STABILITY CRITERIA FOR SYSTEMS WITH
COLORED MULTIPLICATIVE NOISE

by

David Newton Martin

B.E.E. Rensselaer Polytechnic Institute
1967

S.M. Massachusetts Institute of Technology
1967

E.E. Massachusetts Institute of Technology
1973

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
July, 1974

Signature of Author ..............................................
Department of Electrical Engineering,
July 5, 1974

Certified by .............................................. Thesis Supervisor

Accepted by .............................................. Chairman, Departmental Committee on Graduate Students

Archives

OCT 21 1974
STABILITY CRITERIA FOR SYSTEMS
WITH COLORED MULTIPLICATIVE NOISE

by

David Newton Martin

Submitted to the Department of Electrical Engineering
on July 5, 1974 in partial fulfillment of the require-
ments for the degree of Doctor of Philosophy.

ABSTRACT

Often a mathematical model of a dynamical system requires
the consideration of uncertain elements. Historically,
emphasis has been placed on studying systems in which
disturbances enter additively, but many control problems
include disturbances that can be more naturally modeled
as uncertain (time-varying) multiplicative gains. If
the multiplicative gain is of the white noise type,
necessary and sufficient conditions for the stability of
the system have been derived previously. In this thesis,
we develop conditions for the stability of some systems
containing multiplicative colored noise.

In the context of continuous-time systems, three classes
of problems are treated. For the first order system
whose coefficient is an Ornstein-Uhlenbeck process,
necessary and sufficient conditions for mean square
stability and almost sure stability are provided. Using
the concepts of Lie groups, the stability properties
of a class of higher order systems (those evolving on
a solvable Lie group) are analyzed. The third class
of systems for which stability criteria are derived is
that consisting of single-input single-output systems
in which the Ornstein-Uhlenbeck process plays the role
of a feedback gain. We have used the Picard expansion
of the solution to the scalar problem as a bound on the
expansion obtained in the higher order case to obtain
sufficient conditions for mean square stability. The
damped harmonic oscillator problem is used as an example.
An existence and uniqueness proof for the solution of
these systems is included.
In the discrete-time case, systems in which the colored noise is generated nonrecursively are studied. Necessary and sufficient conditions for their mean square stability have been obtained, and numerical solutions are provided for cases in which the correlation time of the colored noise parameter is small. We have also provided a necessary condition for the mean square stability of some systems in which the colored noise parameter is generated recursively.

THESIS SUPERVISOR: Timothy L. Johnson

TITLE: Assistant Professor of Electrical Engineering
ACKNOWLEDGEMENTS

I would like to thank Professor Timothy Johnson for his support, confidence, and assistance during this thesis research. As an advisor and as a friend he provided the encouragement that was essential in the completion of this thesis.

I would like to thank Professor Alan Willsky for his contributions to this research over the past year. Those who are familiar with his research will recognize his strong influence. Professor Arthur Baggeroer served as the other reader on my thesis committee. I am grateful for all his comments and corrections.

Other members of the Electronic Systems Laboratory from whom I have benefitted include Professor Jan Willems who introduced me to this research; Professor Sanjoy Mitter who provided some early advice; and Professor Nils Sandell, who helped me with generalities and details, from start to finish. Special thanks also go to Arthur Giordani for his drafting help.

Miss Cyndi Curdo deserves credit for her accuracy, speed, and skill in typing this manuscript.

Finally, I wish to express special and most sincere thanks to my wife Leslie for her love and understanding over the past two years. It will be many years before she can be repaid for the contributions she has made.

Financial support for this research was provided by the Advanced Studies Program of the Raytheon Corporation.
DEDICATED
TO MY PARENTS

Newton A. Martin
&
Edith J. Martin
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>CHAPTER I</td>
<td>INTRODUCTION</td>
<td>8</td>
</tr>
<tr>
<td>CHAPTER II</td>
<td>BACKGROUND MATERIAL</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2.1 Continuous-Time Systems</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2.2 Discrete-Time Systems</td>
<td>46</td>
</tr>
<tr>
<td>CHAPTER III</td>
<td>CONTINUOUS TIME RESULTS</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>3.1 Scalar, First Order Systems</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>3.2 Applications of Lie Algebra Theory</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>3.3 Extensions to the Approach of I.L. Rabotnikov [R1]</td>
<td>108</td>
</tr>
<tr>
<td>CHAPTER IV</td>
<td>DISCRETE TIME RESULTS</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>4.1 System Definition</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>4.2 A Theorem Not Requiring Zero Correlation Time</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>4.3 Stability Criteria Based Upon Finite Correlation Times</td>
<td>138</td>
</tr>
<tr>
<td>CHAPTER V</td>
<td>SUMMARY AND CONCLUSIONS</td>
<td>160</td>
</tr>
</tbody>
</table>
REFERENCES 165

APPENDIX A APPLICATIONS AND LIMITATIONS OF THE ITO CALCULUS 172

APPENDIX B LIE ALGEBRA DEFINITIONS 189

APPENDIX C LISTINGS OF COMPUTER PROGRAMS 194

BIOGRAPHICAL NOTE 199
CHAPTER I
INTRODUCTION

In this thesis, we analyze systems that contain uncertain elements. Our goal is to determine conditions that will guarantee the stability (in some probabilistic sense) of these systems, making use of the statistical properties of the unknown components. A great deal of research along this line has been completed for systems in which the stochastic elements are random processes that enter additively as control terms or as observation noise. More recently, researchers have attempted to develop a parallel theory in which the stochastic elements appear as multipliers, but the results obtained to date are quite incomplete. In the following chapters, we make some contributions to this theory.

In addition to the mathematical motivation for studying this problem, there is an abundance of physical problems which require the use of multiplicative noise models in their analysis. The easiest example to visualize is a linear dynamical system containing a noisy gain in the feedback path, as in Figure 1.1 Other researchers
INITIAL CONDITION

\[ u = 0 \]

\[ y(t) \text{ or } y(n) \]

**System Model with Random Process Entering as a Multiplier**

*Figure 1.1*
have found many applications of the theory [W3, W7, W9, B2]. Among the examples cited are human operators, whose errors tend to be proportional to a desired control action, noisy voltage-controlled oscillators within a phase-locked loop, machines with a randomly switching element, switching jitter in sampled-data systems and wave transmission through a randomly varying medium.

We will be studying systems of the following forms:

\[
\frac{dx(t)}{dt} = (A+f(t)B)x(t) \quad \text{(continuous time)} \quad (1.1)
\]

\[
x(n+1) = (A+f(n)B)x(n) \quad \text{(discrete time)} \quad (1.2)
\]

In both cases, the initial conditions are given as a random variable, A and B are constant matrices, and \( f(*) \) is a real-valued colored noise process. A more precise interpretation of Equations (1.1) and (1.2) is given in Chapter II.

Several researchers have developed stability criteria for systems identical to Equations (1.1) and (1.2) except with \( f(t) \) chosen to be a white noise process. In Appendix A, we discuss some of their results, most of which require the use of Itô calculus. However, in that context, one is required to use Itô correction
terms to interpret the results in terms of physical systems, in which each signal must have finite power and finite bandwidth.

For linear systems in which the random process enters additively, the use of a white noise model is well justified, especially if state augmentation is used to generate the colored noise processes required. Even without state augmentation, the white noise model can often be used to represent wideband processes because the extra power in the white noise model will be dissipated anyway. For multiplicative noise, however, the situation is not so simple. Using state augmentation and white noise to generate the required colored noise process leads to a non-linear system model which is, in fact, bilinear. Without state augmentation, the only interpretation of the results is in terms of physical systems involving very wideband noise processes. The goal of this research, therefore, is to develop a theory for the stability of systems (1.1) and (1.2) that is valid for colored f(*)

It seems reasonable to expect that the stability of systems (1.1) and (1.2) should depend upon both the amplitude and the bandwidth of the noise process. In Section 2.1, we discuss some published stability criteria that apply to system (1.1), but do not depend
upon the bandwidth of $f(t)$. In the discrete-time con-
text (system (1.2)), a similar state-of-the-art exists, 
although the methods used are quite different.

The main contribution of this research is the 
establishment of some criteria for the stability (or 
instability) of systems (1.1) and (1.2) that explicitly 
involve the power and the bandwidth of the colored noise 
parameter $f(*)$.

In Chapter II, we establish the framework under which 
we will be working, and provide the required definitions 
of stochastic stability for both continuous and discrete 
time systems. In spite of the large number of results 
that have appeared concerning the continuous-time system 
(1.1), we were unable to locate an existence and uniqueness 
proof for its solution, so we have given such a proof in 
Section 2.1. For the discrete-time systems of Section 2.2, 
the Markov property of the augmented state variables is 
established. As a byproduct of that proof, we derive 
the Chapman-Kolmogorov equation associated with Equation 
(1.2). Each of the sections in Chapter II ends with 
a survey of some related literature. The main contri-
bution of Chapter II is the existence and uniqueness proof 
for the solution.
of system (1.1), contained in Lemma 2.1.1, Lemma 2.1.2, Theorem 2.1.1, and Theorem 2.1.2.

Chapter III is devoted to deriving stability criteria for the continuous time system of Equation (1.1). In Section 3.1, we consider first order systems containing a single noise parameter. This analysis is particularly simple but we have included all the details for two reasons. First, as we are able to completely answer many questions about the system's stability, the first order system provides a model for the analysis we are trying to complete for higher order systems. Second, the theorems derived in the scalar case are used in Section 3.2 and 3.3, where we employ the first order system results to bound regions of stability for higher order systems. In Section 3.2, we use the theory of Lie algebras and Lie groups to define and analyze a class of higher order systems for which necessary and sufficient conditions for stability can be obtained. The methods used are analogous to, and are actually a generalization of the techniques defined in Section 3.1. Section 3.3 presents the major contributions of this thesis for continuous time systems. In that section, we develop a technique for bounding the mean square*

* Mean square stability is defined in Chapter III.
response of higher order systems by the mean square response of the scalar system studied in Section 3.1. In Section 3.3 we demonstrate that the bandwidth of the colored noise process is an important consideration in determining the stability properties of system 1.1. We provide a boundary for the known region of stability of system 1.1 that explicitly involves the bandwidth of the colored noise. For large (but finite) bandwidths and for small damping the boundary derived provides the largest known region of stability for the systems considered. Also, the new bound is derived in such a way that it remains valid in the limit as the colored noise parameter becomes white. Any bound that depends only on the mean square value of the noise process will necessarily fail as the bandwidth becomes arbitrarily large. The last part of Section 3.3 defines a procedure for optimizing the strength of the derived bound.

In Chapter IV we attempt to answer an analogous set of questions about the stability properties of the (discrete-time) system (1.2). Unfortunately, the theorems from Chapter III do not apply to discrete-time systems and entirely new procedures must be used. As in the continuous
time case, we are able to demonstrate the importance of the shape of the autocorrelation function of the colored noise process. Suppose system (1.2) represents a first order system. Then, in the case that the colored noise process is generated by passing white noise through a low-order non-recursive system, we are able to obtain a set of necessary and sufficient conditions for the stability of system (1.2). If \( f(n) \) is generated recursively and the autocorrelation function of \( f(n) \) is always positive, then necessary conditions for the mean square stability of system (1.2) have been obtained.

In their algebraic form, the criteria for stability involves the forming of certain large matrices and the finding of the largest eigenvalue of each. These operations have been computerized and the results plotted so that some conclusions can be made and different cases can be compared.

Chapter V is a summary of the research completed and contains some ideas for future research and some speculation about the possibilities of generalizing our results.
CHAPTER II

BACKGROUND MATERIAL

In this chapter, we present the background material for studying the stability of stochastic dynamical systems. In addition to providing a brief survey of the available literature on the subject, this chapter establishes the framework under which we will be working and makes precise the meaning of the formulae used.

We have found it convenient to divide this chapter into two sections, one for continuous-time systems, the other covering discrete-time systems. In both cases, however, we will be considering systems that evolve on a continuous state space.

Section 2.1 provides an existence and uniqueness proof for the continuous-time systems under consideration. It is interesting to note that we were unable to find any applicable proof in the literature, especially considering the number of technical articles that have appeared on the subject. In that section, we also define the types of stochastic stability to be considered, and discuss some of the references to
which we will compare the results derived in Chapter II.

Section 2.2 provides an analogous discussion for the discrete-time systems under consideration. In this case, there is no difficulty with the existence of solutions, but a demonstration of the fact that the augmented system defines a Markov process is included. As part of this demonstration, we define the conditional expected value operators that are required for establishing stability.

For a survey of the available literature beyond the references discussed below, the reader is directed especially to the survey paper by Kozin [K3], and the proceedings of a symposium held at the University of Warwick, [W5].

2.1 Continuous-Time Systems

We will be dealing with vector-valued stochastic differential equations of the form:

\[
\begin{align*}
\text{dx}(t, \omega) &= (A + f(t, \omega)B) \text{x}(t, \omega) \, dt \\
\text{x}(0, \omega) &= x_0(\omega)
\end{align*}
\]  

(2.1.1)
In which $f(t, \omega)$ is a scalar valued, colored noise process, $A$ and $B$ are $(n \times n)$ constant matrices, $x(t) \in \mathbb{R}^n$, and $\omega \in \Omega$, where $(\Omega, A, \mathbb{P})$ is the basic probability space. We will consider $f(t, \omega)$ to have been generated by the stochastic Itô differential equation

$$df(t, \omega) = -af(t, \omega) + \sigma d\beta_t(\omega)$$

(2.1.2)

$$f(0, \omega) = f_0(\omega)$$

where \{\beta_t, A_t, 0 \leq t \leq T < \infty\} is a Brownian motion.

It is known that there exists a separable, measurable process \{\{f(t, \omega), a \leq t \leq T\}\} with the following properties. [W1, page 150]

- $P_1$: for each $t$ in $[0, T]$, $f(t, \omega)$ is $A_t$-measurable

(2.1.3)

- $P_2$: \(\int_0^T \mathbb{E}[f^2(t, \omega)] < \infty\)

(2.1.4)

- $P_3$: \{\{f(t, \omega)\}\} satisfies (2.1.2) with $f(0, \omega) = f_0(\omega)$

(2.1.5)

- $P_4$: \{\{f(t, \omega), 0 \leq t \leq T\}\} is sample continuous

w.p.1.
P_5: \{f(t,\omega), 0 \leq t \leq T\} is unique with probability 1. \hfill (2.1.7)

P_6: \{f(t,\omega), 0 \leq t \leq T\} is a Markov process. \hfill (2.1.8)

It will be assumed that \(f_0(\omega)\) is chosen so that 
\(f(t,\omega)\) is a stationary Gaussian random process for 
\(0 \leq t \leq T\), and \(x_0(\omega)\) is independent of the process 
\(f(t,\omega)\).

Let us now consider the question of existence and uniqueness for solutions of the system (2.1.1). A straightforward approach to the problem is to augment the system, so that the coefficients from Equation (2.1.2) become part of the augmented matrix, which would then have dimension \((n+1)\times(n+1)\). However, with this approach, the coefficient matrix of the right hand side of Equation (2.2.1) contains the term \(x(t,\omega)f(t,\omega)\), and violates the global Lipschitz conditions required in all the documented proofs of existence and uniqueness. The proofs given by Koch [K8, K11] based upon the use of McShane integrals [M3 - M6] also fail because the colored noise processes \(f(t,\omega)\) is neither Lipschitz nor Brownian motion. We will take an approach that mimics the proof given by
Wong [W1, W11, W12] but we will let \( f(t, \omega) \) play the role usually given to a Brownian Motion in such proofs. Properties (1) - (6) of \( f(t, \omega) \) guarantee the success of the proof.

We will use the norms:

\[
|x| = (x^T x)^{1/2} 
\]

\[
|G| = (\sum_{i=1}^{n} \sum_{j=1}^{m} G_{ij}^2)^{1/2} = [\text{tr}(GG^T)]^{1/2} 
\]

where \( x \) and \( G \) represent a vector and an \( nxm \) matrix, respectively.

In Section 3.3, we are often required to interchange the order of (time) integration and expected value of the random process \( f(t, \omega) \). We will establish the validity of changing the order of these operations as part of the proof of the following lemma:

**Lemma 2.1.1**

Let the random process \( f(t, \omega) \) be generated by Equation (2.1.2) so that it is measurable, separable, stationary, and has properties \( P_1 \) through \( P_6 \) given above. Define a sequence of (vector valued) random processes \( \{F_n(t, \omega), 0 \leq t \leq T\} \) as follows:
\[ F_0(t, \omega) = x_0(\omega) \tag{2.1.9} \]

\[ F_{n+1}(t, \omega) = \int_0^t f(\sigma, \omega) \, e^{-A\sigma B} e^{A\sigma} F_n(\sigma, \omega) \, d\sigma \tag{2.1.10} \]

where \( x_0(\omega) \) is \( A_0 \)-measurable and \( E\{x_0^2\} < \infty \).

Then, for each \( n \), \( F_n(t, \omega) \) is jointly measurable in \( (\omega, t) \) and, for each \( t \), it is \( A_t \)-measurable, \( t \in [0, T] \).

Also, the partial sums, \( \sum_{n=0}^{k} F_n(t, \omega) \) converge in quadratic mean, uniformly in \( t \).

**Remark on Lemma 2.1.1**

The factors \( e^{-A\sigma B} e^{A\sigma} \) appear in the differential equation for \( y(t, \omega) = e^{-At} x(t, \omega) \). That is;

\[
\begin{align*}
\frac{dy(t, \omega)}{dt} &= [-Ae^{-At} x(t, \omega) + e^{-At}[A+f(t, \omega)B]x(t, \omega)]dt \\
&= f(t, \omega) e^{-AtB} x(t, \omega)dt
\end{align*}
\]

so

\[
y(t, \omega) - y(0, \omega) = \int_0^t f(\sigma, \omega) e^{-A\sigma B} e^{A\sigma} y(\sigma, \omega) \, d\sigma \tag{2.1.11}
\]
Without loss of generality we are really establishing an existence and uniqueness proof for Equation (2.1.11). We have made this transformation to avoid accumulating an excessive number of terms in equations like (2.1.13) and (2.1.14), below.

Proof of Lemma 2.1.1

By induction, we will show that $F_n(t,\omega)$ is:

(i) A measurable random process,
(ii) $A_t$-measurable, and satisfies
(iii) $E\{F_n^2(t)\} < \infty$ for all $t \in [0,T]$. 

First, we will verify these properties for $n = 0$. Clearly $F_0(t,\omega)$ is a measurable process because $x_0(\omega)$ is an $A$-measurable function, in fact it is $A_0$-measurable. $E\{F_0^2(t)\} < \infty$ by hypothesis. Now assume that (i), (ii), and (iii) are satisfied for $n = 0,1,2,\ldots,k$.

$$F_{k+1}(t,\omega) = \int_0^t f(\sigma,\omega) \left[ e^{-A_\sigma B} e^{A_\sigma} \right] F_k(\sigma,\omega) d\sigma \quad (2.1.12)$$

The integrand in (2.1.12) is jointly measurable in $(\omega,t)$ because each of its three factors have that property.
But

\[ F_{k+1}(t, \omega) = \int_0^t \int_0^{\sigma_1} \ldots \int_0^{\sigma_k} f(\sigma_1, \omega) f(\sigma_2, \omega) \ldots f(\sigma_{k+1}, \omega) \]

\[ e^{-A\sigma_1} B e^{A(\sigma_1 - \sigma_2)} B \ldots e^{A(\sigma_2 - \sigma_3)} B \ldots \]

\[ B e^{A\sigma_{k+1}} x_0(\omega) d\sigma_{k+1} d\sigma_k \ldots d\sigma_1 \quad (2.1.13) \]

\[ |F_{k+1}(t, \omega)|^2 = \int_0^t \int_0^{\rho_1} \int_0^{\sigma_1} \ldots \int_0^{\rho_k} f(\sigma_1, \omega) f(\rho_1, \omega) \cdot f(\sigma_2, \omega) \ldots f(\rho_{k+1}, \omega) \cdot \]

\[ e^{-A\rho_1} B e^{A(\rho_1 - \rho_2)} B \ldots e^{A(\rho_2 - \rho_3)} B \ldots \]

\[ [e e^{-A\rho_1} B e^{A(\rho_1 - \rho_2)} B \ldots e^{A(\rho_2 - \rho_3)} B \ldots e^{A\rho_{k+1}}] \]

\[ x_0(\omega) [e e^{-A\rho_1} B e^{A(\rho_1 - \rho_2)} B \ldots e^{A\rho_{k+1}}] \]

\[ d\sigma_{k+1} d\rho_{k+1} \ldots d\sigma_1 d\rho_1 \quad (2.1.14) \]
Let $K$ denote $\sup_{-T \leq t \leq T} |e^{At}|$. Then,

$$|F_{k+1}(t, \omega)|^2 \leq \int_0^t \int_0^{\sigma_k} \int_0^{\rho_k} f(\sigma_1, \omega) f(\rho_1, \omega) f(\sigma_2, \omega) \ldots f(\rho_{k+1}, \omega) |K^{(2k+4)}| B^{2k+2} .$$

$$\cdot |x_T^T(\omega)x_0(\omega)| d\sigma_{k+1} d\rho_{k+1} \ldots d\sigma_1 d\rho_1$$

(2.1.15)

The integrand in (2.1.15) is positive, so Fubini's theorem allows interchanging the order of integration and expectation. That is,

$$E \left\{|F_{k+1}(t, \omega)|^2\right\} \leq \int_0^t \int_0^{\sigma_k} \int_0^{\rho_k} E \left\{|f(\sigma_1, \omega)\ldots f(\rho_{k+1}, \omega)| x_T^T(\omega)x_0(\omega)|\right\}$$

$$f(\rho_{k+1}, \omega) |x_T^T(\omega)x_0(\omega)|$$

$$\cdot |K^{(2k+4)}| B^{2k+2} d\sigma_{k+1} d\sigma_k \ldots d\sigma_1 d\rho_1$$

(2.1.16)
Next, using the moment factoring theorem for Gaussian random variables, the stationarity of \( f(t, \omega) \), and the fact that \( E\{ f(t_1, \omega) f(t_2, \omega) \} \leq E\{ f^2(t_1, \omega) \} \), we obtain:

\[
E\{|F_{k+1}(t, \omega)|^2\} \leq \int_0^t \cdots \int_0^t E\{|f(0, \omega)|^{2k+2}\}
\]

\[
E\{|x_0^T(\omega)x_0(\omega)|^2\}K^{2k+4}
\]

\[
|B|^{2k+2} \, d\sigma_{k+1} \cdots d\rho_1
\]

\[
= K^{2k+4} |B|^{2k+2} E\{|x_0(\omega)|^2\}
\]

\[
\frac{1 \cdot 3 \cdot 5 \cdots (2k+1)E\{f^2(0, \omega)\}^{2k+2} t^{2k+2}}{(k+1)! (k+1)!}
\]

\[
(2.1.17)
\]

\[
\leq K^2 E\{|x_0(\omega)|^2\} \left[ \frac{\sigma_f^2 K^2 |B|^{2t^2}}{2^{-(k+1)} (k+1)!} \right]^{k+1}
\]

\[
(2.1.18)
\]
which is \( < \infty \), for \( 0 \leq t \leq T \), and
\[
\sigma_f^2 = E\{|f(o, \omega)|^2\}.
\]

By evaluating Equation (2.1.18) at \( t = T \), we see that there exists an \( M < \infty \), such that \( E\{|F_{k+1}(t, \omega)|^2\} < M \) for all \( t \in [0, T] \). This implies:

\[
E\{|F_{k+1}(t, \omega)|\} < M^{1/2} < \infty \text{ for all } t \in [0, T] \quad (2.1.19)
\]

and

\[
\int_0^T E|F_{k+1}(t, \omega)| \, dt < \infty \quad (2.1.20)
\]

To show that \( F_{k+1}(t, \omega) \) is \( A_t \)-measurable we must consider Equation (2.1.12). By our induction hypothesis and Equation (2.1.3), the integrand in (2.1.12) is \( L \times A_t \)-measurable* so Fubini's theorem implies \( F_{k+1}(t, \omega) \) is \( A_t \)-measurable. (Equation (2.1.20) justifies using Fubini's theorem here.)

We will show that \( F_{k+1}(t, \omega) \) is a measurable random process by first showing that it is mean square continuous. From Equation (2.1.13),

* \( L \) stands for the Lebesgue measurable sets in \([0, T]\).
\[ [F_{k+1}(t, \omega) - F_{k+1}(s, \omega)] = \int_s^t \int_0^{\sigma_1} \cdots \int_0^{\sigma_k} f(\sigma_1, \omega) \cdots \]

\[ f(\sigma_{k+1}, \omega) e^{-A\sigma_1 B \sigma_1 \cdots B \sigma_{k+1}} \]

\[ x_0(\omega) d\sigma_{k+1} \cdots d\sigma_1 \quad (2.2.21) \]

Then, (similar to Equation (2.1.17),

\[ E |F_{k+1}(t, \omega) - F_{k+1}(s, \omega)|^2 \leq k^{2k+4} |B|^{2k+2} \]

\[ E \{ |x_0(\omega)|^2 \} \{ 1 \cdot 3 \cdot 5 \cdots (2k+1) \sigma_f^{2k+2} \} \]

\[ \int_s^t \int_0^{\sigma_1} \frac{k}{k!} \frac{\rho_1}{k!} d\sigma_1 d\rho_1 \]

\[ = k^{2k+4} |B|^{2k+2} E \{ |x_0(\omega)|^2 \} \{ 1 \cdot 3 \cdot 5 \cdots (2k+1) \sigma_f^{2k+2} \} \]

\[ \frac{(t^{k+1} s^{k+1})^2}{(k+1)!(k+1)!} \]
This shows that $F_{k+1}(t,\omega)$ is mean-square continuous, and, therefore, a measurable version can be chosen.\[\text{[W1, page 151, page 45].}\]

This completes the induction. It remains to show that $\sum_{n=0}^{k} F_n(t,\omega)$ converges in quadratic mean, uniformly in $t$. We will demonstrate this by showing that $\sum_{n=0}^{k} F_n(t,\omega)$ satisfies the Cauchy criterion* \[
\lim_{m,k \to \infty} E\left\{ \left| \sum_{n=0}^{k} F_n(t,\omega) - \sum_{n=0}^{m} F_n(t,\omega) \right|^2 \right\} = 0 \quad (2.1.24)
\]

For each $t$ and $m > k$,

* [J2, page 58]
\[
E \left\{ \sum_{n=0}^{k} F_n(t, \omega) \right. - \sum_{n=0}^{k} F_n(t, \omega) \right\}^2 = E \left\{ \sum_{n=k+1}^{m} F_n(t, \omega) \right\}^2 \\
\leq E \left\{ \sum_{n=k+1}^{m} |F_n(t, \omega)|^2 \right\} \\
= \sum_{n=k+1}^{m} E \{|F_n(t, \omega)|^2\} \\
\leq K^2 E \{|x_0(\omega)|^2\} \sum_{n=k+1}^{m} \frac{[2 \sigma_f^2 K^2 |B| t^2]^n}{n!} \\
(because\ of\ (2.1.18)) \\
\leq K^2 E \{|x_0(\omega)|^2\} 2^{-k} \sum_{n=k+1}^{m} \frac{[4 \sigma_f^2 K^2 |B| t^2]^n}{n!} \\
\leq 2^{-k} K^2 E \{|x_0(\omega)|^2\} \sum_{n=0}^{\infty} \frac{[4 \sigma_f^2 K^2 |B| t^2]^n}{n!} \\
= 2^{-k} K^2 E \{|x_0(\omega)|^2\} \exp(4 \sigma_f^2 K^2 |B| t^2) \\
\tag{2.1.25} \\
\rightarrow 0\ as\ K \rightarrow \infty \tag{2.1.26}
\]
The convergence is uniform in $t$ on the finite interval $[0,T]$.

This shows that $\sum_{n=0}^{k} F_n(t,\omega)$ converges in quadratic mean, uniformly in $t$, completing the proof of Lemma 2.1.1.

Lemma 2.1.2 will make use of these results.

Lemma 2.1.2

Let $F(t,\omega)$ denote the vector-valued random process to which $\sum_{n=0}^{k} F_n(t,\omega)$ converges in quadratic mean. Then $F(t,\omega)$ is quadratic mean continuous.

Proof of Lemma 2.1.2 - See Jazwinski [J2, page 62]

A further consequence of Lemma 2.1.2 is that there exists a separable, measurable version of $F(t,\omega)$. This is proven by Wong [W1, proposition 2.3.] Henceforth, we will assume $F(t,\omega)$ is a measurable random process.

The following two theorems represent the main results of this section.

Theorem 2.1.1

Let $F(t,\omega) = \sum_{n=0}^{\infty} F_n(t,\omega)$ in the sense given by Lemma 2.1.2. Then $F(t,\omega)$ has the following properties.

(i) $F(t,\omega)$ is $\mathcal{A}_t$-measurable for $t \in [0,T]$. (2.1.27)
(ii) \[ \int_0^T E\{|F(t,\omega)|^2\}dt < \infty \] (2.1.28)

(iii) \( \{F(t,\omega), 0 \leq t \leq T\} \) satisfies Equation (2.1.11) with \( F(0,\omega) = x_0(\omega) \). 

(iv) \( \{F(t,\omega), 0 \leq t \leq T\} \) is unique with probability 1. 

(2.1.30)

Remark: \( \{F(t,\omega), 0 \leq t \leq T\} \) is not a Markov process.

**Proof of Theorem 2.1.1**

For each \( n \) and for every \( t \), \( F_n(t,\omega) \) has been shown to be \( A_t \)-measurable. Then, for every \( k \) and every \( t \),
\[
\sum_{n=0}^{k} F_n(t,\omega) \text{ is } A_t \text{-measurable, so } F(t,\omega) \text{ is } A_t \text{-measurable.}
\]
This proves (i). To prove (ii), we can use (2.1.18).

\[
\int_0^T E\{|F(t,\omega)|^2\}dt = \int_0^T E\{|\sum_{n=0}^{\infty} F_n(t,\omega)|^2\}dt
\]

\[
\leq \int_0^T \sum_{n=0}^{\infty} E\{|F_n(t,\omega)|^2\}dt
\]

\[
\leq TK^2 E\{|x_0(\omega)|^2\} \sum_{n=0}^{\infty} \frac{[2\sigma_f^2K^2|B|^2T^2]^n}{n!}
\]

(2.1.31)
\[
 TK^2 \mathbb{E}\{ |X_0(\omega)|^2 \exp[2\sigma_f 2K^2 |B|^2T^2] \}
\]

\[ (2.1.32) \]

\[
 < \infty
\]

To prove \( F(t,\omega) \) is a solution of \((2.1.11)\) with 
\( F(0,\omega) = x_0(\omega) \), define 
\[
 \Delta_t = F(t,\omega) - F(0,\omega) - \int_0^t f(\sigma,\omega) e^{-A\sigma} Be^{A\sigma} F(\sigma,\omega) d\sigma
\]

\[ (2.1.34) \]

Using \((2.1.10)\), we can rewrite \( \Delta_t \) as 
\[
 \Delta_t = [F(t,\omega) - \sum_{n=0}^{k} F_n(t,\omega)] - \int_0^t f(\sigma,\omega) e^{-A\sigma} Be^{A\sigma}
\]

\[
 \cdot [F(\sigma,\omega) - \sum_{n=0}^{k-1} F_n(\sigma,\omega)] d\sigma
\]

The first term goes to zero in quadratic mean as 
k \to \infty by definition. The second term can be rewritten as:
\[ \int_0^t f(\sigma, \omega) e^{-A\sigma} Be^{A\sigma} \left[ \sum_{n=k}^{\infty} F_n(\sigma, \omega) \right] d\sigma \]

\[ = \sum_{n=k+1}^{\infty} F_n(\sigma, \omega) \quad \text{by Equation (2.1.10)} \]

But

\[ \lim_{k \to \infty} \left| \sum_{n=k+1}^{\infty} F_n(\sigma, \omega) \right|^2 = 0 \quad \text{by Equation (2.1.26)} \]

Therefore \[ \mathbb{E}\{ t^2 \} = 0 \] for every \( t \in [0, T] \), so

\[ F(t, \omega) = F(0, \omega) + \int_0^t f(t, \sigma) e^{-A\sigma} Be^{A\sigma} F(\sigma, \omega) d\sigma \]

(2.1.35)

with probability 1 for each \( t \in [0, T] \).

If we choose a separable version of the integral in Equation (2.1.35), then

\[ \mathbb{P}\{ F(t, \omega) = F(0, \omega) + \int_0^t f(\sigma, \omega) e^{-A\sigma} Be^{A\sigma} F(\sigma, \omega) d\sigma, \]

\[ 0 \leq t \leq T \} = 1. \quad (2.1.36) \]
This proves (iii).

To prove uniqueness, suppose that \( F(t,\omega) \) and \( \tilde{F}(t,\omega) \) are both solutions of Equation (2.1.11), that satisfy (i), (ii), and (iii). Then,

\[
F(t,\omega) - \tilde{F}(t,\omega) = \int_{0}^{t} f(\sigma,\omega) e^{-A\sigma} B e^{A\sigma} [F(\sigma,\omega) - \tilde{F}(\sigma,\omega)] d\sigma
\]

\[
|F(t,\omega) - \tilde{F}(t,\omega)| \leq \int_{0}^{t} |f(\sigma,\omega)||e^{-A\sigma} B e^{A\sigma}||F(\sigma,\omega) - \tilde{F}(\sigma,\omega)| d\sigma
\]

(2.1.37)

But, \( f(t,\omega) \) is separable, so \( \sup_{t \in [0, T]} f(t,\omega) \) is a random variable. For almost every \( \omega \), \( f(t,\omega) \) is continuous on \( 0 \leq t \leq T \) and \( f(0,\omega) \) is finite. Therefore, \( \sup_{t \in [0, T]} f(t,\omega) \) is finite for almost all \( \omega \) and so is \( \sup_{t \in [0, T]} f(t,\omega) \).

[WL, page 42]

Define

\[
\eta(\omega) = \sup_{0 \leq t \leq T} |f(t,\omega)||e^{-At} Be^{At}|
\]

(2.1.38)
A and B are constant matrices, so \( \eta(\omega) \) is finite for almost all \( \omega \). Then (2.1.37) becomes

\[
|F(t,\omega) - \bar{F}(t,\omega)| \leq \eta(\omega) \int_0^t |F(\sigma,\omega) - \bar{F}(\sigma,\omega)| \, d\sigma
\]

(2.1.39)

For each \( \omega \) for which \( \eta(\omega) \) is finite, we can solve (2.1.39) for \(|F(\sigma,\omega) - \bar{F}(\sigma,\omega)|\),

\[
|F(t,\omega) - \bar{F}(t,\omega)| = |F(0,\omega) - \bar{F}(0,\omega)| \exp(\eta(\omega) t)
\]

Therefore, \(|F(t,\omega) - \bar{F}(t,\omega)| = 0 \) for each \( t \in [0,T] \) and for almost every \( \omega \).

(2.1.40)

If \( \bar{F}(t,\omega) \) is chosen to be separable, then

\[
P(F(t,\omega) = \bar{F}(t,\omega) \text{ for all } t \in [0,T]) = 1
\]

This completes the proof of Theorem (2.1.1).
Theorem 2.1.2

There exists a separable and measurable random process \(x(t, \omega)\) with the following properties.

(i) \(x(t, \omega)\) is \(A_t\)-measurable for \(t \in [0, T]\).

(ii) \(\int_0^T E\{|x(t, \omega)|^2\}dt < \infty\).

(iii) \(\{x(t, \omega), 0 \leq t \leq T\}\) satisfies Equation (2.1.1) with \(x(0, \omega) = x_0(\omega)\).

(iv) \(\{x(t, \omega), 0 \leq t \leq T\}\) is unique with probability 1.

Proof

Let

\[\{x(t, \omega), 0 \leq t \leq T\} = \{e^{A_t} F(t, \omega), 0 \leq t \leq T\}\]

Then, properties (i), (ii), and (iv) are a consequence of the properties of \(F(t, \omega)\) proven in Theorem 2.1.1.

To prove (iii), note that \(F(t, \omega)\) solves (2.1.11):
\[ F(t, \omega) - F(0, \omega) = \int_0^t f(\sigma, \omega) e^{-A\sigma} B e^{A\sigma} F(\sigma, \omega) d\sigma \]

\[ = \int_0^t f(\sigma, \omega) e^{-A\sigma} B x(\sigma, \omega) d\sigma \quad (2.1.45) \]

Next, we can multiply both sides of Equation (2.1.45) by \( e^{At} \):

\[ x(t, \omega) - x(0, \omega) = \int_0^t f(\sigma, \omega) e^{A(t-\sigma)} B x(\sigma, \omega) d\sigma \]

(2.1.46)

which is just the integral form of Equation (2.1.1) that can be obtained through the variation of constants formula. Theorem 2.1.2 is now proven.

**Stability Concepts: Continuous Time Systems**

One major difficulty in studying and comparing the results found in the literature is the abundance of concepts and definitions surrounding the phrase "stochastic stability". To avoid any problems of this kind, only the following types of stability will be considered here [K3]:
Definition

The zero equilibrium solution of system (2.1.1) is almost surely stable if

$$\lim_{\delta \to 0} \mathbb{P} \left\{ \sup_{\delta < \delta} \sup_{t \geq 0} |x(t, \omega)| > \varepsilon \right\} = 0 \quad (2.1.47)$$

for any given \( \varepsilon > 0 \).

Definition

The zero equilibrium solution of system (2.1.1) is said to possess exponential stability of the pth mean if there exists positive constants \( \alpha, \beta, \) and \( \delta \), such that \( |x_0(\omega)| < \delta \) implies that for all \( t > 0 \),

$$\mathbb{E} \left\{ \left| x(t, \omega) \right|^p \right\} \leq \beta \left| x(0, \omega) \right|^p \exp(-\alpha t) \quad (2.1.48)$$

where

$$\left| x(t, \omega) \right|^p = \sum_{i=1}^{n} \left| x_i(t, \omega) \right|^p$$
In this thesis, the case $p = 2$ will receive the most consideration and the phrase "exponentially stable in the mean square" will be used.

The following paragraphs discuss some of the results from previous research that have been reported in the literature. By far, the majority of the reports are restricted to the consideration of systems in which the stochastic parameter is a white noise process, but we will defer any discussion of these works until Appendix A.

We are primarily concerned with those systems containing a stochastic parameter that has been generated by passing white noise through a linear system, but other types of non-white parameters have been treated as well. For example, Frish [Fl] and McKenna and Morrison [M1] have studied the stability of systems in which the stochastic parameter is a random telegraph wave. They have obtained exact expressions for various moments of the output process and have isolated examples of systems for which the first moment of the output process decay exponentially but higher moments grow exponentially. We have noted a similar phenomenon for the systems treated in Chapter III.
A great deal of research has also been centered on finding stability criteria for linear systems containing a stochastic parameter that is a finite-state Markov Chain [M1, M3]. These analyses are also directed at finding exact formulae for the moments of the random output function. In his survey paper, Kozin [K3] refers to several other examples of research into these systems.

Among the systems containing a parameter that has been generated by passing white noise through a linear system, there are two categories of major interest. In the first case, the noise parameter is multiplied by a small constant, so the overall effect of the noise is a perturbation. In the general case, we allow the noise to be of arbitrary strength.

Kolovskii and Troitskaia [K5] have studies stability of linear systems with random coefficients by the method of perturbations. They have expanded the solution process into formulae very much like Equation (3.1.10), but they have used a small parameter to justify truncating the series. The danger of using such an approach is that there is no apriori guarantee that powers of the small parameter are sufficiently small to justify eliminating terms involving higher
order moments of the random coefficients. For example, the moments of a Gaussian random variable can not be bounded in such a way.

Reference [B2], by Blankenship, discusses techniques for approximating the behavior of dynamical systems containing a small random parameter. His examples are drawn from the theory of electromagnetic wave transmission and waveguides, which is apparently a rich source of problems of this type. His discussion overlaps some of the scalar case methods and Lie group analysis contained in Chapter III, below, but he chooses the undamped harmonic oscillator for an example. With that example, it is possible to carry the analysis a great deal further than what is possible with the damped harmonic oscillator problem. Blankenship further discusses some limiting methods based upon approximating the solution of the original dynamical system over small time intervals. An assumption required for this approach is that the colored noise parameter be bounded, which includes most physical processes, but is not consistent with the mathematical model we have chosen.

In Reference [W5], Wedig has reported some dramatic results for systems containing a small colored noise
parameter. His research is of particular interest to us because he studies the same example we choose in Section 3.3. The use of a small parameter, however, enables him to linearize his dynamical system so that it looks like the systems treated in Section 3.1 with complex coefficients. With this simplification, he is able to analyze a harmonic oscillator that is excited by bandpass noise. His results show a marked tendency toward instability if the noise center frequency is twice the natural frequency of the unperturbed system. However, after linearizing the original damped harmonic oscillator problem, he is working with a very lightly damped system, so instabilities are to be expected. The methods we use in Chapter III allow systems with bandpass noise parameters to be analyzed only in the first order case. We have no results that can be meaningfully compared with Wedig's.

Several papers [K11, C1, G1, I1] which contain stability theorems for linear dynamical systems whose stochastic parameters are not of the white noise type and which do not involve any small parameters have
appeared in the literature. The theorems obtained all require the stochastic parameter process to be ergodic with almost surely continuous sample functions. The stability boundaries obtained in the paper by Caughey and Grey [Cl] extend Koxin'g [K11] results by using different matrix norms and manipulations. They also extend the theory to allow deterministic inputs to be applied to the systems under investigation.

Through the use of a simple result from the theory of pencils of matrices, Infante [Il] was able to dramatically improve the known regions of stability for the systems studied by Caughey and Grey. The strength of his theorems depends upon the choice of the matrix used in the (quadratic form) Lyapunov functions. However, for second order systems, he was able to obtain bounds for a general quadratic form and optimize over the entries of the matrix at the end. Infante's results establish, for the first time, that the variance of the stochastic parameter can approach infinity as the system damping approaches infinity without causing instability. However, the stability tests obtained by each of these theorems do not involve the bandwidth of the parameter process. In Chapter III, we demonstrate
that there does exist a power-bandwidth tradeoff.

For the damped harmonic oscillator problem studied in Section 3.3, the theorem of Infante is used as the reference to which our results are compared.

Brockett and Willems [B6] have applied the theory of completely symmetric systems to the stability problem for linear systems with a colored noise feedback gain. By assuming that the colored noise process is ergodic, they are able to derive a stability criterion based upon the expected value of the maximum eigenvalue of the closed-loop system matrix. The strength of their method is that there is no requirement that the colored noise be confined to any bounded region for all time. That is, the noise gain can be negative for some time as long as it compensates by being sufficiently positive at other times. This is in distinction to the criteria that would result from direct use of Zames' theorems [Z1]. Their work, however, has two disadvantages. First, it only applies to completely symmetric systems, which implies that the system have only real poles. Second, the criteria derived involve evaluating expected values of very non-linear functions
of the colored noise process which do not even admit analytic expressions.

The proofs used in their theorems rely on the same eigenvalue inequalities used by Infante [Il]. These criteria, therefore, are independent of the autocorrelation function of the colored noise and rely only on the moments of the process, and its ergodic property.

Further extensions of the theory that cover approximation procedures for distributed parameter systems are given in References [M1] and [W6], respectively.

Another reference which is of great importance to this research is Rabotnikov's article [R1]. However, we have deferred discussing his paper until Section 3.3, where it is used.
2.2 **Discrete-Time Systems**

In Chapter IV, we will be considering systems that evolve on a continuous state space in discrete time steps. The most general class of systems to be treated is described by difference equations of the form:

\[ x_{k+1} = F_k(w_k, x_k), \quad x_0 = \text{random variable} \quad (2.2.1) \]

in which \( F_k \) is a nonlinear vector valued function and \( w_k \) represents a sample from a white noise sequence. (More precise definitions will be given below.) The theory by which the stability properties of the system (2.2.1) are derived relies very heavily on establishing that the sequence \( \{x_n\} \) is a discrete-parameter Markov process. It is not clear from the outset that the required conditional probabilities are well defined.

The difficulty in establishing that system (2.2.1) defines a Markov process involves showing that transition probabilities can be used as conditional expectations. Gikhman and Skorokhod [G3] refer to the Markov processes defined by transition probabilities satisfying the Chapman-Kolmogorov equation as "Markov processes in the broad sense". Those processes satisfying the conditional expectation requirements are termed "Markov processes
in the narrow sense". In that reference, the authors establish that a narrow sense Markov process can be constructed from a broad-sense Markov process, but the assumption that system (2.2.1) generates a Markov process in the narrow sense must still be justified. We will take this approach, and the Chapman-Kolmogorov equation will appear as a by-product. System (2.2.1) is given a precise meaning in the following paragraphs.

The initial conditions, \( x_0 \), must be defined on some given probability space \((X_0, \mathcal{X}_0, \mu)\) and each random variable, \( w_k \), will be defined on a space \((\Omega_k, \mathcal{X}_k, P_k)\); \( k = 0, 1, 2, \ldots \). For each integer \( k \geq 0 \), \( X_k \) will denote \( n \)-dimensional Euclidean space, \( \mathbb{R}^n \), the range space of \( F_{k-1} \). The map \( F_{k-1}: (\Omega_{k-1} \times X_{k-1}) \to X_k \) is measurable with respect to the product algebra \((\mathcal{A}_{k-1} \times \mathcal{X}_{k-1})\). The \( \sigma \)-algebra \( \mathcal{X}_k \) defined on \( X_k \) is assumed to be the Borel sets of \( \mathbb{R}^n \) and, in particular, contains the singleton \( \{x_k\} \) for every \( x_k \in X_k \). These assumptions on the functions \( F_k \) are required for the transition probabilities discussed below to be well defined as conditional probabilities. In order to consider the probabilistic properties of sequences \( x_0, x_1, \ldots \), we must first establish the existence of a probability measure on the product measure space \((X, \mathcal{X}) = (\prod X_k, \prod \mathcal{X}_k)\).
A map \( p_{km} : (X_k, X_m) \to [0,1] \) is called a transition probability from \( X_k \) into \( X_m \) if:

1. \( p_{km}(x_k, \cdot) \) is a probability measure on \( (X_m, X_m) \) for every fixed point \( x_k \in X_k \).

2. \( p_{km}(\cdot, A) \) is \( X_k \)-measurable for every \( A \in X_m \).

3. For arbitrary \( k, m, n, x, \) and \( A \) (with \( k < m < n \)), the functions \( p_{km}, (x, A) \) satisfy the Chapman-Kolmogorov equation:

\[
p_{k,n}(x,A) = \int_{X_m} p_{k,m}(x,dy)p_{m,n}(y,A) \quad (2.2.2)
\]

**Notation:**

The spaces \( (\Omega_{k,m}, A_{k,m}, P_{k,m}) \) will denote probability spaces where:

\[
\Omega_{k,m} = \prod_{i=k}^{m} \Omega_i
\]

\[
A_{k,m} = \prod_{i=1}^{m} A_i
\]

and \( P_{k,m} \) is the product measure on \( A_{k,m} \). This is consistent with the above assumption that \( \{w_k\} \) is a white sequence.
\( \chi_A \) will denote the characteristic function of the set \( A \). \( \sigma \{ y_i, i = k, k+1, \ldots m \} \) denotes the \( \sigma \)-algebra generated by the random variables \( y_k, y_{k+1}, \ldots y_m \)

For fixed \( x \in X_k \), define the function:

\[
g_{k,m}^x : ( \prod_{i=k}^{m-1} \Omega_i ) \to X_m \text{ recursively by:}
\]

\[
g_{k,k+1}^x(w_k) = F_k(w_k, x)
\]

\[
g_{k,k+2}^x(w_k, w_{k+1}) = F_{k+1}(w_{k+1}, g_{k,k+1}^x(w_k))
\]

\[
\vdots
\]

\[
g_{k,m}^x(w_k, w_{k+1}, \ldots w_{m-1}) = F_{m-1}(w_{m-1}, g_{k,m-1}^x(w_k, w_{k+1}, \ldots w_{m-2})) \quad (2.2.3)
\]

It will also be convenient to use the function \( g_{k,m}^x(X_k \times \Omega_k \times \Omega_{m-1}) \to X_m \), defined by:

\[
g_{k,m}^x(x_k, w_k, w_{k+1}, \ldots w_{m-1}) = g_{k,m}^x(w_k, w_{k+1}, \ldots w_{m-1}).
\]
Let
\[ \eta_k = \{ x \in X_k | x \notin \Omega_{o,k}^{X_0} \} \]

**Lemma 2.2.1**

The function specified on \((X_k \times X_m)\) by

\[ p_{k,m}(x,A) = p_{k,m-1}(\{ w \in \Omega_{k,m-1}^{X} : g_{k,m}^X(w) \in A \} ) \quad (2.2.4) \]

for each \( A \subset X_m \), \( k = 0,1,2,...; m = k+1,k+2,... \) satisfies (i), (ii), and (iii).

Actually, \( p_{k,m}(x,A) \) can be an arbitrary probability measure on \((X_m \times X_m)\) for \( x \in \eta_k \), but choosing (2.2.4) for all \( x \in X_k \) simplifies the analysis.

**Proof of Lemma 2.2.1**

\( g_{k,m}^X(\cdot) \) is \( A_{k,m-1} \)-measurable because of the way it is constructed from the measurable functions \( F_k, F_{k+1},... F_{m-1} \). Therefore, for any fixed \( x_k \in X_k \), Equation (2.2.4) induces a probability measure on \((X_m,X_m)\). This proves (i).

By construction, \( g_{k,m}^X(x_k,w_k,w_{k+1},...,w_{m-1}) \) is \((X_k \times A_{k,m-1})\)-measurable, so by definition, \( \chi^{(g_{k,m})^{-1}}(A) \) is a \((X_k \times A_{k,m-1})\)-measurable function.
Furthermore, the characteristic function is non-
negative, so, by Fubini's theorem,

\[
\int_{(\Omega_{k,m-1})}^\chi (g_{k,m})^{-1}(A) \quad dP_{k,m-1} \text{ is } X_k\text{-measurable.}
\]

For any \( x \in X_k \), its value is given by:

\[
\int_{(\Omega_{k,m-1})}^\chi (g_{k,m})^{-1}(A) \quad dP_{k,m-1} = P_{k,m-1}(\{ w \in \Omega_{k,m-1}: g_{k,m}(w) \in A \})
\]

\[= P_{k,m}(x,A) \quad \text{(2.2.5)}
\]

This proves (ii).

To prove (iii), we must evaluate the integral

\[
I = \int_{X_m}^p_{k,m}(x,dy)p_{m,n}(y,A)
\]

\[= \int_{X_m}^p_{k,m}(x,dy)p_{m,n-1}(\{ w \in \Omega_{m,n-1}: g_{m,n}(w) \in A \})
\]

\[\text{(2.2.6)}
\]
By Equation (2.2.5), $P_{m,n-1}(\{w \in \Omega_{m,n-1}: g_{m,n}^w(w) \in A\})$ is $X_m$-measurable, and by the first part of this proof, $P_{k,m}(x, B)$ (for $B \in X_m$) is a probability measure on $(X_m, X_m)$, so (2.2.6) is a well defined integral. Again, resorting to the use of characteristic functions, Equation (2.2.6) becomes:

$$I = \int_{X_m} p_{k,m}(x, dy) \int_{\Omega_{m,n-1}} \chi_{(g_{m,n}^y)^{-1}(A)} dP_{m,n-1}^{m,n-1} \quad (2.2.7)$$

$$= \int_{\Omega_{m,n-1}} \int_{X_m} p_{k,m}(x, dy) \chi_{(g_{m,n}^y)^{-1}(A)} dP_{m,n-1}^{m,n-1} \quad (2.2.8)$$

by Fubini's theorem.

The inner integral in (2.2.8) can be evaluated as though $(w_k, w_{k+1}, \ldots, w_{n-1})$ were a fixed point. Let $B(w)$ denote the set $\{y \in X_m: g_{m,n}(y, w_m, w_{m+1}, \ldots, w_{n-1}) \in A\}$. Then, Equation 2.2.8 takes the form:

$$I = \int_{\Omega_{m,n-1}} p_{k,m}(x, B(w)) dP_{m,n-1}^{m,n-1}$$

$$= \int_{\Omega_{m,n-1}} p_{k,m}(x, \{y \in X_m: g_{m,n}(y, w_m, \ldots, w_{n-1}) \in A\}) dP_{m,n-1}^{m,n-1}$$
This completes the proof of (iii), and the lemma, so that $p_{k,m}(x,A)$ is a transition probability for system (2.2.1).

In order to consider the probabilistic properties of sequences $x_0, x_1, x_2, \ldots$, we must now establish the existence of a product measure on the product measurable space

$$(X, \mathcal{X}) = (\prod X_k, \Pi \mathcal{X}_k)$$

The Iterated Conditional Probabilities Theorem of C.I. Tulcea provides the answer. [L1, page 569]. This establishes the existence of a unique probability, $P$, on $(X, \mathcal{X})$ given by:
\[
\mathcal{P}( \prod_{j=0}^{n} A_j \times \prod_{j=n+1}^{\infty} X_j ) = \int_{A_0}^{A_1} \mathcal{P}_{0,1}(x_0, dx_1) \int_{A_2}^{A_n} \mathcal{P}_{1,2}(x, dx_2) \cdots \int_{A_{n-1}}^{A_n} \mathcal{P}_{n-1,n}(x_{n-1}, dx_n)
\]

where \( A_i \in X_i \), \( i = 0, 1, 2, \ldots \).

Therefore, \((X, \mathcal{X}, \mathcal{P})\) is a probability space.

We will let \( \mathcal{P}_{\ell_1, \ell_2, \ldots, \ell_j} \) denote the restriction of \( \mathcal{P} \) to the \( \sigma \)-algebra, \( \sigma\{x_{\ell_1}, x_{\ell_2}, \ldots, x_{\ell_j}\} \). For example, \( \mathcal{P}_k(A) \) will denote \( \mathcal{P}(X_0 \times X_1 \times \cdots \times X_{k-1} \times A \times X_{k+1} \times \cdots) \).

Below, we will be required to evaluate integrals of the form:

\[
I = \int_B f(x_k) \mathcal{P}_k(dx_k),
\]

where \( f(x_k) \) is some \( X_k \)-measurable function, and \( B \in X_k \).

If \( f(x_k) = x_D \) for some \( D \in X_k \), then by Equation (2.2.10),

\[
\int_B x_D \mathcal{P}_k(dx_k) = \mathcal{P}_k(B \cap D) = \int_{X_0}^{X_1} \mathcal{P}_{0,1}(x_0, dx_1) \int_{x_{k-1}}^{x_k} \mathcal{P}_{k-1,k}(x_{k-1}, dx_k)
\]
Then, by the monotone convergence theorem, if \( f(x_k) \geq 0 \), and is \( x_k \)-measurable,

\[
\int_{B} f(x_k) p_k(dx_k) = \int_{B} \mu(dx_0) \int_{B_0} \cdots \int_{B_{k-1}} f(x_0, x_1, \ldots, x_{k-1}, dx_k) dx_0 dx_1 \cdots dx_{k-1}.
\]  

(2.2.11)

More generally, if \( B = B_0 \times B_1 \times \cdots \times B_k \), where \( B_i \in X_i \), and \( f(x_0, x_1, \ldots, x_k) \) is some \( \prod_{i=0}^{k} X_i \)-measurable function, then by identical reasoning,

\[
\int_{B} f(x_0, x_1, \ldots, x_k) p_{0,1,\ldots,k}(dx_0 dx_1 \cdots dx_k) = \int_{B_0} \mu(dx_0) \int_{B_1} \cdots \int_{B_k} f(x_0, x_1, \ldots, x_{k-1}, dx_{k-1}) dx_0 dx_1 \cdots dx_{k-1}.
\]  

(2.2.12)
Let us now show that $p_{k,m}(x,A)$ is actually a version of the conditional probability $P(x_m \in A | x_k)$. First, $p_{k,m}(x,A)$ must be measurable with respect to the $\sigma$-algebra generated by $x_k, \sigma\{x_k\}$. But, $(X,X)$ is defined as a product space with the product $\sigma$-algebra, so $\sigma\{x_k\}$ is the family of sets in the product $\sigma$-algebra, $\{X_0, \phi\} \times \{X_1, \phi\} \times \ldots \times \{X_{k-1}, \phi\} \times X_k \times \{X_{k+1}, \phi\} \times \ldots$, where $\{X_i, \phi\}$ is the trivial field on $X_i$. Property (iii) has established that $p_{k,m}(x,A)$ is $X_k$-measurable for every $A \in X_m$. By trivially extending its domain, we can think of the transition probability as:

$$p_{k,m} = (\Pi X_i, X_m) \to [0,1],$$

but it is constant on $X_0 \times X_1 \times \ldots \times X_{k-1} \times X_{k+1} \times \ldots$. Therefore, $p_{k,m}(\cdot, A)$ is $\sigma\{x_k\}$-measurable.

We must also demonstrate that for each $B \in X_k$ and $A \in X_m$,

$$\int_{p_{k,m}(x,A) \in B} P_k(dx) = P(x_k \in B, x_m \in A)$$
From Equation (2.2.10),

\[
P(x_K \in B, x_m \in A) = \int_{\mathbb{R}} \mu(dx_0) \int_{x_0}^{x_1} p_{0,1}(x_0, dx_1) \cdots \int_{x_{k-1}}^{x_k} p_{k-1,k}(x_{k-1}, dx_k) \int_{x_k}^{x_{k+1}} p_{k,k+1}(x_k, dx_{k+1}) \cdots \int_{x_{m-2}}^{x_{m-1}} p_{m-2,m-1}(x_{m-2}, dx_{m-1}) \int_{x_{m-1}}^{x_m} p_{m-1,m}(x_{m-1}, dx_m)
\]

\[
= \int_{\mathbb{R}} \mu(dx_0) \int_{x_0}^{x_1} p_{0,1}(x_0, dx_1) \cdots \int_{x_{k-1}}^{x_k} p_{k-1,k}(x_{k-1}, dx_k) \int_{x_k}^{x_{k+1}} p_{k,k+1}(x_k, dx_{k+1}) \cdots \int_{x_{m-2}}^{x_{m-1}} p_{m-2,m-1}(x_{m-2}, dx_{m-1}) \int_{x_{m-1}}^{x_m} p_{m-1,m}(x_{m-1}, dx_m)
\]

\[
p_{k,m}(x_k, A) = \int_B p_{k,m}(x_k, A) \mu(dx)
\]

because of the Chapman-Kolmogorov equation.

But \( p_{k,m}(x_k, A) \) is \( x_k \)-measurable, so Equation (2.2.11) applies.

i.e.,

\[
\int_B p_{k,m}(x, A) p_k(dx) = \int_{\mathbb{R}} \mu(dx_0) \int_{x_0}^{x_1} p_{0,1}(x_0, dx_1) \cdots \int_{x_{k-1}}^{x_k} p_{k-1,k}(x_{k-1}, dx_k) \int_{x_k}^{x_{k+1}} p_{k,k+1}(x_k, dx_{k+1}) \cdots \int_{x_{m-2}}^{x_{m-1}} p_{m-2,m-1}(x_{m-2}, dx_{m-1}) \int_{x_{m-1}}^{x_m} p_{m-1,m}(x_{m-1}, dx_m)
\]

\[
\int_B p_{k-1,k}(x_{k-1}, dx_k) p_{k,m}(x_m, A) = \int_{\mathbb{R}} \mu(dx_0) \int_{x_0}^{x_1} p_{0,1}(x_0, dx_1) \cdots \int_{x_{k-1}}^{x_k} p_{k-1,k}(x_{k-1}, dx_k) \int_{x_k}^{x_{k+1}} p_{k,k+1}(x_k, dx_{k+1}) \cdots \int_{x_{m-2}}^{x_{m-1}} p_{m-2,m-1}(x_{m-2}, dx_{m-1}) \int_{x_{m-1}}^{x_m} p_{m-1,m}(x_{m-1}, dx_m)
\]
Comparing (2.2.13) and (2.2.14) yields the desired conclusion.

Therefore, we can henceforth refer to $p_{k,m}(x,A)$ as a conditional probability. In fact, it is a regular version of the conditional probability $P(x_m \in A | x_k)$ because it satisfies (i) and (ii) in the defining properties of a transition probability. [L1, pages 569, 137]. Regularity is important here because it allows conditional expectations to be expressed by integrals using conditional probabilities for measures.

Next, a Markov process in the narrow sense (*) is defined:

**Definition:** A random process $\{x_n\}$ with range in a probability space $(X,X,P)$ is called a Markov process if:

(i) $P(x_m \in A | x_0, x_1, x_2, \ldots, x_k) = P(x_m \in A | x_k) \mod(P)$

(ii) The conditional probabilities satisfy the requirements of transition probabilities enumerated above.

**Theorem:** The system (2.2.1) defines a narrow-sense Markov process with transition probability given by Equation (2.2.3).

(*) We follow Gikhman and Skorokhod [G3, page 344].
Proof: Everything has been proven except to demonstrate that $P(x_m \in A | x_k)$ is equal to $P(x_m \in A | x_0, x_1, \ldots, x_k)$. First, $P(x_m \in A | x_k)$ is $\sigma\{x_0, x_1, \ldots, x_k\}$-measurable, because $\sigma\{x_k\} \subset \sigma\{x_0, x_1, \ldots, x_k\}$ and $P(x_m \in A | x_k)$ is $\sigma\{x_k\}$-measurable. It remains only to demonstrate that for every set $B \in \prod_{i=0}^{k} X_i$ and for every $A \in X_m$,

$$\int_B P(x_m \in A | x_k)P_{0,1,\ldots,k}(dx_0 dx_1 \ldots dx_k) = P((x_0, x_1, \ldots x_k) \in B, x_m \in A) \quad (2.2.15)$$

It is sufficient to demonstrate (2.2.15) for cylinder sets of the form $B = B_0 \times B_1 \times \ldots \times B_k \times \prod_{i=k+1}^{\infty} X_i$, where $B_i \in X_i$. Working first with the right-hand side of (2.2.15), and using (2.2.10), we see:

$$P((x_0, x_1, \ldots, x_k) \in B, x_m \in A) = \int_{B_0} \int_{B_1} \ldots \int_{B_k} P_{0,1}(x_0, dx_1) \ldots \int_{x_{k-2}}^{x_{k-1}} P_{k-1,k}(x_{k-1}, dx_k) \ldots \int_{x_{m-2}}^{x_{m-1}} P_{m-2,m-1}(x_{m-2}, dx_{m-1}) \int_{x_{m-1}}^{A} P_{m-1,m}(x_{m-1}, dx_m)$$
By the Chapman-Kolmogorov equation.

The left-hand side of (2.2.15) is evaluated by Equation (2.2.12). This completes the proof.

The criteria for stability makes use of the difference operator (Lyapunov operator), \( L(\cdot) \) of system (2.2.1) defined by:

\[
L(V(k,x_k)) = \int_{X_k} V(k,y) p_{k,k+1}(x,dy) - V(k,x) \quad (2.2.16)
\]

The domain of definition of \( L, D_L \), consists of all real-values functions which are measurable with respect to the measures \( p_{k,k+1}(x,dy) \) for each \( k \).
Discrete-Time Systems: Stability Theory

Under the conditions given that guarantee that the system (2.2.1) defines a Markov process, the following types of stability will be considered. (We follow Kozin's terminology [K3]). Note that it is only the equilibrium solution, \( x_k \equiv 0 \), whose stability is being considered.

Definition:

The equilibrium solution \( x_k \equiv 0 \) of system (2.2.1) is \textbf{almost surely stable} if for \( \varepsilon > 0 \)

\[
\lim_{|x_0| \to 0} P\left\{ \sup_{k \geq k_0} |x_k| > \varepsilon \right\} = 0 \quad (2.2.17)
\]

(Konstantinov [K4] refers to this property as "probability stable".)

Definition:

The solution \( x_k \equiv 0 \) of system (2.2.1) is \textbf{exponentially stable in the pth mean} if there are positive constants \( A \) and \( \alpha \) such that:

\[
E\left\{ |x_k|^p \right\} \leq A |x_0|^p \exp(-\alpha(k - k_0)) \quad (2.2.18)
\]

\((*)\) The probability measure, \( P \), is defined by Equation (2.2.10)
If (2.2.18) holds for $p = 2$, $x_k \equiv 0$ is said to be \textit{exponentially mean square stable}.

Earlier, the transition probability for the Markov process $\{x_k\}$ and the Lyapunov operator $(L)$ for system (2.2.1) have been defined. Their use in proving stability criteria is well established in the paper by Konstantinov [K4], which contains the following two theorems:

**Theorem:**

The solution $x_k \equiv 0$ of system (2.2.1) is exponentially stable in the pth mean if and only if there exists a function $V(k,x) \in D_L$ such that for $k \geq k_0$ and for $x \neq 0$,

$$C_1 |x|^p \leq V(k,x) \leq C_2 |x|^p$$

$$L[V(k,x)] \leq -C_3 |x|^p$$

(2.2.19)

where $C_1$, $C_2$, and $C_3$ are positive constants.

**Theorem:**

The solution $x_k \equiv 0$ of system (4.1.1) is almost surely stable for $k \geq k_0$ if there exists a function $V(k,x) \in D_L$, which for $k \geq k_0$ satisfies the following conditions.
a.) $V(k, x)$ is continuous at $x = 0$, and $V(k, 0) = 0$.

b.) $\inf_{|x|>\delta} V(k, x) > a(\delta) > 0$ for any $\delta > 0$. \hspace{1cm} (2.2.20)

c.) $L[V(k, x)] \leq 0$ in some neighborhood of $x = 0$.

In (c) the neighborhood is in $\mathbb{R}^n$ under the same norm used in (2.2.19),

These results rely very heavily on the assumption that the quantities $w_k$ of system (2.2.1) are independent random variables. If system (2.2.1) is used to represent a linear system with additive white noise or with additive colored noise, methods are available for determining the transition probability functions of the Markov sequence $x_k$ and the criteria for stability under those conditions are well established. [K4, W3, W4]. If the system under consideration is a linear system with multiplicative white noise, Lyapunov functions for use in Equations (2.2.19) and (2.2.20) have been derived. However, even in the first order case, stability criteria for systems perturbed by multiplicative colored have not been obtained. The integrals (2.2.2) that must be solved to derive the transition probability functions have not been evaluated in closed form, and useful Lyapunov functions are not known.
In addition to the two theorems given above, Konstantinov [K4] has obtained results for the special cases in which the system (2.2.1) assumes either of the forms:

\[ x_{k+1} = B_k(w_k)x_k \]  

(2.2.21)

or

\[ x_{k+1} = B_k(w_k)x_k + F_k(w_k,x_k) \]  

(2.2.22)

In the first case, a suitable Lyapunov function for use in Equations (2.2.19) and (2.2.20) is provided. In the latter case, separate requirements are established for the \( B_k \) and \( F_k \) functions, provided a simplified sufficient condition for the exponential pth mean stability of (2.2.22).

In Reference [K9], Kushner touches upon the stability of discrete-parameter Markov processes. The theorem given there (Chapter 2, Theorem 12), uses a Lyapunov function similar to that of Equation (2.2.19) to establish a set of requirements which guarantee that for \( (0 < m < \infty) \),

\[ P \left\{ \sup_{0 \leq n < \infty} V(x_n) \geq m|x_o \right\} \leq V(x_o)/m \]  

(2.2.23)
The proofs of this theorem and those of Konstantinov rely upon defining the set $Q_m = \{x : V(x) < m\}$ and the stopped process defined by stopping $x_n$ upon exit from $Q_m$. Kushner's theorem is slightly stronger than that of Konstantinov, because it provides a bound for any value of $x_0$, Equation (2.2.23), and not just the bound in the limit as $x_0 \to 0$ which is required in the definition of almost sure stability.

The problem treated by Akhmetkaliev [Al] concerns systems of the type:

$$x_{k+1} - x_k = A(\eta_{k+1})x_k + Q(x_k, \eta_{k+1}) \quad (2.2.24)$$

where $\{\eta_k\}$ is a Markov chain defined on a countable state space. He is able to get some mean square stability results by an investigation that leans very heavily toward Lyapunov methods. The assumption of a countable state space; however, limits the usefulness of his theorems.

The research of Willems and Blankenship [W3] has considered a type of input-output stability for linear systems with a multiplicative (white) noise process playing the role of a feedback gain. Under the criterion
that the system with zero feedback is stable, their concept is to permit controls to be applied to the system and to determine the system's stability properties by examining the resulting output. This research is sufficiently general to allow both time varying systems and a nonstationary noise process, as well as allowing controls to be applied to the system under consideration. The given stability criteria are strong enough to imply mean square exponential stability as defined above. Even though only white noise parameters are treated, Theorem 1 of this reference provides the stability criteria to which the results of Chapter IV are compared. See Figure (4.2.1) and Figures (4.3.2) through (4.3.8). Unfortunately, there are no colored noise results available for a comparison.

A definitive study of the mean square stability properties of time invariant linear systems with white noise parameters has been included in the paper of J.L. Willems [W4]. Necessary and sufficient conditions for mean square stability are given in terms of the system matrices in case a state-space description is used, and in terms of the system's unit pulse response (or its Z-transform) if higher-order difference equations are
used. Moreover, it is proven that a positive definite solution to the discrete-time Lyapunov equation provides a quadratic form that is a useful Lyapunov function. The necessary and sufficient criteria derived here turn out to be equivalent to those given by Theorem 1 of Reference [W3] for the scalar system treated in Chapter IV.
CHAPTER III

CONTINUOUS TIME RESULTS

In this chapter, we present the criteria for stochastic stability that we have derived for continuous time systems. In the three sections of this chapter, an analysis progresses from first order systems to higher order systems. Necessary and sufficient conditions for the almost sure stability and the exponential mean square stability of the null solution of first order systems and for a second order system evolving on a solvable Lie group are derived.

In the last section, the damped harmonic oscillator problem is treated in some detail, and sufficient conditions for the mean square stability of its null solution are given. The damped harmonic oscillator problem was chosen because it has received considerable attention in the literature [11, W3] so results are available for comparison. It is an example of a system that does not evolve on a solvable Lie group.

-68-
3.1 Scalar, First Order Systems

Although first order systems represent only a small fraction of all interesting dynamical systems, they play an important role in the research described below for two reasons.

(i) The formulae can be evaluated easily and provide an example to follow in the higher dimensional cases.

(ii) The formulae found for the first order case lead to a bound for the mean square stability of higher order systems, so the detailed consideration of this case seems warranted.

Colored Noise Results

It turns out that one can solve first order linear stochastic differential equations because the solution can be expressed in terms of a functional of the random processes involved. [F1, M1]. That is, suppose \( f(t) \) is a scalar function of time. Then the equation

\[
\dot{x} = (a + f(t))x \quad x(0) = x_0 \quad (3.1.1)
\]

has for its solution
\[ x(t) = x_0 \exp\left(\int_0^t (a + f(\sigma)) d\sigma\right) \] (3.1.2)

Furthermore, if \( f(t) \) is a Gaussian colored noise process, one can calculate all of the statistical properties of the exponential function in (3.1.2).

In Equation (3.1.1) there is no need to require \((a + f(t))\) to be a real function of time. Allowing complex values means that exact solutions can be obtained for systems as complicated as a damped harmonic oscillator. Making \( f(t) \) purely imaginary leads to the Frequency Modulation model common in communication theory. However, using complex multipliers with the scalar model does not allow one to distinguish between velocity feedback noise and position feedback noise. Therefore, the following analysis cannot be interpreted as the solution to the stability problem addressed in later sections.

References [F1] and [W2] solve the stability problem for system (3.1.1) with \((a + f(t))\) purely imaginary. With little added work, this can be generalized:

Equation (3.1.1) can be rewritten to display explicitly the real and imaginary parts of \((a + f(t))\) as
\[ x(t) = R_e [a + f(t)]x(t) + j I_m [a + f(t)]x(t) \]  

\[ x(0) = x_0 \]  

As in (3.1.2), the solution obtained is

\[ x(t) = \exp \left[ (R_e [a] + j I_m [a])t + \int_0^t (R_e [f(\sigma)] + j I_m [f(\sigma)]) d\sigma \right] x_0 \]  

Moreover, the mean square of \( x(t) \) is given by:

\[ E \{ x(t) x^*(t) \} = E \{ \exp (2at + 2 \int_0^t f(\sigma) d\sigma) \} x_0^2 \]

\[ = e^{2at} E \{ \exp (2 \int_0^t f(\sigma) d\sigma) \} x_0^2 \]  

(3.1.5)
For the purpose of making comparisons with later results, conditions that guarantee exponential mean square stability are derived next. Assume now that \( f(t) \) is a real-valued zero mean, stationary Gaussian random process with autocorrelation function

\[
R_f(t_1, t_2) = \sigma_f^2 e^{-\alpha|t_1-t_2|}
\]

and define the process

\[
\eta(t) = \int_0^t f(\sigma) d\sigma
\]  

(3.1.6)

Then \( \eta(t) \) is a Gaussian random process with zero mean and autocorrelation function given by:

\[
R_\eta(t_1, t_2) = R_f(t_1, t_2) \ast h(t_1) \ast h(t_2)
\]  

(3.1.7)

where \( h(t) \) is the impulse response of an integrator. \( \eta(t) \) is not stationary, for at the very least there are initial transients in the moments of the process.

\( \eta(t) \) as defined by Equation (3.1.6) assumes the existence of a perfect integrator, but in practice there will probably be some saturation present. Examining the equations below that involve \( \eta(t) \), we see that any saturation effects can only tend to make the systems
more stable. In that sense, Equation (3.1.6) is conservative.

Performing the indicated convolutions in Equation (3.1.7), and setting $t_1 = t_2 = t$, one obtains:

$$R_{\eta}(t,t) = \frac{2}{\alpha} \sigma_f^2 t + \frac{\sigma_f^2}{\alpha^2} [2e^{-\alpha t} - 2], \ t \geq 0$$

Also, $\eta(t)$ is Gaussian, so

$$E[\exp(p\eta(t))] = e^{\frac{p^2}{2}R_{\eta}(t,t)}$$

so combining this with (3.1.5) and (3.1.6) yields

$$E\{|x(t)|^2\} = e^{2at} e^{2R_{\eta}(t,t)}$$

$$= \exp(2at + 2 \frac{2\sigma_f^2}{\alpha} t + 2 \frac{\sigma_f^2}{\alpha^2} [2e^{-\alpha t} - 2])$$

(3.1.8)
so

\[ \lim_{t \to \infty} E\{ |x(t)|^2 \} = 0 \text{ if and only if } a + \frac{2\sigma_f^2}{\alpha} < 0 \]

(3.1.9)

and if (3.1.9) is satisfied, the convergence will be exponentially fast.

Equation (3.1.9) is the desired result. It will also be required in the proofs in Sections 3.2 and 3.3.

Other Types of Stability for System (3.1.1)

Knowing Equation (3.1.2), it is straightforward to calculate criteria for the exponential stability of the pth mean of the null solution of System (3.1.1) for \( p \neq 2 \). A special case (\( p = 1 \)) is treated by Brockett [B1, page 58].

It is easy to show that

\[ x^p(t) = x^p(0) \exp \left[ p \int_0^t (a + f(\sigma)) d\sigma \right] \]  

(3.1.10)

so

\[ E\{ |x(t)|^p \} = e^{p \alpha t} E\{ e^{p \eta(t)} \} |x_0|^p \]  

(3.1.11)
where \( \eta(t) \) is as defined in Equation (3.1.6).

But,

\[
E\{e^{n(t)}\} = e^{\frac{\sigma_f^2}{\alpha} R_n(t,t)}
\]

so the null solution of System (3.1.1) is exponentially stable in the pth mean if and only if

\[
a + p \frac{\sigma_f^2}{\alpha} < 0
\]  

(3.1.12)

Figure (3.1.1) shows how the region of exponential stability of the pth mean decreases with increasing p.

These results are easily generalized to the case in which \( f(t) \) is a band-pass process, whose autocorrelation function is given by:

\[
R_f(t_1, t_2) = \sigma_f^2 \exp(-\alpha |t_1 - t_2|) \cos \omega_0(t_1 - t_2)
\]

(3.1.13)
Unstable By All Definitions

A.S. Stable, Entire Quadrant
(See Theorem 3.1.1)

Boundaries for
Stability in
pth Mean
(Stable Region Below Line)

Stability Regions for System (3.3.1)

Figure 3.1.1
Once again defining \( \eta(t) \) as

\[
\eta(t) = \int_{0}^{t} f(\sigma) d\sigma ,
\]

the autocorrelation function of \( \eta(t) \) is given by

\[
R_{\eta}(t_1, t_2) = \int_{0}^{t_1} \int_{0}^{t_2} R_{f}(\sigma_1, \sigma_2) d\sigma_1, d\sigma_2
\]

\[
= \frac{2\sigma_f^2 \alpha}{\alpha^2 + \omega_o^2} \min(t_1, t_2) + \left( \text{terms that grow less than linearly in } t_1 \text{ or } t_2 \right)
\]

Then, as in Equation (3.1.11),

\[
E[x^p(t)] = e^{\text{pat} E[e^{p\eta(t)}]} x_0^p
\]

\[
= e^{\text{pat}} e^{\frac{p^2}{2} R_{\eta}(t,t)} x_0^p
\]

so

\[
\lim_{t \to \infty} E[x^p(t)] = 0 \iff \text{pa} + \frac{p^2}{2} \left( \frac{2}{\alpha} + \frac{\sigma_f^2}{1 + \omega_o^2/\alpha^2} \right) < 0
\]

(3.1.14)
in which case the convergence will be at an exponential rate.

**Almost Sure Stability Criteria for First Order Systems**

Still considering the System (3.1.1), the region of almost sure stability is given by the following:

**Theorem (3.1.1)**

The null solution of System (3.1.1) is **almost surely stable** if $a < 0$.

**Proof**

We must now return to the notation of Chapter II and explicitly display the dependence of the colored noise $f(t, \omega)$ on $\omega$. We have assumed that the colored noise process is stationary and has been generated by passing white noise through a linear filter and is, therefore, ergodic. Then,

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t f(\tau, \omega) \, d\tau = E\{f(t, \omega)\} \quad (3.1.15)$$

with probability one, and the sample functions are almost surely continuous. Property (3.1.15) can be rewritten as follows:
Given \( \beta > 0 \), there exists a random time \( T_\beta(\omega) \), depending upon \( \omega \), such that

\[
\left| \frac{1}{t} \int_0^t f(\sigma,\omega) d\sigma - E\{f(t,\omega)\} \right| < \beta \quad \text{a.s.}
\]

for all \( t > T_\beta(\omega) \). With \( E\{f(t,\omega)\} = 0 \), we have

\[
\left| \frac{1}{t} \int_0^t f(\sigma,\omega) d\sigma \right| < \beta \quad \text{almost surely for all } t > T_\beta(\omega).
\]

\[ \Rightarrow \int_0^t f(\sigma,\omega) d\sigma < \beta t \quad \text{almost surely for all } t > T_\beta(\omega) \]

(3.1.16)

In the following equations we will make use of the time \( T_{|a/2|}(\omega) \), and simplify notation by defining

\[ T(\omega) = T_{|a/2|}(\omega) \]. We must show that

\[
\lim_{\delta \to 0} \mathbb{P}\left\{ \sup_{\|x_0\| < \delta} \sup_{t \geq 0} \|x(t,\omega)\| > \varepsilon \right\} = 0
\].
But

\[ \lim_{\delta \to 0} P \left\{ \sup_{\|x_o\| < \delta} \sup_{t \geq 0} \|x(t, \omega)\| > \epsilon \right\} \]

\[ = \lim_{\delta \to 0} P \left\{ \sup_{\|x_o\| < \delta} \sup_{t \geq 0} \left\| e^{at} \exp \int_0^t f(\sigma, \omega) d\sigma \right\| > \frac{\epsilon}{\delta} \right\} \]

\[ \leq \lim_{\delta \to 0} \left[ P \left\{ 0 < t < T(\omega) \left\| e^{at} \exp \int_0^t f(\sigma, \omega) d\sigma \right\| > \frac{\epsilon}{\delta} \right\} \right] \]

\[ + P \left\{ t > T(\omega) \left\| e^{at} \exp \int_0^t f(\sigma, \omega) d\sigma \right\| > \frac{\epsilon}{\delta} \right\} \]

\[ (3.1.17) \]
But

\[ \lim_{\delta \to 0} \mathbb{P} \left\{ \sup_{0 < t < T(\omega)} \left| e^{at} \exp \int_{0}^{t} f(\sigma, \omega) d\sigma \right| > \frac{\varepsilon}{\delta} \right\} \]

= \lim_{\delta \to 0} \mathbb{P} \left\{ \sup_{0 < t < T(\omega)} \left| \int_{0}^{t} (a + f(\sigma, \omega)) d\sigma \right| > \ln \frac{\varepsilon}{\delta} \right\}

= \lim_{n \to \infty} \mathbb{P} \left\{ \sup_{0 < t < T(\omega)} \left| \int_{0}^{t} (a + f(\sigma, \omega)) d\sigma \right| > \ln n \varepsilon \right\}

= \mathbb{P} \left[ \bigcap_{n=1}^{\infty} \left\{ \omega : \sup_{0 < t < T(\omega)} \left| \int_{0}^{t} a + f(\sigma, \omega) d\sigma \right| > \ln n \varepsilon \right\} \right]

(3.1.18)

= 0 \text{ because of Equation (2.1.38) and } a < 0.
The last term in Equation (3.1.17) can be treated by using Equation (3.1.16).

\[ \lim_{\delta \to 0} P \left\{ \sup_{t \leq T(\omega)} |e^{at} \exp \int_{-\infty}^{t} f(\sigma, \omega) d\sigma| > \frac{\epsilon}{\delta} \right\} \]

\[ \leq \lim_{\delta \to 0} P \left\{ \sup_{t > T(\omega)} |e^{at} e^{-\frac{at}{2}}| > \frac{\epsilon}{\delta} \right\} \]

\[ = \lim_{\delta \to 0} P \left\{ |e^{\frac{aT(\omega)}{\epsilon}}| > \frac{\epsilon}{\delta} \right\} = 0 \quad (3.1.19) \]

Combining Equations (3.1.17), (3.1.18) and (3.1.19) completes the proof.

Referring to Figure (3.1.1) again, it is easy to identify systems that have rather peculiar properties. That is, in the region \( 0 > a > -\sigma_f^2/a \), System (3.1.1) is unstable in the pth mean for \( p > 1 \), yet almost every sample path has been shown to approach zero after large time intervals. Moreover, for any combination of \( a \), \( \sigma_f^2 \) and \( \alpha > 0 \), System (3.1.1) will be unstable in the pth mean if \( p \) is chosen large enough. Similarly, for any given \( p \), \( \sigma_f^2 \), and \( \alpha > 0 \), System 3.1.1 will be stable in the pth mean if \( \alpha \) is chosen large enough.
This last observation has an obvious interpretation; for a given feedback noise power, concentrating that power near zero frequency provides the most destabilizing (in pth mean) influence. For further comments on the tradeoffs of studying the behavior of sample paths (almost sure stability criteria) versus pth mean stability, see Kozin [K3, K6].

3.2 Applications of Lie Algebra Theory

The key to being able to establish necessary and sufficient criteria for each of the types of stability for first order systems is our ability to solve the differential equations (3.1.1) in terms of linear functionals and memoryless transformations of a known random variable. The same property accounts for the ease of obtaining the pth mean criteria once the mean square criteria was established. In this section, we establish stability criteria for higher order systems having a similar property, i.e., those systems whose transition functions are decomposable into elements of a Lie group corresponding to a solvable Lie algebra.* Any first order system is an example of a solvable case.

* Refer to Appendix B for relevant definitions here.
In this section, we will be studying systems of the form:

\[ x(t) = [A_0 + \sum_{i=1}^{N} f_i(t) A_i]x(t) \]  \hspace{1cm} (3.2.1)

where \( A_i \) is a constant \( nxn \) matrix, \( i = 0,1,2,\ldots,N \), and \( f_i(t) \) is a colored and possibly correlated Gaussian random process, \( i = 1,2,\ldots,N \). Henceforth, we will work under the assumption that the Lie algebra, \( L \), generated by \( \{A_1, A_2, \ldots, A_N\} \) is solvable. For this case, Willsky and Marcus [W9] have demonstrated a procedure for determining criteria for stochastic stability of System (3.2.1) based upon the existence of a similarity transformation that will simultaneously transform all the \( A_i \) matrices into upper triangular form. We will take an alternate approach, based upon the paper by Wei and Norman [W8]. It turns out, however, that the calculations required in the two methods are identical, except that here we will start by defining a set of basis matrices for the Lie algebra instead of searching for a similarity transformation.

What we really require for solving Equation (3.2.1) for \( x(t) \) is to determine the transition matrix \( \Phi(t, t_o) \), satisfying
\[
\frac{d}{dt} \varphi(t, t_0) = [A_0 + \sum_{i=1}^{N} f_i(t) A_i] \varphi(t, t_0);
\]

\[
\varphi(t_0, t_0) = I \tag{3.2.2}
\]

We will follow the work of reference [W8]. The major result states that if the smallest Lie algebra containing the \( A_i \) is solvable, then there exists a basis \( \{B_1, \ldots B_\ell\} \) for \( \{A_i\}_{LA} \) and an ordering of that basis, such that

\[
\varphi(t, t_0) = \exp(g_1(t) X_1) \exp(g_2(t) X_2) \ldots \exp(g_\ell(t) X_\ell)
\]

where \( g_i(t) \) are scalar functions of time. Moreover, the \( g_i(t) \) satisfy a set of differential equations which depend only on \( \{A_i\}_{LA} \) and the functions \( f_i(t) \).

The equations used to define the \( g_i(t) \) are derived by differentiating Equation (3.2.3) with respect to time and setting the result equal to:
To simplify matters we can specialize to the case \( l = 3 \), which is sufficient to solve the example given below.

After differentiating both sides of Equation (3.3.2) and multiplying by some identity matrices, we obtain

\[
\frac{d}{dt} \Phi(t, t_0) = \frac{d}{dt} g_1(t)X_1 \Phi(t, t_0) \\
+ \exp(g_1(t)X_1) g_2'(t)X_2 \exp(-g_1(t)X_1) \Phi(t, t_0) \\
+ \exp(g_1(t)X_1) \exp(g_2(t)X_2) g_3'(t)X_3 \\
\exp(-g_2(t)X_2) \exp(-g_1(t)X_1) \Phi(t, t_0)
\]

(3.2.4)

But \( \Phi(t, t_0) \) is non-singular, so it must be that:
\[ [A_0 + \sum_{i=1}^{N} A_i f_i(t)] = \frac{d}{dt} g_1(t)X_1 \]

\[ + \exp(g_1(t)X_1) \frac{d}{dt} (g_2(t)X_2) \exp(-g_1(t)X_1) \]

\[ + \exp(g_1(t)X_1) \exp(g_2(t)X_2) \]

\[ \frac{d}{dt} (g_3(t)X_3) \exp(-g_2(t)X_2) \exp(-g_1(t)X_1) \]

(3.2.5)

The matrices \( X_i \) form a basis for \( L \overset{\triangleleft}{\triangleleft} \{A_i\}_{LA} \), so we can replace the left hand side of Equation (3.2.5) by

\[ [A_0 + \sum_{i=1}^{N} A_i f_i(t)] = \sum_{i=1}^{3} F_i(t)X_i \]  

(3.2.6)

for some set of functions, \( F_i(t) \). Finally, equating the coefficient of each \( X_i \) in Equation (3.2.6) with those in (3.2.5) yields the required differential equations that can be solved for \( g_1(t) \), \( g_2(t) \), and \( g_3(t) \).
The $g_i(t)$ functions are, therefore, random processes because their defining differential equations contain terms involving $f_j(t)$. The assumption that $L$ is solvable is sufficient to guarantee that, under the basis chosen, the $g_i(t)$ are well defined for all $t$, so that Equation (3.2.3) leads to a meaningful representation of $\Phi(t,t_0)$. $\Phi(t,t_0)$ is given as a non-linear function of the random processes $g_i(t)$ and the $g_i(t)$ are defined by non-linear differential equations so it is not clear at the outset that Equation (3.2.3) will be useful. However, as the following example demonstrates, this approach can be fruitful for establishing criteria for both mean square stability and almost sure stability of the System (3.2.1).

Example:

Consider the following dynamical system:

$$
\begin{bmatrix}
\dot{x}_1(t) \\
\dot{x}_2(t)
\end{bmatrix} = \begin{bmatrix}
\lambda_1 + \lambda_2 & \lambda_1 \\
-\lambda_2 & 0
\end{bmatrix} + f(t) \begin{bmatrix}
1 & 1 \\
0 & 0
\end{bmatrix} \begin{bmatrix}
x_1(t) \\
x_2(t)
\end{bmatrix}
$$

(3.2.6)

$$
x_1(0) = x_{10}, \quad x_2(0) = x_{20}
$$
where it is assumed that $\lambda_1, \lambda_2 < 0$, and $f(t)$ is described prior to Equation (3.1.6).

The set of matrices

\[
L = \left\{ X_1 = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}, X_2 = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}, X_3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \right\}
\]

(3.2.7)

form a Lie algebra. $\Phi(t, t_0)$, the transition matrix for system (3.2.7) will propagate on the associated Lie group, $\{\exp(L)\}_G$. To see that $L$ is solvable, we can form the table of commutator products for $L$:

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>0</td>
<td>$X_2$</td>
<td>0</td>
</tr>
<tr>
<td>$X_2$</td>
<td>$-X_2$</td>
<td>0</td>
<td>$X_2$</td>
</tr>
<tr>
<td>$X_3$</td>
<td>0</td>
<td>$-X_2$</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE OF PRODUCTS $[X_i, X_i]$

TABLE 3.2.1
From the table it is clear that \( L^{(2)} = 0 \), so \( L \) is solvable. Table (3.2.1) will also be useful in evaluating the matrix exponentials of Equation (3.2.5).

In terms of the basis matrices of \( L \), the matrix in Equation (3.2.7) is given by:

\[
\begin{bmatrix}
\lambda_1 + \lambda_2 & 1 \\
-\lambda_2 & 0
\end{bmatrix} + f(t)\begin{bmatrix}
1 & 1 \\
0 & 0
\end{bmatrix} = \lambda_2 X_1 - \lambda_2 X_2 + (f(t) + \lambda_1) X_3
\]

(3.2.8)

The other formulae needed in Equation (3.2.5) are given by the Baker-Hausdorff Lemma (Equation (B4)).

\[
\exp(g_1(t)X_1)X_2 \exp(-g_1(t)X_1)
\]

\[
= X_2 + g_1(t)[X_1, X_2] + \frac{g_1^2(t)}{2!} [X_1, [X_1, X_2]] + \ldots
\]

\[
= X_2 + g_1(t)X_2 + \frac{g_1^2(t)}{2!} X_2 + \ldots
\]

\[
= X_2 \exp(g_1(t))
\]

(3.2.9)
Similarly,

\[
\exp(g_1(t)X_1)\exp(g_2(t)X_2)X_3 \exp(-g_2(t)X_2)\exp(-g_1(t)X_1) = \exp(g_1(t)X_1) [X_3 + g_2(t)X_2] \exp(-g_1(t)X_1)
\]

\[
= X_3 + g_2(t)X_2 \exp(g_1(t)) \tag{3.2.10}
\]

Therefore, Equation (3.2.5) becomes*

\[
\lambda_2 X_1 - \lambda_2 X_2 + (f(t) + \lambda_1)X_3 = g_1'(t)X_1
\]

\[
+ g_2'(t)X_2 \exp(g_1(t))\]

\[
+ g_3'(t) [X_3 + g_2(t)X_2 \exp(g_1(t))] \tag{3.2.11}
\]

* (') denotes time derivative.
Equating the coefficients of each $X_i$ yields the required equations for $g_1(t)$ (with $g_1(0) = 0$).

\begin{align*}
g_1'(t) &= \lambda_2 \quad (3.2.12a) \\
g_2'(t) &= -\exp(-g_1(t))\lambda_2 - g_3'(t)g_2(t) \quad (3.2.12b) \\
g_3'(t) &= \lambda_1 + f(t) \quad (3.2.12c)
\end{align*}

These equations can be solved sequentially to yield:

\begin{align*}
g_1(t) &= \lambda_2 t \\
g_2(t) &= \int_0^t \exp\left(\int_\tau^t (\lambda_1 + f(\sigma))d\sigma\right)\lambda_2 e^{-\lambda_2 \tau}d\tau \\
g_3(t) &= \int_0^t (f(\sigma) + \lambda_1)d\sigma
\end{align*}

(3.2.13)

The matrix exponentials involving the basis matrices are given by:

\[
\exp(g_1(t)X_1) = \begin{bmatrix} 1 & 1 - e^{-g_1(t)} \\ 0 & e^{-g_1(t)} \end{bmatrix} \quad (3.2.14a)
\]
\[
\exp(g_2(t)X_2) = \begin{bmatrix}
1 - g_2(t) & -g_2(t) \\
g_2(t) & 1 + g_2(t)
\end{bmatrix}
\] (3.2.14b)
\[
\exp(g_3(t)X_3) = \begin{bmatrix}
eg_3(t) & e_3(t) - 1 \\
0 & 1
\end{bmatrix}
\] (3.2.14c)

It is straightforward to verify that \( \Phi(t, t_0) \) in Equation (3.2.3) is given by:
\[
\Phi(t, t_0) = \begin{bmatrix}
-g_2 e^{g_1 + g_3 + e_{g_3}} & e^{g_3} - g_2 e^{g_1 + g_3 + e_{g_3}} & e \n-g_2 e^{g_1 + g_3} & g_2 e^{g_1 + g_3 + e_{g_3}}
\end{bmatrix}
\]
(3.2.15)

Therefore,
\[
x_1(t) = [e^{g_3(t)} - g_2(t)e^{g_1(t) + g_3(t)}] [x_{10} + x_{20}] - e^{g_1(t)} x_{20}
\]
\[
x_2(t) = [g_2(t)e^{g_1(t) + g_3(t)}] [x_{10} + x_{20}] + e^{g_1(t)} x_{20}
\]
(3.2.16)
or

\[ x_2(t) = -\lambda_2^2 e^{\lambda_2 t} \int_0^t [\exp(\lambda_1 \tau - \lambda_2 \tau) \exp \int_0^\tau f(\sigma) d\sigma] d\tau \]

\[ \cdot (x_{10} + x_{20}) - e^{\lambda_2 t} x_{20} \quad (3.2.17) \]

\[ x_1(t) = \exp \int_0^t (f(\sigma) + \lambda_1) d\sigma \left( x_{10} + x_{20} \right) + x_2(t) \]

\[ (3.2.18) \]

From this point we are able to derive criteria for both the mean square stability and the almost sure stability of the null solution of System (3.2.7).

To demonstrate that \( E[x_2^2(t)] \) approaches zero exponentially fast, it is sufficient to demonstrate that the square of each of the terms of the right hand side of Equations (3.2.17) and (3.2.18) goes to zero exponentially fast.
The last term in Equation (3.2.17) goes to zero exponentially fast, obviously. The mean square value of \[ \exp \int_0^t (f(\sigma) + \lambda_1) \, d\sigma \] approaches zero at an exponential rate if and only if:

\[
\lambda_1 + \frac{2\sigma_f^2}{\alpha} < 0 \quad (3.2.19)
\]

This is precisely the result derived in Section (3.1). (Specifically, Equation (3.1.9)). The difficult aspect of the problem is to demonstrate an exponentially decaying bound for the term

\[
E [-\lambda_2 \int_0^t \exp(\lambda_1 \tau - \lambda_2 \tau) \exp \int_0^\tau f(\sigma) \, d\sigma \, d\tau]^2
\]

Equation (3.1.7) is a formula for the autocorrelation function of the random process

\[
\exp(\eta(t)) = \exp \int_0^t f(\sigma) \, d\sigma
\]

That implies that the autocorrelation function of the process
\[ \rho(t) = \exp(\lambda_1 t - \lambda_2 t) \exp \int_0^t f(\sigma) d\sigma \]

is given by

\[ R_{\rho}(t_1, t_2) = \exp[(\lambda_1 - \lambda_2)(t_1 + t_2) + R_{\eta}(t_1, t_2) + 1/2 R_{\eta}(t_1, t_1) + 1/2 R_{\eta}(t_2, t_2)] \]

Now, defining \( \psi(t) = \int_0^t \rho(\sigma) d\sigma \), we obtain

\[ R_{\psi}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} R_{\rho}(\sigma_1, \sigma_2) d\sigma_2 d\sigma_1 \]

\[ = \int_0^{t_1} \int_0^{t_2} \exp[(\lambda_1 - \lambda_2 + \frac{\sigma_f^2}{\alpha})(\sigma_1 + \sigma_2) + \frac{2\sigma_f^2}{\alpha} \min(\sigma_1, \sigma_2)] \]

\[ \times \exp[\frac{\sigma_f^2}{\alpha^2}(2e^{-\alpha\sigma_1} + 2e^{-\alpha\sigma_2} - e^{-\alpha|\sigma_1 - \sigma_2|} - 3)] d\sigma_2 d\sigma_1 \]

(3.2.20)
But,

\[-4 \leq 2e^{-\alpha \sigma_1} + 2e^{-\alpha \sigma_2} - e^{-\alpha |\sigma_1 - \sigma_2|} - 3 \leq 0,\]

for \(\sigma_1, \sigma_2 > 0\)

Therefore,

\[
\exp\left(-\frac{4\sigma_f^2}{\alpha^2}\right) \int_0^{t_1} \int_0^{t_2} \exp\left[(\lambda_1 - \lambda_2 + \frac{\sigma_f^2}{\alpha})(\sigma_1 + \sigma_2)\right]
\]

\[+ \frac{2\sigma_f^2}{\alpha} \min(\sigma_1, \sigma_2)]d\sigma_2 d\sigma_1\]

\[\leq R_\psi(t_1, t_2)\]

\[\leq \int_0^{t_1} \int_0^{t_2} \exp\left[(\lambda_1 - \lambda_2 + \frac{\sigma_f^2}{\alpha})(\sigma_1 + \sigma_2) + \frac{2\sigma_f^2}{\alpha} \min(\sigma_1, \sigma_2)\right]d\sigma_2 d\sigma_1\]

\[(3.2.21)\]

After some lengthy calculations and setting \(t_1 = t_2 = t\),

the integrals in Equation (3.2.21) are found to have

the form
where $K_1$ and $K_2$ are positive constants. Therefore, for large $t$, the mean square behavior of the random process

$$
\lambda_2 e^{\frac{\lambda_2}{2}t} \int_0^t [\exp(\lambda_1 - \lambda_2) t \exp \int_0^t f(\sigma) d\sigma] dt = \lambda_2 e^{\frac{\lambda_2}{2}t} \psi(t)
$$

is determined by the algebraic signs of the two coefficients $(\lambda_1 + \lambda_2 + \frac{\sigma_f^2}{\alpha})$ and $(2\lambda_1 + 4 \frac{\sigma_f^2}{\alpha})$ that are obtained by multiplying both sides of Equation (3.2.22) by $e^{\frac{\lambda_2}{2}t}$.

We have previously assumed that $\lambda_1$ and $\lambda_2$ are less than zero. By Equation (3.2.19),

$$\lambda_1 + \frac{2\sigma_f^2}{\alpha} < 0,$$

so $\lambda_1 + \lambda_2 + \frac{\sigma_f^2}{\alpha} < 0$ also.

Therefore, we have shown that the System (3.2.7) is exponentially mean square stable if and only if
Criteria for Almost Sure Stability

Theorem

The null solution of System (3.2.7) is almost surely stable if \( \lambda_1 < 0 \) and \( \lambda_2 < 0 \).

Proof

Again it is necessary to explicitly display any dependence upon \( \omega \). From Equations (3.2.17) and (3.2.18) we have

\[
||x(t,\omega)||^2 = x_1^2(t,\omega) + x_2^2(t,\omega)
\]

\[
\leq 2\left[\exp \int_0^t (f(\sigma,\omega) + \lambda_1) d\sigma\right] \left[x_{10} + x_{20}\right]^2 + 4\left[\exp \int_0^t f(\sigma,\omega) d\sigma\right] \left[x_{10} + x_{20}\right]^2
\]

\[
+ 4[-\lambda_2 \exp^t \int_0^\tau \left[\exp(\lambda_1 - \lambda_2) \exp \int_0^\tau f(\sigma,\omega) d\sigma\right] d\tau] \left[x_{10} + x_{20}\right]^2
\]

\[
- \left[x_{10} + x_{20}\right]^2 \quad \text{(3.2.24)}
\]

\[
(3.2.23)
\]
Clearly

\[ [x_{10} + x_{20}]^2 < 2 |x_o|^2 \]

and

\[ x_{20}^2 < |x_o|^2 \]

so

\[ |x(t, \omega)|^2 \leq 4 |x_o|^2 \left[ \exp \int_{0}^{t} f(\sigma, \omega) + \lambda_1 d\sigma \right]^2 \]

\[ + 4 \left[ e^{\lambda_2 t} \right]^2 |x_o|^2 + 8 |x_o|^2 \left[ \lambda_2 e^{\lambda_2 t} \right]^2 \]

\[ \cdot \int_{0}^{t} \left[ \exp(\lambda_1 - \lambda_2) \tau \exp \int_{0}^{\tau} f(\sigma, \omega) d\sigma d\tau \right]^2 \quad (3.2.25) \]
Therefore,

\[ \lim_{\delta \to 0} P\left\{ \left| x_0 \right| < \delta \sup_{t \geq 0} \left| x(t, \omega) \right| > \varepsilon \right\} \]

\[ = \lim_{\delta \to 0} P\left\{ \left| x_0 \right| < \delta \sup_{t \geq 0} \left| x(t, \omega) \right|^2 > \varepsilon^2 \right\} \]

\[ \leq \lim_{\delta \to 0} P\left\{ \left| x_0 \right| < \delta \sup_{t \geq 0} 4 \left| x_0 \right|^2 \left[ \exp \int_0^t f(\sigma, \omega) + \lambda_1 d\sigma \right]^2 \right\} \]

\[ + \sup_{t \geq 0} 4 \left| x_0 \right|^2 \left[ \exp \lambda_2 t \right]^2 \]

\[ + \sup_{t \geq 0} 8 \left| x_0 \right|^2 \left[ \lambda_2 e^{\lambda_2 t} \right]^2 \left[ \exp (\lambda_1 - \lambda_2) \tau \right] \exp \int_0^\tau f(\sigma, \omega) d\sigma d\tau \right]^2 > \varepsilon^2 \}

\[ \leq \lim_{\delta \to 0} P\left\{ \left| x_0 \right| < \delta \sup_{t \geq 0} 4 \left| x_0 \right|^2 \left| \exp \int_0^t f(\sigma, \omega) + \lambda_1 d\sigma \right|^2 \right\} \]

\[ > \frac{\varepsilon^2}{3} \} \]
\[
\lim_{\delta \to 0} P \left\{ \sup_{|x_o| < \delta} \sup_{t \geq 0} 4|\lambda|^{2} [\exp \left( \lambda^{2} t \right)]^{2} > \frac{\varepsilon^{2}}{3} \right\} \\
\lim_{\delta \to 0} P \left\{ \sup_{|x_o| < \delta} \sup_{t \geq 0} 8|\lambda|^{2} \lambda^2 t^{2} \int_{0}^{t} \exp(\lambda_1 - \lambda_2) \tau \exp \int_{0}^{\tau} f(\sigma, \omega) d\sigma d\tau \right\}^{2} > \frac{\varepsilon^{2}}{3} \right\}
\]

(3.2.26)

We will show that each term in last sum in Equation (2.3.26) is zero, which will justify interchanging the order of addition and taking the limit as \( \delta \to 0 \), and simultaneously complete the proof:

The first of the terms is treated in Theorem (3.1.1) of Section 3.1. It is equal to zero. For the second term, there is really nothing to prove because

\[
\lim_{\delta \to 0} P \left\{ \delta > \frac{\varepsilon^{2}}{12} \right\} = 0
\]

The last term requires all the work.
\[
\lim_{\delta \to 0} P \left\{ \left| \sup_{x_0} \sup_{t > 0} \left| \lambda_2 t \int_0^t \left[ \exp(\lambda_1 - \lambda_2) \tau \right] \right| \right| > \frac{\epsilon}{3} \right\}
\]

\[
\exp \int_0^\tau f(\sigma, \omega) d\sigma d\tau \left| \left| \left| x_0 \right| \right| ^2 > \frac{\epsilon^2}{3} \right\}
\]

\[
\leq \lim_{\delta \to 0} \left\{ \sup_{t > 0} \left| \lambda_2 e^{\lambda_2 t} \int_0^t \left[ \exp(\lambda_1 - \lambda_2) \right] \right| \right\}
\]

\[
\exp \int_0^\tau f(\sigma, \omega) d\sigma d\tau \left| \left| \left| x_0 \right| \right| ^2 > \frac{\epsilon}{5\delta} \right\}
\]

\[
\leq \lim_{\delta \to 0} P \left\{ \sup_{0 < t < T(\omega)} \left| \lambda_2 e^{\lambda_2 t} \int_0^t \left[ \exp(\lambda_1 - \lambda_2) \right] \right| \right\} (\ast)
\]

\[
\exp \int_0^\tau f(\sigma, \omega) d\sigma d\tau \left| \left| \left| x_0 \right| \right| ^2 > \frac{\epsilon}{5\delta} \right\} (\ast)
\]

\[
+ \lim_{\delta \to 0} P \left\{ \sup_{t > T(\omega)} \left| \lambda_2 e^{\lambda_2 t} \int_0^{T(\omega)} \left[ \exp(\lambda_1 - \lambda_2) \right] \right| \right\}
\]

\[
\exp \int_0^\tau f(\sigma, \omega) d\sigma d\tau \left| \left| \left| x_0 \right| \right| ^2 > \frac{\epsilon}{10\delta} \right\}
\]

\(\ast\) T(\omega) is taken to mean \(T_\lambda(\omega)\) as defined by Equation (3.1.6)
Again we will take the last three probabilities and show that each is zero in the limit.

\[
\lim_{\delta \to 0} \mathbb{P} \left\{ \sup_{0 \leq t \leq T(\omega)} | | \lambda_2 e^{\lambda_2 t} \int_0^t \exp(\lambda_1 - \lambda_2) \tau \right. \\
\exp \left. \int_0^T f(\sigma, \omega) d\sigma \right| | > \frac{\epsilon}{10} \right\}
\]

(3.2.27)
\[
\lim_{\delta \to 0} P \left\{ \left| \lambda_2 \right| \int_0^\tau \exp \int_0^{T(\omega)} (f(\sigma, \omega) + \lambda_1 d\sigma) d\tau > \frac{\epsilon}{5\delta} \right\} \\
\leq \lim_{\delta \to 0} P \left\{ \left| \lambda_2 \right| |T(\omega)| \exp \int_0^\tau |f(\sigma, \omega) + \lambda_1| d\sigma > \frac{\epsilon}{5\delta} \right\} \\
\to 0 \text{ because of Equation (3.1.18)}
\]

The second term of Equation (3.2.27) is treated identically and is zero in the limit. The last term in Equation (3.2.27) is evaluated with the help of Equation (3.1.6), which states

\[
\int_0^\tau f(\sigma, \omega) d\sigma < -\frac{\lambda_1}{2} \tau \text{ almost surely for all } \tau > T(\omega).
\]

Therefore,

\[
\lim_{\delta \to 0} P \left\{ \sup_{t \geq T(\omega)} \left| \lambda_2 e^{\lambda_2 t} \int_{T(\omega)}^T \exp(\lambda_1 - \lambda_2) \tau \right| \exp \int_0^\tau f(\sigma, \omega) d\sigma d\tau \right\} > \frac{\epsilon}{10\delta}
\]
This completes the proof of the theorem.

As in the scalar case, we are able to identify values of \( \lambda \), for which the null solution of System (3.2.7) is almost surely stable but the mean square value of \( x(t) \) diverges. The parameters \( \lambda_1 \) and \( \lambda_2 \) in System (3.2.7) are the eigenvalues of the matrix \( A_0 \) of Equation (3.2.1). This observation provides some additional insight into the results for this system. We know of no other examples in the literature for which criteria for almost sure stability in the colored noise case have been determined.
Unfortunately, multidimensional systems whose transition functions propagate on a solvable Lie algebra represent only a small fraction of linear dynamical systems. In the next section, we consider a system that is not in the solvable category.
3.3 Extensions to the Approach of I.L. Rabotnikov [R1]

In Reference [R1], Rabotnikov derives a criterion for the mean square stability of a single-input single-output linear dynamical equation containing a white noise parameter. The strength of his method lies in the fact that the white noise hypothesis is not used until the end of his proof, and then only to solve some integrals involving the noise autocorrelation function. Without the white noise assumption, the series of integrals obtained is too hard to evaluate. However, the example studied in Section (3.1) provides a useful bound for this series of integrals and leads to a simple stability criterion for the example treated in this section.

In Section (3.1), we derived a criterion for the exponential mean square stability of the system

\[ \dot{x}(t) = (a + f(t))x(t), \quad x(0) = x_0 \]  \hspace{1cm} (3.3.1)

where \( f(t) \) is described prior to Equation (3.1.6). Suppose now that we try to derive that same criterion using the formulation of Rabotnikov.
Let $Z(t)$ be the solution of the deterministic system

$$Z(t) = -a Z(t) \quad (3.3.2)$$

$$Z(0) = x_0$$

Then, the solution to Equation (3.3.1) can be given in terms of the solution to Equation (3.3.2) by treating the term $f(t)x(t)$ as though it were a feedback control. Equation (3.3.1) can be rewritten as

$$x(t) = Z(t) + \int_0^t w(t - \sigma)f(\sigma)x(\sigma)d\sigma \quad (3.3.3)$$

where $w(t) = e^{-at}$, the impulse response of System (3.3.2). Repeatedly substituting Equation (3.3.3) back into itself yields

$$x(t) = Z(t) + \int_0^t e^{-a(t-\sigma)}f(\sigma)Z(\sigma)d\sigma$$

$$+ \int_0^t e^{-a(t-\sigma)}f(\sigma) \int_0^\sigma e^{-a(\sigma-q)}f(q)x(q)dqd\sigma$$

$$= Z(t) + \int_0^t e^{-a(t-q)}f(q)Z(q)dq$$
The convergence of this series is implied by the analysis in Section 2.1.

Next, by multiplying the series in Equation (3.3.4) by itself, it is not difficult to see that $E\{x^2(t)\}$ can be expressed as:
\[ E\{x^2(t)\} = Z^2(t) + \text{terms involving } E\{f(t)\} \]

\[ + E \int \int e^{-a(t-q)}e^{-a(t-p)}f(q)f(p)Z(q)Z(p)dqdp \]

\[ + 2Z(t)E \int \int e^{-a(t-q_1)}e^{-a(q_1-q_2)}f(q_1)f(q_2)Z(q_2)dq_2dq \]

\[ + \text{terms involving third moments of } f(t) \]

\[ + 2Z(t)E \int dq_1 \int dq_2 \int dq_3 \int dq_4 e^{-a(q_2+q_3+q_4)}f(q_1)f(q_2)f(q_3)f(q_4)Z(q_4) \]

\[ + 2E \int dq_1 \int dp_1 \int dq_2 \int dq_3 e^{-a(q_1+q_2+q_3)}f(p_1)f(q_1)f(q_2)Z(q_3)Z(p_1) \]

\[ + E \int dq_1 \int dq_2 \int dp_1 \int dp_2 e^{-a(p_1+p_2)}f(p_1)f(p_2)f(q_1)f(q_2)Z(q_2)Z(p_2) \]

\[ + \ldots \quad (3.3.5) \]
Equations (3.3.5) and (3.1.8) both give expressions for $E\{x^2(t)\}$, so the right hand side of Equation (3.1.8) is an evaluation of the series given in (3.3.5). This is important because the series expansion (similar to Equation (3.3.5)) for the second order system considered next can be bounded termwise by the right hand side of Equation (3.3.5). Therefore, Equation (3.1.9) can be employed as a stability criterion for the second order system, but now only as a sufficient condition.

For the remainder of this section, we will be considering the system

\[
\begin{bmatrix}
\dot{x}_1(t) \\
\dot{x}_2(t)
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
-1 & -2\zeta
\end{bmatrix} \begin{bmatrix}
x_1(t) \\
x_2(t)
\end{bmatrix} + \begin{bmatrix}
0 \\
g(t)x_1(t)
\end{bmatrix} \tag{3.3.6}
\]

\[x(0) = x_0\]

\[y(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} x(t) \tag{3.3.7}\]

Also we will assume that we are interested only in the stability of $x_1(t)$. From System (3.3.6) we will define the matrices
\[
A = \begin{bmatrix} 0 & 1 \\ -1 & -2\varepsilon \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad c = [1, 0] \tag{3.3.8}
\]

It is also convenient to define

\[r(t) = e^{At} x_0 \tag{3.3.9}\]

and let \(g(t)\) be a stationary colored noise process, Gaussian with zero mean, variance \(\sigma_g^2\) and autocorrelation function

\[R_g(\tau) = \sigma_g^2 e^{-\alpha|\tau|} \tag{3.3.10}\]

Combining Equations (3.3.6) through (3.3.9), \(x_1(t)\) is given by the integral equation

\[
x_1(t) = cr(t) + c \int_{0}^{t} e^{A(t-\sigma)} b g(\sigma)x_1(\sigma) d\sigma \tag{3.3.11}
\]

\[
= r_1(t) + \int_{0}^{\tau} w_\zeta(t-\sigma) g(\sigma)x_1(\sigma) d\sigma \tag{3.3.12}
\]
where

\[ w_\zeta(t) = ce^{Atb} = \frac{e^{-\xi t}}{2\sqrt{\xi^2 - 1}} (e^{\sqrt{\xi^2 - 1} t} - e^{-\sqrt{\xi^2 - 1} t}) \]

(3.3.13)

Substituting Equation (3.3.12) into itself yields

\[ x_1(t) = r_1(t) + \int_0^t w_\zeta(t-\sigma)g(\sigma)x_1(\sigma)d\sigma \]

\[ = r_1(t) + \int_0^t w_\zeta(t-\sigma)g(\sigma)r_1(\sigma)d\sigma \]

\[ + \int_0^t \int_0^{q_1} w_\zeta(t-q_1)w_\zeta(q_1-q_2)g(q_1)g(q_2)r_1(q_2)dq_2dq_1 \]

\[ + ... \]

(3.3.14)
So

\[ E\{x_1(t)\} = r_1^2(t) + \text{terms involving first moments of } g \]

\[ + E \int_0^t \int_0^t w_\zeta(t-q)w_\zeta(t-p) g(q)g(p)r_1(q)r_1(p) dq dp \]

\[ + 2r_1(t) E \int_0^t \int_0^{q_1} w_\zeta(t-q_1)w_\zeta(q_1-q_2) g(q_1)g(q_2) \]

\[ r_1(q_2) dq_2 dq_1 \]

\[ + \text{Higher order terms as in Equation (3.3.5)} \]

(3.3.15)

Turning our attention back to Equation (3.3.5), and bringing the expected value operator inside the integrals in that equation, we see that every factor of every term is positive. This allows us to prove the following theorem.
Theorem 3.3.1

Consider System (3.3.6). Suppose there exist two functions of $\zeta$, say $\beta(\zeta)$ and $\gamma(\zeta)$ and a constant $\delta > 0$, such that

$$r_1(t) \leq \delta e^{-\beta(\zeta)t} \quad t \in [0, \infty] \quad (3.3.16)$$

and

$$w_{\zeta}(t) \leq \gamma(\zeta)\exp(-\beta(\zeta)t) \quad t \in [0, \infty] \quad (3.3.17)$$

Then,

$$\lim_{t \to \infty} E\{x_1^2(t)\} = 0 \quad \text{if} \quad \frac{2\sigma^2}{\alpha} < \frac{\beta(\zeta)}{\gamma^2(\zeta)} \quad (3.3.18)$$

and the convergence will be at an exponential rate, with time constant $-\beta(\zeta) + (2/\alpha)\sigma^2 \gamma^2(\zeta)$.

Proof

Using Equations (3.1.16) and (3.1.17) in (3.1.15) yields
\[ E\{x_1^2(t)\} \leq \delta^2 z_2(t) + \text{terms involving first moments of } g \]

\[ + \delta^2 E \int_0^t \int_0^t \gamma^2(\xi)e^{-\beta(\xi)(t-q)}e^{-\beta(\xi)(t-p)} \]

\[ g(q)g(p)Z(q)Z(p)dqdp \]

\[ + \delta^2 Z(t)E \int_0^t \int_0^{q_1} \gamma^2(\xi)e^{-\beta(\xi)(t-q_1)}e^{-\beta(\xi)(q_1-q_2)} \]

\[ g(q_1)g(q_2)Z(q_2)dq_2dq_1 \]

\[ + \text{Higher order terms as in Equation (3.3.5)} \]

\[ (3.3.19) \]

The factor, \( \delta^2 \) does not influence the result, because the limit of the right hand side will be shown to be equal to zero.
Let us now assume the following relations between the parameters in Equation (3.3.5) and those in Equation (3.3.19). Let

\[ f(t) = \gamma(\xi)g(t) \quad \sigma_f^2 = \gamma^2(\lambda)\delta_g^2 \]  

(3.3.20)

so that the right hand sides of Equations (3.3.19) and (3.3.5) are the same except for the factor \( \delta^2 \).

Therefore, Equation (3.1.9), which says

\[ \lim_{t \to \infty} E\{x_2(t)^2\} = 0 \text{ if } -a + \frac{2\sigma_f^2}{\alpha} < 0, \]

must imply that

\[ \lim_{t \to \infty} E\{x_1(t)^2\} = 0 \text{ if } -\beta(\xi) + \frac{2}{\alpha} \sigma_g^2 \gamma^2(\xi) < 0 \]  

(3.3.21)

This completes the proof.

We now direct our attention to making judicious choices for \( \beta(\xi) \) and \( \delta(\xi) \).
Evaluating $\beta(\zeta)$ and $\gamma^2(\zeta)$

If $\zeta^2 < 1$, one can use

$$\beta(\zeta) = \zeta$$

$$\gamma(\zeta) = (1-\zeta^2)^{-1/2}$$

(3.3.22)

If $\zeta^2 > 1$, use

$$\beta(\zeta) = \zeta - \zeta^2 - 1$$

$$\gamma(\zeta) = \frac{1}{2} (\zeta^2 - 1)^{-1/2}$$

(3.3.23)

The stability region defined by using Equations (3.3.22) and (3.3.23) in Equation (3.3.21) is plotted in Figure 3.3.1 for various values of $\alpha$, the proven region of mean square stability being represented by the area below the curves. The behavior of the graphs about the point $\zeta^2 = 1$ is misleading because the dip in the curves is caused by a poor choice of $\beta(1)$ and $\gamma(1)$, not by any properties of the system (3.3.36). We will now remove this trouble by optimizing the strength of Theorem 3.3.1.

Optimizing the Choice of $\beta$ and $\gamma$

There are no conditions that $\beta(\zeta)$ and $\gamma(\zeta)$ must satisfy as functions of $\zeta$. Therefore, we are free to choose $\beta(\zeta)$ and $\gamma(\zeta)$ independently for each value of $\zeta$. The following constrained optimization problem can,
Stability Regions for System (3.3.6)

Figure 3.3.1
therefore, be solved in order to optimize the strength of Theorem 3.1.1.

\[
\text{Maximize } \frac{\beta}{\gamma^2} \text{ subject to } w_\zeta(t) \leq \gamma \exp(-\beta t) \text{ for all } t \in [0, \infty). \quad (3.3.24)
\]

The values of $\beta$ and $\delta$ at which the maximum is achieved for each value of $\zeta$ form the functions $\beta(\zeta)$ and $\delta(\zeta)$. Using these values of $\beta(\zeta)$ and $\delta(\zeta)$ leads to the strongest version of Theorem 3.3.1.

The maximization procedure has been accomplished with a small computer program (listed in Appendix C) which produced the graphs given on Figure 3.3.2. The improvement over the graphs given in Figure 3.3.1 is most dramatic for \( 0.5 < \zeta^2 < 2 \), which is a region that occurs often in practice.

In the literature [Cl, G1, K3] there have been a series of stability boundaries derived for System (3.3.6) that apply to the colored noise case, but none of which depend upon the bandwidth of the noise. The boundary derived by Infante [Il] is the best of those reported, and is superimposed onto Figure 3.3.2. For very low $\zeta$, the curves given by Equation (3.3.21)
Stability Regions for System (3.3.6) Optimized by Equation (3.3.24)
Figure 3.3.2
are parabolic and, hence, greater than the Infante curve for any value of $\alpha$. However, the Infante boundary is comparable or better than Theorem 3.3.1 for values of $\zeta$ up to about 8. For wider bandwidths, the new boundary is superior, except for $\zeta$ corresponding to greatly overdamped systems.

In the literature [K6], there have appeared plots of the mean square stability criteria for System (3.3.6) superimposed on the criteria for the corresponding white noise parameter system (Itô sense). This practice is misleading because, in the Infante case, the boundary is given for $E\{g^2(t)\}$. In the Itô case, (see Appendix A) the bound is given in terms of $\sigma_g^2$, where $g(t)$ is white noise with autocorrelation function

$$R_g(\tau) = \sigma_g^2 \delta(\tau)$$

In the following paragraphs the result of Theorem 3.3.1 is recalculated in such a way that the comparison to the Itô result is justified.
The power spectral density of $g(t)$ corresponding to Equation (3.3.10) is given by

$$S_g(\omega) = \frac{2/\alpha \sigma_g^2}{(1 + \frac{\omega^2}{\alpha^2})}$$

Suppose now that we change the vertical axis on Figure 3.3.2 so that the stability regions are plotted in terms of $S_g(0)$. If we do this, all the graphs scale into the same graph, which is independent of $\alpha$. This procedure leads to Figure 3.3.3. The data presented in this way indicates how the region of stability changes as $\alpha \to \infty$. It does not change! It is known that there are no Ito correction terms required for System (3.3.6), so the boundary in Figure 3.3.3 is also a boundary in the limit as $\alpha \to \infty$, i.e., the white noise case. Superimposed on Figure 3.3.3 is the necessary and sufficient condition for the mean square stability of System (3.3.6) for the white noise case [W3]. We see that as $\xi^2$ becomes large, the boundary of Equation (3.3.21) differs from the Ito result by just a factor of 2. It is reasonable to expect the two answers to differ by such an amount because the inequality
Data from Figure 3.3.2 Replotted for Comparison to White Noise Case

Figure 3.3.3
(3.3.17) is really quite conservative. We remark in passing that the result of Infante can not be superimposed onto Figure 3.3.3 because his result would appear as a straight line whose slope is proportional to $\alpha^{-1}$. So as $\alpha \to \infty$, Infante's boundary would approach the horizontal axis.

Clearly, the idea of using the results of Section 3.1 for bounding the mean square response of higher order systems applies to more than just second order system. It works for any single-input, single-output system whose impulse response can be bounded by a decaying exponential, and whose noisy parameter can be treated as a feedback gain.
The type of systems to be considered in this chapter are those that evolve in a continuous state space in discrete-time steps. In addition, after properly augmenting the state, all the systems treated have the Markov property (see Section (2.2)). As in the continuous time case, there are many definitions of stochastic stability which can be considered. In the discrete-time setting, however, one does not have to worry about the interpretation of white noise. A random sequence \{w(n)\} is said to be white if it is a Markov sequence whose conditional density function satisfies:

\[ p(w(k)|w(n)) = p(w(k)) \quad (k > n) \]

In other words, the w(k)'s form a mutually independent set of random variables. There is no Itô-type calculus required, because a white sequence is a physically
realizable process of finite average power. We are not treating the discrete-time systems as a means of simulating the continuous-time systems, but rather as a separate problem of equal importance.

A major difference found between the theories of continuous-time systems and of discrete-time systems is that the stability question for first order systems with a colored noise parameter has not been resolved in the discrete-time setting, because there is no useful analog to the solvable case of continuous-time systems. Solutions to difference equations appear as products of (possibly) correlated random variables (Equation (4.2.2)), not as a memoryless function of one random variable, as in Equation (3.1.5). The use of the logarithmic function in Section (4.2) "fixes" this problem to look like the solvable case, at the price of being conservative.

4.1 System Definition:

In what follows we will restrict our attention to the scalar system (with an obvious change in notation).
\[ x(n+1) = A x(n) + f(n) x(n) \] \hspace{1cm} (4.1.1)

\[ x(0) = x_0 \]

In Equation (4.1.6), \( \{f(n)\} \) represents a zero-mean, Gaussian random sequence with variance \( \sigma_f^2 \) and autocorrelation function

\[ E\{f(n)f(m)\} = \sigma_f^2 \beta^{|n-m|} \quad (|\beta| < 1) \] \hspace{1cm} (4.1.2)

By itself, \( \{x(n)\} \) is not a Markov sequence. That is, the statistics of \( x(n+1) \) depend upon more than the value of \( x(n) \) because of the memory of the \( \{f(n)\} \) sequence. By augmenting the state of the system so that \( \{f(n)\} \) is generated by passing white noise through a linear system and then considering the state to be the two dimensional vector \( (x(n), f(n)) \), a random sequence having the Markov property is obtained.

The augmented system is given by

\[ x(n+1) = A x(n) + f(n)x(n) \] \hspace{1cm} (4.1.3)

\[ f(n+1) = \beta f(n) + w(n) \]
\[
x(0) = x_0
\]

\(w(n) = \) White Gaussian sequence with zero mean, and

\[
E\{w(n)w(m)\} = \begin{cases} 
\sigma_w^2 & \text{if } n = m \\
0 & \text{otherwise}
\end{cases} \quad (4.1.4)
\]

and \(f(0) = \) Normal random variable with zero mean and variance

\[
\sigma_f^2 = \frac{\sigma_w^2}{1 - \beta^2}. \quad [S1] \quad (4.1.5)
\]

Note that the probability density function of \(f(0)\) is chosen so that \(f(n)\) is a stationary random sequence. (See Figure 4.1.1.)

For this system, the transition probability density function (equivalent to knowing the transition probability function) is determined by the following procedure:
Block Diagram of System (4.1.3)

Figure 4.1.1
\[
p(x(n+1), f(n+1) | x(n), f(n)) = p(x(n+1) | f(n+1), x(n), f(n))
\]

\[
\cdot p(f(n+1) | x(n), f(n))
\]

\[
= \delta(x(n+1) - A(x(n) - f(n)x(n)) p(f(n+1) | x(n), f(n))
\]

But \( p(f(n+1) | x(n), f(n)) = p_{W(n)}(f(n+1) - f(n)) \),

where \( p_{W(n)}(\cdot) \) is the probability density function of

the random variable \( w(n) \). So

\[
p(x(n+1), f(n+1)) = \int_{-\infty}^{\infty} dx(n) \int_{-\infty}^{\infty} df(n) \ p(x(n+1), f(n+1) | x(n), f(n))
\]

\[
\]

\[
= \int_{-\infty}^{\infty} dx(n) \int_{-\infty}^{\infty} df(n) \ (x(n+1) - A x(n) - f(n) x(n)) p_w(n) (f(n+1) - f(n)) p(x(n), f(n))
\]

\[
= \int_{-\infty}^{\infty} dx(n) \int_{-\infty}^{\infty} df(n) \ (x(n+1) - A x(n) - f(n) x(n)) p_w(n) (f(n+1) - f(n)) p(x(n), f(n)) \ (4.1.6)
\]

\([\ast]\) \( \delta = \text{Dirac delta function} \)
If $\beta \neq 0$, $f(n)$ is not a white random sequence, no closed form solution to (4.1.5) is known to the author. In particular, $x(n)$ and $f(n)$ are not jointly Gaussian random variables for $n \geq 1$ even if $x(0)$ and $f(0)$ are jointly Gaussian. Therefore, the function $P_{k,k+1}(x;dy)$ required in Equation (2.2.16) can not be evaluated. Clearly an alternate approach to the question of stability for System (4.1.1) is required.

4.2 A Theorem Not Requiring Zero Correlation Time

In this section, we follow a line of thinking that does not require us to evaluate the transition probabilities cited in the last section. Instead, our work will be based upon the simple property [L1]

$$\ln \mathbb{E}\{y\} \geq \mathbb{E}\{\ln y\}$$  \hspace{1cm} (4.2.1)

which is a direct application of Jensen's inequality [R2]. Because of the direction of the inequality (4.2.1), the resulting stability criteria is only a necessary condition. Being such a simple technique, it seems remarkable that the results come so close to the necessary and sufficient condition derived
by Willems [W3] for the white noise case. The following theorem does not depend upon the autocorrelation function of the \( \{f(n)\} \) sequence so it is necessarily conservative. However, it does hold for the non-white case.

**Theorem 4.2.1**

Let \( x(n) \) be generated by the System (4.1.1). Then

\[
\lim_{n \to \infty} E\{|x^2(n)|\} < \infty \quad \text{only if} \quad E\{ln |A + f(i)|\} < 0 \tag{4.2.2}
\]

**Proof of Theorem 4.2.1**

Using Equation (4.1.1), it is obvious that

\[
x^2(n+1) = (A+f(n))^2 \, x^2(n)
\]

\[
= (A+f(n))^2 (A+f(n-1))^2 \, x^2(n-1)
\]

\[
= \prod_{i=0}^{n} (A+f(i))^2 \, x_o^2 \tag{4.2.3}
\]

Then, using (4.2.1), one obtains
\[
\ln E\{x^2(n+1)\} \geq E\{\ln \prod_{i=0}^{n} (A+f(i))^2 \} x_0^2
\]

\[
= E\left\{ \sum_{i=0}^{n} \ln(A+f(i))^2 \right\} + E\{\ln x_0^2\}
\]

\[
= 2 \sum_{i=0}^{n} E\{\ln|A+f(i)|\} + E\{\ln x_0^2\}
\]

\[
= 2(n+1) E\{\ln|A+f(i)|\} + E\{\ln x_0^2\} \quad (4.2.4)
\]

because \(\{f(n)\}\) is a stationary sequence. The conclusion follows from Equation (4.2.4).

The region in the \((Ax\sigma_f)\) plane specified by (4.2.2) with \(\{f(n)\}\) assumed Gaussian has been determined numerically and is plotted in Figure 4.2.1. For comparison purposes, the result of Willems and Blankenship, which holds only for the special case, \(\beta = 0\), is also given. Notice also that the necessary condition (4.2.2) holds for any \(\{f(n)\}\) sequence having zero mean regardless of any correlation that may exist between \(f(n)\) and any delayed version of itself. Figure 4.2.1, however, holds only for the case in which \(\{f(n)\}\) is Gaussian.
Stability Region Specified by Equation (4.2.2)

Figure 4.2.1
Although the use of Figure 4.2.1 is not restricted to the case of white noise, there does not appear to be a way to improve the bound by exploiting any correlation properties that may be known. In particular, Equation (4.2.2) can be replaced by the requirement

\[ E\{\ln|A+f(i)|(A+f(i-1))|\} < 0 \quad (4.2.5) \]

in which the expected value is taken over the joint density of \( f(i) \) and \( f(i-1) \) which are (possibly) correlated.

The question still unanswered is:

Is there a \( \beta \in (-1,1) \) as defined by Equation (4.1.2) such that for some values of \( A \) and \( \sigma_w \),

\[ E\{\ln|A+f(i)|(A+f(i-1))|\} < 0 < E\{\ln|A+f(i)|\}? \quad (4.2.6) \]

The answer is no. To see this, we must simply note that
Therefore, there is no $\beta$ satisfying the question posed by Equations (4.2.6), and the bound given by Figure 4.2.1 is all that can be achieved via this technique.

In the literature [T1, F3, K7] the logarithmic function has been used to study properties of products of random matrices. However, each of those works requires the matrices to be statistically independent and independent of the state variables, so they are not applicable here.

4.3 Stability Criteria Based Upon Finite Correlation Times

Our goal is still to determine conditions that are necessary and/or sufficient for the stability of System (4.1.1), in which \( \{f(n)\} \) is a non-white noise sequence. Let us make the following definition:

**Definition**

The zero-mean sequence \( \{f(n)\} \) will be said to have **correlation time** \( M \) if \( M \) is an integer and
Clearly, a white sequence has zero correlation time. So, starting with the known case (zero correlation time) it seems reasonable to attempt to solve the problem in which the stochastic parameter has a finite, but non-zero correlation time. It turns out that the answer to this problem provides a necessary condition for the exponential mean square stability of systems with infinitely correlated noise sequences (e.g., System (4.1.1)).

Suppose again that the system under consideration is given by Equation (4.1.1), and the correlation function of $f(n)$, given by Equation (4.1.2), is

$$E[f(n)f(\ell)] = [\sigma_w^2 \beta |\ell-n| \, , \, |\beta| < 1] \quad (4.3.2)$$

The approach to determining conditions for the stability of (4.1.1) will be to study the systems

$$x_M(n+1) = Ax_M(n) + f_M(n)x_M(n) \quad x_M(0) = x_0 \quad (4.3.3a)$$
\[ f_M(n) = \sum_{i=0}^{M} \beta^i w(n-i) \]  

(4.3.3b)

for many values of \( M \). (See Figure 4.3.1.)

The behavior of the solutions of (4.3.3) as \( M \) becomes large will be the result of interest.

There are situations in which a colored noise sequence could be generated non-recursively, by passing a white noise sequence into a linear (non-recursive) filter. Such noise sequences actually have finite correlation time, so analyzing System (4.3.3) provides results that are useful in their own right. See [G2] for a discussion of the terminology.

Comparing Figures (4.3.1) and (4.1.1) provides insight into the scheme of things. For each \( M \), \( \{f_M(n)\} \), as defined by Equation (4.3.3), is a stationary Gaussian zero mean sequence such that for all \( n \),

\[ \lim_{M \to \infty} f_M(n) \to f(n) \]  
in the mean square sense.

\( f_M(n) \), however, is only a parameter in the System (4.3.3). What is really desired is to show that \( x_M(n) \to x(n) \)
System with Colored Noise Gain Having Finite Correlation Time

Figure 4.3.1
in some sense. This has not been achieved because the probability density function of \( x(n) \) is not known, and only very few of its statistical properties have been obtained.

To put things in a more concrete setting, consider the case \( M = 1 \), and \( x_0 = 1 \). Then, Equation (4.3.3) becomes

\[
x_1(n+1) = (A+w(n)+\beta w(n-1))x_1(n), \quad x_1(0) = 1
\]

(4.3.4)

So \( E\{x_1^2(n)\} \) must satisfy the equation

\[
E\{x_1^2(n+1)\} = E\{[A^2+2Aw(n)+2A\beta w(n-1)+w(n)^2
+ 2\beta w(n)w(n-1)+\beta^2 w(n-1)^2]x_1^2(n)\}
\]

(4.3.5)

Then, using the facts that \( w(n) \) and \( x(n) \) are statistically independent and that expectation is a linear operator, one obtains
\[ E\{x_1^2(n+1)\} = A^2 E\{x_1^2(n)\} + E[w^2(n)E[x_1^2(n)]] \]
\[ + 2A_\beta E[w(n-1)x_1^2(n)] + \beta^2 E[w(n-1)x_1^2(n)] \]

\[(4.3.6)\]

Writing expressions similar to (4.3.7) for the propagation of \( E\{w(n-1)x_1^2(n)\} \), and of \( E[w^2(n-1)x_1^2(n)] \)
leads to equations involving only scalar multiples of \( E[w(n-2)x_1^2(n-1)] \), \( E[w^2(n-1)] \), and \( E[w^2(n-2)x_1^2(n-1)] \).
Therefore, \( E\{x_1^2(n)\} \) can be propagated using the three dimensional system:

\[
\begin{bmatrix}
  E\{x_1^2(n+1)\} \\
  E\{w(n)x_1^2(n+1)\} \\
  E\{w^2(n)x_1^2(n+1)\}
\end{bmatrix} =
\begin{bmatrix}
  A^2 + \sigma_w^2 & 2A_\beta & \beta^2 \\
  2A_\sigma_w^2 & 2\beta_\sigma_w^2 & 0 \\
  A^2 \sigma_w^2 + 3\sigma_w^4 & 2A_\beta \sigma_w^2 & \beta^2 \sigma_w^2
\end{bmatrix}
\begin{bmatrix}
  E\{x_1^2(n)\} \\
  E\{w(n-1)x_1^2(n)\} \\
  E\{w^2(n-1)x_1^2(n)\}
\end{bmatrix}
\]

\[(4.3.7)\]

A computer program has been written to do the calculations required to generalize Equation (4.3.7)
to larger values of \( M \). Prior to describing that program,
however, two theorems will be considered.
Lemma 4.3.1

All the eigenvalues of a matrix \( \Gamma \) which contains only non-negative entries are smaller than unity in modulus if and only if all the successive principal minors of the matrix \((I - \Gamma)\) are positive.

Proof of Lemma 4.3.1

See Gantmakker, page 370.

Theorem 4.3.1

For each \( \beta > 0, A > 0, x_0 > 0 \), and for each \( M \),

\[
E\{x_M^2(n)\} \leq E\{x_{M+1}^2(n)\} \leq E\{x^2(n)\}
\]

\[n = 1, 2, \ldots \quad M = 0, 1, 2, \ldots \quad (4.3.8)\]

so that necessary conditions for the mean square stability of \( x(n) \) can be derived by studying the stability of \( x_M(n); M = 1, 2, \ldots \).

Proof of Theorem 4.3.1

From Equation 4.3.3, we can write \( x_M(n+1) \) as

\[
x_M(n+1) = Ax_M(n) + \sum_{i=0}^{M} \beta^i w(n-i)x_M(n)
\]
Therefore,

\[ E\{x_M^2(n+1)\} = A^2E\{x_M^2(n)\} + 2A \sum_{i=0}^{M} \beta^i E\{x_M(n)w(n-i)\} \]

\[ + \sum_{i=0}^{M} \sum_{j=0}^{M} \beta^{i+j} E\{w(n-i)w(n-j)\} \]

(4.3.9)

Each term in Equation (4.3.9) is non-negative and increasing M only adds some terms to the right hand side of (4.3.9). Therefore,

\[ E\{x_M^2(n)\} \leq E\{x_{M+1}^2(n)\} \]

(4.3.10)

In Equations (4.1.3) and (4.1.5) we have generated the (infinitely correlated) sequence \{f(n)\} with an initial condition chosen to make it stationary. An equivalent procedure would be to generate \( f(n) \) via

\[ f(n+1) = \sum_{i=0}^{\infty} \beta^i w(n-i) \]

(4.3.11)
and assume that \( \{w(n)\} \) started in the infinite past.

Using Equation (4.3.11) in deriving a formula for

\[ E\{x^2(n)\} \]

amounts to setting \( M = \infty \) in the right hand side of Equation (4.3.9) so that

\[ E\{x^2(n)\} \geq E\{x_{M+1}^2(n)\} \]

follows from Equation (4.3.10).

**Remark 1**

The sign of \( A \) in Equation (4.3.7) does not matter.

To see this define \( \Delta(A) \) as the determinant of the matrix

\[
\Delta(A) = \begin{vmatrix}
A^2 + \sigma_w^2 - S & 2A\beta & \beta^2 \\
2A\sigma_w^2 & 2\beta\sigma_w^2 - S & 0 \\
A^2\sigma_w^2 + 3\sigma_w^4 & 2A\beta\sigma_w^2 & \beta^2\sigma_w^2 - S
\end{vmatrix}
\]

Without changing the value of \( \Delta(A) \) one can multiply

the second row of the matrix by \((-1)\) and multiply the

second column of the matrix by \((-1)\). The result is

\[
\Delta(A) = \begin{vmatrix}
A^2 + \sigma_w^2 - S & -2A\beta & \beta^2 \\
-2A\sigma_w^2 & 2\beta\sigma_w^2 - S & 0 \\
A^2\sigma_w^2 + 3\sigma_w^4 & -2A\beta\sigma_w^2 & \beta^2\sigma_w^2 - S
\end{vmatrix}
\]
The result can be recognized as $\Delta(-A)$, so $\Delta(A) = \Delta(-A)$ and this is true regardless of the sign of $\beta$. The results of the computer programs described below indicate that the sign of $A$ does not matter even for higher values of $M$.

The following theorem and its implications represent the main results of this section.

**Theorem 4.3.2**

For any finite value of $M$ ($M > 0$ as defined by Equation (4.3.3)), the mean square value of $x_M(n)$ (also defined by Equation (4.3.3) can be propagated (exactly) by a linear system having dimension at most $1 \cdot 3 \cdot 5 \ldots (2M + 1)$.

**Proof of Theorem 4.3.2**

From (4.3.3) one can immediately write

$$
x_M^2(n+1) = \left[ A^2 + 2A \sum_{i=0}^{M} \beta^i w(n-i) \right. \\
\left. + \sum_{i=0}^{M} \beta^i \rho w(n-i) w(n-\rho) \right] x_M^2(n) \tag{4.3.12}
$$
Then, for $j_k = 0, 1, 2, \ldots, 2(M-k+1)$ and for $k = 1, 2, \ldots, M$, one can write each of the states of the system to propagate $E\{x_M^2(n)\}$ as:

$$E\{w^1(n)w^2(n-1) \ldots w^M(n-M+1)x_M^2(n+1)\}$$

$$= E\{A^2w^1(n)w^2(n-1) \ldots w^M(n-M+1)x_M^2(n)\}$$

$$+ 2A E\left\{ \sum_{i=0}^{M} \beta^i w(n-i)w^1(n)w^2(n-1) \ldots w^M(n-M+1)x_M^2(n) \right\}$$

$$= E\left\{ \sum_{\ell=0}^{M} \sum_{i=0}^{M} \beta^{i+\ell} w(n-i)w(n-\ell)w^1(n)w^2(n-1) \ldots w^M(n-M+1)x_M^2(n) \right\}$$

(4.3.13)

Careful observation of (4.3.13) reveals that at each step, the propagation of each of the states, as specified by the left hand side of (4.3.13) requires knowledge of only each of the states at the present time. Q.E.D.
Remark 2

The bound on the number of states is not overly conservative. The results of the computer program show that propagation of $x_M^2(n)$ requires 1, 3, and 12 states for $M = 0, 1, 2$, respectively. One reason for the abundance of white noise ($M = 0$) results is now apparent.

Remark 3

The number of states required increases very rapidly with $M$. It appears that the System (4.3.3) with $M = 3$ is all that can be propagated (using equations like (4.3.7)) on a large computer. Even with that limitation, however, a substantial decrease in the possible region of parameter space representing stable systems has been achieved.

Remark 4

The above proof is constructive in that it specifies what the required states are. Also, the three terms on the right hand side of (4.3.13) provide the formulae used in the computer programs.
A computer program (based on the formula (4.3.10)), has been written to determine what combinations of values of \( \beta, A, M, \) and \( \sigma_w \) imply that the System (4.3.3) is stable. For \( M = 0, 1, 2 \), and for various values of \( \beta \), the following figures show the results of the program. For \( M = 0 \), the curve is identical to that given earlier by Willems and Blankenship [W3]. The curves for \( M > 0 \) and \( \beta > 0 \) indicate a decrease in the possible area of stability in the \((A, \sigma_w)\) plane as the noise correlation time \((M)\) is increased. The bottom figure on each page contains the same information as the top figure, plotted as a function of the steady-state variance of \( f(n) \). The plot was made to show that the change in the area of stability of the System (4.3.3) is not solely due to the increase in the variance of the \( \{f_M(n)\} \) sequence as \( M \) is increased.

The computer program results for the \( \beta < 0 \) cases indicate some more interesting phenomena. In this case, we are unable to show that increasing \( M \) provides successively better boundaries for the known region of instability and, in fact, the computer output shows that the opposite is true. Therefore, for the cases characterized by \( \beta < 0 \), the results do not have the
desired interpretation of providing a necessary stability
criteria for the infinite correlation-time case (e.g.,
the system with f(n) generated recursively, Figure
(4.1.1)).

From Figures (4.3.5), (4.3.6), and (4.3.7), it is
easy to pick a counterexample to the hypothesis that,
for a given noise power, coloring always provides a
destabilizing influence. One such case is defined by
the parameter values $\beta = -0.5$, $|A| = 0.9$, and $\sigma_f = 0.5$.
According to Figure (4.3.6b), that system is unstable
if $\{f(n)\}$ is a white sequence, but it is exponentially
mean square stable if $\{f(n)\}$ is correlated over one or
two time steps. This is not to say that an unstable,
deterministic system can be stabilized by adding zero-
mean colored noise to one of its parameters; the boundary
at $A = 1$ in each of the graphs rules out that possi-
bility.
Stability Regions Defined by Equation (4.3.13)

Figure 4.3.2
Stability Regions Defined by Equation (4.3.13)
Figure 4.3.3
Stability Regions Defined by Equation (4.3.13)

Figure 4.3.4
Stability Regions Defined by Equation (4.3.13)

Figure 4.3.5
Stability Regions Defined by Equation (4.3.13)
Figure 4.3.6
Stability Regions Defined by Equation (4.3.13)
Figure 4.3.7
For $M = 2$ and for each value of $\beta$ considered, the stability boundaries of System (4.3.3) are all superimposed in Figure (4.3.8). Each of these curves appeared in a previous figure, but are repeated here for comparison purposes. One can easily draw some conclusions about the effects of increasing or decreasing $\beta$, but an even more important conclusion is possible: The stability of a linear dynamical system perturbed by a stationary, non-white multiplicative noise process depends upon the detailed correlation properties of the noise as well as the power of the noise and the specification of its white or non-white property. This phenomenon has not previously been demonstrated.
Comparison of Stability Regions for Various Values of $\beta$

Figure 4.3.8
CHAPTER V

SUMMARY AND CONCLUSIONS

In this chapter, we have the opportunity to discuss the results presented in previous chapters, and to take the liberty of speculating on the possibilities of some future research. As in most investigations of this type, we have raised some new questions while answering the old ones.

In the context of continuous time systems, we have provided an existence and uniqueness proof for the solution of a linear system with a colored noise parameter. In each of the examples studied in Chapter III, we assumed that only one random parameter was present and it is only in that case that the existence proof applies. Except for the possibility of incurring a growing number of terms in the Picard expansion of the solution, extending the proof to allow for more than one noise parameter should be straightforward. In the discrete-time context, however, our demonstration
of the Markov property for the augmented state applies even in the case that multiple noise parameters are present.

In Chapter III, we have established the usefulness of the Lie algebra theory in proving the stability properties of certain systems. The results were presented in such a way as to be easily interpreted in terms of the eigenvalues of the system matrix of the noise-free system. However, the reader is cautioned against generalizing this interpretation to systems whose transition matrices do not evolve on a solvable Lie group. In the non-solvable case, there is considerably more coupling between the state variables of the system, making its properties more difficult to establish.

Section 3.3 provides an exponentially decaying bound for the mean square response of a non-solvable example, the damped harmonic oscillator with noisy feedback. The key to proving the mean square stability of the damped harmonic oscillator was our ability to solve the first order case exactly and then to recognize that the Picard expansion of the solution
to the scalar problem bounded the Picard expansion for the higher order problem. The following question then arises: Can the solution to some larger class of stochastic systems be bounded in the mean square by the solution of a solvable (Lie sense) problem? We leave this question as a topic for further research.

Further generalizations of Section 3.3 are also possible. For example, only the single-input, single-output case has really been treated and the extension to multiple inputs and outputs remains to be done. Also, in line with the comments concerning the existence proof in Chapter II, stability criteria for systems containing more than one random parameter have not been derived.

In Chapter IV, we have addressed the stability problem for discrete-time, continuous-state-space systems. In spite of the apparent analogy to the continuous-time problem we were able to make progress on the discrete-time problem only by using completely different tactics. We were able to derive some necessary criteria for the mean square stability of some discrete-time systems, but the scope of our results is considerably more restricted than that achieved in the continuous-time case.
In the discrete-time case we have found that very few stability results for systems with colored noise parameters are available. Therefore, we were forced to concentrate on first order systems to make any progress. Analysis of those systems containing multiple inputs, multiple outputs and multiple stochastic parameters has not yet been attempted.

We wish to re-emphasize that the information in Figures (4.3.4) through (4.3.10) did not result from a Monte Carlo simulation of the system studied in Chapter IV. Quite to the contrary, such a simulation would probably show that for \( |A| < 1 \) and for \( \mathbb{E}\{f(n)\} = 0 \) for all \( n \), then almost every sample sequence \( \{x(n)\} \) would approach zero, so the estimated variance would also.

During the course of this research we tried to isolate any special cases that seemed to behave counter to our intuition. We were unable to identify any unstable system that could be stabilized by adding zero-mean colored noise to one of its parameters. However, in the discrete-time setting, systems containing a random parameter with oscillatory autocorrelation
function did have a peculiar feature. For values of \(|A|\) close to unity, increasing the correlation of the noise process had a stabilizing influence. No such phenomenon was found in the continuous-time case.
REFERENCES


M7 J.A. Morrison and J. McKenna, "Analysis of Some Stochastic Ordinary Differential Equations"; Bell Laboratories, to be Published.


We have repeatedly made reference to the results of previous research that considered systems with multiplicative white noise. In this appendix we discuss the theory developed for analyzing such systems and indicate why we were unable to make greater use of these ideas in our analysis.

In this appendix, we will be considering equations of the form

\[
\frac{dx(t)}{dt} = A(x(t), t) + G(x(t), t)w(t) \tag{A.1}
\]

\[x(t_0) = x_0\]

where \(G\) is a real valued, \(nxm\) matrix,
\(w(t)\) is a white Gaussian process,
and \(A\) is an \(n\)-vector.

It is known that a white noise process is not mean square integrable and its sample functions are not integrable (with probability one), so Equation (A.1)
has no meaning as a differential equation. However, a mathematical theory for the analysis of Equation (A.1) has been developed by Itô. The Itô calculus, which involves significant changes in the usual rules of differentiation, has proven to be very successful. The basis for this calculus comes from the properties of Brownian motions, and the representation of a white noise process as the (formal) derivative of a Brownian motion.

**Definition:**

A continuous parameter process \( \{ \beta(t), \ t \geq 0 \} \) is a **Brownian motion** process if

(i) \( \{ \beta(t), \ t \geq 0 \} \) has stationary independent increments; \( \text{(A.2)} \)

(ii) \( \beta(t) \) is a Gaussian random variable for every \( t \geq 0 \). \( \text{(A.3)} \)

(iii) \( \mathbb{E}\{\beta(t)\} = 0 \) for every \( t \geq 0 \). \( \text{(A.4)} \)

(iv) \( \beta(0) = 0 \) w.p.1 \( \text{(A.5)} \)

It follows that \( \mathbb{E}\{\beta(t)\beta(\tau)\} = \sigma^2 \min(t, \tau) \), \( \text{(A.6)} \)

where \( \sigma^2 \) is an empirical positive constant called the variance parameter. [J2, page 72] Using the notation
\( \beta_t = \beta(t) \) will simplify what follows.

In order to interpret Equation (A.1), we will need the first-order and second-order stochastic integrals as defined by Itô:

Let \( E|g_t(\omega)|^2 \ll \) for all \( t \in [a,b] \) and suppose \( g_t(\omega) \) is mean square continuous on \( t \), then

\[
\int_a^b g_t(\omega) d\beta_t = \lim_{\rho \to 0} \sum_{i=0}^{n-1} g_{\tau_i}(\omega)(\beta_{\tau_{i+1}} - \beta_{\tau_i})
\]

(A.7)

and

\[
\int_a^b g_t(\omega) d\beta_t^2 = \lim_{\rho \to 0} \sum_{i=0}^{n-1} g_{\tau_i}(\beta_{\tau_{i+1}} - \beta_{\tau_i})^2
\]

(A.8)

where \( a = t_0 < t_1 < \ldots < t_n = b \),

and \( \rho = \max_{i}(t_{i+1} - t_i) \).

Reference [J2, pages 99, 102] establishes that (A.7) and (A.8) are well defined. Equation (A.1) will be given a precise meaning by interpreting it to be the integral equation.
\[ x_t - x_0 = \int_0^t A(x_t, t) \, dt + \int_0^t G(x_t, t) \, d\beta_t \quad (A.9) \]

where the last term is a stochastic integral. In Chapter II, Equations (2.1.2) through (2.1.8), we have stated some properties of the solution of Equation (A.9) under suitable conditions on the matrices \( A(x,t) \) and \( G(x,t) \). A proof of existence and uniqueness for its solution is given by Wong, [W2].

The reason that the standard calculus can not be used in these equations comes from a peculiar property of Brownian motion processes, i.e., the squared sample increments \((\Delta \beta_t)^2\) are of order \( O(\Delta t) \), not \( O((\Delta t)^2) \).

The consequence of this is that some second derivative terms (those that involve the second power of \( d\beta \)) in a Taylor series expansion of a function must be included in the stochastic differential of the function. In the formula for stochastic differentials given below, the extra term that would not appear using the standard calculus is a consequence of the property

\[ (d\beta_t)^2 \sim dt \quad (A.10) \]
The important consequence of the Itô interpretation of Equations (A.1) and (A.9) is the following differentiation rule that holds under the same conditions required in the existence and uniqueness proofs: [J2, page 112].

Let \( x_t \) be the unique solution of the vector Itô stochastic differential equation

\[
dx_t = A(x_t,t)dt + G(x_t,t)d\beta_t
\]

where \( \{\beta_t, t \geq t_0\} \) is an \( m \)-vector Brownian motion process with \( E[d\beta_t d\beta_t^T] = Q(t)dt \).

Let \( \phi(x_t,t) \) be a scalar-valued real function, continuously differentiable in \( t \) and having continuous second mixed partial derivatives with respect to the elements of \( x \). Then the stochastic differential \( d\phi \) of \( \phi \) is

\[
d\phi = \phi_t dt + \phi_x^T dx_t + \frac{1}{2} \text{tr} GQQ^T \phi_{xx} dt
\]

\[
= \phi_t dt + \phi_x^T dx_t + \frac{1}{2} \sum_{i,j=1}^{n} \left(GQQ^T\right)_{ij} \frac{\partial^2 \phi}{\partial x_i \partial x_j} dt
\]

(A.12)
where

\[ \phi_t = \frac{\partial \phi}{\partial t} \quad \phi_x^T = \left[ \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n} \right] \]

\[ \phi_{xx} = \begin{bmatrix} \frac{\partial^2 \phi}{\partial x_1^2} & \frac{\partial^2 \phi}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 \phi}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \phi}{\partial x_n \partial x_1} & \frac{\partial^2 \phi}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 \phi}{\partial x_n^2} \end{bmatrix} \]  

(A.13)

Again let \( x_t \) be the solution of the differential equation (A.11). The \( \{x_t\} \) process is characterized by the evolution of its probability density function \( p(x_t) \) (or \( p(x,t) \)), which is known to satisfy the Fokker-Planck equation [J2, page 130], if it exists.

\[ \frac{\partial p(x,t)}{\partial t} = - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left( p(x,t)A_i(x,t) \right) \]

\[ + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 (p(x,t) [GG^T]_{ij})}{\partial x_i \partial x_j} \]

(A.14)
where $A_i(\cdot, t)$ is the $i$th component of $A(\cdot, t)$ and $G$ and $Q$ are given above.

Equation (A.14) is also called Kolmogorov's forward equation. Kolmogorov's backward equation is the formal adjoint of the forward equation, given by

$$
\frac{-\partial p(x, t)}{\partial t} = \sum_{i=1}^{n} A_i(x, t) \frac{\partial}{\partial x_i} p(x, t) + \frac{1}{2} \sum_{i,j=1}^{n} [GQG^T]_{ij} \frac{\partial^2 p(x, t)}{\partial x_i \partial x_j} \quad (A.15)
$$

which is used to define the backward diffusion operator.

$$
L(\cdot) = \sum_{i=1}^{n} A_i(x, t) \frac{\partial}{\partial x_i} (\cdot) + \frac{1}{2} \sum_{i,j=1}^{n} [GQG^T]_{ij} \frac{\partial^2 (\cdot)}{\partial x_i \partial x_j} \quad (A.16)
$$
Using $L(\cdot)$, we can rewrite Equation (A.12) as

$$d\phi = \phi_t dt + \phi_x^T (A(x_t,t) dt + G(x_t,t) d\beta_t)$$

$$+ \frac{1}{2} \sum_{i,j=1}^n [GQG^T]_{ij} \frac{\partial^2 \phi}{\partial x_i \partial x_j}$$

$$= \phi_t dt + L(\phi) + \sum_{i,j=1}^n G_{ij}(x_t,t) (d\beta_t) j \frac{\partial \phi}{\partial x_i}$$

(A.17)

If we were using ordinary calculus here, only the first two terms on the right hand side of Equation (A.17) would have appeared. [K3]

The differential operator associated with Equation (A.11) is important in the study of the stability of linear system with multiplicative white noise. For example, the following theorem is given by Nevelson and Khas'minskii [N2].

Theorem

Let $V(x,t)$ be a scalar valued real function continuously differentiable in $t$ and having continuous mixed second partial derivatives with respect to the
elements of $x$. Let $V(x,t)$ satisfy

$$C_1 ||x||_M^M \leq V(x,t) \leq C_2 ||x||_M^M$$

$$L(V(x,t)) \leq -C_3 ||x||_M^M$$

$$\left| \frac{\partial V}{\partial x_i} \right| \leq C_4 ||x||_M^{M-1}$$

$$i = 1, 2, \ldots, n$$

where $C_1$, $C_2$, $C_3$, and $C_4$ are any positive constants. Then, the equilibrium solution of (A.11) possesses exponential stability of the $M$th moments.

$$(||x||_M^M = \sum_{i=1}^{n} |x_i|^M)$$

**Example**

Suppose Equation (A.11) represents the linear scalar system

$$dx_t = ax dt + gx d\beta_t$$

(A.21)

The diffusion operator associated with Equation (A.21) is
The function $V(x,t) = x^2(t)$ satisfies Equations (A.18) and (A.20) of the theorem. Using (A.22) on $x^2(t)$, one obtains

$$L(V) = ax(t) \frac{\partial}{\partial x} x^2(t) + \frac{g^2}{2} x^2(t) \frac{\partial^2 x^2(t)}{\partial x^2}$$

$$= 2a x^2(t) + \frac{g^2}{2} x^2(t)$$

So $L(V) \leq C_3 |x|^2$ for some $C_3 > 0$ if and only if

$$2a + g^2 < 0 \quad (A.23)$$

and, if that is true, condition (A.19) of the theorem is satisfied so the System (A.21) is exponentially mean square stable.

If the results of Section 3.1 are applied to the system

$$\frac{dx(t)}{dt} = [a + \sigma f_\alpha(t)] x(t) \quad (A.24)$$
where \( f_\alpha(t) \) is a real, stationary, zero-mean colored noise process having autocorrelation function and power spectral density

\[
R_{f_\alpha}(\tau) = \frac{\alpha}{2} \ e^{-|\alpha|\tau}, \quad S_{f_\alpha}(\omega) = \frac{1}{1 + \omega^2/\alpha^2}
\]

respectively, the following is obtained.

The System (A.24) is stable in the mean square sense for any finite \( \alpha > 0 \) if and only if

\[
a + \sigma^2 < 0. \quad (A.25)
\]

Equation (A.23) appears to contradict this result if one tries to let \( \alpha \to \infty \) and identify the limit of Equation (A.24) with the \( \text{Itô} \) Equation (A.21),

\[
dx = ax \ dt + gx \ d\beta_t
\]

where \( \beta_t \) is a Brownian motion. That is, the stability region of (A.21) has been identified by the inequality

\[
2a + g^2 < 0 \quad (A.26)
\]
The apparent discrepancy between (A.25) and (A.23) lies in the Itô interpretation of a white noise process. The answer is contained in the following result [W1, W2].

If, under certain smoothness conditions on \( m \) and \( \sigma \), the white-noise driven equation

\[
\frac{d}{dt} x(t) = m(x(t), t) + \sigma(x(t), t) \zeta(t) \quad (A.27)
\]

is to be interpreted as the limit as \( \alpha \to \infty \) of a sequence of equations like (A.24), then (A.27) is equivalent to the equation

\[
dx = m(x, t) + \frac{1}{2} \sigma(x, t) \frac{\partial}{\partial x} \sigma(x, t) + \sigma(x, t) d\beta
\]

(A.28)

interpreted in the sense of Itô. That is, the solution of a sequence of equations like (A.24) converges to the Itô equation

\[
dx = (\alpha + \frac{1}{2} \sigma^2) xdt + \sigma xd\beta
\]

(A.29)
The added term, $\frac{1}{2} \sigma^2 x dt$, is called the Wong-Zakai correction term [W11, W12]. Then applying formula (A.25) to System (A.29) one obtains the requirement

$$2(a + \frac{1}{2} \sigma^2) + \sigma^2 < 0$$

or

$$a + \sigma^2 < 0,$$

which agrees with (A.25)

The formulae for the correction terms required for higher order systems is given in [W1, page 159] and [W9].

Further results about the properties of the solution of the Itô equation (A.11) have been obtained. In [H3], Haussmann has considered the existence question for moments of the random process defined by Equation (A.11). Brockett [B5] has obtained equations that govern the propagation of arbitrary moments of the solution of Equation (A.11) for the white noise multiplier case. The results are easily expressed in terms of Kronecker products. In the special case that second moments are being considered we refer the reader to Reference [B6].
The sample path behavior of the solution to the Itô equation (A.11) has been studied by Kozin and Prodromou [K12]. They have obtained necessary and sufficient conditions for the almost sure stability of the solution process for two special cases of Equation (A.11), one of which is the damped harmonic oscillator problem.

By using state augmentation and Itô correction terms, we are able to represent systems containing colored noise parameters as Itô systems with multiplicative white noise. Unfortunately, the systems obtained through this procedure contain quadratic terms which violate the global Lipschitz hypothesis of many of the theorems given above.

For example, the damped harmonic oscillator problem studied in Section 3.3 can be rewritten as

\[
\begin{align*}
\frac{dx_1}{dt} &= x_2 dt \\
\frac{dx_2}{dt} &= -x_1 dt - x_3 x_1 dt - 2\zeta x_2 dt \\
\frac{dx_3}{dt} &= -bx_3 dt + \sigma d\beta_t
\end{align*}
\]

(A.30)
Formally applying the formula (A.15) to Equation (A.30), we obtain

\[
\frac{\partial p}{\partial t} = -x_2 \frac{\partial p}{\partial x_1} + x_1 \frac{\partial p}{\partial x_2} + x_3 x_1 \frac{\partial p}{\partial x_2} + 2\zeta x_2 \frac{\partial p}{\partial x_2} + 2\zeta p + bx_3 \frac{\partial p}{\partial x_3} + bp + \frac{\sigma^2}{2} \frac{\partial^2 p}{\partial x_3^2}
\]  

(A.31)

Harris [H2] has studied the Fokker-Planck equation for nonlinear stochastic systems by defining a transformation to bring the system into the form of an equivalent deterministic system with additive white stochastic perturbations. Unfortunately, Equation (A.31) does not fall into the category of systems for which his theory is applicable. The work of Evans [E1], which investigates the asymptotic properties of a similar Fokker-Planck equation, is also not applicable. Kozin [K12] has noted similar difficulties in trying to solve the Fokker-Planck equation corresponding to the damped harmonic oscillator with white multiplicative noise. In that case, singular diffusions are generated, making solution of the Fokker-Planck equation impossible, at least if strong-sense solutions are required.
Our attempts to investigate the properties of the solution to Equation (A.31) have similarly failed. Two approaches were tried. In the first case we have tried to represent $p$ from Equation (A.31) in the form

$$p = \sum_{i=0}^{\infty} \rho_i(x_1,x_2,t) \eta(x_3)$$

(A.32)

Then (A.31) can be rewritten as

$$\sum_{i=0}^{\infty} \left\{ \frac{\sigma^2}{2} \rho_i \eta_i''(x_3) - \eta_i(x_3) \left[ \frac{\partial \rho_i}{\partial t} + x_2 \frac{\partial \rho_i}{\partial x_1} \right. \right. \right.$$  

$$- \left( x_1 + x_3 x_1 + 2 \xi x_2 \right) \frac{\partial \rho_i}{\partial x_2} \right\} + (2 \xi + b) \rho_i \eta_i(x_3) + bx_3 \rho_i \eta_i'(x_3) \right\}$$

(A.33)

Denoting the terms inside the brackets as $A_{x_3}(\rho)$, we tried to find eigenvalues of the operator $A_{x_3}$ (i.e., solutions to $A_{x_3}(\rho_i) = \lambda_{x_3} \rho_i$) by the method of characteristics and then solve the remaining equation for $\eta(x_3)$:
\[ \sum_{i=0}^{\infty} \rho_i \left[ \frac{\sigma^2}{2} \eta_i''(x_3) - \lambda x_3 \eta_i(x_3) + (2\xi + b) \eta_i(f) \right] + b f \eta_i'(f) \right] = 0 \quad (A.34) \]

Unfortunately, we were unable to find any set of orthogonal polynomials that solved Equation (A.34). A second approach was to assume \( p \) has the form

\[ p(x_1, x_2, x_3, t) = \sum_n \rho_n(x_1, x_2, t) H_n(x_3)e^{-\frac{x_3^2}{2}} \]

where \( H_n(*) \) are Hermite polynomials. This also proved unsuccessful.

The major difficulty in attempting to solve Equation (A.31) is that \( x_1 = x_2 = 0 \) is a trap. Hence, the invariant measure will be concentrated on the line corresponding to \( x_1 = x_2 = 0 \) (the \( x_3 \)-axis). [K12]. We will make no further attempt to investigate the behavior of the solution of Equation (A.31) here.
APPENDIX B

LIE ALGEBRA DEFINITIONS

In this appendix, we describe some of the concepts of Lie Algebra Theory that are required for an understanding of Section 3.2. The usefulness of Lie Algebra Theory in the study of time varying multidimensional differential equations has been previously exploited by Niemeyer [N4] and by Wei and Norman [W8]. Brockett [B3, B4, B5] has also used Lie Algebra Theory in his studies of bilinear systems, particularly where the controllability of such systems is concerned. More recently, Willsky and Marcus [W9] have noted the usefulness of Lie Algebra Theory in the study of stochastic stability for bilinear systems.

For information on Lie Algebra Theory beyond what is required in Chapter III, the reader is referred to references [S2] and [J1].

An algebra, $A$, is a triple $(S, +, \cdot)$ for which $(S, +)$ is a vector space over some given field, and $\cdot$ denotes multiplication. If the multiplication is defined to be associative, the algebra $A$ is called an associative
A Lie Algebra, \( L \), is an algebra in which the multiplication, denoted \([ , ]\), satisfies \([x,y] = -[y,x]\), and the Jacobi Identity,

\[
[[x,y],z] + [[y,z],x] + [[z,x],y] = 0 \quad (B.1)
\]

We will be concerned only with Lie Algebras for which the elements are \( n \times n \)-matrices with real entries; multiplication being the commutator product \([A,B] = AB - BA\). A Lie Algebra, \( L \), is called abelian if \([A,B] = 0\) for all \( A, B \in L \).

We will use the standard notation, \( L' = [L,L] \) to denote the subspace spanned by the set of elements of \( L \) that are the commutator products of two elements of \( L \). Let \( A, B \) denote two elements of \( L' \). Then \([A,B] \in L'\) and \( A+B \in L' \), so \( L' \) satisfies the defining properties of a subalgebra and we write \( L' \subseteq L \). Continuing with the process of taking commutator products leads to the derived series*:

\[
L'' = [L',L'] \subseteq \\
\vdots \subseteq L^{(k)} = [L^{(k-1)},L^{(k-1)}] \quad (B.2)
\]

* Reference [J1, page 23]
The original Lie Algebra, \( L \), is called **solvable** if 
\( L^{(h)} = 0 \) for some positive integer \( h \). Suppose we now consider the subalgebras defined by the **lower central series**:

\[
L^2 = L' = [L, L] \\
\subseteq L^3 = [L^2, L] \\
\vdots \\
\subseteq L^k = [L^{k-1}, L]
\]

\( L \) is termed **nilpotent** if \( L^k = 0 \) for some positive integer \( k \). An important example of a nilpotent Lie Algebra is the set of strictly upper-triangular matrices. It is immediate from the definitions that every abelian Lie Algebra is also nilpotent. Also, for every \( k \), \( L^{(k)} \subseteq L^k \), so the nilpotent property implies that \( L \) is also solvable.

A **matrix group** is a set of \( nxn \) matrices that is a group under matrix multiplication. If \( \mathcal{B} \) is a set of \( nxn \) matrices, \( \{ \mathcal{B} \}_G \) denotes the smallest matrix group containing \( \mathcal{B} \). If \( L \) is any Lie Algebra, and \( \exp(L) \) denotes \( \{ \exp(L) : L \in L \} \), then \( G = \{ \exp(L) \}_G \) defines a **matrix Lie group**, every element of which is invertible.
Suppose $A, B \in G$, then $ABA^{-1}B^{-1}$ is called a **commutator**, and the set of all the commutators (i.e., for every pair of elements in $G$) is called the **commutator group** $C(G)$ of $G$. Iterating the process of forming the commutator group defines the **commutator sequence** of $G$. $G$ is called **solvable** if its commutator sequence leads to (the group identity) $\{I\}$ after a finite number of steps.

The tangent space of a group, $G$, evaluated at $\{I\}$, is a Lie Algebra, the **infinitesimal algebra** of $G$. For the case $G = \{\exp(L)\}_G$, $L$ is the infinitesimal algebra of $G$.

The similarity of the definitions of a solvable Lie Algebra and a solvable Lie group is no coincidence, for matrix Lie group is solvable if and only if its infinitesimal algebra is solvable. [F4, page 64].

Those (possibly time-varying) systems whose transition functions evolve on a solvable Lie group can be analyzed by a special set of techniques that are not applicable in more general cases. Refer again to Equation (3.2.2).

$$\frac{d}{dt} \phi(t, t_0) = [A_0 + \sum_{i=1}^{N} f_i(t)A_i] \phi(t, t_0)$$

$$\phi(t_0, t_0) = I \quad \text{(B.3)}$$
and assume \( \{A_i\}_{i=0}^{N} \), the smallest Lie Algebra containing
\( A_i \) for \( i = 0,1,2,...,N \) is solvable. In this case, it
is known that there exists a similarity transformation
that will simultaneously transform all the \( A_i \) matrices
into upper triangular form. In this form it can easily
be seen that we can sequentially solve the transformed
version of (B.3), starting with the bottom row. Also,
if the \( \{A_i\} \) matrices are transformed to upper triangular
form, the \( \phi \) will evolve on the group of non-singular,
upper triangular matrices, which is a solvable group.

The basic tool we will use in the Lie Algebra work
is the Baker-Hausdorff Lemma:

If \( A,B \in L \), then \( \exp(A)B \exp(-A) \in L \) and is given
explicitly by the formula

\[
\exp(A)B \exp(-A) = B + [A,B] + \frac{1}{2!}[A[A,B]] + \frac{1}{3!} [A,[A,[A,B]]] + ... \quad (B.4)
\]

A proof is given by Magnus [M2]. If the matrices \( A \) and
\( B \) commute, then \( \exp(A)B \exp(-A) = B \), otherwise \( \exp(A) \)
and \( B \) don't commute.
APPENDIX C

LISTINGS OF COMPUTER PROGRAMS

We have included listings of the following computer programs.

Listing 1. Program to evaluate $E\{\ln|A + f(i)|\}$ of Equation (4.2.4).

Listing 2. Program to evaluate Equation (4.3.10) for $M = 1$.

Listing 3. Program to evaluate Equation (4.3.10) for $M = 2$.

Listing 4. Program to search for optimum value in Equation (3.3.24).
PI=3.1415927
A=0.5
B=0.5
DO 41 ISIGW=1.5
SIGW=0.25*FLOAT(ISIGW)+0.75
SIGF=SIGW/(1.-B*R)
SWSO=SIGW*SIGW
SF2=0.5*(1.-B*R)/SWSQ
SUM=0.
DO 40 IF2=2A,76
F2=FLOAT(IF2-51)*SIGW/10.
G=A*F2
GSO=G*G
DELF=0.1*SIGW
S2=2.*DELF*(ALOG(DELF)-1.)*EXP(-.5*GSO/SWSQ)
S1=-.5*ALOG(DELF)*EXP(-.5*(G+DELF)**2/SWSQ)
S3=-.5*ALOG(DELF)*EXP(-.5*(G-DELF)**2/SWSQ)
DO 1000 I=1,50
G1=G-FLOAT(I)*DELF
G2=G*FLOAT(I)*DELF
S1=S1+ALOG(ABS(G1+G))*EXP(-0.5*G1*G1/SWSQ)
S3=S3+ALOG(ABS(G2+G))*EXP(-0.5*G2*G2/SWSQ)
1000 CONTINUE
S=S2+S1*DELF+S3*DELF
WRITE(6,100) S,S1,S2,S3,SIGF,F2
100 FORMAT(1P6F15.8)
SUM=SUM+S*EXP(-SF2+F2*S2)
40 CONTINUE
SUM=SUM*SQRT(1.-B*R)/(PI*SWSQ)
WRITE(6,100) SF2,F2,SUM
41 CONTINUE
STOP
END

Listing 1
PROGRAM TO PROPAGATE EXP VAL OF X**2

DIMENSION P(3,3),F(7),X(3),Y(3)
READ(5,21) F
21 FORMAT(7F5.1)
DO 103 I=1,3
A=FLOAT(I)*0.25
B=1.01
DO 101 SIG=1.20
SIG=FLOAT(SIG)*0.05
SIGF=SIG*SORT(1.0)
X(1)=1.0/(4.0*A*SIGF)
X(2)=0.
X(3)=0.
DO 20 J=1,3
P(J,J)=0.
DO 30 J=1,3
30 CONTINUE
WRITE(6,31)((P(I,J),J=1,3),I=1,3)
31 FORMAT(16,5F15.8)
CONTINUE
101 CONTINUE
100 CONTINUE
103 CONTINUE
STOP
END

Listing 2
PROGRAM TO PROPAGATE EXP VAL OF X**2

DIMENSION F(7), X(5,3), Y(5,3), P(5,3,5,3)
READ (5,21) F

21 FORMAT (7F5.2)

DO 10 I=1,5
DO 10 J=1,3
DO 20 K=1,5
DO 20 L=1,3

20 P(I,J,K,L)=0.
X(I,J)=1./((A*A+SIG*SIG)
DO 30 J1=1,3
DO 25 J2=1,3
K=J2+1
L=J2+2

C THESE FORMULAE COME FROM LINE (A)
P(J1,J2,J1,J2)=P(J1,J2,J1,J2)*A*A*F(J1)*SIG**2*(J1-1)

C THESE FORMULAE COME FROM LINE (B)
P(J1,J2,J2,J1)=P(J1,J2,J2,J1)*A*A*F(J1)*SIG**2*(J1-1)
P(J1,J2,K1,J1)=P(J1,J2,K1,J1)*2.*A*A*F(J1)*SIG**2*(J1-1)
P(J1,J2,J2,K1)=P(J1,J2,J2,K1)*2.*A*A*F(J1)*SIG**2*(J1-1)

C THESE FORMULAE COME FROM LINE (C)
P(J1,J2,J2,J1)=P(J1,J2,J2,J1)*F(J1)*SIG**2*(J1-1)
P(J1,J2,J2,K1)=P(J1,J2,K1,J1)*2.*A*A*F(J1)*SIG**2*(J1-1)
P(J1,J2,J2,K2)=P(J1,J2,K2,J1)*2.*A*A*F(J1)*SIG**2*(J1-1)
P(J1,J2,J2,K3)=P(J1,J2,K3,J1)*2.*A*A*F(J1)*SIG**2*(J1-1)
P(J1,J2,J2,J3)=P(J1,J2,J2,J3)*2.*A*A*F(J1)*SIG**2*(J1-1)
25 CONTINUE

30 CONTINUE

DO 32 J1=1,5
32 WRITE (5,31) ((P(J1,J2,J3,J4), J4=1,3) , J3=1,5 , J2=1,3)
31 FORMAT (1H1 , 1P15.2)

C THE MATRIX P PROPAGATES EXP X**2

C FIRST STATE IS EXP X**2

DO 40 N=1,20
DO 50 K=1,5
DO 50 KK=1,3
Y(K, KK)=0.
DO 50 JJ=1,5
DO 50 JK=1,3
50 Y(K, KK)=Y(K, KK)+P(K, KK, J, JJ)*X(J, JJ)

IF (Y(1,1) .GE. 10.) GO TO 102
DO 60 K=1,5
DO 40 KK=1,3

40 WRITE (5,41) (J, K) , X(J, K) , A*SIG*SIG*B
41 FORMAT (1H1 , 1P15.2)

42 CONTINUE

101 CONTINUE
102 CONTINUE
100 CONTINUE

Listing 3
DO 10 IZ=1,9
   Z=FLOAT(IZ)/10.*
   ZZ=1.-7.*Z
   ZT=SQR(7Z)
DO 15 IA=1,100
   B=FLOAT(IB)/100.*
   IF(B.GE.Z) GO TO 10
   T=ATAN(7T/(Z-B))/ZT
   ANS=B*EXP((7-B)*2.)*(7Z+(Z-B)**2)
   ANS2=ANS*4.*
   ANS1=ANS*8.*
   WRITE(6,21) 5,7,T,ANS,ANS1,ANS2
10 CONTINUE
10 CONTINUE
   DO 20 IZ=3,20
      Z=FLOAT(IZ)/2.*
      ZZ=Z*2.-1.
      ZT=SQR(TZ)
   DO 30 IA=1,100
      B=FLOAT(IB)/100.*
      IF(B.GE.ZT) GO TO 20
      T=ALOG((ZT+B)/(ZT-B))/2.*ZT
      ANS=4.*RRZT*(EXP((H-ZT)*T)-EXP((H-ZT)*T)**(-2.))
      ANSI=ANS*8.
      ANS2=ANS*4.*
      WRITE(6,21) 9,7,T,ANS,ANSI,ANS2
20 FORMAT(1PE15.8)
30 CONTINUE
20 CONTINUE
STOP
END

Listing 4
BIOGRAPHICAL NOTE

David N. Martin was born in Boston, Massachusetts on August 6, 1945. He attended the public schools of Belmont, Massachusetts, graduating in 1963. He received the B.E.E. degree from Rensselaer Polytechnic Institute in 1967 and the Master of Science degree from M.I.T. in 1968 under the sponsorship of the National Science Foundation.

From 1968 through the present, Mr. Martin has been employed as an engineer at the Missile Systems Division of Raytheon Corporation and continued his graduate studies at M.I.T. on a part time basis. Under the support of Raytheon, he returned to M.I.T. in 1972 to complete the Doctoral program.

Mr. Martin is a member of Eta Kappa Nu and Tau Beta Pi honorary societies.

Mr. Martin is married to the former Leslie Sharon Holt.