the conditional mathematical expectation for given \( c_1, c_2, \ldots, c_j \), we obtain
\[
E \left\{ \left| c_{j+1} - c^* \right|^2 \mid c_1, c_2, \ldots, c_j \right\} = \left| c_j - c^* \right|^2 - 2\gamma_j E \left\{ (c_j - c^*)^T (\overline{vQ_1} + \overline{vQ_2}) \right\} + \gamma_j^2 E \left\{ (\overline{vQ_1} + \overline{vQ_2})^T (\overline{vQ_1} + \overline{vQ_2}) \right\}
\] (A.30)

From condition (c), (A.30) becomes
\[
E \left\{ \left| c_{j+1} - c^* \right|^2 \mid c_1, \ldots, c_j \right\} < \left| c_j - c^* \right|^2 - 2\gamma_j E \left\{ (c_j - c^*)^T (\overline{vQ_1} + \overline{vQ_2}) \right\} + \gamma_j^2 d(c^T c^* + c_j^T c_j)
\] (A.31)

Using condition (B), (A.31) is reduced to
\[
E \left\{ \left| c_{j+1} - c^* \right|^2 \mid c_1, \ldots, c_n \right\} < \left| c_j - c^* \right|^2 (1 + 2\gamma_j^2 d) + 2\gamma_j^2 d c^T c^*
\] (A.31a)

Using condition (B), we can reduce (A.31a) to
\[
E \left\{ \left| c_{j+1} - c^* \right|^2 \mid c_1, \ldots, c_n \right\} < \left| c_j - c^* \right|^2 (1 + 2\gamma_j^2 d) + 2\gamma_j^2 d c^T c^*
\] (A.32)

Let
\[
\tilde{z}_j = \left| c_j - c^* \right|^2 \sum_{k=j}^{\infty} (1 + \gamma_k^2 d) + \sum_{k=j}^{\infty} 2d \gamma_k^2 c_k^T c^* \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d)
\] (A.33)

Then
\[
\tilde{z}_{j+1} = \left| c_{j+1} - c^* \right|^2 \sum_{k=j+1}^{\infty} (1 + \gamma_k^2 d) + \sum_{k=j+1}^{\infty} 2d \gamma_k^2 c_k^T c^* \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d)
\] (A.34)
Taking the conditional mathematical expectation for given $c_1, c_2, \ldots, c_j$, we have

\[
E \left\{ Z_{j+1} \mid c_1, \ldots, c_n \right\} = E \left\{ \left| c_{j+1} - c^* \right|^2 \mid c_1, \ldots, c_n \right\} \sum_{k=j+1}^{\infty} (1 + \gamma_k^2 d) + \sum_{k=j+1}^{\infty} 2d \gamma_k^2 c^T c^* \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d) \\
\leq \left[ \left| c_j - c^* \right|^2 (1 + d \gamma_1^2) + 2 \gamma_j^2 d c^T c^* \right] \sum_{k=j+1}^{\infty} (1 + \gamma_k^2 d) + \sum_{k=j+1}^{\infty} 2d \gamma_k^2 c^T c^* \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d) \\
= Z_j
\]

or

\[
E \left\{ Z_{j+1} \mid c_1, \ldots, c_n \right\} \leq Z_j
\]  

(A.35)

Next taking the conditional mathematical expectation for given $Z_1, \ldots, Z_j$ on both sides of (A.35) we have

\[
E \left\{ Z_{j+1} \mid Z_1, \ldots, Z_j \right\} \leq Z_j
\]  

(A.36)

Inequality (A.36) shows that $Z_j$ is a semi-martingale, where

\[
E \frac{Z_j}{Z_{j+1}} < E \frac{Z_j}{Z_j} < \cdots < E \frac{Z_1}{Z_1}
\]  

(A.37)

so that, according to the theory of semi-martingales, the sequence $Z_j$ converges with probability one, and hence by virtue of (A.33) and (A.23c) the sequence $(c_j - c^*)$ also converges with probability one to some random number $\xi$. It remains to show that $P(\xi = 0) = 1$.

It is seen that from (A.37), (A.35) and (A.23c) the sequence $E(c_j - c^*)$ is bounded. Now taking the mathematical expectation on both side of the inequality (A.32),

\[
E \left\{ \left| c_{j+1} - c^* \right|^2 \right\} < E \left\{ \left| c_j - c^* \right|^2 \right\} - 2\gamma_j E \left\{ (c_j - c^*)^T \xi \right\} + \gamma_j^2 d^2 [c^* c^* + E(c_j^T c_j)]
\]

D-80
and adding the first \( j \) inequalities together, we have by deduction

\[
E\left( \left| c_{j+1} - c^* \right|^2 \right) \leq E\left( \left| c_1 - c^* \right|^2 \right) + \sum_{k=1}^{j} \left( c_k^T c_k^* + d \gamma_k^2 E(c_j^T c) \right) - \sum_{k=1}^{j} 2 \gamma_k E\left( (c_j - c^*)^T c_q \right) \tag{A.38}
\]

Since \( E\left( \left| c_j - c^* \right|^2 \right) \) is bounded and condition (A.23c) is fulfilled, from Eq. (A.38) it follows that

\[
\sum_{k=1}^{\infty} \gamma_k E\left( (c_j - c^*)^T c_q \right) < \infty \tag{A.39}
\]

Using condition (A.23b), i.e., \( \sum_{j=1}^{\infty} \gamma_j = \infty \)

and noting (A.24)

\[
\varepsilon < \inf \left( |c - c^*| \right) < \frac{1}{\varepsilon} \tag{A.40}
\]

We deduce from (A.39) that

\[
E\left( (c_N - c^*)^T c_q \right) \to 0 \quad \text{with probability one for some sequence } N \tag{A.40-}
\]

Now taking \( E\left( |c_j - c^*| \right) \to 0 \) with probability 1, and comparing (A.40) with (A.24), we obtain

\[
\xi = c \quad \text{with probability 1} \tag{A.41}
\]

Therefore, algorithm (A.26) converges w.h.p. probability one

\[
P\left\{ \lim_{j \to \infty} (c_j - c^*) = 0 \right\} = 1 \tag{A.42}
\]

as well as in mean square sense, i.e.,

\[
\lim_{j \to \infty} E\left( \left| c_j - c^* \right|^2 \right) = 0 \tag{A.43}
\]
Appendix B

Some properties of Gamma function

Since $\Gamma(a + n) = (a + n - 1) \Gamma(a + n - 1)$

\[= (a + n - 1) (a + n - 2) \Gamma(a + n - 2)\]

\[= \ldots \ldots \]

\[= (a + n - 1) (a + n - 2) \ldots \alpha \Gamma(a)\]

We have

\[\sum_{k=1}^{n} (\alpha + k - 1) = \alpha (\alpha + 1) \ldots (\alpha + n - 1)\]

\[= \frac{\Gamma(a + n)}{\Gamma(a)} \quad (B.1)\]

Thus Eq. (4.29) becomes

\[\sum_{k=1}^{j} \left(1 - \frac{\lambda}{j+1}\right) = \sum_{k=1}^{j} \left(\frac{\lambda}{j+1} - \lambda\right) = \frac{\Gamma(j+2 - \lambda)}{(j+1)! \Gamma(2 - \lambda)} \quad (B.2)\]

Eq. (B.2) can be approximated by using the formula*

\[\Gamma(x) = \frac{1}{x} \left(x^\frac{-1}{2} \left(\frac{2\pi}{x}\right)^\frac{1}{2}\right) \left\{1 + \frac{1}{12x} + \frac{1}{288x^2} - \frac{139}{51840x^3}\right\} + \frac{571}{2488320x^4} \quad (B.3)\]

\[\approx e^{-x} x^{\frac{1}{2}} \left(\frac{2\pi}{x}\right)^\frac{1}{2} \quad \text{for } x \gg 1\]

From Eq. (B.3) we can write for $j \gg 1$,

\[\Gamma(j+2 - \lambda) = e^{-(j+2-\lambda)} (j+2-a) \frac{1}{2} \frac{1}{(2\pi)^2} \]

\[\approx e^{-(j+2-\lambda)} (j+2-a) \frac{3}{2} \frac{1}{(2\pi)^2} \quad (B.4)\]

* Whittaker and Watson, Modern Analysis, p. 253

D-82
\[ (j+1)! = \Gamma(j+2) = e^{-(j+2)} (j+2)^{\frac{3}{2}} \sqrt{2\pi} \] (B.5)

Since
\[ j + 2 - \alpha \approx j + 2 \text{ if } j \gg \alpha \]
we obtain from (B.4) and (B.5)
\[
\frac{\Gamma(j+2-\alpha)}{(j+1)!} \approx \frac{(j+2-\alpha)^{-\alpha}}{\Gamma(j+2)} = \frac{1}{(j+1)^\alpha} \text{ if } j \gg 1 \text{ and } j \gg \alpha
\] (B.6)

Therefore, combining (B.2) and (B.6) gives
\[
\frac{1}{\Gamma(2-\lambda)(j+1)^\lambda} \sum_{k=1}^{j} \left( 1 - \frac{\lambda}{j+1} \right) \approx \frac{1}{\Gamma(2-\lambda)(j+1)^\lambda}
\] (B.7)

and furthermore,
\[
\sum_{j=1}^{m} \left( 1 - \frac{A}{j+1} \right) \approx \frac{m^A}{(n+1)^A}
\] (B.8)
REFERENCES


