INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
KOLESZAR, George Edmund, 1941-
THE STOCHASTIC PROPERTIES OF THE WEIGHTS
IN AN ADAPTIVE ANTENNA ARRAY.

The Ohio State University, Ph.D., 1975
Engineering, electronics and electrical

Xerox University Microfilms, Ann Arbor, Michigan 48106
THE STOCHASTIC PROPERTIES OF THE WEIGHTS IN AN ADAPTIVE ANTENNA ARRAY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

George Edmund Koleszar, B.S.E.E., M.S.E.E.

*****

The Ohio State University
1975

Reading Committee:
Professor A. A. Ksienski
Professor C. H. Walter
Professor R. T. Compton, Jr.

Approved By:

[Signature]
Adviser
Department of Electrical Engineering
ACKNOWLEDGMENTS

I wish to express my sincere appreciation to my adviser, Dr. R. T. Compton, Jr., not only for suggesting the topic area for this dissertation, but also for his technical advice and encouragement during the course of my graduate program. Without his expert guidance, I feel this research would not have reached a fruitful conclusion within the rather severe time constraints under which I had been working. The help and counseling offered by the other members of my reading committee, Dr. A. A. Ksierski and Dr. C. H. Walter are also deeply appreciated.

To the United States Air Force and in particular the Air Force Institute of Technology, my thanks for giving me the opportunity to further my education through their sponsorship of my program.

Finally and most importantly, I want to thank my wife Irene for her patience and understanding during the long months of course work and research and to her this work is dedicated.

The work reported in this dissertation was supported in part by Contract NO0019-75-C-0179 between Naval Air Systems Command and The Ohio State University Research Foundation.
VITA

November 22, 1941...... Born - Newark, New Jersey
1963............... B.S.E.E., Newark College of Engineering, Newark, New Jersey
1963-1975............. Electronics Developmental Engineer, United States Air Force
1969............... M.S.E.E., University of Missouri, Columbia, Missouri

FIELDS OF STUDY

Major Field. Electrical Engineering

Studies in Communication Theory. Professor C.E. Warren
Studies in Control Systems. Professor R.E. Fenton
Studies in Quantum Physics. Professor J.H. Hopps
Studies in Applied Mathematics. Professor J.T. Scheick
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF SYMBOLS</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter I: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Chapter II: THE ANTENNA</td>
<td>3</td>
</tr>
<tr>
<td>Chapter III: STOCHASTIC</td>
<td>14</td>
</tr>
<tr>
<td>Chapter IV: WEIGHT</td>
<td>32</td>
</tr>
<tr>
<td>Chapter V: CONCLUSIONS</td>
<td>95</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>98</td>
</tr>
</tbody>
</table>

**Chapter II: THE ADAPTIVE ANTENNA ARRAY**

- A. General Description: 3
- B. The Feedback Algorithm: 6
- C. The Weight Equation: 9
- D. The Reference Signal: 12

**Chapter III: STOCHASTIC DIFFERENTIAL EQUATIONS**

- A. Case I: Constant Coefficient, Random Forcing Function: 17
- B. Case II: Random Coefficient, Random Forcing Function: 20
- C. Case III: Ito Stochastic Differential Equation: 27

**Chapter IV: THE WEIGHT SOLUTION PROCESS**

- A. Introduction: 32
- B. Direct Approach: 38
- C. Ito Equation Approach: 48
- D. Stochastic Perturbation Approach: 61

**Chapter V: CONCLUSIONS**

- BIBLIOGRAPHY: 98
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Basic Adaptive Array Structure</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Implementation of Feedback Algorithm</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>Typical Power Spectrum of Input Signal</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>Multiplier Filtering of Product Term Power Spectrum</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>One Element &quot;Array&quot;</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>Graphical Representation of Random Process $w(t, \xi)$</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>Linear System Interpretation</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>Power Spectrum of Input cw Signal</td>
<td>34</td>
</tr>
<tr>
<td>9</td>
<td>$E{w(t)}$ vs. Noise Power</td>
<td>57</td>
</tr>
<tr>
<td>10</td>
<td>Weight Variance vs Loop Product Term $\varepsilon^2 k N_0$</td>
<td>60</td>
</tr>
<tr>
<td>11</td>
<td>Power Spectrum of Constant Weight Term $w_0$</td>
<td>65</td>
</tr>
<tr>
<td>12</td>
<td>Power Spectrum of Input Noise Process</td>
<td>67</td>
</tr>
<tr>
<td>13</td>
<td>Power Spectrum of Quadrature Noise Process</td>
<td>68</td>
</tr>
<tr>
<td>14</td>
<td>Power Spectrum of $w_1(t)$ Process</td>
<td>69</td>
</tr>
<tr>
<td>15</td>
<td>Normalized Average Power of $w_1(t)$ Process vs Noise Bandwidth</td>
<td>71</td>
</tr>
<tr>
<td>16</td>
<td>Approximate Power Spectrum of $w_1(t)$ Process</td>
<td>72</td>
</tr>
<tr>
<td>17</td>
<td>$E{w(t)}$ Variation with Noise Bandwidth</td>
<td>77</td>
</tr>
<tr>
<td>18</td>
<td>Definition of $q_{BW}(\omega)$</td>
<td>83</td>
</tr>
<tr>
<td>19</td>
<td>Definition of $\overline{q}_{BW+X}(\omega)$</td>
<td>83</td>
</tr>
<tr>
<td>20</td>
<td>Normalized Power Spectrum of $w_2(t)$ Process</td>
<td>85</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>a(t)</td>
<td>random coefficient process</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>amplitude of desired signal</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>amplitude of reference signal</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Wiener process covariance matrix</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>expectation operator</td>
<td></td>
</tr>
<tr>
<td>f(t)</td>
<td>random forcing function</td>
<td></td>
</tr>
<tr>
<td>h(t)</td>
<td>system impulse response</td>
<td></td>
</tr>
<tr>
<td>H(jω)</td>
<td>system function</td>
<td></td>
</tr>
<tr>
<td>I₀</td>
<td>zero order Bessel function</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>fixed gain constant</td>
<td></td>
</tr>
<tr>
<td>n(t)</td>
<td>narrowband, Gaussian noise process</td>
<td></td>
</tr>
<tr>
<td>p( )</td>
<td>probability density function</td>
<td></td>
</tr>
<tr>
<td>q(t)</td>
<td>envelope squared process</td>
<td></td>
</tr>
<tr>
<td>r(t)</td>
<td>reference signal</td>
<td></td>
</tr>
<tr>
<td>Rₓₓ(τ)</td>
<td>autocorrelation function of x(t) process</td>
<td></td>
</tr>
<tr>
<td>s(t)</td>
<td>array output; desired input signal</td>
<td></td>
</tr>
<tr>
<td>Sₓₓ(ω)</td>
<td>power density spectrum of x(t) process</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>time</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>transpose of matrix</td>
<td></td>
</tr>
<tr>
<td>wᵢ</td>
<td>ith weight</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>weight vector</td>
<td></td>
</tr>
</tbody>
</table>
\begin{itemize}
\item $x_i(t)$: $i$th Wiener process; $i$th input signal
\item $\hat{x}(t)$: Hilbert transform of $x(t)$ process
\item $y(t), z(t)$: quadrature components of $n(t)$ noise process
\item $\delta(\cdot)$: delta function
\item $e(t)$: error signal
\item $\mu_x$: mean of $x(t)$ process
\item $\phi$: covariance matrix
\item $\phi_{xx}(\cdot)$: characteristic function of $x(t)$ process
\item $\rho$: correlation coefficient
\item $\omega$: frequency in rad/sec
\item $\omega_0$: carrier frequency
\item $\omega_c$: one sided noise bandwidth
\item $\sigma_x^2$: variance of $x(t)$ process
\item $\nabla$: gradient operator
\item $\text{cov}(x(t))$: covariance of $x(t)$ process
\end{itemize}
CHAPTER I
INTRODUCTION

Research into adaptive antenna arrays has been vigorously undertaken since the first papers appeared on the subject during the middle 1960's. Credit for the earliest work in this area is usually given to Applebaum [1] and Shor [2] who suggested an algorithm for weight adaptation based on maximization of output signal to noise ratio. Another early contribution was the paper by Widrow et al [3] who recommended an automatic weight adjustment scheme based on minimization of the mean square error between the array output and some desired or local reference signal. The test results of an early experimental adaptive array for signal interference rejection were reported by Riegler and Compton [4]. Berni [5] has suggested using the deep nulls of an adaptive array for angle of arrival estimation. The behavior of adaptive antenna arrays in coded communications systems has been studied by Reinhard [6]. Other studies have examined different algorithms for adjustment of the weights [7, Chap. 5], the use of differently derived reference signals and their effect on array performance [7,18] and the effects of control loop noise [9]. Various experimental arrays have been constructed and test results reported in the literature - see for example [4,10,18]. A host of different applications of adaptive arrays, especially for radar systems and in the communications field have been proposed. The report by Compton [11] on adaptive arrays for airborne communications systems is an example of one such application.

The purpose of this research was to perform a theoretical study of the statistics of the adaptive array weights when an incident signal is corrupted by additive random noise. The difficulty in treating this subject arises from the presence of both a random coefficient and random forcing function in the differential equation obeyed by the weights. The analysis assumes the use of the minimization of mean square error criteria or so-called LMS algorithm for the weights [3].

One way to circumvent the random differential equation problem has been proposed by Berni [12]. He suggests modeling the array weights as a Markov process. This allows the use of Fokker-Planck diffusion equations to solve for the steady state probability density function. From this, the mean and variance can then immediately be calculated.
In this report, several methods for solving this problem are investigated. First, a direct approach using characteristic functions is described. Equations for the mean and autocorrelation function of the weights are developed. In a second approach, we extend Berni's work to the case of a random coefficient, random forcing term differential equation, by formulating the problem in terms of an Ito equation. We use formulae developed from the Fokker-Planck equation to derive expressions for the mean and variance of the weights. Finally, in a third approach, we assume the additive random noise portion of the signal be small compared to a desired cw signal. This assumption allows the use of stochastic perturbation theory - a technique which has found wide application for treating stochastic differential equations [13]. This method has been most widely employed in the treatment of second order stochastic wave equations resulting from the study of propagation of waves through random media [14].

The application of stochastic perturbation theory to the random weight problem is shown to greatly simplify the analysis. Still, it gives a good description of the stochastic properties of the weights which would be sufficient for many applications. Furthermore, it provides some understanding of the effect of the bandwidth of the noise process on the weights, a result not evident when employing the Markov characterization approach.

Chapter II will review adaptive arrays and in particular develop the differential equations satisfied by the weights when the LMS algorithm is used. This material can be found in numerous reports on adaptive arrays but is included here for completeness and to provide a base upon which to begin. The equation for a simple one element "array" is selected for analysis in this study. Chapter III examines some of the mathematical difficulties and techniques for treating stochastic differential equations. Chapter IV applies these techniques using the three approaches mentioned earlier to study the statistical properties of the weights.
CHAPTER II
THE ADAPTIVE ANTENNA ARRAY

A. General Description

The operation of an adaptive antenna receiving array is based on an internal feedback mechanism which controls the gain and phase processing of received signals. It is done in such a way as to automatically change the antenna pattern to meet some design criteria. For example, the array pattern may be shaped to maximize the gain in the direction of a desired signal. Alternatively, it may be desirable to minimize the antenna pattern in a direction from which an interfering signal is emanating. The ability of a receiving array to accomplish this automatically under a varying signal environment is what makes the adaptive array particularly valuable.

A block diagram of the basic N-element adaptive antenna array is shown in Fig. 1. Incoming signals impinge on the N-array elements creating the input signals denoted $y_i(t)$. These signals are then split into in-phase and quadrature components using 90° hybrids. The resulting signal components, designated $x_i(t)$, are then multiplied by the weights $w_i$ whose characteristics are controlled by the feedback processor. The feedback processor compares the received signal components with an error signal $\epsilon(t)$ which has been derived by comparing the array output $s(t)$ with some reference signal $r(t)$. After comparing the error signal and received signal, the processor operates on this data using some preconceived algorithm to generate a control signal for adjustment of the weights.

From Fig. 1, the array output can be written as

$$ s(t) = \sum_{i=1}^{2N} w_i x_i(t) $$

The error signal is then

$$ \epsilon(t) = r(t) - \sum_{i=1}^{2N} w_i x_i(t) $$

Fig. 1. Basic adaptive array structure.
and the error squared is

\[ e^2(t) = r(t) - 2r(t) \sum_{i=1}^{2N} w_i x_i(t) + \sum_{i=1}^{2N} \sum_{j=1}^{2N} w_i w_j x_i(t) x_j(t) \]

The mean squared error, for some given set of weights is then

\[ \overline{e^2(t)} = \overline{r^2(t)} - 2 \sum_{i=1}^{2N} w_i \overline{r(t)x_i(t)} + \sum_{i=1}^{2N} \sum_{j=1}^{2N} w_i w_j \overline{x_i(t)x_j(t)} \]

where the exact meaning of the overbar will be defined later. In matrix form,

\[ \overline{e^2(t)} = \overline{r^2(t)} - 2W^T S + W^T \Phi W \]

where

\[ W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{2N} \end{bmatrix} \]

\[ S = \begin{bmatrix} x_{1}(t) r(t) \\ x_{2}(t) r(t) \\ \vdots \\ x_{2N}(t) r(t) \end{bmatrix} \]
and

$$\begin{bmatrix}
x_1(t) x_1(t) & x_1(t) x_2(t) & \cdots & x_1(t) x_{2N}(t) \\
x_2(t) x_1(t) & x_2(t) x_2(t) & \cdots & x_2(t) x_{2N}(t) \\
\vdots & \vdots & & \vdots \\
x_{2N}(t) x_1(t) & x_{2N}(t) x_2(t) & \cdots & x_{2N}(t) x_{2N}(t)
\end{bmatrix}$$

(8)  $\Phi =$

The superscript $T$ in (5) denotes the transpose of the matrix.

B. The Feedback Algorithm

The particular feedback algorithm used in this study is a form of the LMS or Least Mean Square algorithm originally proposed by Widrow et al [3]. Its philosophy is to adjust the weights so that the time rate of change of the weights is proportional to the slope of the mean square error surface (the surface obtained by plotting $\varepsilon^2(t)$ versus the weights). In mathematical terms, the weights are to be adjusted so that

$$\frac{dw_i(t)}{dt} = -k \nabla_{W_i} \varepsilon^2(t)$$

(9) is satisfied. $\nabla_{W_i}$ denotes the gradient operator and $k$ is a positive constant. Since

$$\nabla_{W_i} \varepsilon^2(t) = \frac{\partial \varepsilon^2(t)}{\partial W_i}$$

(10) then using (5), we get in matrix form

$$\nabla_{W_i} \varepsilon^2(t) = -2S + 2\Phi W$$

(11) or alternatively,

$$\nabla_{W_i} \varepsilon^2(t) = -2 x_i(t) r(t) + 2 \sum_{i=1}^{2N} w_j x_i(t) x_j(t)$$

(12)
Using the definition for $\epsilon(t)$, (12) is just

\begin{equation}
\mathbb{E}_w \epsilon^2(t) = -2 \mathbb{E}_x \epsilon(t)
\end{equation}

So the feedback rule becomes

\begin{equation}
\frac{d w_i(t)}{dt} = 2k \mathbb{E}_x \epsilon(t)
\end{equation}

The implementation of this rule is shown in Fig. 2.

Fig. 2. Implementation of feedback algorithm.

We are now at a point when we can discuss the meaning of the overbar defined in (4). In principle, this symbol can have whatever meaning we wish to give it. The definition we adopt for the overbar in (4) will then also define the operation required in the box in Fig. 2. From the standpoint of array performance, we might wish to define the overbar to mean the statistical expectation. However, from Fig. 2, this would mean we would somehow
have to compute $E\{x_i(t) \varepsilon(t)\}$ - an operation that is impractical to implement in a real array. An infinite time average might also be considered but again practical considerations rule out this option. A low-pass filtering operation, however, is feasible to implement. Moreover, the low-pass characteristics of the multipliers used to form the product $x_i(t) \varepsilon(t)$ do provide this type of filtering action. For this reason and to best match the theory to what has actually been built, we will define the overbar to mean a low pass filtering operation. For a further discussion of this point, see Compton [15] and Di Carlo [16].

To see the effects of this filtering action, consider a typical communications application in which $x_i(t)$ and the reference signal are both bandlimited processes with spectral components centered around some carrier frequency $\omega_0$ - see Fig. 3. We also require $\omega_0 >> \omega_c$. The spectrum of the product terms $x_i(t) r(t)$ and $x_i(t) x_j(t)$ will contain power in a band around zero frequency ($\omega=0$) and twice the carrier frequency ($\omega=\pm 2\omega_0$) - see Fig. 4. The effect of the low pass filtering (which we assume to be ideal) is to filter out the spectral components around $\omega=\pm 2\omega_0$ but to leave the components in the band $-2\omega_c<\omega<+2\omega_c$. The upper cutoff frequency of the filter is assumed to be high enough to completely pass the input narrow-band signal $x_i(t)$; hence a separate low pass filter box is not needed after the multiplier preceding the output summer in Fig. 2.

![Fig. 3. Typical power spectrum of input signal.](image-url)
C. The Weight Equation

Returning to (14) and substituting for \( e(t) \) from (2), we get

\[
\frac{d w_i(t)}{dt} = 2k x_i(t)[r(t) - \sum_{j=1}^{2N} w_j x_j(t)]
\]

or putting this in the standard form,

\[
\frac{d w_i(t)}{dt} + 2k \sum_{j=1}^{2N} x_i(t) x_j(t) w_j = 2k x_i(t) r(t).
\]

In matrix form,

\[
\frac{d}{dt} W + 2k \Phi W = 2k S
\]

or
This is the desired system of differential equations obeyed by the weights.

There are two major obstacles to solving this system of equations. The first is the problem of uncoupling this system of equations (they are coupled due to the covariance matrix $\phi$) and the second is the problem of solving the resulting equations when both the coefficient term and forcing term are random processes. The thrust of this dissertation is to study the second problem; however, before proceeding, a few words about the uncoupling problem are in order.¹

If some restrictive assumptions on the matrix $\phi$ are made, it would be possible to solve (18). For example, in the case when $\phi$ is a constant matrix (there are no time varying components present), the method used to uncouple the system of equations is to perform a coordinate rotation into the principal axes of the $\phi$ matrix [4,15,16,17]. Since $\phi$ is a real and symmetric matrix, an orthogonal change of basis matrix $Q$ exists such that

\begin{equation}
W = QW' \tag{19}
\end{equation}

¹In Chapter III, we will, however, describe a technique that in principle can satisfy both problems simultaneously.
where \( W' \) are the new "normal coordinates" for the weights and such that

\[
(20) \quad Q^{-1} \Phi Q = \text{diag} (\lambda_1, \lambda_2, \ldots, \lambda_{2N})
\]

where the \( \lambda_i \) denote the eigenvalues of \( \Phi \). After making this coordinate rotation, the system of differential equations obeyed by the weights becomes

\[
(21) \quad \frac{d}{dt} W' + 2k [Q^{-1} \Phi Q] W' = 2k Q^{-1} S.
\]

Since \( Q \) has been chosen so that \([Q^{-1} \Phi Q]\) is a diagonal matrix, the system of differential equations is now uncoupled.

The case of a constant \( \Phi \) matrix discussed above is very restrictive. In practice, the products \( x_i(t)x_j(t) \) generated in the array are time varying because the spectral components in the band \( -2\omega_c < \omega < 2\omega_c \) are still present in the output (ref. discussion p. 8). In this study, we will not pursue the coupled equation problem further but will try to obtain an intuitive understanding of the weight statistics by considering the special case of a one dimensional version of the weight equation, i.e.

\[
(22) \quad \frac{d w_i(t)}{dt} + 2k x_i^2(t) w_i(t) = 2k x_i(t) r(t)
\]

where we have reverted to the original notation. With the input \( x_i(t) \) composed of a desired signal component \( s_i(t) \) and a random noise component \( n_i(t) \) so that

\[
(23) \quad x_i(t) = s_i(t) + n_i(t)
\]

then (22) becomes

\[
(24) \quad \frac{d w_i(t)}{dt} + 2k [s_i(t) + n_i(t)]^2 w_i(t) = 2k [s_i(t) + n_i(t)] r(t)
\]

In our model, we will assume that the noise component is a zero mean, narrowband Gaussian random process with power \( \sigma^2 \).
D. The Reference Signal

Turning our attention to the reference signal, we note that in the adaptive array, \( r(t) \) should approximate the desired signal as closely as possible. Then the error signal \( e(t) \) will consist of any interference or noise present in the array output (see Fig. 1). The closer \( r(t) \) is to the desired signal, the better the array performance will be. Of course in practice, \( r(t) \) can never be made exactly coherent with the desired signal. But in many cases, \( r(t) \) can be made to closely approximate the desired signal - see Schwegman and Compton [18]. In this study, we will simply assume that the reference signal is equal to the desired signal \( s(t) \). We will further assume that the reference signal is "clean" in that it has not been corrupted with additive noise such as that present with the incoming signal. Equation (24) can then be rewritten as

\[
\frac{dw_i(t)}{dt} + 2k [s_i(t)+n_i(t)]^2 w_i(t) = 2k[s_i(t)+n_i(t)]s_i(t)
\]

or if the subscript is dropped,

\[
\frac{dw(t)}{dt} + 2k[s(t)+n(t)]^2 w(t) = 2k[s(t)+n(t)]s(t)
\]

This equation would describe the behavior of an "adaptive array" with a single control loop as shown in Fig. 5.

\[ s(t)+n(t) \]
\[ \rightarrow \]
\[ [s(t)+n(t)]w(t) \]
\[ \rightarrow \]
\[ w(t) \]
\[ \rightarrow \]
\[ \frac{dw(t)}{dt} \]
\[ \rightarrow \]
\[ \sum \]
\[ 2K \]
\[ \rightarrow \]
\[ + r(t) = s(t) \]

Fig. 5. One element "array".
The problem addressed by the remainder of this report is then: knowing the properties of the deterministic incoming signal \( s(t) \) and the band-limited Gaussian noise process \( n(t) \), what are the statistical properties of the weight solution process \( w(t) \)? Note that since \( n(t) \) is a random process, then \( w(t) \) must also be a random process. The equation (26) is said to be a first order stochastic differential equation with a random forcing function and a random coefficient. Here the random forcing term is \( 2k[s(t)+n(t)]s(t) \) and the random coefficient is \( 2k[s(t)+n(t)]^2 \). Before analyzing the properties of the weight process \( w(t) \), the general first order stochastic differential equation

\[
\frac{dw(t)}{dt} + a(t)w(t) = f(t)
\]

will be reviewed in Chapter III.
CHAPTER III
STOCHASTIC DIFFERENTIAL EQUATIONS

The theory of stochastic or random differential equations has become extremely important in the past few years due to its widespread application to many engineering and physical science problems. By a stochastic differential equation, we mean a differential equation of the form

\begin{align*}
(28) \quad L w(t) &= a_n(t) \frac{d^n w(t)}{dt^n} + a_{n-1}(t) \frac{d^{n-1} w(t)}{dt^{n-1}} + \ldots \\
&+ a_0(t) w(t) = f(t)
\end{align*}

where the coefficients \( a_i(t) \) and the forcing term \( f(t) \) may all be random processes. Since the equation obeyed by the weights was given by

\begin{align*}
(29) \quad \frac{dw(t)}{dt} + 2k[s(t)+n(t)]^2 w(t) &= 2k [s(t)+n(t)]s(t)
\end{align*}

we will restrict this discussion to first order stochastic differential equations of the form

\begin{align*}
(30) \quad \frac{dw(t)}{dt} + a(t)w(t) &= f(t)
\end{align*}

and initial condition

\begin{align*}
(31) \quad w(t=0) &= w(0).
\end{align*}

It is not the purpose of this chapter to present the mathematical theory or to detail proofs for this general class of differential equations. Rather we will briefly review some basic
analysis techniques that will be needed in the development of the weight solution process \( w(t) \) in Chapter IV. The reader is referred to the texts by Saaty [19], Bharucha-Reid [20], Tsokos and Padgett [21], Soong [22] and Srinivasan and Vasudevan [23] which treat this subject area in detail.

Proceeding as we would with a deterministic solution, we multiply both sides of (30) by

\[
\int_0^t a(u) du
\]

(32)

\[
\text{giving}
\]

\[
\frac{dw(t)}{dt} e^0 + a(t)w(t)e^0 = f(t)e^0
\]

(33)

which can be written as

\[
\frac{d}{dt} \left[ w(t)e^0 \right] = f(t) e^0
\]

(34)

or finally

\[
w(t) = w(0)e^0 + \int_0^t f(v)e^{-v} dv.
\]

(35)

This is the desired solution to the differential equation (30) and it will be used extensively in the sequel. First, however, we make a few comments about the integrals in Eq. (35).

There is no problem in interpretation when \( a(t) \) and \( f(t) \) are deterministic functions. In this case, the integrals are simply the normal Riemann integrals of ordinary calculus and are based on the fundamental idea of a limit of a sequence. However, when \( f(t) \) and/or \( a(t) \) are random processes, which we can denote by
and $a(t,\xi)$, we then have a whole family of solutions $w(t,\xi)$, one for each of the sample functions of the ensemble — see Fig. 6. Thus in order for the integrals in (35) to exist when $a(t,\xi)$ and $f(t,\xi)$ are random processes, they must exist for every $\xi$. Because of the probabilistic nature of random processes, this requirement is held to be too restrictive. In some cases, the integrals may not exist for every possible $\xi$. For this reason, stochastic differentiation and integration based on mean square calculus is usually invoked. Thus given the integral

\[(36) \quad s(t) = \int_0^t a(u)du\]

where $a(t)$ is a random process, we will redefine $s(t)$ to be that function for which

\[(37) \quad \lim_{\Delta t_i \to 0} E\left\{\left[s(t) - \sum_{i=1}^m a(t_i)\Delta t_i\right]^2\right\} = 0\]

where $E$ is the usual expectation operator and $\Delta t_i$ is a small, incremental time interval. When such an $s(t)$ can be found, we say that the integral exists in the mean square sense. Analogous to ordinary Riemann integration, this definition is based on the convergence of a sequence in the mean square sense. For a more detailed discussion of stochastic convergence concepts, see Papoulis [24, pp. 260-263].

In summary, for the remainder of this study, we will assume that the solution process for the weights $w(t)$ given by (35) exists, where the integrals of the random processes $a(t)$ and $f(t)$ are defined in the mean square sense.

In the development of the weight solution process, three special cases of stochastic differential equations will be discussed. These are: (I) constant coefficient, random forcing function; (II) random coefficient, random forcing function and (III) Itô equation. These are briefly discussed below.

---

2When describing a random process, we will no longer indicate the specific dependency on the ensemble of sample functions $\xi$. Thus the random process $a(t)$ will be understood to mean $a(t,\xi)$.
A. **Case I**: Constant Coefficient, Random Forcing Function

The simplest stochastic differential equation to analyze is that having a constant coefficient and a random forcing function

\[(38) \quad \frac{dw(t)}{dt} + a_0 w(t) = f(t) .\]

The solution to (38) is

\[(39) \quad w(t) = w(o) e^{-a_0 t} + \int_0^t f(v)e^{-a_0 (t-v)} \, dv\]

where \(w(o)\) is the value of \(w(t)\) at \(t = 0\) (initial condition).

As is well known, the solution to (38) can be interpreted as the output of a linear system (see Fig. 7) with input \(f(t)\) and a system function

\[(40) \quad H(j\omega) = \frac{1}{j\omega + a_0}\]
or alternatively an impulse response

\begin{equation}
    h(t) = e^{-\alpha_0 t} u(t)
\end{equation}

By the convolution theorem for linear systems,

\begin{equation}
    w(t) = \int_{-\infty}^{\infty} f(t-v)h(v)dv = \int_{-\infty}^{\infty} f(v) h(t-v)dv.
\end{equation}

Using (42) and taking the expected value of both sides of the equation (we assume we can interchange the order of integration and expected value), we find

\begin{equation}
    \mathbb{E}\{w(t)\} = \int_{-\infty}^{\infty} \mathbb{E}\{f(t-v)\} h(v)dv
\end{equation}

\begin{equation}
    = \mathbb{E}\{f(t)\} \int_{-\infty}^{\infty} h(v)dv
\end{equation}

\begin{figure}
    \centering
    \includegraphics[width=0.5\textwidth]{linear_system_interpretation.png}
    \caption{Linear system interpretation.}
\end{figure}
Similarly, the autocorrelation function of the output process can be found to be (Papoulis [24])

\[ R_{WW}(\tau) = R_{ff}(\tau) \otimes h^*(-\tau) \otimes h(\tau) \]

where \( \otimes \) denotes the convolution and * superscript the complex conjugate.

Finally since the power density spectrum is the Fourier transform of the autocorrelation function, then (47) can be put in the more familiar form

\[ S_{WW}(\omega) = S_{ff}(\omega)|H(j\omega)|^2 \]

Thus for the constant coefficient, random forcing function case, direct methods are available to solve for the statistics of the solution process \( w(t) \). In the special case when \( f(t) \) is Gaussian, the output will also be Gaussian since the system is linear.

The above formulae, (46), (47) and (48), apply to the case when the output process is stationary. Consider the following modifications to the convolution formula (42): for a real, causal system we have

\[ w(t) = \int_{0}^{\infty} f(t-v)h(v)dv = \int_{-\infty}^{t} f(v)h(t-v)dv \]

With the driving function applied at \( t=0 \), (49) becomes

\[ w(t) = \int_{0}^{t} f(t-v)h(v)dv = \int_{0}^{t} f(v)h(t-v)dv \]
Thus if a wide sense stationary input is applied at \( t=0 \), the output is only wide sense stationary in the asymptotic or steady state \( (t \to \infty) \) sense (Papoulis [24, p. 347]). For example, consider the mean of the weight solution \( w(t) \). Using (50) to calculate the mean gives

\[
E[w(t)] = \int_{0}^{t} E\{f(v)\} h(t-v)dv
\]

Substituting for \( h(t) \) from (41) gives

\[
E[w(t)] = E\{f(v)\} \int_{0}^{t} e^{-a_0(t-v)} dv
\]

Substituting for \( h(t) \) from (41) gives

\[
E[w(t)] = E\{f(v)\} \int_{0}^{t} e^{-a_0(t-v)} dv
\]

Substituting for \( h(t) \) from (41) gives

\[
E[w(t)] = E\{f(v)\} \int_{0}^{t} e^{-a_0(t-v)} dv
\]

Note that \( E[w(t)] \) is not a constant but contains a decaying exponential term and only approaches stationarity as \( t \to \infty \). For large \( t \), the result is the same as that calculated by (43), namely (46). When using the formulae (46), (47) and (48) then, we will assume that the input process \( f(t) \) has been applied for a long time to enable the output to achieve stationarity.

**B. Case II. Random Coefficient, Random Forcing Function**

The second case is the most complex and difficult to analyze. Here in the differential equation

\[
\frac{dw(t)}{dt} + a(t) w(t) = f(t)
\]

both \( a(t) \) and \( f(t) \) are assumed to be random processes. From the solution
we can see that the determination of the probability density function of \( w(t) \) is rather complicated. Even with zero initial conditions and both \( f(t) \) and \( a(t) \) Gaussian random processes, the \( w(t) \) process transformation involves the integral of the product of a Gaussian and Log-normal random process.

Applying the expectation operator directly to both sides of (56) and assuming zero initial conditions, gives

\[
E[w(t)] = \int_0^t \int_v E[f(v) e^{-\int_v^t a(u)du}] \, dv
\]

for the mean.

Similarly, the autocorrelation function is

\[
R_{ww}(t,t+\tau) = \int_0^t \int_0^{t+\tau} E[f(v_1)f(v_2)]
\]

\[
-\int_v^t a(u_1)du_1 - \int_v^{t+\tau} a(u_2)du_2
\]

\[
e^{\int_v^t v_1 du_1 - \int_v^{t+\tau} v_2 du_2}
\]

As can be seen above, without any further simplifying assumptions, even the determination of the first two moments is extremely difficult. If \( a(t) \) and \( f(t) \) are assumed to be statistically independent, the above reduces to

\[
E[w(t)] = \int_0^t E[f(v)] E[e^{-b(t)-b(v)}] \, dv
\]

where

\[
b(t) = \int a(t) \, dt.
\]
Also,

\begin{equation}
R_{WW}(t, t+\tau) = \int_0^t \int_0^{t+\tau} R_{ff}(v_1, v_2) \ d v_1 \ d v_2
\end{equation}

Furthermore, if \( a(t) \) is Gaussian, then \( b(t) \) will also be Gaussian and \( E[e^{-b(t)}] \) may be associated with the characteristic function of \( b(t) \). This point will be elaborated on in Chapter IV.

Notwithstanding the simplification made above, the calculation of even the simple moments of the solution process is a formidable task. For this reason, approximation methods must be employed, the most powerful and widely used being stochastic perturbation theory. This theory assumes that the solution process and the coefficient (in the case of a random coefficient problem) can be written as a convergent series in powers of \( \varepsilon \) where \( \varepsilon \) is a small parameter.

If we assume that the random coefficient \( a(t) \) in (55) can be written as

\begin{equation}
a(t) = a_0(t) + \varepsilon a_1(t) + \varepsilon^2 a_2(t) + \ldots
\end{equation}

where \( a_0(t) \) represents a deterministic function (this requirement will be elaborated on later) and the solution process can be written as

\begin{equation}
w(t) = w_0(t) + \varepsilon w_1(t) + \varepsilon^2 w_2(t) + \ldots
\end{equation}

then inserting (62) and (63) into (55) gives

\begin{equation}
\frac{d}{dt} [w_0(t) + \varepsilon w_1(t) + \varepsilon^2 w_2(t) + \ldots] + [a_0(t) + \varepsilon a_1(t) + \varepsilon^2 a_2(t) + \ldots] \\
\times [w_0(t) + \varepsilon w_1(t) + \varepsilon^2 w_2(t) + \ldots] = f(t),
\end{equation}
We will call the terms $a_0(t)$ and $w_0(t)$ the unperturbed portions of the coefficient and solution respectively. The small, additive disturbance terms $e_1(t), e^2a_2(t), ...$ and $e_1(t), e^2w_2(t), ...$ will be called the perturbed portions of the coefficient process and solution. Collecting terms in (64) we get

\begin{equation}
\frac{d}{dt} w_0(t) + e \frac{d}{dt} w_1(t) + e^2 \frac{d}{dt} w_2(t) + ... + [a_0(t)w_0(t) - f(t)] \\
+ e[w_0(t)a_1(t) + a_0(t)w_1(t)] + e^2[a_0(t)w_2(t) + a_1(t)w_1(t) + a_2(t)w_0(t)] \\
+ ... = 0
\end{equation}

Equating like powers of $\varepsilon$ results in the following system of differential equations:

\begin{equation}
\frac{dw_0(t)}{dt} + a_0(t)w_0(t) = f(t)
\end{equation}

\begin{equation}
\frac{dw_1(t)}{dt} + a_0(t)w_1(t) = -a_1(t)w_0(t)
\end{equation}

\begin{equation}
\frac{dw_2(t)}{dt} + a_0(t)w_2(t) = -a_1(t)w_1(t) - a_2(t)w_0(t)
\end{equation}

\begin{equation}
\frac{dw_3(t)}{dt} + a_0(t)w_3(t) = -a_1(t)w_2(t) - a_2(t)w_1(t) \\
- a_3(t)w_0(t)
\end{equation}

By solving these equations in order, the terms in the solution for $w(t)$ may be found

\begin{equation}
w(t) = w_0(t) + \varepsilon w_1(t) + \varepsilon^2 w_2(t) + ...
\end{equation}
In utilizing this method, we have reduced the stochastic differential equation (55) with a random coefficient to a set of stochastic differential equations with a deterministic coefficient \( a_0(t) \). The reason for the requirement that \( a_0(t) \) be a deterministic function mentioned in the discussion of (62) should now be clear. In the special case when the unperturbed portion of the random coefficient is a constant

\[
(71) \quad a_0(t) = a_0
\]

then the set of differential equations are all constant coefficient, random forcing function stochastic equations of the type already discussed under Case I (p.17). This technique can be considered particularly successful if only the first few terms in the expansion (70) are sufficient to adequately describe the solution process. This is the principal method which will be used in the development of the weight solution process for the adaptive antenna array.

There is a second beneficial property of the stochastic perturbation approach which deserves mention here. Recall the initial vector formulation of the weight differential equation (17) was

\[
(72) \quad \frac{d}{dt} W + 2k\phi W = 2kS.
\]

Suppose now that the covariance matrix \( \phi \) defined in (8) can be written as

\[
(73) \quad \phi = \phi_0 + \epsilon \phi_1 + \epsilon^2 \phi_2 + \ldots
\]

where \( \phi_0 \) is a real, symmetric constant matrix and the time varying portion of \( \phi \) is included in the additional terms \( \phi_1, \phi_2, \ldots \). Also we assume that

\[
(74) \quad S = S_0 + \epsilon S_1 + \epsilon^2 S_2 + \ldots
\]

and

\[
(75) \quad W = W_0 + \epsilon W_1 + \epsilon^2 W_2 + \ldots
\]
where $S$ and $W$ are defined by (7) and (6) respectively. Upon substituting (73), (74) and (75) into (72), we get

\begin{equation}
\frac{d}{dt} W_0 + \varepsilon \frac{d}{dt} W_1 + \varepsilon^2 \frac{d}{dt} W_2 + \cdots + 2k[\phi_0 + \varepsilon \phi_1 + \varepsilon^2 \phi_2 + \cdots] \\
\times [W_0 + \varepsilon W_1 + \varepsilon^2 W_2 + \cdots] = 2k[S_0 + \varepsilon S_1 + \varepsilon^2 S_2 + \cdots]
\end{equation}

Collecting terms, we get

\begin{equation}
\frac{d}{dt} W_0 + \varepsilon \frac{d}{dt} W_1 + \varepsilon^2 \frac{d}{dt} W_2 + \cdots + 2k[\phi_0 W_0 + \varepsilon (\phi_0 W_1 + \phi_1 W_0)] \\
+ \varepsilon^2 (\phi_0 W_2 + \phi_1 W_1 + \phi_2 W_0) + \cdots = 2k[S_0 + \varepsilon S_1 + \varepsilon^2 S_2 + \cdots]
\end{equation}

Finally, equating like powers of $\varepsilon$, we get the following system of vector differential equations:

\begin{equation}
\frac{dW_0}{dt} + 2k \phi_0 W_0 = 2kS_0
\end{equation}

\begin{equation}
\frac{dW_1}{dt} + 2k \phi_0 W_1 = 2k(S_1 - \phi_1 W_0)
\end{equation}

\begin{equation}
\frac{dW_2}{dt} + 2k \phi_0 W_2 = 2k(S_2 - \phi_1 W_1 - \phi_2 W_0)
\end{equation}

\vdots

These vector equations are each coupled due to the covariance matrix $\phi_0$. But in our perturbation expansion (73), we chose $\phi_0$ to be a real, symmetric constant matrix. Hence we can use a coordinate rotation, as discussed in Chapter II (pp.10-11), to diagonalize the $\phi_0$ matrix. This results in the following system of uncoupled equations:
\[ \frac{dW'_0}{dt} + 2k[Q^{-1}\phi_0 Q]W'_0 = 2k Q^{-1}S_0 \]

(82) \[ \frac{dW'_1}{dt} + 2k[Q^{-1}\phi_0 Q]W'_1 = 2k Q^{-1}[S_1 - \phi_1 W'_0] \]

(83) \[ \frac{dW'_2}{dt} + 2k[Q^{-1}\phi_0 Q]W'_2 = 2k Q^{-1}[S_2 - \phi_1 W'_1 - \phi_2 W'_0] \]

\cdots

Finally, from (19), the weights in the two coordinate systems are related by

(84) \[ W'_0 = Q W''_0 \]

and

(85) \[ W'_1 = Q W''_1 \]

\cdots

Substituting these into the system of equations above gives

(86) \[ \frac{dW''_0}{dt} + 2k[Q^{-1}\phi_0 Q]W''_0 = 2k Q^{-1}S_0 \]

(87) \[ \frac{dW''_1}{dt} + 2k[Q^{-1}\phi_0 Q]W''_1 = 2k[Q^{-1}S_1 - [Q^{-1}\phi_1 Q]W''_0] \]

(88) \[ \frac{dW''_2}{dt} + 2k[Q^{-1}\phi_0 Q]W''_2 = 2k[Q^{-1}S_2 - [Q^{-1}\phi_1 Q]W''_1 - [Q^{-1}\phi_0 Q]W''_0] \]

\cdots

26
By solving these equations in order, the complete vector solution for the weights $W'$ in the new "normal coordinates" can be found as

$$(89) \quad W' = W'_0 + \varepsilon W'_1 + \varepsilon^2 W'_2 + \cdots$$

or reverting back to the original coordinate system

$$(90) \quad W = QW' = QW'_0 + \varepsilon QW'_1 + \varepsilon^2 QW'_2 + \cdots$$

$$(91) \quad = W'_0 + \varepsilon W'_1 + \varepsilon^2 W'_2 + \cdots$$

Thus by using stochastic perturbation theory, not only are we able to solve the random coefficient problem but also we can overcome the uncoupling problem as well. We will not pursue the vector problem further in this study but mention here to point out some of the desirable features of the stochastic perturbation approach.

C. **Case III: Ito Stochastic Differential Equation**

The third kind of stochastic differential equation that will arise in Chapter IV is the Ito equation. This special differential equation has been perhaps the most widely studied and used in the analysis of systems driven by random inputs. Its popularity is due to both its mathematical simplicity and the fact that the concept of "white noise" has proven to give very good results in various circuit theory, control system and communication theory problems.

Returning to the original random coefficient differential equation (55), if we let the random processes $a(t)$ and $f(t)$ both be white noise (they may be correlated), we have

$$(92) \quad a(t) = \xi_1(t)$$

$$(93) \quad f(t) = \xi_2(t)$$
where the \( \xi_1(t) \) are zero mean, Gaussian, stationary random processes with a constant power density spectrum over the entire real frequency axis. Then we have

\[
\frac{dw(t)}{dt} + \xi_1(t) w(t) = \xi_2(t)
\]

Rewriting this in a different form, we get

\[
dw(t) = -w(t) \xi_1(t) dt + \xi_2(t) dt
\]

Now recall that

\[
x_1(t) = \int_0^t \xi_1(\tau) d\tau
\]

defines a Wiener process. It has independent increments, a zero mean and autocorrelation function

\[
R_{x_1x_1}(t_1, t_2) = \begin{cases} 
\sigma_1^2 t_2 & t_1 \geq t_2 \\
\sigma_1^2 t_1 & t_2 \geq t_1 
\end{cases}
\]

and hence a probability density function

\[
p(x_1, t) = \frac{1}{\sqrt{2\pi\sigma_1^2 t}} e^{-x_1^2/2\sigma_1^2 t}; \quad -\infty < x_1 < +\infty
\]

From (96), the Wiener process satisfies

\[
dx_1(t) = \xi_1(t) dt
\]

so (95) can be written
(100) \[ dw(t) = -w(t)dx_1(t) + dx_2(t) \]

This equation is a special case of an Ito stochastic differential equation given more generally by

(101) \[ dw(t) = \hat{f}(w(t),t)dt + \hat{G}(w(t),t)dx(t) \]

where

- \( \hat{w}(t) \) is an \( n \)-dim column vector
- \( \hat{f}(w(t),t) \) is an \( n \)-dim real column vector function of \( w(t) \) and \( t \)
- \( \hat{G}(w(t),t) \) is an \( nxm \) real vector function of \( w(t) \) and \( t \)

and

- \( \hat{x}(t) \) is an \( m \)-dim Wiener process vector

Equivalently, the integral representation is

(102) \[ \hat{w}(t) = \hat{w}_0 + \int_0^t \hat{f}(w(\tau),\tau)d\tau + \int_0^t \hat{G}(w(\tau),\tau)\hat{x}(\tau) \]

and is known as an Ito stochastic integral equation. In the special case of (100), we see with \( n=1, m=2 \) that

(103) \[ \hat{f}(w(t),t) = 0 \]

(104) \[ \hat{G}(w(t),t) = [-w(t) 1] \]

and

(105) \[ \hat{x}(t) = \begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} \]

In many texts, the term stochastic differential equation is synonymous with Ito stochastic differential equation. For example, a recent text by Arnold [25] titled "Stochastic Differential Equations" treats only equations of the Ito type (101).
In this study, we will use the term stochastic differential equation to mean the more general class of equations given by (30) where \( a(t) \) and \( f(t) \) are not necessarily white noise processes.

We have shown that Case II equations given by (55) are a special case of an Ito equation when both \( a(t) \) and \( f(t) \) are white noise processes. Clearly, Case I equations given by (38) are also Ito equations when the driving function \( f(t) \) is white noise. In this case, we have

\[
\text{(106)} \quad dw(t) = -a_0 w(t) dt + dx(t)
\]

where \( x(t) \) is a Wiener process. In correspondence to (101), we have

\[
\text{(107)} \quad f(w(t), t) = -a_0 w(t)
\]

and

\[
\text{(108)} \quad G(w(t), t) = 1
\]

and so once again we have an Ito equation - this time a scalar version.

The key feature of the Ito equation is that the solution process \( w(t) \) is Markov \([22,25,26,27]\). As such the probability density function for \( w(t) \) satisfies the Fokker-Planck diffusion equation (also called Kolmogorov's forward equation) given by (see Jazwinski \([26]\) or Viterbi \([27]\))

\[
\text{(109)} \quad \frac{\partial p(w,t)}{\partial t} = - \sum_i \frac{\partial}{\partial w_i} [p(w,t)f_i(w(t), t) + \frac{1}{2} \sum_i \sum_j \frac{\partial^2}{\partial w_i \partial w_j} \left[G(w(t), t) D G(w(t), t)^T \right] p(w, t)]
\]

where the matrix \( D \) is defined by the relation

\[
\text{(110)} \quad E\{dx(t)dx(t)^T\} = Ddt
\]
The scalar form reduces to

\[ \frac{\partial p(w,t)}{\partial t} = -\frac{\partial}{\partial w} [p(w,t)f(w(t),t)] + \frac{\sigma^2}{2} \frac{\partial^2}{\partial w^2} [p(w,t)G^2(w(t),t)] \]

where we have let

\[ D = \sigma^2 \]

The Fokker-Planck equation, (109) or (111), describes the evolution of the probability density function of the Markov process \( w(t) \) generated by the Ito equation (101). Of course, knowing the probability density function of \( w(t) \) allows one to calculate the desired statistics of the solution process. Alternatively, one can show (Jazwinski [26]) that the mean and covariance function satisfy

\[ \frac{dE[w(t)]}{dt} = E[f(w(t),t)] \]

and

\[ \frac{d}{dt} \text{Cov}(w(t)) = 2 [E[w(t)f(w(t),t)] - E[w(t)]E[f(w(t),t)]] + E[G(w(t),t)D G(w(t),t)^T] \]

where

\[ \text{Cov}(w(t)) = E[(w(t) - E[w(t)])[w(t) - E[w(t)]]^T] \]

Equations (113) and (114) provide a direct method to calculate the statistics of the weight solution without having to derive the probability density function. We will make use of these equations in the following chapter.
CHAPTER IV
THE WEIGHT SOLUTION PROCESS

A. Introduction

In Chapter II, we showed that the weight in a one-loop "adaptive array" obeyed the equation

\[ \frac{dw(t)}{dt} + 2k [s(t) + n(t)]^2 w(t) = 2k r(t)[s(t)+n(t)] \]

where

- \( w(t) \) is the weight
- \( s(t) \) is the incident desired signal
- \( r(t) \) is a reference signal of the same form as \( s(t) \)
- \( n(t) \) is a narrowband, Gaussian noise process

and

\( k \) is a loop gain constant.

Since \( n(t) \) is a random process, (116) is a first order stochastic differential equation with a random coefficient and a random forcing function. Its solution process, discussed in general in Chapter III, is given by

\[ w(t) = w(0) e^{\int_0^t 2k[s(u)+n(u)]^2 du} + \int_0^t \frac{2k[s(v)+n(v)]s(v)e^{-\int_v^t 2k[s(u)+n(u)]^2 du}}{2k[s(v)+n(v)]s(v)e^{-\int_v^t 2k[s(u)+n(u)]^2 du}} \]
As a simplification, we will assume zero initial conditions \((w(0)=0)\) so (117) becomes

\[
(118) \quad w(t) = \int_0^t 2k[s(v)+n(v)]s(v) e^{-\int_v^t 2k[s(u)+n(u)]^2 du} dv
\]

In a communications application, the adaptive array would normally be receiving signals centered at some carrier frequency \(\omega_0\). Let us assume initially that the desired input signal is a constant amplitude cw signal with a power spectrum

\[
(119) \quad S_{ss}(\omega) = \frac{\pi A^2}{2} \left[ \delta(\omega-\omega_0) + \delta(\omega+\omega_0) \right]
\]

as shown in Fig. 8. The total signal appearing at the input to the feedback processor is then

\[
(120) \quad s(t) + n(t) = A \cos \omega_0 t + n(t)
\]

where \(n(t)\) is the input noise at the feedback control loop. We assume \(n(t)\) to be a narrowband, Gaussian noise process with zero mean and variance \(\sigma^2\). We will delay discussion of the spectral characteristics of \(n(t)\) until later. Finally, to differentiate between the desired input signal and the local reference signal (which we have assumed to be of the same form), we will let the amplitude of the reference signal be denoted by \(B\):

\[
(121) \quad r(t) = B \cos \omega_0 t,
\]

Substituting (120) and (121) into (116) gives

\[
(122) \quad \frac{dw(t)}{dt} + 2k[A \cos \omega_0 t + n(t)]^2 w(t) = 2k B \cos \omega_0 t[A \cos \omega_0 t + n(t)]
\]

for the differential equation obeyed by the weight \(w(t)\).
Before proceeding, we will determine the statistical properties of the random forcing function and the random coefficient. The random forcing term is

\[(123) \quad f(t) = 2kB \cos\omega_0 t [A \cos\omega_0 t + n(t)]\]

\[(124) \quad = 2kAB \cos^2\omega_0 t + 2kB n(t) \cos\omega_0 t\]

\[(125) \quad = kAB + 2kB n(t) \cos\omega_0 t\]

where, as discussed earlier, the overbar represents the action of a low pass filter that removes the components at \(\pm 2\omega_0\).

Since we have specified \(n(t)\) to be a narrow-band Gaussian noise process, it may be written as [24,28]

\[(126) \quad n(t) = y(t) \cos\omega_0 t - z(t) \sin\omega_0 t\]
where

\begin{align}
(127) \quad y(t) &= n(t) \cos \omega_0 t + \hat{n}(t) \sin \omega_0 t \\
(128) \quad z(t) &= \hat{n}(t) \cos \omega_0 t - n(t) \sin \omega_0 t
\end{align}

and \( \hat{n}(t) \) denotes the Hilbert transform of \( n(t) \).

\begin{equation}
(129) \quad \hat{n}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{n(\tau)}{t-\tau} \, d\tau
\end{equation}

Then (125) becomes

\begin{equation}
(130) \quad f(t) = k_{AB} + 2kB \cos \omega_0 t [y(t) \cos \omega_0 t - z(t) \sin \omega_0 t]
\end{equation}

or

\begin{equation}
(131) \quad f(t) = k_{AB} + kBy(t)
\end{equation}

The driving process \( f(t) \) then consists of a constant term and a Gaussian distributed term. Note that \( y(t) \) (also \( z(t) \)) is Gaussian, zero mean with variance \( \sigma^2 \), having the same power (variance) as the input noise process \( n(t) \).

Thus the random forcing term is Gaussian distributed with mean

\begin{equation}
(132) \quad \mathbb{E}\{f(t)\} = k_{AB}
\end{equation}

and variance

\begin{equation}
(133) \quad \sigma_f^2 = k^2B^2\sigma^2
\end{equation}

Hence, its probability density function is

\begin{equation}
(134) \quad p_f(f) = \frac{1}{\sqrt{2\pi} k B \sigma} e^{-\frac{(f - k_{AB})^2}{2k^2B^2\sigma^2}}, \quad -\infty < f < +\infty
\end{equation}
Turning to the random coefficient \(a(t)\), we have

\[(135)\quad a(t) = 2k \left[ A \cos \omega_0 t + n(t) \right]^2\]

\[(136)\quad = kA^2 + 4kA \cos \omega_0 t [y(t) \cos \omega_0 t - z(t) \sin \omega_0 t] + 2k [y(t) \cos \omega_0 t - z(t) \sin \omega_0 t]^2\]

\[(137)\quad = kA^2 + 2kAy(t) + ky^2(t) + kz^2(t)\]

\[(138)\quad = k([A + y(t)]^2 + z^2(t)).\]

This is just a special case of the envelope squared of a randomly phased input signal plus narrowband noise process given by

\[(139)\quad q(t, \theta) = [A \cos \theta + y(t)]^2 + [A \sin \theta + z(t)]^2\]

Multiplying through by \(k\) and letting \(\theta = 0^\circ\) results in the random coefficient \(a(t)\) given by (138). The probability density function for \(q(t, \theta)/\sigma^2\) (dividing (139) by \(\sigma^2\) normalizes the Gaussian processes \(y(t)\) and \(z(t)\) to unit variance) is derived in several texts (for example, see Whelen [28, pp. 113-118]) and is

\[(140)\quad p_{q'}(q') = \frac{1}{2} \int_{q'}^{\infty} \left( \frac{A^2}{\sigma^2} + q' \right)^{\frac{1}{2}} I_0 \left( \frac{A}{\sigma \sqrt{q'}} \right) \, dq', \quad q' \geq 0\]

where

\[(141)\quad q' = q/\sigma^2\]
and

\[(142) \quad I_0(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{2^n n!^2}\]

is the modified zero order Bessel function. The distribution above is a special case of the non-central \(x^2\) (chi-squared) distribution. Note that since the result is independent of \(\theta\), it is valid for \(\theta=0^\circ\).

To get the exact density function for \(a(t)\) we note that since

\[(143) \quad a(t) = k\sigma^2 q^r(t,0^\circ)\]

and

\[(144) \quad p_a(a) = \frac{1}{|k\sigma^2|} p_{q^r}\left(\frac{a}{k\sigma^2}\right)\]

then

\[(145) \quad p_a(a) = \frac{1}{2k\sigma^2} \ e^{-\frac{1}{2k\sigma^2}} (kA^2+a)\]

Thus \(a(t)\) has the form of a modified non-central \(x^2\) distribution with 2 degrees of freedom. From the probability density function (145) or alternatively from the definition of \(a(t)\) (138), the random coefficient has a mean

\[(146) \quad E\{a(t)\} = kA^2 + 2k\sigma^2\]

and variance

\[(147) \quad \sigma_a^2 = 4k^2A^2\sigma^2 + 4k^2\sigma^4\].

To summarize, the differential equation satisfied by the weights for the case of a CW signal corrupted by narrowband, Gaussian noise \(n(t)\) is

\[(148) \quad \frac{dw(t)}{dt} + k \{[A+y(t)]^2+z^2(t)\} w(t) = kB[A+y(t)]\]
which is a stochastic, first order differential equation with a random coefficient

\[(149) \quad a(t) = k\{[A+y(t)]^2 + z^2(t)\}\]

and a random forcing function

\[(150) \quad f(t) = kB[A+y(t)]\]

with random processes \(y(t)\) and \(z(t)\) given by (127) and (128) respectively. Furthermore, we have shown that \(a(t)\) has the form of a non-central \(\chi^2\) distribution with 2 degrees of freedom and with mean and variance given by (146) and (147) respectively. Also \(f(t)\), the random forcing term, is Gaussian distributed with a mean given by (132) and variance by (133). Thus we have defined the problem as a Case II type of equation given by (55).

For the remainder of the chapter, we turn our attention to the weight solution process \(w(t)\) and determine its statistical properties.

In outline form, the following steps will be performed:

1. **A Direct Approach**: A Gaussian approximation for the coefficient process \(a(t)\) will be made and a direct approach using the characteristic function (see p. 22) will be used to develop expressions for the mean and autocorrelation function of \(w(t)\).

2. **Ito Equation Approach**: An alternate method using a Markov characterization for the weight process will be reviewed. The mean and variance of \(w(t)\) will be derived using Eqs. (113) and (114) referenced in Chapter III.

3. **A Stochastic Perturbation Approach**: Stochastic perturbation theory as developed in Chapter III (pp. 22-27) will be applied and will be shown to greatly simplify the analysis. Statistics for the weight solution will be derived for both a two term and three term perturbation expansion.

**B. A Direct Approach**

We have shown the equation obeyed by the weight process is

\[(151) \quad \frac{dw(t)}{dt} + k\{[A+y(t)]^2 + z^2(t)\} w(t) = kB[A+y(t)]\].
The solution process to this differential equation is

\[ w(t) = \int_0^t k[B + y(u)] e^{-\int_y^t k[A + y(v)] e^{-\int_y^t k[A + y(u)]} du} dv \]  

where zero initial conditions have been assumed. The presence of the modified non-central $x^2$ distributed random process in the coefficient term of (151) complicates the analysis considerably. As pointed out in Chapter III, without any simplifying assumptions, even the calculation of the simple statistical properties of $w(t)$ becomes almost impossible. In such cases, various approximation methods must be used; but before turning to these in the next sections, it will be instructive to look at a direct method that can be used to obtain the mean and autocorrelation function of $w(t)$.

In Chapter III, we showed (Eqs. (57) and (58)) that the mean and autocorrelation function for $w(t)$ could be written as

\[ E\{w(t)\} = \int_0^t E\left\{ f(v) e^{-\int_v^t a(u) du} \right\} dv \]

and

\[ R_{ww}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} E\left\{ f(v_1)f(v_2)e^{-\int_{v_1}^{t_1} a(u) du - \int_{v_2}^{t_2} a(u) du} \right\} dv_1 dv_2 \]

respectively. Also in the adaptive array problem we are considering, we showed from (149) and (150) that

\[ a(t) = k[A + y(t)]^2 + z^2(t) \]

and

\[ f(t) = kB[A + y(t)] \]
Now let us introduce the process (Soong [22])

\[(157) \quad s(t) = \int_0^t e^{-\alpha f(v)} - b(v) \, dv \]

where

\[(158) \quad b(v) = \int_v^t a(u) \, du \]

Then

\[(159) \quad w(t) = -\left[ \frac{\partial}{\partial \alpha} s(t) \right]_{\alpha=0} \]

Applying the expectation operator, we get

\[(160) \quad E\{w(t)\} = -\left[ \frac{\partial}{\partial \alpha} E\{s(t)\} \right]_{\alpha=0} \]

or upon substitution of (157) into (160), we get

\[(161) \quad E\{w(t)\} = -\int_0^t \frac{\partial}{\partial \alpha} \left[ E\{e^{-\alpha f(v)} - b(v)\} \right]_{\alpha=0} \, dv \]

Similarly, the autocorrelation function is

\[(162) \quad R_{ww}(t_1, t_2) = \left[ \frac{\partial^2}{\partial \alpha \partial \beta} E\{s(t_1)s(t_2)\} \right]_{\alpha=0, \beta=0} \]
or

\[
R_{WW}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \frac{e^{2}}{\beta a \beta} \left[ E\{e^{-\alpha f(v_1) - \beta f(v_2) - b(v_1) - b(v_2)}\} \right]_{\alpha=0}^{\beta=0} dv_1 dv_2
\]

But the integrands of (161) and (163) with a proper change of variables are in a familiar form. The integrand in (161) can be expressed in terms of the characteristic function of the random processes \(f(v)\) and \(b(v)\). Similarly, the integrand in (163) can be expressed in terms of the joint characteristic function of the random processes \(f(v_1)\), \(f(v_2)\), \(b(v_1)\) and \(b(v_2)\). Thus with

\[
\phi_{fb}(\omega_1, \omega_2) = E\{e^{j\omega_1 f + j\omega_2 b}\}
\]

and

\[
\phi_{ffbb}(\omega_1, \omega_2, \omega_3, \omega_4) = E\{e^{j\omega_1 f + j\omega_2 f + j\omega_3 b + j\omega_4 b}\}
\]

defining the joint characteristic functions, then

\[
\phi_{fb}(\alpha, j) = E\{e^{-\alpha f(v) - b(v)}\}
\]

and

\[
\phi_{ffbb}(\alpha, \beta, j, j) = E\{e^{-\alpha f(v_1) - \beta f(v_2) - b(v_1) - b(v_2)}\}
\]

are the integrands of (161) and (163). We can make further progress if \(f(t)\) and \(b(t)\) are both Gaussian random processes. In the adaptive array problem, this assumption will hold when the noise is small compared with the desired signal.
To see this, let the input signal be represented by

\[(168) \quad x(t) = A \cos \omega_0 t + \epsilon n'(t)\]

where \(\epsilon\) is a small parameter and \(n'(t)\) is a zero mean Gaussian noise process with variance \(\sigma^2\). Let us suppose that

\[(169) \quad A^2 \gg \epsilon^2 \sigma^2\]

or that the power of the desired input signal is very large compared to the power of the noise process. Then the random coefficient \(a(t)\) becomes (see (155))

\[(170) \quad a(t) = k[A + \epsilon y'(t)]^2 + \epsilon^2 z(t)^2\]

or

\[(171) \quad a(t) = kA^2 + \epsilon^2 k \epsilon y'(t) + \epsilon^2 k (y'(t)^2 + z(t)^2)\]

If the terms containing \(\epsilon^2\) are neglected, the random coefficient becomes

\[(172) \quad a(t) = kA^2 + \epsilon^2 kA y'(t)\]

Similarly, the random forcing term (156) is

\[(173) \quad f(t) = kAB + \epsilon kB y'(t)\]

Since \(y'(t)\) is a Gaussian, wide sense stationary random process with zero mean and variance \(\sigma^2\) (see pp.34-35), then \(a(t)\) and \(f(t)\) are also Gaussian with means

\[(174) \quad \mu_a = E[a(t)] = kA^2\]

\[(175) \quad \mu_f = E[f(t)] = kAB\]
and variances

(176) \[ \sigma_a^2 = \varepsilon 2k^2A^2a^2 \]

(177) \[ \sigma_f^2 = \varepsilon 2k^2B^2f^2 \]

We note that \( a(t) \) and \( f(t) \) given by (172) and (173) are both derived from the same input random process \( n'(t) \) and hence are not statistically independent. They are correlated and have a correlation coefficient

(178) \[ \rho_{af} = \frac{E\{[a(t_1)-\mu_a][f(t_2)-\mu_f]\}}{\sigma_a \sigma_f} \]

where \( \mu_a, \mu_f, \sigma_a \) and \( \sigma_f \) are defined above. Expanding this expression, we have

(179) \[ \rho_{af} = \frac{E\{a(t_1)f(t_2)\} - \mu_a \mu_f}{\sigma_a \sigma_f} \]

(180) \[ = \frac{E\{[kA^2+\varepsilon 2kAy'(t_1)][kAB+\varepsilon kBy'(t_2)]-(kA^2)(kAB)\}}{(\varepsilon 2kA \sigma)(\varepsilon kB \sigma)} \]

(181) \[ = \frac{\varepsilon 2k^2AB R_y'y_1'(\tau)}{\varepsilon 2k^2AB \sigma^2} \]

(182) \[ = \frac{R_y'y_1'(\tau)}{\sigma^2} = \frac{R_y'y_1'(\tau)}{R_y'y_1'(0)} = \rho_{y'y_1'(\tau)} \]
If \( a(t) \) is Gaussian and we assume the integral in (158) exists, then the \( b(t) \) process will also be Gaussian. So with the restriction that the noise power be small compared to the desired signal power, we have shown that \( b(t) \) and \( f(t) \) are both correlated Gaussian random processes. Hence their joint characteristic function is

\[
\phi_{fb}(\omega_1, \omega_2) = e^{-\frac{1}{2} \left( \sigma_f^2 \omega_1^2 + 2 \rho_{fb} \sigma_f \sigma_b \omega_1 \omega_2 + \sigma_b^2 \omega_2^2 \right)}
\]

or using the change of variable in (166), we get

\[
\phi_{fb}(\alpha_f \omega, \omega_b) = e^{-\frac{1}{2} \left( \alpha_f^2 \sigma_f^2 + 2 \rho_{fb} \sigma_f \sigma_b + \sigma_b^2 \right)}
\]

where \( \mu_f \) and \( \sigma_f^2 \) have been previously defined in (175) and (177)

\( \mu_b \) is the mean of the \( b(t) \) process defined by (158)

\( \sigma_b^2 \) is the variance of the \( b(t) \) process

and

\( \rho_{fb} \sigma_f \sigma_b = E[(f(t)-\mu_f)(b(t)-\mu_b)] \) is the covariance function of the \( f(t) \) and \( b(t) \) processes.

Substituting (184) into (161), we get for the mean of the \( w(t) \) process

\[
E\{w(t)\} = \int_0^t - \frac{2}{\alpha} \left[ e^{-\left( \alpha \mu_f + \mu_b \right) + \frac{1}{2} \left( \alpha \sigma_f^2 + 2 \rho_{fb} \sigma_f \sigma_b + \sigma_b^2 \right)} \right] dv
\]

Carrying out the differentiation and letting \( \alpha = 0 \), we get

\[
E\{w(t)\} = \int_0^t [\mu_f - \rho_{fb} \sigma_f \sigma_b] e^{-\mu_b} + \frac{1}{2} \sigma_b^2] dv
\]

We have already defined \( \mu_f \) and \( \sigma_f^2 \) in (175) and (177) respectively. Also, from (158) and (174), we have for the mean of the \( b(t) \) process
\[ \mu_b(t,v) = E\{b(v)\} \]
\[ = \int_v^t \mu_a \, du \]
\[ = kA^2(t-v) \]

Similarly, from (158) and (176), the variance of \( b(t) \) is

\[ \sigma^2_{b}(t,v) = E\{b^2(v)\} - E\{b(v)\}^2 \]
\[ = \sigma_a^2 \int_v^t \int_v^t E\{y'(u_2)y'(u_1)\} \, du_2 \, du_1 \]
\[ = \epsilon^2 4k^2 A^2 \int_v^t \int_v^t R_{y' y'}(u_2-u_1) \, du_2 \, du_1 \]

The covariance function \( \rho_{fb} \sigma_f \sigma_b \) defined in (184) becomes

\[ \rho_{fb}(t,v) \sigma_f \sigma_b(t,v) = E\{[f(t)-\mu_f][b(v)-\mu_b]\} \]
\[ = E\{f(t)b(v)\} - \mu_f \mu_b \]

From (189) and (158), (194) becomes

\[ \rho_{fb}(t,v) \sigma_f \sigma_b(t,v) = \int_v^t E\{f(t)a(u)\} \, du - \mu_f \mu_a(t-v) \]
(196) \[ \int_v^t \mathcal{E}\{[f(t) - u_f][a(u) - u_a]\} \, du \]

(197) \[ \sigma_f \sigma_a \int_v^t \rho_{fa}(t-u) du \]

or finally, using (176), (177) and (182), we get

(198) \[ \rho_{fb} \sigma_f \sigma_b = e^{2k^2AB} \int_v^t R_{y'y'}(t-u) du \]

Now substituting (175), (189), (192) and (198) into (186), we get finally for the mean of the weight process

(199) \[ E\{w(t)\} = \int_0^t [kAB - e^{2k^2AB} \int_v^t R_{y'y'}(t-u_3) du_3] \]

\[ -kA^2(t-v) + e^{2k^2A^2} \int_v^t \int_v^t R_{y'y'}(u_2-u_1) du_2 du_1 \]

\[ \times e^{kA^2(t-v)} \varepsilon^{2k^2A^2} \int_v^t \int_v^t R_{y'y'}(u_2-u_1) du_2 du_1 \]

Thus knowing the autocorrelation function \( R_{y'y'}(t) \) of the Gaussian random process \( y'(t) \) allows one to calculate the mean of the weight solution process. Note that if the terms containing \( \varepsilon^2 \) are neglected (recall \( \varepsilon \) is a small parameter), then the expression for the mean reduces to the completely deterministic case

(200) \[ E\{w(t)\} = \int_0^t kAB e^{-kA^2(t-v)} dv \]

(201) \[ = \frac{B}{A} (1 - e^{-kA^2t}) \]
It is also interesting to note that although we assumed the additive noise process was on an order of ε compared to the input signal (ref. (168)), the expected value of the weight shows a noise effect on the order of ε^2. This result can be attributed to the squaring and filtering action in the feedback loop. Thus the effect of the additive noise n'(t) with the input signal is minimal in the behavior of the weight in the array when ε is small.

Similarly, one can go through the same type of analysis for the autocorrelation function R_{ww}(t_1,t_2) using the four dimensional characteristic function (167) and the defining equation (163). Rather than go through the lengthy analysis, we will quote the results of Soong [22, p. 221]:

\begin{equation}
R_{ww}(t_1,t_2) = \int_0^{t_1} \int_0^{t_2} \sigma_f^2 \rho_{ff}(v_1-v_2) + \left[ \mu_f - \sigma_f \sigma_b(t_1,v_1) \right] \rho_{af}(v_1-s_1) ds_1 \\
- \sigma_f \sigma_b(t_2,v_1) \left[ \int_0^{t_2} \rho_{af}(v_1-s_2) ds_2 \right] \left[ \mu_f - \sigma_f \sigma_b(t_1,v_2) \right] \rho_{af}(v_2-s_1) ds_1 \\
- \sigma_f \sigma_b(t_2,v_2) \left[ \int_0^{t_2} \rho_{af}(v_2-s_2) ds_2 \right] \exp[-\mu_a(t_1+t_2-v_1-v_2)] + \frac{1}{2} \sigma_b^2(t_1,v_1) \\
+ \frac{1}{2} \sigma_b^2(t_2,v_2) + \sigma_b(t_1,v_1) \sigma_b(t_2,v_2) \int_0^{t_1} \int_0^{t_2} \rho_{aa}(s_2-s_1) ds_2 ds_1 \right] dv_1 dv_2
\end{equation}

where μ_f, σ_f^2, \mu_a, σ_b^2 and ρ_{af} were previously defined in (175), (177), (174), (192) and (182) respectively. Also, clearly from (172) and (173)

\begin{equation}
\rho_{aa}(s_2-s_1) = \frac{R_{yy}(s_2-s_1)}{\sigma_a^2}
\end{equation}
and

\begin{equation}
(204)\quad \rho_{ff}(v_1-v_2) = \frac{R_{y'y'}(v_1-v_2)}{\sigma^2}
\end{equation}

As with the mean, if those terms containing \( \epsilon^2 \) are neglected, the above reduces to the deterministic case

\begin{equation}
(205)\quad R_{WW}(t_1,t_2) = \int_0^{t_1} \int_0^{t_2} (kAB)^2 e^{-kA^2(t_1+t_2-v_1-v_2)} \, dv_1 \, dv_2
\end{equation}

or

\begin{equation}
(206)\quad R_{WW}(t_1,t_2) = \frac{B^2}{A^2} \left(1 - e^{-kA^2t_1}\right)\left(1 - e^{-kA^2t_2}\right)
\end{equation}

Thus in the special case of a Gaussian random coefficient, it is possible to obtain the mean and autocorrelation function of the weight process \( w(t) \) by using Eqs. (199) and (202). However, the complexity of these expressions is evident in this direct approach and could be difficult to evaluate depending on the autocorrelation function of the \( y'(t) \) process. For this reason and the fact that in many cases the random coefficient is not Gaussian, approximation methods must be employed. We will now turn to a second method to obtain the weight statistics, namely that employing a Markov characterization approach for the weight process.

C. Ito Equation Approach

In the previous section, we examined a direct approach using the basic definition for the weight solution process \( w(t) \) (152) to obtain equations for the mean and autocorrelation function. We further showed that if the random coefficient was Gaussian, we could use the well known joint characteristic function for Gaussian random processes to derive expressions (199) and (202) (although complicated) for these statistics in terms of the autocorrelation of the \( y'(t) \) process defined in (127). A second technique involves modelling the weight differential equation in the
form of an Ito equation defined by (101). As discussed in Chapter III, the Ito equation is defined in terms of the Wiener process (96). Therefore we must assume that the random portions of the defining differential equation for the weights are "white noise" processes with flat power spectra extending over all real frequencies. This requirement can be seen more clearly by studying the example discussed on pp. 28-29 in Chapter III. If the weight process can be modelled as an Ito equation, then \( w(t) \) is Markov and we can use the Fokker-Planck diffusion equation (109) to calculate the probability density function \([22, 25, 26, 27]\). We may also use the moment equation derivatives (113) and (114) to calculate the mean and variance.

This approach to the adaptive array problem was first proposed by Berni [12] and he solved the weight equation given by

\[
\frac{dw(t)}{dt} + [k_1 x^2(t) + k_2] w(t) = w_0
\]

where \( k_1 \) and \( k_2 \) are constants
\( x(t) \) is the input signal process
\( w_0 \) is a constant offset weight
and
\( w(t) \) is the weight

This is a somewhat different problem than we have been studying in that Berni uses a zero reference signal \( r(t)=0 \) but includes a constant offset weight \( w_0 \) and the gain term \( k_2 \). Note that there is no random forcing term in (207) which simplifies the analysis.

In our problem, we have shown the weight equation is

\[
\frac{dw(t)}{dt} = -k[[A+y(t)]^2+z^2(t)]w(t) + kB[A+y(t)]
\]

In the special case when the input noise process is small compared to the desired signal (see (168)), we see that from (171) and (173) that (208) can be written as

\[
\frac{dw(t)}{dt} = -k[A^2+e2Ay'(t)+e^2(y'2(t)+z'2(t))]w(t)
\]

\[
+ kB[A + eY'(t)].
\]
The requirement that the noise term be small is not necessary when using the Ito approach but is used here to facilitate the comparison of the three methods described in this report.

In order to model (209) in terms of an Ito equation, the general form of which is

\[ \text{dw}(t) = f(w(t), t) dt + G(w(t), t) \, dx(t) \]

where \( x(t) \) is a Wiener process and \( f \) and \( G \) are real functions of the weight \( w(t) \) and time \( t \), we must replace the random portions of both the coefficient and forcing term with a zero mean white noise process. For the random portion of the forcing term (which is already zero mean), then let

\[ e^{kBy'}(t) + e^{2k(t)} \]

where \( y_1(t) \) is a white noise process with constant power spectrum

\[ S_{y_1y_1}(\omega) = N_1, \quad -\infty \leq \omega \leq +\infty. \]

Similarly, let

\[ e^{2kA}y''(t) + e^{2k(y''(t) + z^2(t))} + e^{2y_2(t)} + e^{2k\sigma^2} \]

where \( y_2(t) \) is a white noise process with constant power spectrum

\[ S_{y_2y_2}(\omega) = N_2, \quad -\infty \leq \omega \leq +\infty. \]

Note that we have separated out the non-zero mean portion \( e^{2k\sigma^2} \) of the random process. At this point we may question the validity of the two substitutions (211) and (213). In reality, they mean we are eliminating the low pass filter in the feedback loop of the array (see Fig. 5). Practically speaking, that change will not affect the results much. But we are also extending the bandwidth of the narrowband process \( n'(t) \) in (168) to infinity. Finally by using (213) we are changing the probability density function of the coefficient term \( a(t) \). These last two changes are important in the array problem. For example, we can no
longer determine the effect of varying noise bandwidth on the weight process. But even with these limitations, the Ito approach provides a solution to the random coefficient problem. For wideband noise, it will provide the weight statistics which show how the noise term in (168) affects the weights. We now proceed with the Ito approach.

If we substitute (211) and (213) into the weight differential equation (209), we have

$$\frac{dw(t)}{dt} = -\{kA^2 + \epsilon_2 y_2(t) + \epsilon_2^2 2k\sigma^2\} w(t) + kAB + \epsilon_1 y_1(t).$$

This can be put in the form

$$dw(t) = -\{kA^2 + \epsilon^2 2k\sigma^2\} w(t) dt + \epsilon_1 dx_1(t)$$

$$- \epsilon_2^2 w(t) dx_2(t)$$

where $x_1(t)$ and $x_2(t)$ are Wiener processes defined by (96). But this is just the form of an Ito equation (210) with

$$f(w(t), t) = -\{kA^2 + \epsilon_2 2k\sigma^2\} w(t) - kAB$$

$$G(w(t), t) = [\epsilon_1 - \epsilon_2 w(t)]$$

and

$$dx(t) = \begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix}.$$  

As discussed earlier (p.30), the solution process to the Ito equation is Markov and its probability density function satisfies the Fokker-Planck equation (109).
\[
\frac{ap(w,t)}{\partial t} = - \frac{a}{\partial w} [p(w,t)f(w(t),t)] + \frac{1}{2} \frac{a^2}{\partial w^2} [p(w,t)\hat{G}(w(t),t)D\hat{G}(w(t),t)^T]
\]

where the matrix \( \hat{D} \) satisfies

\[
E\{dx(t) dx(t)^T\} = \hat{D} dt.
\]

Equation (220) gives the time varying probability density function \( p(w,t) \) for the weights. In the steady state, when \( w(t) \) has achieved stationarity and since \( f \) and \( \hat{G} \) are functions only of the weights in our problem, the probability density function satisfies

\[
0 = - \frac{d}{dw} [p(w)f(w)] + \frac{1}{2} \frac{d^2}{dw^2} [p(w)\hat{G}(w)\hat{G}(w)^T]
\]

or

\[
\frac{d}{dw} [p(w)\hat{G}(w)\hat{G}(w)^T] - 2p(w)f(w) = C_1
\]

where \( C_1 \) is a constant of integration and \( p(w) \) is the steady state probability density function.

The remaining unknown in the problem involves the matrix \( \hat{D} \) defined by (221). Let us write \( D \) as

\[
\hat{D} = \begin{bmatrix}
N_1 & \rho\sqrt{N_1N_2} \\
\rho\sqrt{N_1N_2} & N_2
\end{bmatrix}
\]

where \( N_1 \) and \( N_2 \) were defined in (212) and (214) and \( \rho \) is a correlation coefficient for the two processes \( x_1(t) \) and \( x_2(t) \). (We assume here that the \( D \) matrix is non-time varying and that \( \rho_{21} = \rho_{12} = \rho \)). Then using (218), we have
\[ (225) \quad \hat{G}(w) \hat{D}G(w)^T = \epsilon_1^2 N_1 - 2\epsilon_1 \epsilon_2 p \sqrt{N_1 N_2 w} + \epsilon_2^2 N_2 w^2 \]

Note that if \( x_1(t) \) and \( x_2(t) \) were statistically independent, then \( \rho = 0 \) and

\[ (226) \quad \hat{G}(w) \hat{D}G(w)^T = \epsilon_1^2 N_1 + \frac{2}{1} N_2 w^2 \]

Substituting (217) and (225) into (223) gives

\[ (227) \quad \frac{dp(w)}{dw} + \left[ \frac{(2\epsilon_2^2 N_2 + 2kA^2 + \epsilon_2^2 4k\sigma^2)w - (2\epsilon_1 \epsilon_2 p \sqrt{N_1 N_2 w} + 2kAB)}{\epsilon_1^2 N_1 - 2\epsilon_1 \epsilon_2 p \sqrt{N_1 N_2 w} + \epsilon_2^2 N_2 w^2} \right] p(w) = \frac{C_1}{\epsilon_1^2 N_1 - 2\epsilon_1 \epsilon_2 p \sqrt{N_1 N_2 w} + \epsilon_2^2 N_2 w^2} \]

or using a simpler notation

\[ (228) \quad \frac{dp(w)}{dw} + \left[ \frac{a_w w + b_o}{a + bw + cw^2} \right] p(w) = \frac{C_1}{a + bw + cw^2} \]

where

\[ (229) \quad a_o = 2\epsilon_2^2 N_2 + 2kA^2 + 4\epsilon_2^2 k\sigma^2 \]

\[ (230) \quad b_o = -(2\epsilon_1 \epsilon_2 \sqrt{N_1 N_2} + 2kAB) \]

\[ (231) \quad a = \epsilon_1^2 N_1 \]
\[(232) \quad b = -2 \varepsilon_1 \varepsilon_2 \rho \sqrt{N_1 N_2}\]

\[(233) \quad c = \varepsilon_2^2 N_2\]

The solution to (228) is

\[(234) \quad p(w) = \frac{1}{R} \int \frac{c_1 R}{a + bw + cw^2} \, dw + \frac{c_2}{R}\]

where the integrating factor is

\[(235) \quad R = e^{\frac{a_0 w + b_0}{a + bw + cw^2}}\]

\[(236) \quad = (a + bw + cw^2)^{\frac{a_0}{2c}} e^{-\left(\frac{a_0 b}{2c} - b_0\right) \frac{2}{\sqrt{4ac - b^2}}} \tan^{-1} \frac{2cw + b}{\sqrt{4ac - b^2}}\]

where \(c_1, c_2\) are constants of integration. The conditions necessary to solve for the constants of integration are

\[(237) \quad p(\infty) = 0\]

\[(238) \quad \int_0^\infty p(w) \, dw = 1\]

Although a first order system, the solution for \(p(w)\) given by (234) is still a formidable problem. [Note that in the special case]
\( e_1 = 0 \) in (227), the problem reduces to the same form as (207) solved by Berni [12]. The steady state probability density function is then

\[
(239) \quad p(w) = \frac{1}{\beta I(\alpha-1)} \left( \frac{\beta}{w} \right)^\alpha e^{-\frac{\beta}{w}}, \ w \geq 0
\]

where

\[
(240) \quad \beta = \frac{2kAB}{\epsilon_2^2N_2}
\]

\[
(241) \quad \alpha = \frac{2\epsilon_2^2N_2 + 2kA^2 + 4\epsilon_2^2k\sigma^2}{\epsilon_2^2N_2}
\]

and

\( \Gamma(\cdot) \) is the Gamma function.

Fortunately, one of the attributes of the Ito approach is that the statistics for \( w(t) \) can be solved for directly without a knowledge of the probability density function. Since our interest in this study are the weight statistics, we will proceed in this way.

The specific formulae developed for calculating the mean and variance of the solution \( w(t) \) to the Ito equation (210) were given in Chapter III, Eqs. (113) and (114). For the mean we have

\[
(242) \quad \frac{dE[w(t)]}{dt} = E[f(w(t), t)]
\]
Substituting in for \( f(w(t), t) \) from (217), we get

\[
\frac{dE[w(t)]}{dt} = E[-(kA^2 + \varepsilon^2 2k\sigma^2)w(t) + kAB]
\]

or

\[
\frac{dE[w(t)]}{dt} + (kA^2 + \varepsilon^2 2k\sigma^2) E[w(t)] = kAB
\]

The solution to this equation is

\[
E[w(t)] = \frac{AB}{A^2 + \varepsilon^2 2\sigma^2} [1 - e^{-(kA^2 + \varepsilon^2 2k\sigma^2)t}]
\]

For \( t \) large, \( E[w(t)] \) has a constant value given by

\[
E[w(t)]_{\text{Ito}} = \frac{AB}{A^2 + \varepsilon^2 2\sigma^2}
\]

\[
= \frac{B}{A} - \frac{\varepsilon^2 2B\sigma^2}{A^3 + \varepsilon^2 2A\sigma^2}
\]

Note that since \( A^2 >> \varepsilon^2 \sigma^2 \) (see (169)), the mean is approximately the result for the deterministic case

\[
E[w(t)] \approx \frac{B}{A}
\]

This result agrees with the discussion (ref. p. 46) of the mean using the direct approach. Figure 9 shows a plot of \( E[w(t)] \) versus noise variance \( \varepsilon^2 \sigma^2 \) for \( A=B=1 \). The addition of the noise results in a mean weight \( E[w(t)] \) below that obtained in the deterministic case. Also, increasing the noise power causes the mean to deviate farther below the deterministic value. Finally,
from (244), we see the mean weight depends only on the mean of the coefficient process \(a(t)\) and random forcing term \(f(t)\). It does not depend explicitly on the random portion of either process, here assumed to be white noise with power density spectrum \(N_0\). Since we are assuming \(A^2 \gg \epsilon^2 \sigma^2\), the amount of deviation below the deterministic value is small enough to be neglected in most cases. We will therefore use (248) for the mean value of the weight process.

![Graph](image)

**Fig. 9.** \(E\{w(t)\}\) vs. noise power.
Turning to the variance, we have from (114)

\[
\frac{d\sigma_{w}^2}{dt} = 2[E\{w(t)f(w(t))\} - E\{w(t)\}E\{f(w(t))\}] + E\{\hat{G}(w(t))\} \hat{D} \hat{G}(w(t))
\]

Substituting for \(f(w(t))\) from (217) and for \(\hat{G}(w(t))\) \(\hat{D} \hat{G}(w(t))\) from (225), we have

\[
\frac{d\sigma_{w}^2}{dt} = 2[E\{(-kA_w^2 - \epsilon^2 2k\sigma_w^2 + kAB)w\} - E\{w\}E\{-(kA_w^2 + \epsilon^2 2k\sigma_w^2)w + kAB\}] + E\{\epsilon_1^2 N_1 - 2\epsilon_1 \epsilon_2 \rho \sqrt{N_1 N_2} + \epsilon_2^2 N_2 \}
\]

\[
= 2[(-kA^2 + 2k\sigma^2)E\{w^2\} + (kA^2 + \epsilon^2 2k\sigma^2)E\{w\}^2]
\]

\[
+ \epsilon_1^2 N_1 - 2\epsilon_1 \epsilon_2 \rho \sqrt{N_1 N_2} E\{w\} + \epsilon_2^2 N_2 E\{w^2\}
\]

\[
- \epsilon_0^2 N_2 E\{w\}^2 + \epsilon_2^2 N_2 E\{w\}^2
\]

or

\[
\frac{d\sigma_{w}^2}{dt} + 2(kA^2 + \epsilon^2 2k\sigma^2)\sigma_{w}^2 = \epsilon_1^2 N_1 - 2\epsilon_1 \epsilon_2 \rho \sqrt{N_1 N_2} E\{w\}
\]

\[
+ \epsilon_2^2 N_2 E\{w\}^2 + \epsilon_2^2 N_2 \sigma_{w}^2
\]

Finally, rearranging we have

\[
\frac{d\sigma_{w}^2}{dt} + [2(kA^2 + \epsilon^2 2k\sigma^2) - \epsilon_2^2 N_2] \sigma_{w}^2
\]

\[
= \epsilon_1^2 N_1 - 2\epsilon_1 \epsilon_2 \rho \sqrt{N_1 N_2} E\{w\} + \epsilon_2^2 N_2 E\{w\}^2
\]
The steady state solution is

\[ \sigma_w^2 = \frac{\varepsilon_1^2 N_1 - 2 \varepsilon_1 \varepsilon_2 \rho \sqrt{N_1 N_2} E[w] + \varepsilon_2^2 N_2 E[w]^2}{2 k A^2 + 4 \varepsilon_2^2 k \sigma^2 - \varepsilon_2^2 N_2} \]  

Let us assume that the random coefficient noise and forcing term noise are fully correlated (ref. pp. 42-43). Then we have \( \rho = 1 \). To facilitate comparison with the stochastic perturbation results of the next section, let us replace \( \varepsilon_1 \) with \( \varepsilon k B \) (from (211)) and \( \varepsilon_2 \) with \( \varepsilon_2 k A \) (from (213)). Finally, assuming \( N_1 = N_2 = N_0 \), (254) becomes

\[ \sigma_w^2 = \frac{\varepsilon k^2 B^2 N_0 - 2 \varepsilon k^2 A \tilde{B} N_0 E[w] + 4 \varepsilon k^2 A^2 N_0 E[w]^2}{2 k A^2 + 4 \varepsilon_2^2 k \sigma^2 - 4 \varepsilon_2^2 k^2 A^2 N_0} \]

Using \( B/A \) for \( \tilde{B} \) from (248), we get

\[ \sigma_w^2 = \frac{3 \varepsilon k^2 B^2 N_0}{2 k A^2 + 4 \varepsilon^2 k \sigma^2 - 4 \varepsilon^2 k^2 A^2 N_0} \]

With \( \varepsilon^2 k N_0 \ll A^2 \) and \( A = B = 1 \), Fig. 10 shows the weight variance plotted for different values of the loop gain—noise power spectral density product \( \varepsilon^2 k N_0 \). For \( \varepsilon \) small, increasing \( k N_0 \) results in nearly a linear relationship with increasing weight jitter (variance). We will discuss these results further at the end of the next section.

The key drawback with this technique is that we have no measure of its validity for finite bandwidth noise inputs. The next section on stochastic perturbation theory overcomes this problem.
Fig. 10. Weight variance vs loop product term $\varepsilon^2 k N_0$. 

$A = B = 1$
D. Stochastic Perturbation Approach

Returning once again to the differential equation obeyed by the weights

\[
\frac{dw(t)}{dt} + k([A+y(t)]^2+z^2(t))w(t) = kB[A+y(t)]
\]

we have stated that the principal obstacle in obtaining the solution process for \( w(t) \) is the fact that the coefficient term is a random process - in this case, non-central \( x^2 \) distributed. In the previous sections, we showed that of the coefficient process were assumed to be Gaussian, a direct but cumbersome approach utilizing characteristic functions could be used to obtain expressions for the mean and autocorrelation function. A second approach was to remove all bandwidth constraints and assume the input noise to be white. This allowed us to formulate the problem in terms of an Ito equation and calculate the statistics using standardized formulae. We now consider a third approach using the stochastic perturbation theory developed in Chapter III.

To do this, we first consider the following: assume that the incoming signal plus noise process is characterized by a signal component perturbed by a small, random Gaussian distributed noise term; that is let

\[
x(t) = A \cos \omega_0 t + \varepsilon n'(t)
\]

with \( \varepsilon \) being a small parameter. Note that this is the same approximation that was made using the direct approach (ref. p. 42). Hence the comments made as to the interpretation of (168) are to apply here also. From (169), we require \( A^2 \gg \varepsilon^2 \sigma^2 \) where \( \sigma^2 \) is the variance of the noise process. The coefficient term \( a(t) \) can now be written as

\[
a(t) = k[A^2+\varepsilon^2A y^2(t) + \varepsilon^2(y^2(t)+z^2(t))]
\]

\[
= a_0 + \varepsilon a_1(t) + \varepsilon^2 a_2(t)
\]

where

\[
a_0 = kA^2
\]
(262) \[ a_1 = 2kAy'(t) \]

(263) \[ a_2 = k(y'^2(t) + z^2(t)) \]

The random forcing term is

(264) \[ f(t) = kB[A + \varepsilon y'(t)] \]

(265) \[ = f_0 + \varepsilon f_1(t) \]

where

(266) \[ f_0 = kAB \]

(267) \[ f_1(t) = kBy'(t) \]

We note that (262) is a truncated version of (62) with \( a_0 \) being an unperturbed constant term and \( a_1(t) \) and \( a_2(t) \) being perturbing random processes.

Following the procedure developed in Chapter III, pp. 22-24 we arrive at the following set of differential equations:

(268) \[ \frac{d}{dt}w_0(t) + kA^2 w_0(t) = kAB \]

(269) \[ \frac{d}{dt}w_1(t) + kA^2 w_1(t) = kBy'(t) - 2kAy'(t)w_0(t) \]
where in the stochastic perturbation method, we have assumed the series expansion

\( w(t) = w_0(t) + \epsilon w_1(t) + \epsilon^2 w_2(t) + \cdots \)

Initially we will assume a two term perturbation expansion for \( w(t) \). By determining the spectral characteristics and auto-correlation function for \( w_0(t) \) and \( w_1(t) \) from (268) and (269), these same statistics can then be developed for \( w(t) \). The outstanding feature of this approach is that we can analyze the weight statistics using narrowband noise. This differs from the white noise approximation used in the Itô approach.

After comparing the results of the two methods, we will then proceed to a three term perturbation expansion by analyzing (270). Again, the results will be compared.

To analyze each of the constant coefficient equations (268), (269), etc., we will assume each of the \( w_i \) are the output from a linear system with system function (see (40))

\[ H(j\omega) = \frac{1}{a_0 + j\omega} \]
For a stationary driving function, the output power density spectrum is then

\[ (275) \quad S_{w_iw_i}(\omega) = S_{r_ir_i}(\omega) |H(j\omega)|^2. \]

\[ (276) \quad S_{r_ir_i}(\omega) = \frac{S_{r_ir_i}(\omega)}{a_0^2 + \omega^2}. \]

\(S_{r_ir_i}(\omega)\) denotes the power density spectrum of the input forcing function. Thus, it is the power spectrum of the right hand side of each of the equations (268), (269), etc., in the set. Similarly, from (46), the mean of the \(w_i(t)\) are

\[ (277) \quad E{w_i(t)} = \frac{E{r_i(t)}}{a_0}. \]

where \(E{r_i(t)}\) is the mean of the forcing function or right hand side of (268), (269) etc. Finally using (261), we have

\[ (278) \quad S_{w_iw_i}(\omega) = \frac{S_{r_ir_i}(\omega)}{k^2A^4 + \omega^2}. \]

for the output power spectrum and

\[ (279) \quad E{w_i(t)} = \frac{E{r_i(t)}}{kA^2}. \]

for the mean. We must use the steady state value for each of the \(w_i(t)\) since the output statistics given by (278) and (279) assume stationarity has been achieved (ref discussion pp. 19 - 20). 3

3 A similar analysis can be made without this restriction using two dimensional Fourier transform pairs, \(R_{ww}(t_1, t_2)\) and \(I(\omega_1, \omega_2)\). This analysis would be required to study the transient behavior of the \(w_i(t)\) processes -see Papoulis [24, Chapter 12].
Turning to the first equation (268), we have

\[
\frac{dw_0(t)}{dt} + kA^2w_0(t) = kAB
\]

This equation is completely deterministic and the solution is

\[
w_0(t) = \frac{B}{A} \left( 1 - e^{-kA^2t} \right)
\]

In the steady state, we have

\[
w_0(t) = w_0 = \frac{B}{A}
\]

The power density spectrum of a constant or D.C. term is a delta function $\delta(\omega)$ at zero frequency. Here it will have a weight $2\pi B^2/A^2$ - see Fig. 11. Hence the $w_0(t)$ process has an average power of

\[
P_{w_0} = \frac{B^2}{A^2}
\]

![Diagram](attachment:image.png)

**Fig. 11.** Power spectrum of constant weight term $w_0$.  

65
Turning to the second differential equation (269) which was obtained by equating terms containing \( t \), we have

\[
(284) \quad \frac{dw_1(t)}{dt} + kA^2 w_1(t) = kBy'(t) - 2kAy'(t)w_0(t)
\]

Using the steady state value of \( B/A \) for \( w_0(t) \), (284) becomes

\[
(285) \quad \frac{dw_1(t)}{dt} + kA^2 w_1(t) = -kBy'(t) = r_1(t)
\]

where \( y'(t) \) is a zero mean, Gaussian random process with variance \( \sigma^2 \) and we define the forcing function as \( r_1(t) \).

From (279), we see \( w_1(t) \) is a zero mean random process. Also, \( w_1(t) \) will be Gaussian distributed since the driving function \(-kBy'(t)\) is Gaussian and the system is linear. The autocorrelation function of the forcing process \( r_1(t) \) is

\[
(286) \quad R_{r_1r_1}(\tau) = E[r_1(t)r_1(t+\tau)]
\]

\[
(287) \quad = k^2B^2R_{y'y'}(\tau)
\]

We have not as yet specified the autocorrelation function of the \( y'(t) \) process; only that \( R_{y'y'}(\tau) = \sigma^2 \). We also know that \( y'(t) \) is the quadrature component of the narrowband noise process \( n'(t) \) (ref discussion p. 34). Let us assume that the original input noise process \( n'(t) \) has a power density spectrum (see Fig. 12).

\[
(288) \quad S_{n'n'}(\omega) = \begin{cases} 
N_o/2 & -\omega_c < \omega < -\omega_c, \omega - \omega_c < \omega < \omega + \omega_c \\
& \text{elsewhere}
\end{cases}
\]
Since \( n(t) \) has an even, symmetrical power spectrum about the carrier frequency \( \omega_0 \), then its power spectrum can be written in terms of the power spectrum of the quadrature noise process \( y'(t) \) as [24,28]:

\[
S_{n'n'}(\omega) = \frac{1}{2} \left[ S_{y'y'}(\omega + \omega_0) + S_{y'y'}(\omega - \omega_0) \right]
\]

The power spectrum of the \( y'(t) \) process will then be (see Fig. 13)

\[
S_{y'y'}(\omega) = \begin{cases} 
N_0 & |\omega| \leq \omega_c \\
0 & \text{elsewhere},
\end{cases}
\]

The corresponding autocorrelation function for \( y'(t) \) is

\[
R_{y'y'}(\tau) = N_0 \frac{\sin \omega_c \tau}{\pi \tau}
\]
The variance is then

\begin{equation}
R_{y'y'}(0) = \sigma^2 = \frac{N_0 \omega_c}{\pi}
\end{equation}

From (289), we have

\begin{equation}
R_{r_1r_1}(\tau) = k^2 B^2 N_0 \frac{\sin \omega_c \tau}{\pi \tau}
\end{equation}

and the corresponding spectrum of the forcing term is

\begin{equation}
S_{r_1r_1}(\omega) = \begin{cases} 
  k^2 B^2 N_0 \frac{\sin \omega_c \tau}{\pi \tau} & |\omega| \leq \omega_c \\
  0 & \text{elsewhere.}
\end{cases}
\end{equation}
From (278), the power spectrum of the $w_1(t)$ process is (see Fig. 14)

$$ S_{w_1 w_1}(\omega) = \begin{cases} \frac{k^2 B^2 N_0}{k^2 A^4 + \omega^2} & |\omega| \leq \omega_c \\ 0 & \text{elsewhere} \end{cases} $$

Fig. 14. Power spectrum of $w_1(t)$ process.

The total average power in the $w_1(t)$ process is,

$$ P_{w_1} = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} S_{w_1 w_1}(\omega) \, d\omega $$

69
(297) \[ P_{R} = \frac{k^2B^2N_o}{\pi} \int_{-\infty}^{\infty} \frac{1}{kA^2 + \omega^2} d\omega. \]

(298) \[ P_{R} = \frac{k^2B^2N_o}{\pi kA^2} \tan^{-1} \frac{\omega_c}{kA^2} \]

The total average power of the \( w_1(t) \) process, normalized to \( P_R = N_0/\pi A^2 \), is plotted in Fig. 15 for various values of loop gain \( kA^2 \). There are two observations which can be made by studying these curves. First, increasing the bandwidth \( \omega_c \) of the noise process \( n'(t) \) increases the average power of the \( w_1(t) \) process, but only up to approximately \( \omega_c = kA^2 \), which is the half power bandwidth. Above that frequency, increasing \( \omega_c \) has little effect on \( w_1(t) \). Secondly, increasing the loop gain for a given noise bandwidth \( \omega_c \) increases the average power of the \( w_1(t) \) weight process. We will say more about these observations later.

One immediate result of these curves is that we can approximate the power spectrum of Fig. 14 for \( w_1(t) \) with the idealized rectangular spectrum of Fig. 16. This approximation is justified by the arguments stated above. Since almost all of the average power is concentrated within the \( |\omega| \leq kA^2 \) band, restricting ourselves to this bandwidth will not greatly affect the \( w_1(t) \) process. We also assume, the noise bandwidth \( \omega_c \) is greater than the loop gain \( kA^2 \), which is true in most array applications. The corresponding autocorrelation function is

(299) \[ R_{W_1W_1}(\tau) = \frac{B^2N_o}{A^4} \frac{\sin kA^2 \tau}{\pi \tau}. \]

The variance is

(300) \[ R_{W_1W_1}(0) = \sigma^2_{W_1} = \frac{k^2B^2N_o}{\pi kA^2}. \]
Fig. 15. Normalized average power of \( w_1(t) \) process vs. noise bandwidth.
We have now completely specified the $w_1(t)$ process. Combining our results into a two term perturbation expansion, we have

$$w(t) = \frac{B}{A} + \varepsilon w_1(t)$$  \hspace{1cm} (301)$$

where $w_1(t)$ is a zero mean, Gaussian random process with autocorrelation function given by (299). The weight $w(t)$ then is also Gaussian, to first order in $\varepsilon$, with a mean

$$E[w(t)] = \frac{B}{A}$$  \hspace{1cm} (302)$$

and autocorrelation function

$$R_{ww}(\tau) = \frac{B^2}{A^2} + \frac{\varepsilon^2 B^2 N_0}{A^4} \frac{\sin kA^2 \tau}{\pi \tau}$$  \hspace{1cm} (303)$$

Fig. 16. Approximate power spectrum of $w_1(t)$ process.
The variance is

\[ (304) \quad \sigma_w^2 = \frac{\varepsilon^2 k^2 B^2 N_0}{\pi k A^2} \]

In terms of the variance of the input narrowband noise process \( n'(t) \), we have from (292)

\[ (305) \quad \sigma_w^2 = \frac{\varepsilon^2 k^2 B^2 \sigma^2}{k A^2 \omega_c} \]

We see from (302) that to first order in \( \varepsilon \), the weight process has the mean of the completely deterministic case. It is a constant and not dependent on the bandwidth \( \omega_c \) of the perturbing noise process. This agrees with the earlier results of the direct and Ito approaches that the additional noise process causes a second order effect in \( \varepsilon \) on the mean weight.

For the variance, we see from (304) that \( \sigma_w^2 \) increases linearly as \( k N_0 \) for given values of \( A \) and \( B \). If we write the weight variance from the Ito approach (256) as

\[ (306) \quad \sigma_w^2(\text{Ito}) = \frac{3\varepsilon^2 k^2 B^2 N_0}{2k A^2} \]

we see the perturbation result (304) gives a smaller variance by a factor \( 2/3\pi \). The Ito result assumes an infinite bandwidth white noise process while the perturbation result was derived using narrowband noise.

For small \( \varepsilon \), the representation (301) is sufficient to describe the statistical characteristics of \( w(t) \). However, we now extend these results to the three term perturbation expansion

\[ (307) \quad w(t) = \frac{B}{A} + \varepsilon w_1(t) + \varepsilon^2 w_2(t) \]

For the third equation (270), we have
\[
\frac{dw_2(t)}{dt} + kA^2w_2(t) = -2kA\dot{y}(t)w_1(t) - \frac{kB}{A}(y^2(t) + z^2(t))
\]

Let us define the random forcing term
\[
\begin{align*}
\mathbf{r}_2(t) &= \mathbf{r}_{21}(t) + \mathbf{r}_{22}(t) = -2kA\dot{y}(t)w_1(t) \\
&\quad - \frac{kB}{A}(y^2(t) + z^2(t))
\end{align*}
\]

The mean is
\[
(310) \quad \mathbb{E} \{ \mathbf{r}_2(t) \} = -2kA \mathbb{E} \{ \dot{y}(t)w_1(t) \} - \frac{2kB\sigma^2}{A}
\]

where we have used (ref. p. 35).

\[
(311) \quad \mathbb{E} \{ y^2(t) \} = \mathbb{E} \{ z^2(t) \} = \sigma^2
\]

We now make use of the approximation for \( w_1(t) \) discussed on p. 70. The power spectrum of Fig. 16 is obtained by passing the driving function \(-kB\dot{y}(t)\) through an ideal low pass filter with an impulse response

\[
(312) \quad h(\tau) = \frac{1}{kA^2} \frac{\sin kA^2\tau}{\pi\tau}
\]

\( w_1(t) \) is then from the convolution theorem,

\[
(313) \quad w_1(t) = -kB \int_{-\infty}^{\infty} y'(\tau) \frac{1}{kA^2} \frac{\sin kA^2(t-\tau)}{(t-\tau)} \, d\tau
\]

Multiplying through by \( y'(t) \) and taking the expectation of both sides we get

\[
(314) \quad \mathbb{E} \{ y'(t)w_1(t) \} = \frac{\pi A^2}{kA^2} \int_{-\infty}^{\infty} \mathbb{E} \{ y'(t)y'(\tau) \} \frac{\sin kA^2(t-\tau)}{(t-\tau)} \, d\tau
\]
But \( y'(t) \) is a stationary random process with an autocorrelation function given by (291). So we have

\[
(315) \quad E\{y'(t)w_1(t)\} = -\frac{BN_0}{\pi A^2} \int_{-\infty}^{+\infty} \frac{\sin \omega_c(t-\tau)}{(t-\tau)^2} \sin kA^2(t-\tau) \, d\tau
\]

Carrying out the integration, we get

\[
(316) \quad E\{y'(t)w_1(t)\} = -\frac{kBN_0}{\pi}
\]

Substituting back in (310), we have

\[
(317) \quad E\{r_2(t)\} = \frac{2k^2ABN_0}{\pi} - \frac{2kB\sigma^2}{A}
\]
or using (292)

\[
(318) \quad E\{r_2(t)\} = \frac{2k^2AB\sigma^2}{\omega_c} - \frac{2kB\sigma^2}{A}
\]

From (46), the mean of the \( w_2(t) \) process is

\[
(319) \quad E\{w_2(t)\} = \frac{2kB\sigma^2}{A} \left( \frac{1}{\omega_c} - \frac{1}{kA^2} \right)
\]

With a three term perturbation expansion, we have then for the mean of the weight \( w(t) \)

\[
(320) \quad E\{w(t)\} = \frac{B}{A} + \epsilon^2 \frac{2kB\sigma^2}{A} \left( \frac{1}{\omega_c} - \frac{1}{kA^2} \right)
\]

Since \( \omega_c \) is assumed to be larger than \( kA^2 \), the presence of the \( w_2(t) \) term has added a small negative bias to the mean. For \( \omega_c >> kA^2 \), the mean reduces to
This result agrees well with the mean value of $w(t)$ calculated using the Ito approach (see (247)). Figure 17 shows a plot of $E[w(t)]$ as a function of noise bandwidth $\omega_c$ using the values $A=B=1$ and $\epsilon^2a^2=10^{-4}$. The weight mean for the deterministic case and that calculated using the Ito method are also shown for comparison.

Next, we want to determine the spectrum of the $r_2(t)$ process. From (309), we have

\begin{equation}
E[r_2(t)r_2(t+\tau)] = E[r_{21}(t)r_{21}(t+\tau)] + E[r_{21}(t)r_{22}(t+\tau)] + E[r_{22}(t)r_{21}(t+\tau)] + E[r_{22}(t)r_{22}(t+\tau)]
\end{equation}

where

\begin{equation}
r_{21}(t) = -2KA \gamma'(t)w_1(t)
\end{equation}

\begin{equation}
r_{22}(t) = -\frac{KB}{A} (\gamma'^2(t) + z'^2(t))
\end{equation}

Using the expression (313) for $w_1(t)$, we have for the first term in (322)

\begin{equation}
E[r_{21}(t)r_{21}(t+\tau)] = 4k^4A^2B^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{E[y'(t)y'(t+\tau)y'(v_1)y'(v_2)]}{k^2A^4} \\
\times \sin kA^2(t-v_1) \sin kA^2(t+\tau-v_2) \frac{dv_1}{\pi(t-v_1)} \frac{dv_2}{\pi(t+\tau-v_2)}
\end{equation}
Fig. 17. $E\{w(t)\}$ variation with noise bandwidth.
We now state the following theorem: Given the zero mean, Gaussian random process \( y(t) \), then

\[
E\{y(t_1)y(t_2)\ldots y(t_n)\} = \begin{cases} 
0 & \text{n odd} \\
\sum_{\text{all pairs}} E\{y_jy_k\}E\{y_p\} \ldots E\{y_qy_r\} & n=2m
\end{cases}
\]

where the number of terms in the sum is \( \frac{2m!}{m!2^m} \). Applying the theorem to the expectation in the integrand we have

\[
E\{y'(t)y'(t+\tau)y'(v_1)y'(v_2)\} = E\{y'(t)y'(t+\tau)\} \\
E\{y'(v_1)y'(v_2)\} + E\{y'(t)y'(v_1)\}E\{y'(t+\tau)y'(v_2)\} \\
+ E\{y'(t)y'(v_2)\} E\{y'(t+\tau)y'(v_1)\}
\]

Substituting (327) back into (325) and using (291) for the autocorrelation function of \( y'(t) \), we get

\[
E\{r_{21}(t)r_{21}(t+\tau)\} = 4kA^2B^2 \int_{-\infty}^{+\infty} \frac{R_{y'y'}(\tau)}{2k^2A^4} \frac{N_0 \sin \omega_c (v_2-v_1)}{\pi (v_2-v_1)}
\]

\[
\frac{\sin kA^2(t-v_1)}{\pi (t-v_1)} \frac{\sin kA^2(t+\tau-v_2)}{\pi (t+\tau-v_2)} \, dv_1 dv_2
\]

\[
+ \int_{-\infty}^{+\infty} \frac{N_0}{kA^2} \frac{\sin \omega_c (t-v_1)}{\pi (t-v_1)} \frac{\sin kA^2(t-v_1)}{\pi (t-v_1)} \, dv_1
\]

\[
\cdot \int_{-\infty}^{+\infty} \frac{N_0}{kA^2} \frac{\sin \omega_c (t+\tau-v_2)}{\pi (t+\tau-v_2)} \frac{\sin kA^2(t+\tau-v_2)}{\pi (t+\tau-v_2)} \, dv_2
\]
The double integral in the first term of (328) is a double convolution; the integrals in the second term are evaluated as \( N_0/\pi \) and the integrals in the third term are single convolutions. Using (290) and defining \( H(j\omega) \) as the system function for \( h(t) \) in (312), the spectrum \( S_{21,21}(\omega) \) is

\[
S_{21,21}(\omega) = \frac{4kA^2B^2}{2\pi} \left[ S_{y'y',(\omega)} \otimes S_{y'y',(\omega)} \cdot H(j\omega) + \frac{8kA^2B^2N_0^2}{\pi} \delta(\omega) + \frac{4kA^2B^2}{2\pi} S_{y'y',(\omega)} \cdot H(j\omega) \otimes S_{y'y',(\omega)} \cdot H(j\omega) \right]
\]

where \( \otimes \) denotes the convolution. For the second term in (322), we have

\[
E\{r_{21}(t)r_{22}(t+\tau)\} = 2k^2B \left[ E[y'(t)w_1(t)[y'(t+\tau)+z'^2(t+\tau)] \right]
\]

\[
= -2k^3B^2 \left[ \int_{-\infty}^{\infty} \frac{E[y'(t)y'(t+\tau)y(v)]}{kA^2} \right]
\]

\[
\sin \frac{kA^2(t-v)}{\pi(t-v)} \, dv + \frac{\sigma^2N_0}{kA^2} \left[ \int_{-\infty}^{\infty} \frac{\sin \omega_c(t-v)}{\pi(t-v)} \, dv \right]
\]
Again, using the theorem (326) and evaluating the integrals we have

\[
E\{r_{21}(t)r_{22}(t+\tau)\} = \frac{-4k^3B^2\sigma^2N_o}{\pi} - 4k^3B^2R_{y',y'}(\tau)
\]

\[
\int_{-\infty}^{+\infty} \frac{N_0}{kA^2} \frac{\sin \omega_c(t+\tau-v)}{\pi(t+\tau-v)} \frac{\sin kA^2(t-v)}{\pi(t-v)} \, dv
\]

The corresponding spectrum is

\[
S_{r_{21},r_{22}}(\omega) = -8k^3B^2\sigma^2N_0\delta(\omega) - \frac{4k^3B^2}{2\pi} S_{y',y'}(\omega) \otimes S_{y',y'}(\omega) \cdot H(j\omega)
\]

The evaluation of the third term in (322) is identical to the second term and we get

\[
E\{r_{22}(t)r_{21}(t+\tau)\} = \frac{-4k^3B^2\sigma^2N_o}{\pi} - 4k^3B^2R_{y',y'}(\tau)
\]

\[
\int_{-\infty}^{+\infty} \frac{N_0}{kA^2} \frac{\sin \omega_c(t-v)}{\pi(t-v)} \frac{\sin kA^2(t+\tau-v)}{\pi(t+\tau-v)} \, dv
\]

Taking the inverse Fourier transform, we get

\[
S_{r_{22},r_{21}}(\omega) = -8k^3B^2\sigma^2N_0\delta(\omega) - \frac{4k^3B^2}{2\pi} S_{y',y'}(\omega) \otimes S_{y',y'}(\omega) \cdot H(j\omega)
\]

For the fourth term in (322), we have
\( E[r_{22}(t)r_{22}(t+\tau)] = \frac{k^2 B^2}{A^2} E[[y'(t)^2 + z'(t)^2][y'(t+\tau)^2 + z'(t+\tau)^2]] \)

\( = \frac{k^2 B^2}{A^2} \left[ E[y'(t)^2 y'(t+\tau)^2] + E[z'(t)^2 z'(t+\tau)^2] \right. \\
+ \left. E[y'(t)^2 z'(t+\tau)^2] + E[z'(t)^2 y'(t+\tau)^2] \right] \)

But \( y'(t) \) and \( z'(t) \) are both zero mean, Gaussian, statistically independent random processes with variance \( \sigma^2 \). Also applying (326), we get

\( E[r_{22}(t)r_{22}(t+\tau)] = \frac{k^2 B^2}{A^2} [4\sigma^4 + 4R_{y'y'}(\tau)] \)

The corresponding spectrum is

\( S_{r'_{22}, r'_{22}}(\omega) = \frac{8\pi k^2 B^2 \sigma^4}{A^2} \delta(\omega) + \frac{4k^2 B^2}{2\pi A^2} S_{y'y'}(\omega) \otimes S_{y'y'}(\omega) \)

With

\( S_{y'y'}(\omega) \cdot H(j\omega) = N_0 H(j\omega) \)

and combining the results for the four terms in (322), we have

\( S_{r'_{22}, r'_{22}}(\omega) = \frac{4k^4 A^2 B^2 N_0}{2\pi} S_{y'y'}(\omega) \otimes H^2(j\omega) + \frac{8k^4 A^2 B^2 N_0}{\pi} \delta(\omega) \)

\( + \frac{4k^4 A^2 B^2 N_0}{2\pi} H(j\omega) \otimes H(j\omega) - 16k^3 B^2 \sigma^2 N_0 \delta(\omega) \)

\( - \frac{8k^3 B^2 N_0}{2\pi} S_{y'y'}(\omega) \otimes H(j\omega) + \frac{8\pi k^2 B^2 \sigma^4}{A^2} \delta(\omega) \)

\( + \frac{4k^2 B^2}{2\pi A^2} S_{y'y'}(\omega) \otimes S_{y'y'}(\omega) \)
or

\begin{equation}
S_{r_2 r_2}(\omega) = \frac{4k^4 A^2 B^2 N_0^2}{2\pi} \left\{ \left( 4 - \frac{8\omega_c}{kA^2} + \frac{4\omega_c^2}{k^2 A^4} \right) \delta(\omega) + H(j\omega) \otimes H(j\omega) \right\},
\end{equation}

\begin{align*}
&- \frac{2}{kA^2 N_0} S_{\gamma,\gamma'}(\omega) \otimes H(j\omega) + \frac{1}{k^2 A^4 N_0^2} S_{\gamma,\gamma'}(\omega) \otimes S_{\gamma,\gamma'}(\omega) \}
\end{align*}

If we define

\begin{equation}
H(j\omega) \otimes H(j\omega) = \frac{2}{kA^2} q_{2k}A^2(\omega)
\end{equation}

\begin{equation}
S_{\gamma,\gamma'}(\omega) \otimes S_{\gamma,\gamma'}(\omega) = 2N_0^2 \omega_c q_{2\omega_c}(\omega)
\end{equation}

and

\begin{equation}
S_{\gamma,\gamma'}(\omega) \otimes H(j\omega) = 2N_0 \bar{q}_{\omega_c+kA^2}(\omega)
\end{equation}

where \( q_{BW}(\omega) \) denotes a triangular power spectrum centered at zero frequency with one sided bandwidth BW (see Fig. 18) and \( q_{BW+Y}(\omega) \) denotes a trapezoidal power spectrum as shown in Fig. 19. Using this notation, (342) becomes

\begin{equation}
S_{r_2 r_2}(\omega) = \frac{4k^4 A^2 B^2 N_0^2}{2\pi} \left\{ \left( 2 - \frac{2\omega_c}{kA^2} \right) \delta(\omega) + \frac{2}{kA^2} q_{2kA^2}(\omega) 
\right. \\
\left. + \frac{2\omega_c}{k^2 A^4} q_{2\omega_c}(\omega) - \frac{4}{kA^2} \bar{q}_{\omega_c+kA^2}(\omega) \}
\end{equation}

From (278), the \( w_c(t) \) power spectrum is then
Fig. 18. Definition of $q_{BW}(\omega)$.

Fig. 19. Definition of $q_{BW+x}(\omega)$. 
\[
S_{w_2w_2}(\omega) = \frac{4k^2B^2N^2}{\pi A^2} \left\{ \frac{2 \left( 1 - \frac{\omega}{kA^2} \right)^2 \delta(\omega)}{\omega} + \frac{kA^2q2kA^2(\omega) + \omega c q2\omega c(\omega) - 2kA^2c q2\omega c(\omega) + kA^2(\omega)}{k^2A^4 + \omega^2} \right\}
\]

\[
S_{w_2w_2}(\omega) \text{ normalized to } \frac{4k^3B^2N^2}{\pi} \text{ is shown in Fig. 20 for the case } \omega_c^2 = 10kA^2. \text{ It consists of a delta function at zero frequency and a small, narrow band component centered about zero frequency. Unlike the } w_1(t) \text{ process, the amplitude of the power spectrum of } w_2(t) \text{ is dependent directly on the noise bandwidth } \omega_c. \text{ For large } \omega_c, \text{ the delta function becomes dominant and the power contribution of } w_2(t) \text{ to the weight is almost entirely DC in nature.}

The total average power of the } w_2(t) \text{ process is}

\[
P_{w_2} = \frac{2k^2B^2N^2}{\pi^2A^2} \left\{ \frac{2 \left( 1 - \frac{\omega}{kA^2} \right)^2}{\omega} + \int_{-\infty}^{+\infty} \frac{kA^2q2kA^2(\omega) + \omega c q2\omega c(\omega) - 2kA^2c q2\omega c(\omega) + kA^2(\omega)}{k^2A^4 + \omega^2} d\omega \right\}
\]

\[
= \frac{2k^2B^2N^2}{\pi^2A^2} \left\{ \frac{2 \left( 1 - \frac{\omega}{kA^2} \right)^2}{\omega} + \int_{0}^{2\omega_c} \frac{2kA^2(1-\omega/2kA^2)}{k^2A^4 + \omega^2} d\omega + \int_{0}^{2\omega_c} \frac{2\omega_c(1-\omega/2kA^2)}{k^2A^4 + \omega^2} d\omega - \int_{0}^{\omega_c} \frac{4kA^2}{k^2A^4 + \omega^2} d\omega \right\}
\]

where we have approximated the trapezoidal spectrum \( \bar{\omega}_c + kA^2(\omega) \) with a rectangular spectrum with bandwidth \( \omega_c \). Evaluating the integrals, we get

84
Fig. 20. Normalized power spectrum of $w_2(t)$ process.

$$P_{w_2} = \frac{2k^3 B^2 N_0^2}{\pi^2} \left\{ \frac{2}{kA^2} \left( 1 - \frac{\omega_c}{kA^2} \right)^2 + \frac{1.865}{kA^2} \right. \right.$$  
\[ + \frac{2\omega_c}{kA^4} \tan^{-1} \frac{2\omega_c}{kA^2} - \frac{1}{2kA^2} \log(k^2A^4 + 4\omega_c^2) \] 
\[ + \frac{1}{2kA^2} \log k^2A^4 - \frac{4}{kA^2} \tan^{-1} \frac{\omega_c}{kA^2} \left\} \right.$$
In the evaluation of the $w_2(t)$ process, we have assumed $\omega_c > kA^2$ (see p. 70 and Fig. 14). Also, the loop gain $kA^2$ which has the dimensions of rad/sec is assumed to be much greater than 1. Examination of (350) shows there are only two dominant terms in the enclosed brackets—the first and third. Figure 21 compares the contribution of these terms with the total $P_{W_2}$ value for the case $kA^2 = 1000$. It shows the first and third terms of (350) accurately represents the $w_2(t)$ average power except over a narrow range of noise bandwidths $\omega_c$ with values near the loop gain $kA^2$. Therefore let

$$
P_{W_2} \approx \frac{2k^3B^2N_0^2}{\pi^2} \left\{ \frac{2}{kA^2} \left( 1 - \frac{\omega_c}{kA^2} \right)^2 + \frac{2\omega_c}{k^2A^4} \tan^{-1} \frac{2\omega_c}{kA^2} \right\}
$$

The second term in (351) is of the same form as the average power $P_{W_1}$ of the $w_1(t)$ process (see (298)). In treating the $w_1(t)$ case, we approximated the spectrum of $w_1(t)$ with an ideal low pass filter with cutoff frequency $kA^2$. Let us make the same approximation for the second term in (351) and use a cutoff frequency of $kA^2/2$. The power contributed by the first term remains unchanged and is the result of a D.C. term which can be represented as a delta function at zero frequency. The modified spectrum for the $w_2(t)$ process is shown in Fig. 22.

The corresponding autocorrelation function for the $w_2(t)$ process is

$$
R_{W_2W_2}(\tau) = \frac{4k^3B^2N_0^2}{\pi^2kA^2} \left( 1 - \frac{\omega_c}{kA^2} \right)^2 + \frac{2B^2N_0^2}{\pi A^6} \frac{\omega_c}{\pi \tau} \sin kA^2/2\tau
$$

We will now combine the results found for the $w_2(t)$ process into our three term perturbation expansion. From this, we can calculate the spectrum and autocorrelation function for the weight process $w(t)$. The mean of $w(t)$ has already been derived (see (320)). For the autocorrelation function, we have
Fig. 21. Power contribution of terms in Eq. (350).

\[ P_R = \frac{2k^3B^2N_0^2}{\pi^2} \]

\[ kA^2 = 10^3 \]

FIRST AND THIRD TERMS ONLY

ALL TERMS OF (350)
Fig. 22. Power spectrum approximation for $w_2(t)$ process.

\[
E[[w_0 + \epsilon w_1(t) + \epsilon^2 w_2(t)][w_0 + w_1(t+\tau) + \epsilon^2 w_2(t+\tau)]] \\
= w_0^2 + \epsilon[w_0E[w_1(t)] + w_0E[w_1(t+\tau)]] \\
+ \epsilon^2[w_0E[w_2(t)] + w_0E[w_2(t+\tau)] + E[w_1(t)w_1(t+\tau)]] \\
+ \epsilon^3[E[w_2(t)w_1(t+\tau)] + E[w_1(t)w_2(t+\tau)]] \\
+ \epsilon^4E[w_2(t)w_2(t+\tau)]
\]

We note that the terms in odd powers of $\epsilon$ are zero since $E[w_1(t)]=0$ and from (326) $E[y(t_1)y(t_2)y(t_3)]=0$. Using $w_0 = B/A$, we have
\[ R_{WW}(\tau) = \frac{B^2}{A^2} + \epsilon^2 \left[ \frac{2B}{A} \left( E[w_2(t)] + R_{W_1W_1}(\tau) \right) \right] \]

\[ + \epsilon^4 [R_{W_2W_2}(\tau)] \]

Substituting in for \( E[w_2(t)] \) from (319), \( R_{W_1W_1}(\tau) \) from (299) and \( R_{W_2W_2}(\tau) \) from (352), we have

\[ R_{WW}(\tau) = \frac{B^2}{A^2} + \epsilon^2 \left[ \frac{4k^2 B^2 N_0 \omega_c}{\pi k A^2} \left( \frac{1}{\omega_c} - \frac{1}{k A^2} \right) \right] \]

\[ + \frac{B^2 N_0}{A^4} \sin k A^2 \tau + \epsilon^4 \left\{ \frac{4k^3 B^2 N_0^2 \omega_c}{\pi^2 k A^2} \left( 1 - \frac{\omega_c}{k A^2} \right)^2 \right\} \]

\[ + \frac{2B^2 N_0^2 \omega_c}{\pi A^6} \frac{\sin k A^2/2 \tau}{\pi} \]

The weight variance is, to second order in \( \epsilon \),

\[ \sigma_w^2 = \frac{\epsilon^2 k B^2 N_0}{\pi k A^2} + \frac{\epsilon^4 k^3 B^2 N_0^2 \omega_c}{\pi^2 k A^4} \]

or

\[ \sigma_w^2 = \frac{\epsilon^2 k B^2 N_0}{\pi k A^2} \left( 1 + \frac{\epsilon^2 k N_0 \omega_c}{\pi k A^2} \right) \]

We see from (357) the use of a three term perturbation expansion has added a term to the variance calculated using a two term expansion (304). Figure 23 shows a plot of the weight variance (357) for various values of \( \epsilon^2 k N_0 \). The variance calculated using the Ito Approach (256) is shown superimposed on the same graph.

For the parameters shown in Fig. 23, the weight variance calculated
Fig. 23. Comparison of weight variance $\sigma_w^2$ derived using Ito method and perturbation method.

- $\omega_c = \infty$
- $\omega_c = 10^6$
- $kA^2 = 10^3$
- $A = B$

**WEIGHT VARIANCE $\sigma_w^2$**

**Ito Method**

**Perturbation Method**

(3 Term Expansion)
by the Ito approach is approximately four times as great as that calculated using the perturbation method. Only when the noise becomes "large" do the weight variances cross.

Figure 24 shows another comparison of the two approaches. Here the value of $\varepsilon^2 k N_0$ is fixed and the noise bandwidth $\omega_c$ is allowed to vary over a wide range ($10^2$ to $10^8$ rad/sec). The graph shows the weight variance to be essentially independent of noise bandwidth over most of the range of values of $\omega_c$. At values of $\omega_c$ below the loop gain $kA^2$, the weight variance will fall off rapidly. This fact was previously demonstrated in Fig. 17 and Fig. 21 by showing the average power in each of the perturbation terms as a function of noise bandwidth. The point at which $\sigma^2_\omega$ begins to roll off is determined by the loop gain $kA^2$. For wide noise bandwidths ($10^8$ rad/sec), Fig. 24 shows the weight variance rapidly increasing. This is not a valid result. Although there is a gradual increase in weight variance with increasing noise bandwidth, at values near $\omega_c=10^8$ rad/sec, we begin to violate the assumptions made in employing the perturbation method. For example, with the values shown in Fig. 24, let $A=B=1$, $k=10^3$ and $\varepsilon^2 k N_0=10^{-4}$. This implies $\varepsilon^2 N_0 c=10^1$. But we have previously assumed $A^2>>\varepsilon^2 N_0 c$ (169) where

$$\sigma^2 = \frac{N_0 \omega_c}{\pi}$$

Therefore, near $\omega_c = 10^8$, we begin to violate this assumption. However, for smaller values of $\varepsilon^2 k N_0$, the perturbation method is valid over these increased noise bandwidths.

Before concluding, we mention some factors on the convergence of the perturbation expansion

$$(359) \quad w(t) = w_0(t) + \varepsilon w_1(t) + \varepsilon^2 w_2(t) + \cdots$$

In the deterministic case, the convergence of (359) is well known (see for example Nayfeh [13]). However when (359) represents a stochastic expansion, no proof is yet available to show that $w(t)$, given by (359), is convergent in mean square or in any other sense (Soong [22, p. 206]). Thus we must assume that the random process $w(t,\xi)$ is convergent for each of its sample functions $\xi_i$.

---

4Reference discussion p. 16 and Fig. 6 for notation.
Fig. 24, Weight variance as function of noise bandwidth.
An alternative way to look at the convergence problem in the adaptive array problem is to examine (359) in terms of the average power contributed by each of the terms in the expansion. If we can generate an asymptotic expansion for $P_w$, valid for large $A^2$, of the form

$$P_w = \frac{1}{A^2} h_1 + \frac{1}{A^4} h_2 + \frac{1}{A^6} h_3 + \ldots$$

then we can assume the first few terms of the series are adequate to describe $P_w$. $P_w$ is defined to be the total average power of the weight process. In turn, based on this power argument, we can then say that the first few terms of (359) are adequate to represent the weight process $w(t)$.

Recall (283) that

$$P_{w_0} = \frac{B^2}{A^2}$$

and (298)

$$P_{w_1} = \frac{k^2 B^2 \sigma^2}{k A^2 w_c} \tan^{-1} \frac{\omega_c}{k A^2}$$

where we have used (292) for $\sigma^2$. Also (351)

$$P_{w_2} = \frac{2 k^3 B^2 \sigma^4}{w_c^2} \left\{ \frac{2}{k A^2} \left(1 - \frac{\omega_c}{k A^2}\right)^2 + \frac{2 \omega_c}{k^2 A^4} \tan^{-1} \frac{2 \omega_c}{k A^2} \right\}$$

But in our analysis of $w_1(t)$ and $w_2(t)$ we have assumed

$$\omega_c > k A^2$$

So let

$$\omega_c = c k A^2$$
where $c$ is a positive constant greater than 1. Then (362) and (363) can be written as

$$\text{(366)} \quad P_w = \frac{k^2B^2\sigma^2}{c^2k^2A^4} \tan^{-1} c$$

and

$$\text{(367)} \quad P_w = \frac{2k^3B^2\sigma^4}{c^2k^3A^6} \{2(1-c)^2 + \frac{2c}{kA^2} \tan^{-1} 2c\}$$

The total average power of the weight with $c$ assumed to be a fixed constant is then

$$\text{(368)} \quad P_w = \frac{B^2}{A^2} + \frac{\epsilon^2B^2\sigma^2}{cA^4} \tan^{-1} c + \frac{\epsilon^2B^2\sigma^4}{c^2A^6} \{2(1-c)^2 + \frac{2c}{kA^2} \tan^{-1} 2c\} + \ldots$$

But from (169)

$$\text{(369)} \quad A^2 >> \epsilon^2\sigma^2$$

so (368) is in the desired form (360) of an asymptotic expansion valid for large $A^2$. We can thus assume that the series (359) is convergent, at least in the asymptotic sense.
CHAPTER V
CONCLUSIONS

The purpose of this study was to determine the statistical properties of the weights in an adaptive antenna array. To this end, the differential equations obeyed by the weights were developed in Chapter II. They were a set of coupled, first order stochastic differential equations with random coefficients and random forcing terms. The two principal problem areas in treating these equations were: (1) they were coupled and (2) the coefficients were random processes. The approach taken in this dissertation was to consider a one element adaptive "array". This reduced the complexity of the analysis considerably and removed the coupling problem as well. The thrust of the study was therefore to concentrate on the random coefficient problem.

For the one element case, the distribution and statistics for the forcing function and the random coefficient were developed in Chapter IV, Section A. Assuming an incident cw signal, Acos\omega_0t, corrupted by narrowband, Gaussian noise n(t), the forcing term was determined to be Gaussian with mean and variance given by Eqs. (132) and (133) respectively. Similarly, the distribution of the random coefficient term was shown to be a special case of the non-central chi-squared ($\chi^2$) with a mean given by (146) and variance by (147).

To solve the random coefficient problem, three approaches were examined in this study. The first approach, called the direct approach, operated directly on the expression for the solution to the weight differential equation (35). Using the joint characteristic function of the random coefficient and forcing term with a Gaussian approximation for the random coefficient, expressions were developed for the weight mean (199) and autocorrelation function (202). These equations, although they provided a general method to solve the random coefficient problem, were long and rather complicated. For this reason, particular solutions to these equations were abandoned in favor of two approximation methods.
The first approximation method replaced the random portions of both the coefficient and forcing terms with white noise. This allowed the array weight differential equation to be modeled as an Ito equation. In this form, formulae (reference Eqs. (118) and (114)) were available to solve directly for the weight mean $E\{w(t)\}$ and weight variance $\sigma^2_w$. These results were given by equations (245) and (256) and plotted in Fig. 9 and 10, respectively for the mean and variance. We delay summarizing these results until later.

The second approximation method applied stochastic perturbation theory to the weight differential equation. By assuming a small additive noise process, represented as $\epsilon n'(t)$ with $\epsilon$ a small parameter, this technique expanded the random coefficient and weight solution in a convergent series in powers of $\epsilon$. The study showed there were several important advantages in using the perturbation method to study the array weights. First, it reduced the random coefficient weight differential equation to a set of constant coefficient equations. This allowed the use of well known, classical methods to determine the weight statistics. Second, compared to the Ito approach which assumed an infinite power white noise process, the perturbation method was used with narrowband noise which is more representative of the actual physical system. Third, although valid only for the small noise case, the method provides results which give a satisfactory description of the weight statistics sufficient for many applications. Additional terms can be added to the perturbation expression to improve the accuracy of the results. Fourth, not only does the perturbation method overcome the random coefficient problem, but it also solves the weight coupling problem as well. In Chapter III, Section C, a set of uncoupled, vector equations for the N-element array were developed using the perturbation method. Although the vector case was not treated further in this report, these equations provide a foundation for further research into this phase of the weight problem.

To determine the weight statistics, both a two term and three term perturbation expansion were derived in Chapter IV, Section D for the weight process. In each case, the mean and variance was calculated and results compared to the Ito approach. The two term results were given on pp. 72-73 and the three term results on pp. 75-89 and in Figs. 17, 23 and 24. For the weight mean $E\{w(t)\}$, the following facts were ascertained:

1. For an incident signal plus noise process $A \cos \omega_0 t + \epsilon n'(t)$, a small parameter, the disturbance effect on the mean is reflected in a second order term in $\epsilon$. For $\epsilon$ small, the effect is thus minimal.
(2) With the noise present, the steady state mean weight is reduced to a value below that attained in the deterministic case \((B/A)\). \(B\) is the amplitude of the reference signal.

(3) The amount of deviation below the deterministic value increases with increasing noise power.

(4) For large noise bandwidths \(\omega_c\), the Ito approach and perturbation approach give nearly identical results (see (247) and (321)).

For the weight variance \(\omega_w^2\), the following conclusions were reached:

(1) \(\omega_w^2\) varies as the square of the deterministic mean \((B^2/A^2)\) times the loop gain product \(kN_0\) where \(k\) is the loop gain constant and \(N_0\) is the amplitude of the noise power spectral density. Thus an increase in the power of the reference signal or a reduction in the received power of the desired signal causes an increase in the weight jitter. Also increasing \(k\) or \(N_0\) causes a corresponding increase in the weight variance.

(2) An explicit dependence of the variance on the noise bandwidth \(\omega_c\) is not observed in the two term perturbation expansion, only in the three term expansion. Thus the weight variance increases on the order of \(\varepsilon^2\) with increasing noise bandwidth. Since \(\varepsilon\) is assumed small, the variance then remains nearly constant over a broad range of values for \(\omega_c\).

(3) Except for a constant of proportionality, the weight variance as calculated by the Ito approach and perturbation approach are similar. The Ito technique, employing infinite power white noise, gives the higher value (see Fig. 24).
BIBLIOGRAPHY


