NGL-22-009-124

STABILITY OF UNCERTAIN SYSTEMS

by

GILMER LEROY BLANKENSHIP

(NASA-CR-131372) STABILITY OF UNCERTAIN SYSTEMS Ph.D. Thesis (Massachusetts Inst. of Tech.) 126 p HC $8.50 CSCL 09C

N73-20260

Unclas

G3/10 67101

Electronic Systems Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

Department of Electrical Engineering

June, 1971

Report ESL-R-448

STABILITY OF UNCERTAIN SYSTEMS

by

Gilmer Leroy Blankenship

This report is based on the unaltered thesis of G. L. Blankenship, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in May, 1971. The research was conducted at the Massachusetts Institute of Technology, Electronic Systems Laboratory with support extended partially by NASA under Grant NGL-22-009-124.

Electronic Systems Laboratory
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

I

STABILITY OF UNCERTAIN SYSTEMS

by

GILMER LEROY BLANKENSHIP

S.B., MASSACHUSETTS INSTITUTE OF TECHNOLOGY
(1967)

S.M., MASSACHUSETTS INSTITUTE OF TECHNOLOGY
(1969)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
MAY 21, 1971

Signature of Author Gilmer Leroy Blankenship
Department of Electrical Engineering, May 21, 1971

Certified by ______________________
Thesis Supervisor

Accepted by ______________________
Chairman, Departmental Committe on Graduate Students

STABILITY OF UNCERTAIN SYSTEMS

by

GILMER LEROY BLANKENSHIP

Submitted to the Department of Electrical Engineering on May 21, 1971, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

This research is concerned with the asymptotic properties of feedback systems containing uncertain parameters and subjected to stochastic perturbations. The approach is functional analytic in flavor and thereby avoids the use of Markov techniques and auxillary Lyapunov functionals characteristic of the existing work in this area. The results are given for the probability distributions of the accessible signals in the system and are proved using the Prohorov theory of the convergence of measures and some recent work on the preservation of convergence under operations. For general nonlinear systems a result similar to the Small Loop-Gain Theorem of deterministic stability theory is given that is sufficient to guarantee that totally bounded stochastic inputs give rise to totally bounded outputs. Here boundedness is a property of the induced distributions of the signals and not the usual notion of boundedness in norm. For the special class of feedback systems formed by the cascade of a white noise, a sector nonlinearity, and a convolution operator conditions are given to insure the total boundedness of the overall feedback system. These conditions are expressed in terms of the Fourier transform of the convolution kernel, the sector parameters of the nonlinearity, and the mean and the variance parameters of the noise. Their form is reminiscient of the familiar Nyquist Criterion and the Circle Theorem for deterministic systems. Applications of the criteria to analyze rounding errors in machine computations and to study control systems containing human operators are suggested.

THESIS SUPERVISOR: Sanjoy K. Mitter
TITLE: Associate Professor of Electrical Engineering

ACKNOWLEDGEMENTS

It is a considerable pleasure to acknowledge the contributions of my teachers and friends to this work and my education.

First I would like to express my sincere appreciation to Professor Sanjoy Mitter for his patient and scholarly guidance of this work during the past year. His encouragement and confidence have been essential elements in the completion of this thesis.

I would like to thank Professor Jan Willems who served as my thesis advisor during the first of two years spent on this research and as a reader for the second year. It was he who suggested this topic and kindled my enthusiasm for it.

Professor Arthur Baggeroer served as the other reader and I am grateful for his comments and criticisms. There is no one with a finer eye for detail.

I would like to thank Professor Roger Brockett at Harvard for his spirited comments on this research and other aspects of my education. His remarks have strongly influenced my attitude toward research in systems theory.

I have benefitted from many stimulating discussions on this research and others with several of my colleagues in the Electronic Systems Laboratory, notably Keith Glover, Jon Davis, and Dimitri Bertsekas.

Finally, this work was done under a research assistantship supported by the MIT Lincoln Laboratory, and was also supported by the National Aeronautics and Space Administration under Research Grant No. NGL-22-009-124.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Stability of Dynamical Systems:

The study of dynamical systems has evolved along two paths essentially distinct in mathematical formulation. The first, which is based in the theory of differential equations, uses the concept of a dynamical system as a semigroup of states and thus has an algebraic flavor. For autonomous systems (no forcing function) this approach was already well formulated fifty years ago [8]. For physical systems accurately described by a finite dimensional set of states which have interpertations as physical variables (electrical voltages and currents, for example) powerful and precise conclusions may be drawn about the properties of the system. However, when the physical system admits no accurate finite dimensional model, the general state theory is at this time rather formal and, except in specific cases, the precision attained in the finite dimensional case is lost in technical difficulties.

The use of dynamical systems as models for control processes has led to a second method of analysis based simply on the input-output properties of the system. In this formulation the input and output of a system are considered as points in a set of functions and the system itself as an operator on this function space. Thus, functional analysis replaces the theory of differential equations as the source of analytic tools. Problems associated with selecting a suitable representation for the internal structure of a dynamical element are avoided and large classes of complex systems may be treated qualitatively with simple techniques.

Originating only within the past decade, the operator theoretic treatment of systems has been developed only for the easiest problem associated with feedback systems-stability. Restricting the set of inputs to be perturbations of the system, that is bounded in some sense, a system is defined to be input-output stable if bounded inputs are mapped into bounded outputs or equivalently if the system is represented as a bounded operator. In this context boundedness of a signal may mean the usual boundedness in amplitude or in some more sophisticated sense such as total energy or power. In the state theory stability is defined as asymptotic convergence of the system state to the zero state. Perturbations are introducted by initial displacements of the state from zero. For those systems permitting a simple state representation it is usually easy to commute between the concepts of input-output stability and state stability [63].

Stability theory in the state space setting relies on the use of Lyapunov functionals, certain auxillary functions of the state. These functionals completely specify the asymptotic behavior of the state when they can be found and determined to be positive definite and have negative definite time derivatives in a neighborhood of an equilibrium state. As there at present exists no constructive method of generating Lyapunov functions, the general theory remains in a static condition at present.

By contrast the operator theoretic approach to stability casts the problem into a very active area of mathematical research-the invertibility of operators. To see that this is the case, consider the equation

$$x + KGx = u$$

as the description of a feedback system. Here K and G represent generally nonlinear elements in the feedback loop, u is a perturbation input, and the output x is to be studied. If u is an element of some normed, linear space of functions, then x is bounded (an element of that space) if $I + KG$ has a bounded inverse on that space. Hence, the stability question may be resolved using the mathematical theories relating to the invertibility of operators on normed or metric spaces. Indeed many new as well as some familiar results have been developed using spectral theory and Banach algebras, two of the basic tools in invertibility studies.

It is the presence of an active and well-founded theory for the analysis of deterministic systems in operator form that motivates this research which attempts to extend the theory in such a way as to preserve its essential elements and yet account for stochastic signals and uncertain parameters in the analysis.

1.2 Stochastic Systems:

Efforts to model increasingly complex control systems have led to the study of some systems which simply cannot be modelled accurately with perfect certainty. Uncertainties are introduced either by phenomena that are so complicated as to defy reduction to a tractable deterministic model or are in essence random. As an example of the former consider the generation of roundoff errors in a digital computation. Restricted by finite register size the machine must of necessity round-off stored variables at each stage in a computation. Being a design choice the rounding mechanism is not uncertain, and in any given computation of limited complexity the rounding errors could be monitored exactly. However, in a computation of even moderate complexity the register size will be exceeded at many points in the calculation and monitoring the errors may become a more formidable task than the original computation. In such a case it is reasonable to assume that the evolution of rounding errors is a statistical process in order to appraise their average magnitude.

As an example of the introduction of essentially random phenomena into a physical experiment consider the problem of maintaining the orientation of a rigid body in orbit around the earth. Primary sources of error are sensor errors and propulsion jet errors (in firing and cutoff times). A secondary source of error,but a very important one in very precise applications, is the fluctuations in the earth's gravitational field along the path of the orbit due to surface irregularities and local variations in the density of the earth. Because the sensor errors make an exact determination of position impossible no model apart from a statistical one can accurately (within the usually rigid specifications of these experiments) account for other than the most prominant aberrations. This problem reduces to design of a feedback control law capable of precisely orienting a satellite in the presence of essentially random perturbations. Moreover, the controllers (combining sensors and propulsion units) are themselves subject to stochastic errors that cannot be deterministically approximated within the tolerances fixed for these projects. It is therefore appropriate in a general analysis of systems subjected to uncertainties to consider not only random external perturbations but to permit random parameter variations as well.

One of the major problems faced at the outset of an analysis of a stochastic system is to determine accurate probability distributions for the quantities considered as random in the experiment. In general some method of hypothesis testing must be applied to the available data and distributions deduced from this procedure. Although the possibility of several empirical distributions is permitted in the definitions of a stochastic system in section 3.1 below, in the major portions of the analysis to follow it is assumed that the process of likelihood testing has been completed and that an optimal distribution has been selected. For an interesting and important alternate approach for optimal control problems see the papers and thesis of Witsenhausen [67],[68], and [69].

Following the pattern observed in deterministic systems theory, the first problem to be considered for stochastic systems was stability. Moreover, the framework was that of a state space formulation using Lyapunov like techniques. The reasons in both cases were compelling. First, problems like optimal control of stochastic systems must proceed in two intimately connected steps. Because the state of the system in most cases may be observed only in the presence of uncertainties, it must first be estimated. Only then may optimal controls be selected. See for instance the work of Kushner [48], Wonham [70], Fleming [24],[25], and Benes [3],[4] for discussions of the problems arising from restricted information on the state.

The reason for studying stochastic systems with a state realization is motivated by the powerful and comprehensive mathematical apparatus available for the analysis of Markov processes (see for instance Dynkin [19]). Assuming no more than causality, any system may be shown to have a Markovian state decomposition (see Willems [63] for a similar theorem which may be easily extended to permit stochastic variables), and for those systems with a finite dimensional state space the analytic theory of Markov processes combined with the theory of stochastic differential equations completely determines the system behavior. Using potential functions of the state (like Lyapunov functionals), the stability of a stochastic system with finite dimensional state may be completely determined. This program is developed comprehensively in Kushner's book [48].

However, in contrast to the deterministic case there is a very real confusion over the meaning of stability in a stochastic system. The confusion stems largely from the numerous distinct varieties of probabilistic convergence available. Thus, almost sure convergence, convergence in n^{th} moment, convergence in probability and others have been used to study the asymptotic properties of perturbed stochastic systems. However, for systems defined by stochastic differential equations it is straight forward to commute between these equations for the trajectories (samples) of the signals in the system and the Chapman-Kolmogorov equations for the distributions of the state and the Fokker-Planck equation for its density function (see Ito [40]).

By examining the asymptotic properies of the solution of the Fokker-Planck equation, those of the state may be completely determined. It is in fact entirely appropriate to regard the density function as the state of the system and to describe the behavior of the system in terms of its evolution. In this manner Markovian stochastic systems form an important class of distributed parameter systems (systems whose state satisfies a partial differential equation)-a class somewhat more amenable to analysis than most because of its special nature (particularly the boundary conditions) and the additional interpertation afforded by probabilistic considerations and the differential equation representation. Important work within this interpertation has been dome by Kushner [47], Dym [17], Elliot [22], Il'in and Khas'minskii [38] on stability and Fleming [25] on control.

In the setting provided by a state realization of a stochastic system the natural way to examine the asymptotic properties of the state is to introduce Lyapunov functionals of the state and consider their properties. This has been the approach adopted in almost all of the references mentioned above. Because of certain relationships between Markov processes and potential theory (Meyer [51], Hunt [37], Doob [13]) which seem to account for the restrictions imposed on Lyapunov functionals, the subject is deserving of further study. For example stochastic Lyapunov functions were observed by Kushner [49] and Bucy [10] to be positive supermartingales [51]. However, a supermartingale is a potential subject to certain restrictions [51]. See Dynkin [20] for a discussion of the position of harmonic functions and potentials in the analytic theory of Markov processes and comments on the construction of harmonic functions for a process.

What one hopes would come of an investigation of these relationships is a procedure for generating Lyapunov functionals for interesting systems. At present the obstacle encountered in the deterministic Lyapunov theory is present in the stochastic setting: that is, there exists no systematic method in general of constructing the functionals. Moreover, in specific cases the construction process is far more difficult in the latter case (stochastic systems) because of certain technical aspects of the Markovian structure [48]. For instance deterministic Lyapunov functions must satisfy a first order partial differential inequality constraining the time derivative of the functional to be negative definite. In the stochastic case this inequality involves a second order operator [48,p.39].

Clearly an alternative approach for the analysis of the asymptotic properties of stochastic systems is desireable. The development of such an alternative along the lines of the operator theoretic stability theory is the subject of this research.

Continuing the analogy with the deterministic theory it would seem desireable to have available a kind of "probabilistic functional analysis" so that the input-output results of the deterministic theory may be easily rederived in a probabilistic setting. Such a mathematical theory is available, due largely to a group of Czechoslovak mathematicians headed by Spacek and Hanš [31], [32], [33]. The concepts of random operators equations defined in those papers are presented here in section 2.3 and used in section 3.1 to prove some moment bounds for the signals in a general stochastic system. It is important to note that these bounds are obtained for signals which need not be Markov processes.

However, it is only in combination with another recent collection of work in the general theory of probability that this formulation of a stochastic system as a random operator is able to yield results in terms of the distributions of the processes involved. This work is concerned with topologies for random processes.

Though introduced by Kolmogorov over thirty years ago, the study of the convergence of probability distributions has only recently returned to popularity. The papers of Prohorov [53] and Skorokhod [56] in 1956 were instrumental in generating this revival of interest. Since that time the study of topologies for random processes has evolved in a series of papers summarized and extended in the books of Billingsly [7], Parthasarathy [52], and Topsøe [60]. The basic ideas are the following: for any metric space (X,d) let PM(X) be the set of probability measures on X, then PM(X) may be regarded as a subset of the dual space of BC(X), the bounded, continuous functionals on X [16]. A natural weak topology is then induced on PM(X), and it is this topology that is suitable for determining the distributions of functions of a random process (see section 3.1 for further motivational discussions of this point and [29, Chapter IX]).

A key point in the analyis of convergence of distributions is a description of the compact subsets of PM(X). Under certain conditions on the basic space X a set of distributions is relatively sequentially compact (has sequentially compact closure) if and only if there exists a compact subset of X on which the distributions are concentrated. That is, let $\Lambda \subset$ PM(X) be the subset under consideration, then Λ is relatively compact if for every $\alpha \in (0,1)$ there exists a compact subset $K(\alpha)$ of X such that $\mu[K(\alpha)] > 1 - \alpha$ for every $\mu \in \Lambda$. If X is a space of functions, suitably metrized, the result says that the distributions of the stochastic process taking its values in this set of functions are relatively compact if and only if the values of the process are in a compact set almost surely. This recurrence condition is familiar in ergodic theory and in a sense indicates the possibility of interpertations in that setting.

By assuming X to be the space of continuous functions or piecewise continuous functions, the compact subsets of X may be easily characterized. Sufficient conditions may then be established to assure relative compactness of a set of distributions defined on X. These are summarized in section 2.4 for continuous functions and in section 3.3 for piecewise continuous functions. These conditions are used in sections 3.2 and 3.3 to prove that feedback systems subjected to inputs with relatively compact distributions give rise to outputs which also have relatively compact distributions. In section 3.4 these results are used to analyse the behavior of systems described by stochastic differential equations subjected to input processes in this class.

Implicit in these proofs (3.2 and 3.3) and explicit in section 3.1 is the transformation of weakly convergent sequences of distributions by operators. That is, a key point in the analysis is contained in the question: if a convergent sequence of distributions is mapped by an operator (in some well-defined manner) into another sequence under what conditions on the operator is the latter sequence convergent as well? Finding these conditions forms the heart of the arguments in Chapter 3. The general results that indicate the line of proof were developed by Billingsly [7] and Topsøe [61] among others. These conditions are essentially continuity of the operator on the underlying space X, and in this sense relate back to the operator stability theory of deterministic systems where continuity of the system as a map on a function space is a central concept of stability. It is further in this way , since the feedback equation defines the variable of interest implicitly, that the mathematical theory relating to invertibility of operators is once again identified as a crucial aspect of the framework for the analysis of the asymptotic properties of systems, in this instance stochastic in nature.

CHAPTER 2

MATHEMATICAL PRELIMINARIES AND BACKGROUND MATERIAL

2.1 Remarks and Some Notation:

The purpose of this chapter is to recall some of the basic notions in the operator theoretic treatment of feedback systems and to summarize those aspects of the theory of the convergence of probability measures used in Chapter 3. Although the summaries here are rather concise, appropriate references are given where more thorough treatments may be found. As used here, only the most basic results from each of these theories is required and in this sense the background material necessary for the derivations in Chapter 3 is minimal. The only new results in this chapter are a modification of the usual definition of a random operator and a result on the effect of such operators on convergent probability distributions (section 2.4).

Though most of the notation and definitions from mathematics used here are standard, a few conventions may be unfamiliar. Symbols such as $R = (-\infty,\infty)$, $R^+ = [0,\infty)$, and Z for the set of integers are standard and are freely used. The notation $C(R^+;R)$ indicates the set of real-valued continuous functions on R^+ and is typical of the form used to designate function spaces. Other common notations are:

(i) $(L_p(R^+), ||\cdot||_p)$ the Lebesgue spaces on R^+;

(ii) $(\Omega, \mathcal{F}, P)$ a basic probability space;

(iii) $(X, ||\cdot||)$ a normed, linear space;

(iv) $(x, \mathcal{B}(x))$ a Borel measureable space, $\mathcal{B}(x)$ the Borel σ-algebra of x;

(v) $F(\Omega;X)$ the set of X-valued random variables on Ω;

(vi) $PM(X)$ the set of probability measures on X;

(vii) $BC(X)$ the set of bounded, continuous (real-valued) functionals on X;

(viii) $\mathcal{S}(X)$ the set operators mapping $X \to X$, and

(ix) $\{\pi_t\}_{t \epsilon R^+}$ the set of truncation operators on some function space.

Operators on sets of functions are usually denoted by F, G, H or some other upper case letter. These points are representative of the standard conventions used here.

As a consequence of the mixture of engineering material and some mathematics a few compromises in notation have been necessary. For example the symbol P is reserved for the basic probability measure on Ω, and so $\{\pi_t\}$ is used to denote a set of truncation operators on functions - usually denoted by $\{P_t\}$ in the engineering literature (see for instance Willems [64]).

The terms "stochastic process," "random process," and "function valued random variable" are to be considered equivalent here. The concept of a random variable as a measureable function is used, and when the random variable takes its values in some set of functions, one of the above terms is used to indicate this case. The concept of a stochastic process as an "indexed set of random variables" [29] is not used. The words distribution and probability measure are used interchangeably and should be considered equivalent. Thus, the more common meaning of the former term is never employed here.

Finally, real-valued functions are used almost exclusively in this work, though it is acknowledged that nearly all the results are true for R^n-valued processes. The methods used in the paper [66] to extend the theorems there to the multi-dimensional case may be applied to the results here at the cost of some complication of the notation. The only exception to this voluntary restriction to real-valued functions occurs in section 3.4 where some earlier work is summarized and compared to that given here. The concept of state is fundamental in the differential equation formulation used in the earlier work, and so multi-dimensional variables must be used for the state to thoroughly illustrate the theorems.

2.2 Input-Output Stability of Deterministic Feedback Systems:

In this section a brief summary of the operator-theoretic analysis of feedback systems is presented. The purpose here is to recall a familiar class of problems and indicate an appropriate framework for their analysis. The concept of a feedback system as an operator on function spaces is introduced and stability of the feedback system defined in terms of the continuity properties of the operator. Appropriate references are the original papers of Sandberg [54], and Zames [74], and the papers [63] [65] and monograph [64] of Willems. The thesis of Davis [12] gives a rather complete treatment of the input-output theory of general linear systems.

Let X_e be a vector space of V-valued functions on the set R^+; that is, each element x of X_e maps R^+ into V where V is some given vector space. Let G be an operator mapping X_e into itself such that $G0 = 0$. For $u \in X_e$ as an "input" consider the following equation

$$(1) \qquad x(t) + (Gx)(t) = u(t), \quad t \in R^+,$$

as descriptive of a deterministic feedback system. The operator G represents the cascade of all the elements of the open-loop system and x the "feedback error." As a model of physical elements the operator G must be causal and the solution x must be bounded over finite time intervals (bounded subsets of R^+). These requirements are made precise using the truncation operators $\{\pi_t\}$ defined by

$$(\pi_t x)(s) \triangleq \begin{cases} x(s) & s \leq t \\ 0 & s > t \end{cases} \quad ; \quad s, t \in R^+.$$

Assume that X_e is closed under these truncations. The operator G is <u>causal</u> if (pointwise)

$$\pi_t Gx = \pi_t G \pi_t x, \quad t \in R^+, \quad x \in X_e .$$

Assume that X_e has a normed subspace $(X, ||\cdot||)$ and that the truncations are the projections, $\pi_t : X_e \to X$ for every t. The existence of a locally bounded solution to (1) is assumed in the following fashion: every (input) element u of X_e gives rise to a unique (output) element x of X_e such that

$$(\pi_t x)(s) + (\pi_t \, G(\pi_t x))(s) = (\pi_t u)(s) \quad ; \quad s, t \in R^+ .$$

Thus, by the projection property of π_t the function $\pi_t x$ is an element of X and is thus bounded for every $t \in R^+$.

Having assumed the existence of solutions bounded over finite intervals, the system is said to be stable if bounded (on R^+) inputs give rise to outputs bounded over the entire time set. Or precisely:

<u>Definition 1</u>: The feedback system (1) under consideration is said to be <u>X-stable</u> if for any $u \in X$ the conditions hold:

(i) the solution x corresponding to u is actually an element of X,

and (ii) there exists a $K < \infty$ independent of $u \in X$ such that $||x|| \leq K||u||$.

The nature of the definition is clarified by the following restatement:

<u>Theorem 2</u>: [65, Theorem 4.1] Assume that I + G has a causal inverse on X_e, then a necessary and sufficient condition that (1) be X-stable is that $(I + G)^{-1}$ be bounded on X.

The theorem indicates clearly that the correct mathematical framework for the investigation of input-output stability is in that theory relating to the invertibility of (causal) operators. For example if the operator G is linear then the invertibility of I + G requires that -1 not be an element of the spectrum of G [12]. For linear, time-invariant convolution operators on several Banach spaces [11] the spectrum of G is the set (assuming $g \in L_1(R^+)$)

$$\sigma(G) = \bigcup_{Re(s)\geq 0} \int_0^\infty e^{-st} g(t)dt \ .$$

The stability condition on G in this case is the familiar Nyquist Criterion.

For the abstract equation (1) the need for general invertibility criteria led to the following theorem whose proof was perhaps initially motivated by some similar inequalities in the theory of Banach algebras (see Bachman [1, p. 34]).

Theorem 3: (Small-Gain Theorem) For the equation (1) under the existence and causality assumptions suppose that G is a contraction on X; i.e. there exists a constant $\alpha < 1$ such that

$$\sup_{f \in X} \frac{||Gf||}{||f||} \leq \alpha < 1 .$$

Then for any $u \in X_e$ the inequality

$$||\pi_t x|| \leq (1-\alpha)^{-1} \, ||\pi_t u||$$

holds for every $t \in R^+$. Hence, $u \in X$ implies that $x \in X$ and $||x|| \leq (1-\alpha)^{-1}||u||$, and so, that (1) is X-stable.

The power of this simple and obvious result is only fully realized in its special cases, one of which is the Circle Theorem, a striking generalization of the Nyquist Criterion. Let the vector space V be R and define the (nonlinear) operator G as

$$(Gx)(t) = \int_0^t g(t-s)f(s,x(s))ds$$

where $f : R^+ \times R \to R$ is continuous (separately) and the kernel g is locally $L_1(R^+)$ (absolutely integrable on finite intervals). Assume that the feedback equation $u = x + Gx$ is well-posed (has a unique solution) on the space $X_e = L_{\infty e}$ (the extended space with normed subspace $(L_\infty(R^+), ||\cdot||_\infty)$).

Theorem 4: (L_∞-Circle Theorem [75]) Assume the following:

(i) $u \varepsilon (L_\infty(R^+), ||\cdot||_\infty)$.

(ii) For some $r_o > 0$

$$\int_0^\infty e^{r_o t} |g(t)| dt < \infty \quad .$$

(iii) For some constants a, b εR^+

$$0 \leq a \leq \frac{f(t,x)}{x} \leq b < \infty, \text{ for every } t \ \varepsilon \ R^+, x \ \varepsilon \ R \ .$$

(iv) For $\hat{G}(s)$ the Laplace transform of g, and some $r \ \varepsilon \ (0,r_o)$ the exclusion holds

$$\{-[\tfrac{1}{2}(a+b)]^{-1}, j0\} \notin \bigcup_{Re(s) \geq -r} \hat{G}(s) \ .$$

(v) For some $r \ \varepsilon \ (0,r_o)$

$$\sup_{\xi \varepsilon R} |\hat{G}^{-1}(-r+j\xi) + \tfrac{1}{2}(a+b)| > \tfrac{1}{2}(b-a) \ .$$

Then $x \ \varepsilon \ L_\infty(R^+)$ and $||x||_\infty \leq K||u||_\infty$ for some $K < \infty$ independent of u.

Remark 5: Conditions (iv) and (v) mean that the r-shifted Nyquist locus of G does not encircle (iv) or intersect (v) the closed disc in the complex plane centered at $\{-[\tfrac{1}{2}(a+b)]^{-1}, j0\}$ with radius $\tfrac{1}{2}(b-a)$.

The theorem is valid on, for instance, $L_2(R^+)$ with $r_o = 0$; however, in the L_∞ version to be used here (Theorems 3.2.3, 3.3.6) the assumption of "decaying memory" (ii) for the convolution seems necessary in the proof [75]. Note that if a = b the theorem reduces to the Nyquist Criterion which is necessary as well.

In section 3.1 a theorem similar to the Small-Gain Theorem is used to establish general conditions for the asymptotic invariance of the probability distribution of the solution of a general random equation. In sections 3.2 and 3.3 conditions like the Circle Theorem and the Nyquist Criterion are used to guarantee asymptotic invariance for the solutions of random convolution equations. Before proving this result it is necessary to describe precisely the structure of a random operator equations, and introduce a topology suitable for the analysis of probability dsitributions induced by random processes. These topics are discussed in the next two sections.

2.3 Random Operator Equations:

A. Probability spaces:

In this section the concept of a random operator as a model of a physical element with random parameters is rendered precise by defining it to be an operator valued random variable. Certain properties of random operators are noted and the nature of random operator equations investigated. Appropriate references for this section are the papers of Hanš [31], [32], [33] and the survey of Bharucha-Reid [6].

Let $(\Omega, \mathcal{F}, P)$ denote a basic probability space. When this triple occurs in the sequel, the assumptions below will be implicit:

(i) (Ω, τ) is a topological space, always separable,* τ denotes the topology of the set Ω.

---

*See [7, Appendix III] for the implications of this constraint.

(ii) $\mathcal{F}$ is the Borel σ-algebra generated by the topology τ. That is, the least sub-class of 2^{Ω} (the class of all the subsets of Ω) closed under countable unions, finite intersections, and containing τ.

(iii) P is a probability measure, by definition a complete (subsets of sets of measure zero have measure zero), countably additive ($P\{\bigcup_{i=1}^{\infty} A_i\} = \sum_{i=1}^{\infty} P(A_i)$, $A_i \in \mathcal{F}$, $A_i \cap A_j = \phi$, $i \neq j$), finite ($P(\Omega) < \infty$), set function mapping $\mathcal{F}$ into R^+, normalized so that $P(\Omega) = 1$.

For any measureable space $(X, \mathcal{B}(X))$, here $\mathcal{B}(X)$ indicates the Borel σ-algebra of X, let $F(\Omega;X)$ denote the set of X-valued random variables on Ω; that is, the set of functions $f : \Omega \to X$ and f is measureable in the sense that $f^{-1}\mathcal{B}(X) \subset \mathcal{F}$, or that the inverse image of every measureable set is measureable.

Example 1: (Gaussian measure) Let $(X, \mathcal{B}(X)) = (R, \mathcal{B}(R))$ the real line with $\mathcal{B}(R)$ generated by the open intervals of R. Let $f : R \to R$ be a continuous function (hence $\mathcal{B}(R)$ measureable) and assume that the measure μ_f is defined by

$$\mu_f(A) = P\{\omega \in \Omega : f(\omega) \in A \in \mathcal{B}(R)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 1_A(x) e^{-x^2/2} dx \, ^*$$

Then $(R, \mathcal{B}(R), \mu_f)$ is a probability space and f is a Gaussian random variable.

---

*1_A denotes the characteristic function of the set A.

Example 2: (Wiener measure) Let $(X, \mathcal{B}(X)) = (C(R^+), \mathcal{B}(C))$ where $C(R^+)$ is the set of real-valued continuous functions on R^+ topologized by uniform convergence on compact intervals. $\mathcal{B}(C)$ is the least σ-algebra containing sets of the form

$$A(t;a,b) = \{f \in C : f(t) \in [a,b) \subset R\} \quad , \quad t \in R^+ .$$

A measure P is induced on $\mathcal{B}(C)$ by its definition on such sets A.

$$\begin{aligned} P\{A(t;a,b)\} &= P\{f \in C : f(t) \in [a,b) \mid f(r)\ r \leq s < t\} \\ &= \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} \int_a^b e^{-[x-f(s)]^2/2\sigma^2(t-s)}\,dx \end{aligned}$$

P is in this instance the Wiener measure [7]. Note that for $s = 0$, the assumption $f(0) = 0$ is standard.

B. Random Equations:

The following definition was given by Hans [31].

Definition 3. Let $(\Omega, \mathcal{F}, P)$ and $(X, \mathcal{B}(X))$ be given, then a map $T : \Omega \times F(\Omega;X) \to X$ is a random operator if $T(\cdot,x(\cdot)) = y(\cdot)$ is a random variable (an element of $F(\Omega;X)$).

Example 4: (Deterministic operators) Let G be a continuous map X into X. Then it is routine to verify that $G : F(\Omega;X) \to F(\Omega;X)$ and that every continuous deterministic operator is a random operator according to Definition 3.

Example 5: (Linear convolution) Let $(X, \mathcal{B}(X)) = (C(R^+), \mathcal{B}(C))$ and let $g \in C(R^+)$. Let w denote the Wiener process, and $x \in F(\Omega;C)$

independent of w. Then the C-valued function y (on Ω) defined by (its value at t)

$$y(t,\omega) = \int_0^t g(t-s)x(s,\omega)dw(s,\omega) , \quad t \in R^+ \quad \omega \in \Omega$$

is an element of $F(\Omega;C)$. The integral on the right is defined in the Itô sense; its properties and a proof of the assertion here are given in Itô [40]. The convolution above defines a linear random operator on the space of C-valued random variables.

An alternate formulation of the notion of a random operator may be given as follows: Let (X,d) be a separable metric space and $\mathcal{G}(X)$ the set of all continuous maps $X \to X$. Give to the set $\mathcal{G}(X)$ the (strong) topology τ generated by the convergence $G_n \xrightarrow{\tau} G$ if and only if

$$d(G_n x, Gx) \xrightarrow{\tau} 0 \quad \text{for every } x \in X .$$

Let $\mathcal{B}(\mathcal{G})$ denote the Borel σ-algebra of subsets of $\mathcal{G}(X)$ generated by this topology. Then for any $(\Omega, \mathcal{F}, P)$ given, let $F(\Omega;\mathcal{G})$ denote the set of $\mathcal{G}$-valued random variables. That is, each element $\tilde{G}$ of $F(\Omega;\mathcal{G})$ maps Ω into $\mathcal{G}(X)$ such that $\tilde{G}(\omega)[.] = G_\omega(.) \in \mathcal{G}(x)$. Thus, for every $\omega \in \Omega$ $\tilde{G}(\omega)[\cdot]$ is a continuous map $X \to X$; and so, this definition coincides with Definition 3 on the continuous operators. Moreover, it is clear that probability distributions may be introduced on $\mathcal{B}(\mathcal{G})$ and convergence arguments made for random operators as well as for random variables in the usual manner. In the next section this possibility is investigated further and the preservation of probabilistic convergence under random operations discussed.

The use of random operator equations in section 3.1 necessitates a discussion of the nature of a solution to such an equation.

Definition 6 : For $(\Omega, \mathcal{F}, P)$, $(X, \mathcal{B}(X))$, $\tilde{G} \in F(\Omega; \mathcal{B}(X))$, and $y \in F(\Omega;X)$ given, then every element x of $F(\Omega;X)$ satisfying

$$P\{\omega : \tilde{G}(\omega)[x(\omega)] = y(\omega)\} = 1$$

is a solution of the equation $\tilde{G}x = y$.

Thus, a solution is required to be a random variable; that is, to have certain measureability properties. This qualification has been the source of a considerable amount of research on the nature of solutions to random equations (see for instance Hanš [32], Bharucha-Reid [6]). Most of this has been a consequence of the ambiguous nature of Definition 3.

Assume that $(X, ||\cdot||)$ is a Banach space. An element G of $F(\Omega;\mathcal{B}(x))$ is said to be a random contraction if there exists a real-valued random variable c such that $c(\omega) < 1$, for every $\omega \in \Omega$, and

$$||\tilde{G}(\omega)[x_1] - \tilde{G}(\omega)[x_2]|| \leq c(\omega)||x_1-x_2|| \quad .$$

The analog of the Banach-Cacciopoli fixed point theorem [42, p. 627] in this setting is:

Theorem 7: [33] Let $(X,||\cdot||)$ be a Banach space, $\tilde{G} \in F(\Omega; \mathcal{B}(X))$ a random contraction, then there exists an element x of $F(\Omega;X)$ such that

$$P\{\omega:\ \mathcal{G}(\omega)[x(\omega)] = x(\omega)\} = 1 \quad .$$

The random variable x is unique almost everywhere (P), and may be obtained by the process of successive approximations starting at any initial element x_o of $F(\Omega;X)$.

This result is then the basis of the proof of existence and uniqueness of solutions of random operator equations. For the equation $\tilde{G}x = y$, given $\tilde{G} \in F(\Omega; \mathcal{J}(X))$ and $y \in F(\Omega;X)$ if $\tilde{G}$ may be shown to be a contraction, then Theorem 7 may be invoked to assure the existence of a unique element of $F(\Omega;X)$ (as a set of equivalence classes under P) as the solution. Moreover, the classical scheme of Picard iterations may be used to approximate the solution. This is a result of somewhat more subtlety than is apparent at first reading as it implies that the Picard iterates are at each step random variables, and they approach almost surely a random element which is the desired solution. In most cases of course only local existence and uniqueness may be established in this manner.

C. Moment spaces:

As the convergence arguments used in Chapter 3 utilize certain moment bounds, it is appropriate at this point to introduce a few definitions of "moment spaces" and consider operators on these spaces. Let $(\Omega,\mathcal{F},P)$ and $(X, \mathcal{B}(X),||\cdot||)$ be given and denote by $E(\cdot)$ the usual expectation operator on the subset of $F(\Omega;X)$ for which

$$Ex = \int_{\Omega} x(\omega)P(d\omega)$$

is well-defined as a Bochner integral [72, p. 219].

In particular define the sets (of equivalence classes)

$$\mathcal{L}_q(\Omega;X) \triangleq \{x \in F(\Omega;X) : \ |x|_q = (E\{||x||^q\})^{1/q} < \infty; \quad q \in [1,\infty)\}.$$

And in the case that $(X, ||\cdot||)$ is a Banach space of real-valued functions on R^+, the spaces

$$\mathcal{E}_q(\Omega;X;\ell) = \{x \in F(\Omega;X) : \ ||x||_{q,\ell} = (\ell[E\{|x(t,\omega)|^q\}])^{1/q} < \infty \ ; \ q\epsilon[1,\infty)\}.$$

Here ℓ is any sub-additive linear functional* on real-valued functions. Typical examples used here are

$$\ell_1(f) = \operatorname*{ess\,sup}_{t\epsilon R^+} |f(t)|$$

$$\ell_2(f) = \int_0^\infty |f(t)|dt$$

Under these restrictions on ℓ it is clear that $||\cdot||_{q,\ell}$ is a norm and $(\mathcal{E}_q, ||\cdot||_{q,\ell})$ a normed linear space. Under the choices ℓ_1 and ℓ_2, $\mathcal{E}_q$ is a Banach space as well, Thus, elements of $\mathcal{E}_q(\Omega;X;\ell_1)$ are (almost surely) bounded in q^{th} absolute moment. Elements of $\mathcal{E}_q(\Omega;X;\ell_2)$ have absolutely integrable q^{th} moments. See Itô [39] or Skorokhod [55] where similar spaces are defined and used in existence arguments for stochastic differential equations.

Assuming that $(X, ||\cdot||)$ is a space of functions closed under the truncation operation (π_t, see section 2.2), the "extended

---

* $\ell(x+y) \leq \ell(x) + \ell(y)$, $\ell(ax) = |a|\ell(x) \quad x,y \in R^+, a \in R$.

spaces" $\mathcal{E}_{q,\ell} = \{x \in F(\Omega;X) : \pi_t x \in \mathcal{E}_q,\ t \in R^+\}$ are convenient for certain statements pertaining to the existence of solutions in feedback equations.

Let $\mathcal{G}(X)$ denote again the set of operators mapping X into itself continuously. For those elements G of $F(\Omega;\mathcal{G})$ for which the supremum is finite define

$$|G|_q = \sup_{\substack{x \in \mathcal{L}_q \\ |x|_q \neq 0}} \frac{|Gx|_q}{|x|_q}$$

And under the assumption that X is a function space and $\mathcal{G}(X)$ consists of causal operators (see section 2.2), then for $G \in F(\Omega;\mathcal{G})$ set

$$||G||_q = \sup_{\substack{x \in \mathcal{E}_q \\ ||x||_q \neq 0}} \frac{||Gx||_q}{||x||_q} .$$

Note that in this case $||G||_q$ depends on ℓ.

A few examples are given below to illustrate the definitions.

Example 8: Consider the space $\mathcal{E}_2(\Omega;X;\ell_1)$ and the (deterministic) operator G on $X = C(R^+)$, the continuous functions, given by

$$(Gx)(t,\omega) = y(t,\omega) = \int_0^t g(t-s)x(s,\omega)ds$$

Then

$$Ey^2(t) = \int_0^t \int_0^t g(t-s)g(t-r)E\{x(s)x(r)\}dsdr$$

$$\leq (\int_0^t g(t-s)(E\{x^2(s)\})^{1/2}ds)^2$$

Hence,

$$||y||_{q,\ell_1} \leq \int_0^\infty |g(t)|dt \; ||x||_{q,\ell_1} ,$$

and the bound is attained.

Example 9: Consider $\mathcal{E}_2(\Omega;X;\ell_2)$ and the operator G on $X = L_2(R^+)$, the space (deterministic) of square-integrable real-valued functions. Then, for $y = Gx$,

$$y(t,\omega) = \int_0^t g(t-s)x(s,\omega)ds \; , \; t \in R^+, \quad \omega \in \Omega$$

where

$$\int_0^T E\{y^2(t)\}dt = \int_0^T E\{(\int_0^t g(t-s)x(s,\omega)ds)^2\}dt$$

$$\leq \int_0^T \int_0^t |g(r)|dr \int_0^t |g(t-s)|E\{x^2(s)\}ds \; dt$$

$$\leq (\int_0^\infty |g(t)|dt)^2 \cdot ||x||_{2,\ell_2}$$

This bound, however, may be improved by taking into account the fact that for each $\omega \in \Omega$ the integral $\int_0^\infty |x(t,\omega)|^2 dt < \infty$. Hence, each sample function $x(\omega)$ admits a Fourier transform,

$$\hat{x}(j\nu,\omega) \approx \int_0^\infty x(t,\omega)e^{-j\nu t}dt \quad , \omega \in \Omega \quad \nu \in R .$$

Here equality holds in the $L_2(R^+)$ sense. Assuming that $g \in L_2(R^+)$ has a transform $\hat{G}(j\nu)$, then for each $\omega \in \Omega$

$$\int_0^\infty y^2(t,\omega)dt \leq \sup_{\nu \in R} |\hat{G}(j\nu)|^2 \int_0^\infty x^2(t,\omega)dt$$

and use of the Lebesque Dominated Convergence Theorem [16, p.151] for E(.) permits the conclusion

$$||y||_{2,\ell_2} \leq \sup_{\nu \in R} |\hat{G}(j\nu)| \quad ||x||_{2,\ell_2} .$$

Moreover, this bound is attained.

Example 10: Again for the case $X = C(R^+)$, consider the random operator G on $F(\Omega;C)$ given by

$$(G(\omega)[x(\omega)])(t)=y(t,\omega) = \int_0^t g(t-s)x(s,\omega)dw(s,\omega), \quad t \in R^+, \ \omega \in \Omega$$

on non-anticipating functions (i.e. $\pi_t x$ is independent of $(I-\pi_t)w$ for all t, see section 3.2) in $E_2(\Omega;X;\ell_1)$. The following calculations

$$(1) \qquad E\{y^2(t)\} = E\{(\int_0^t \{g(t-s)x(s)dw(s))^2\}$$

$$= \sigma^2 \int_0^t g^2(t-s)E\{x^2(s)\}ds$$

(here $E\{(dw(t))^2\} = \sigma^2 dt$) permit the conclusion

$$||y||_{2,\ell_1} < \sigma(\frac{1}{2\pi}\int_{-\infty}^{\infty}|\hat{G}(j\nu)|^2 d\nu)^{1/2} \ ||x||_{2,\ell_1} \qquad ,$$

where $\hat{G}$ is the Fourier transform of the kernel g. See, for instance, McKean [50] for details of the reduction of (1) which makes use of the decisive property of orthogonal increments of w. Extension of this idea is the basis of several moment inequalities proven and used in sections 3.2 and 3.3 below.

2.4 Topologies for Random Processes:

The appropriate topology for the convergence arguments of the next chapter is introduced in this section. The topology is the usual one generated by weak convergence on a set of measures and, following a brief discussion of the general case, its properties are discussed for certain random function spaces including the continuous functions. The preservation of weak convergence under mappings is the final subject of this section.

Consider the following question: If the random process $x(t)$ is the limit in some precise sense of the sequence of processes $\{x_n(t)\}$, then for some functional f is it possible to determine the distribution of $f(x)$ if those of $f(x_n)$ are known? In other words is the distribution of $f(x)$ the limit in some sense of the distributions of $f(x_n)$? It is clear that some regularity assumptions must be placed on f to make these questions meaningful. Typical examples of functionals f are

$$f(x) = \int_{t_1}^{t_2} g(x(t))dt$$

$$f(x) = \sup_{t_1 \leq t \leq t_2} |x(t)| \quad .$$

The techniques introduced below have been developed to answer questions such as these.

Let (X,d) be a complete, separable metric space and let $\mathcal{B}_d(X)$ denote the class of Borel subsets of X generated by the d-topology. Let $C(X)$ denote the set of continuous (in d) functionals on X. Let $(\Omega, \mathcal{F}, P)$ be the basic probability space and let $x, x_n: \Omega \to X$ be random variables. The distribution of $x(x_n)$ is defined as

$$\mu_{(n)}(A) = P\{\omega \in \Omega: x_{(n)}(\omega) \in A \in \mathcal{B}_d(X)\} \quad .$$

Then a necessary and sufficient condition for convergence of the sequence of distributions of $f(x_n)$ to that of $f(x)$ for all $f \in C(x)$

is that

$$\lim_{n\to\infty} \int_X h(x)\mu_n(dx) = \int_X h(x)\mu(dx)$$

For all bounded, continuous functionals h. This answers the question posed above subject to the restrictions imposed and makes further study of the limiting operation above of interest.

On $\mathcal{B}_d(X)$ let PM(X) denote the set of probability measures, and let BC(X) denote the set of bounded, continuous functionals. If for elements μ_n, μ of PM(X)

$$\int_X hd\mu_n \to \int_X hd\mu \text{ , for every } h \in BC(X),$$

then μ_n converges weakly to μ, or $\mu_n \underset{w}{\to} \mu$. This convergence is determining by the following:

Theorem 1: [7, p.9] Elements μ, ν of PM(X) coincide if $\int_X hd\mu = \int_X hd\nu$ for every $h \in BC(X)$. Other implications are given in [7, Theorem 2.1, p. 11].

Let a subset $\Lambda \subset PM(X)$ be called relatively compact if every sequence in Λ has a weakly convergent subsequence (whose limit need not be in Λ, though in PM(X)). This compactness definition will be used in Chapters 3 and 4 to prove the existence of invariant distributions for stochastic processes. The criteria for determining relative compactness in general metric spaces are due largely to Prohorov and are given below. A family of probability measures $\Lambda \subset PM(X)$ is called tight if for every $\varepsilon > 0$ there exists a compact set $K(\varepsilon) \subset X$ such that $\mu(K(\varepsilon)) > 1-\varepsilon$ for every $\mu \in \Lambda$ [7, p.37].

Theorem 2: [7, p. 37] Let $(\mathbb{X},d)$ be a complete, separable metric space, $\Lambda \subset \mathbb{PM}(\mathbb{X})$, then Λ is tight if and only if it is relatively compact.

On $\mathbb{PM}(\mathbb{X})$ define a neighborhood system via the following sets: for $\mu \in \mathbb{PM}(\mathbb{X})$, $\varepsilon > 0$, $k \in Z^+$

$$N_{k,\nu}(\mu) = \{\nu \in PM(X): \left| \int_X h_i d\nu - \int_X h_i d\mu \right| < \varepsilon,$$

$$h_i \in BC(X),\ i=1,2,\ldots,k\}$$

Call the topology $\mathcal{W}$ generated by these neighborhoods the topology of weak convergence; clearly $\mu_n \underset{w}{\to} \mu$ if and only if $\mu_n \to \mu(\mathcal{W})$. A natural question to pose is: When is $\mathcal{W}$ metrizable?

For μ, $\nu \in PM(X)$ let

$$\varepsilon_1 = \inf \{\varepsilon > 0: \mu(A) \leq \nu(N_\varepsilon(A)) + \varepsilon\}$$

where $N_\varepsilon(A) = \{x \in X: d(x,A) < \varepsilon\}$, and $A \subset X$ is closed. Let ε_2 be defined by reversing the roles of μ and ν. Define

$$L(\mu,\nu) = \max(\varepsilon_1,\varepsilon_2)$$

Theorem 3: [7, p. 238] The function L is a metric on PM(X) called the Prohorov metric. Moreover, the L-topology is equivalent to $\mathcal{W}$ if the set X is separable.

By defining the distance between two random variables to be the L-distance between their distributions a metric (L) may be defined on $F(\Omega;X)$ the set of X-valued random variables. It is routine to verify

<u>Proposition</u>: If $\{x_n\}$, x are elements of $F(\Omega;X)$ then

$$P\{\omega\varepsilon\ \Omega: d(x_n(\omega),x(\omega)) \to 0\} = 1$$

implies $L(x_n,x) \triangleq L(\mu_{x_n},\mu_x) \to 0$.

Conversely,

<u>Theorem 4</u>: [56] If $\{x_n\}$ is an L-Cauchy sequence (possibly defined on different probability spaces), then a sequence $\{y_n\}$ and y may be constructed on $(\Omega, \mathcal{F}, P)$ such that

$$L(x_n,y_n) = 0 \quad \text{and} \quad P\{\omega: d(y_n(\omega),y(\omega)) \to 0\} = 1 .$$

Call a subset $\Lambda = \{x_\alpha,\ \alpha\ \varepsilon\ A\}$ of elements of $F(\Omega;X)$ indexed by A, <u>totally L-bounded</u> in $(F(\Omega;X),L)$ if every infinite sequence $\{x_{\alpha_n}\}_{n=1}^{\infty}$ taken from Λ has an L-Cauchy subsequence. This property is equivalent to the induced distributions of $\{x_\alpha\}$ being relatively compact. Precisely:

<u>Theorem 5</u>: [53] For Λ to be totally L-bounded, it is necessary and sufficient that for every $\varepsilon > 0$, there exists a compact subset $K(\varepsilon) \subset X$ (independent of $\alpha\ \varepsilon\ A$) with

$$P\{\omega : x_\alpha(\omega)\ \varepsilon\ K(\varepsilon)\} > 1-\varepsilon , \qquad \alpha\ \varepsilon\ A$$

Or equivalently, that the induced distributions $\{\mu_{x_\alpha}\}$ be tight.

Assume now that the metric space (X,d) is the space of R-valued continuous functions on R^+ (denoted by $C(R^+)$) with the metric

$$d(f,g) = \sum_{n=1}^{\infty} 2^{-n} \frac{||f-g||_n}{1+||f-g||_n}$$

where $||h||_n = \sup_{0 \leq t \leq n} |h(t)|$. Then (C,d) is a complete, separable metric space. In this case $F(\Omega;C)$ is a space of random functions. The basic compactness result for measures defined on $(C, \mathcal{B}_d(C))$ is given by the following:

Theorem 6: [7, p. 95] A subset $\Lambda \subset F(\Omega;C)$ is totally L-bounded if the following conditions are satisfied for any sequence $\{x_n\} \subset \Lambda$:

(i) the sequence (of distributions induced by) $\{x_n(0)\}$ is tight;

and (ii) there exist constants $\gamma \geq 0$ and $\alpha > 1$ and a non-decreasing, continuous function f on R^+ such that

$$P\{\omega: |x_n(t)-x_n(s)| \geq \lambda\} \leq \frac{1}{\lambda^\gamma} |f(t)-f(s)|^\alpha$$

for all t, $s \in R^+$, $n \in Z^+$, and $\lambda > 0$.

Corollary 7: The moment condition

$$E\{|x_n(t)-x_n(s)|^\gamma\} \leq |f(t)-f(s)|^\alpha$$

implies condition (ii) via Chebyshev's inequality.

Corollary 8: [41, p. 10] A subset $\Lambda \subset F(\Omega;C)$ is totally L-bounded if there exist $c > 0$, $c_n > 0$, $n=1,2,\ldots,A_n \in \mathcal{B}_d(C)$ such that, for every $x \in \Lambda$

(i) $E\{x^2(0)\} \leq c$;

(ii) $E\{|x(t)-x(s)|^2 ; x \in A_n\} \leq c_n|t-s|^2$, $0 \leq s,t \leq n$

(iii) $\sum_{n=1}^{\infty} (1 - P\{\omega:x(\omega) \in A_n\})$ is uniformly convergent on Λ.

These results will be used in Chapter 3 to investigate the behavior of the solutions to stochastic feedback equations. In that setting it is necessary to understand the transformation of distributions by operators on sets of stochastic processes.

Let $(X, \mathcal{B}(x))$ and $(Y, \mathcal{B}(Y))$ be measureable spaces and $G : X \to Y$ a measureable function. for $f \in F(\Omega;X)$ let μ_f be the distribution induced on $\mathcal{B}(X)$ by f. Recall that

$$\mu_f(A) = P\{\omega : f(\omega) \in A \in \mathcal{B}(x)\} = P(f^{-1}\cdot)$$

Then assuming $G : F(\Omega;X) \to F(\Omega; Y)$ for $f \in F(\Omega;X)$, Gf induces in the same way a distribution on $\mathcal{B}(Y)$ according to

$$\begin{aligned}\mu_{Gf}(B) &= P\{\omega : G[f(\omega)] \in B \in \mathcal{B}(Y)\} \\ &= P\{\omega : f(\omega) \in G^{-1}B \in \mathcal{B}(X)\} \\ &= \mu_f(G^{-1}B) \quad .\end{aligned}$$

If G is a random function the transformation is more interesting. Let (X,d_x) and (Y,d_y) be separable metric spaces. Then $\mathcal{G}(X,Y)$ is the set of operators $g : X \to Y$, continuous in the strong topology. Let $\mathcal{G}(X,Y)$ have the strong operator topology [16, p. 475], and let $\mathcal{B}(\mathcal{G})$ be the least Borel σ-algebra induced by this topology. As in section 2.3, $F(\Omega;\mathcal{G})$ denotes the set of $\mathcal{G}(X,Y)$-valued random variables.

A criterion sufficient to guarantee the assumption $G : F(\Omega;X) \to F(\Omega;Y)$ is the following

Theorem 9: [32] Let x be an element of $F(\Omega;X)$ and let $G \in F(\Omega;\mathcal{B})$, then the function $y : \Omega \to Y$ given by

$$y(\omega) = (Gx)(\omega) = G(\omega)[x(\omega)]$$

is a random variable if $G(\omega)[\cdot]$ is continuous $X \to Y$ for almost every $\omega \in \Omega$. Thus, every $G \in F(\Omega;\mathcal{B})$ maps $F(\Omega;X)$ into $F(\Omega;Y)$.

For the random variable $y = Gx$ a distribution is induced on $\mathcal{B}(Y)$ according to

$$\mu_{Gx}(B) = P\{\omega : (Gx)(\omega) \in B \in \mathcal{B}(Y)\}$$

$$= P\{\omega : G(\omega)[x(\omega)] \in B\}$$

Now by assumption (X,d_x) is separable, it follows that X has a countable base [7, Appendix I] that is, a family of open sets such that every other open subset of X is the union of a sub-family of these. Indicate this base by $\mathcal{N} = \{A_i\}_{i=1}^{\infty}$ and assume (without loss of generality) that the A_i are pairwise disjoint. It follows from the Borel property of $\mathcal{B}(X)$ (it is generated by the topology) that $\mathcal{B}(X)$ is generated by $\mathcal{N}$. Returning to the expression for μ_{Gx} for $G \in F(\Omega;\mathcal{B})$, it follows from the last few remarks that

$$\mu_{Gx}(B) = P\{\bigcup_{A_i \in \mathcal{N}} [\{\omega : x(\omega) \in A_i\} \cap \{\omega : G(\omega) \in \mathcal{B}(A_i,B)\}]\}$$

$$= \sum_{i=1}^{\infty} \mu_x(A_i)\mu_G(\mathcal{B}(A_i,B))$$

Here $\mathcal{B}(A_i,B) \subset \mathcal{B}(X,Y)$ is the set of operators g mapping X into Y and A_i into B. (The random variables x and G have been assumed independent under P). The formal expression

$$\mu_{Gx}(B) = \int_X \mu_x(d\eta)\mu_G(\mathcal{L}(\{\eta\},B))$$

follows from above.

Now assume that X and Y are R-valued function spaces on R^+, closed under the truncation operators $\{\pi_t\}_{t\epsilon R^+}$. Let $\mathcal{L}(X,Y)$ be further restricted to include operators causal as well as continuous. Each element x of $F(\Omega;X)$ generates a set of distributions $\{\mu_{\pi_t x}\}_{t\epsilon R^+}$ on $\mathcal{B}(X)$ according to the rule

$$\mu_{\pi_t x}(A) = P\{\omega : \pi_t x(\omega) \ \epsilon \ A\} \quad .$$

And in the same manner as above for $G \ \epsilon \ \mathcal{L}(X,Y)$ and $B \ \epsilon \ \mathcal{B}(Y)$

$$\begin{aligned}\mu_{\pi_t Gx}(B) &= \mu_{\pi_t G\pi_t x}(B) \\ &= \mu_{\pi_t x}[(\pi_t G)^{-1}B] \quad .\end{aligned}$$

Assuming that (X,d_x) and (Y,d_y) as sets of functions are separable, metric spaces, and that the random operator G is an element of $F(\Omega;\mathcal{L})$, then the formal expression below gives

$$\mu_{\pi_t Gx}(B) = \int_X \mu_{\pi_t x}(d\eta)\mu_{\pi_t G}(\mathcal{L}(\{\eta\},B))$$

the distributions induced on $\mathcal{B}(Y)$ by Gx for any element x of $F(\Omega;X)$.

As the final topic of this section consider the questions raised by the transformation of a weakly convergent sequence of distributions by an operator. Precisely, let $\Lambda \subset PM(X)$ be a relatively compact set of measures and let H be a function mapping X into itself: under what restrictions does H preserve weak convergence in PM(X) and relative compactness of Λ? A partial answer is given in

Theorem 10: (Topsøe [61]) Let $(X,d, \mathcal{B}(X))$ be a complete, separable metric space, H a measureable map from X into itself, and $\{\mu_n\}_{n=1}^{\infty}$ a weakly convergent sequence of elements of PM(X) with limit μ. Then the sequence $\{\mu_n(H^{-1}\cdot)\}_{n=1}^{\infty}$ is weakly convergent (to $\mu(H^{-1}\cdot)$) if H is continuous (modulo μ).

Though easily proved by examining the terms

$$\int_X f(H(x))\mu_n(dx)$$

for $f \in BC(X)$ (that is, $f(H\cdot) \in BC(X)$ if H is continuous), generalization of this result to the case where H is random is not straight-forward. For $\mu \in PM(X)$ and $G \in F(\Omega; \mathcal{G}(X,X))$ define

$$\nu(\omega)[\cdot] = \mu(G(\omega)^{-1}(\cdot)) \quad .$$

In general let L denote the Prohorov metric on PM(X) and let $\mathcal{B}(PM)$ be the least Borel σ-algebra generated by the L-topology. For any basic probability space $(\Omega, \mathcal{F}, P)$ then $F(\Omega;PM)$ has the usual interpretation and is well-defined as a consequence of the metric properties of L. Each element ν of $F(\Omega;PM)$ is pointwise a probability measure, $\nu(\omega) \in PM(X)$ for each $\omega \in \Omega$.

Two definitions of convergence of $F(\Omega;PM)$ are given in

<u>Definition</u>: (i) The sequence $\{\nu_n(\omega)\} \subset F(\Omega;PM)$ <u>converges weakly almost surely</u> to $\nu \in F(\Omega;PM)$ if for every $f \in BC(X)$

$$\lim_{n\to\infty} \operatorname*{ess\,sup}_{\omega\in\Omega} \left| \int_X f(x)\nu_n(\omega)[dx] - \int_X f(x)\nu(\omega)[dx] \right| = 0$$

Denote this by $\nu_n \underset{w,L_\infty}{\rightarrow} \nu$.

(ii) The sequence $\{\nu_n(\omega)\} \subset F(\Omega;PM)$ <u>converges weakly in mean</u> to $\nu \in F(\Omega;PM)$ if for every $f \in BC(x)$

$$\lim_{n\to\infty} E\left\{\left| \int_X f(x)\nu_n(\omega)[dx] - \int_X f(x)\nu(\omega)[dx] \right|\right\} = 0$$

Denote this by $\nu_n \underset{w,L_1}{\rightarrow} \nu$.

The next theorem gives conditions on the operator $G \in F(\Omega;\mathcal{B})$ so that convergence in the senses (i) and (ii) above is implied by $\mu_n \rightarrow \mu$ in the Prohorov topology.

<u>Theorem 11</u>: Let $(X,||\cdot||)$ be a separable Banach space, and let $G \in F(\Omega;\mathcal{B}(X))$.

(i) Then $\mu_n \rightarrow \mu$ (in L) implies that

$$\nu_{n,\omega}[\cdot] = \mu_n(G(\omega))^{-1}\cdot) \underset{w,L_\infty}{\rightarrow} \tilde{\nu} \text{ for some } \tilde{\nu} \in F(\Omega;PM) .$$

(ii) Let $G \in F(\Omega;\mathcal{B})$ be such that

$$E\{||G(\omega)[x]||\} \leq K||x||$$

for all $x \in X$ and some finite K independent of x. Then $\mu_n \rightarrow \mu$ (in L) implies that $\nu_n, \underset{w,L_1}{\rightarrow} \tilde{\nu}$ for some $\hat{\nu} \in F(\Omega;PM)$.

<u>Proof</u>: (i) Since $G(\omega) \in \mathcal{B}(X)$ (modulo P), $f(G(\omega)[\cdot])$ is an element of $BC(X)$ for almost every $\omega \in \Omega$ and every $f \in BC(X)$. Hence, for almost every $\omega \in \Omega$

$$\lim_{n\to\infty} \int_X f(G(\omega)[x])\mu_n(dx) - \int_X f(G(\omega)[x])\mu(dx)| = 0$$

which implies the conclusion for $\tilde{\nu}(\omega) = \mu(G(\omega)^{-1}\cdot)$.

(ii) By the hypothesis of (ii) the integral

$\int_{\Omega} G(\omega)[x]P(d\omega)$ is uniformly bounded $(||\cdot||)$

in x. Thus, since the μ_n are probability measures (specifically, they are σ-finite), Fubini's Theorem implies the equality (for every $f \in BC(X)$)

$$\int_{\Omega} \int_{X} f(G(\omega)[x])\mu_n(dx)P(d\omega) = \int_{X} \int_{\Omega} f(G(\omega)[x])P(d\omega)\mu_n(dx) .$$

Since $f \in BC(X)$ and $G(\omega) \in \mathcal{S}(X)$ (almost surely), and by the assumption of (ii), the function

$$\int_{\Omega} f(G(\omega)[\cdot])P(d\omega) : X \to R$$

is an element of BC(X). The conclusion follows using the reasoning in the proof of (i).

QED

Remarks: (1) Thus, continuity of $G(\omega)$ on X, almost surely (P), is sufficient to guarantee (w,L_∞)-convergence for G operating on L-convergent distributions μ_n. Clearly convergence (w,L_1) implies convergence (w,L_∞).

(2) It is useful to think of the elements of $F(\Omega;PM)$ as "random distributions." That is, assume that a number of control policies are available and that each of these is stochastic because of the nature of the task at hand. Then each of these possible policies may be represented by an element of PM(X), and if the control decision is made

at random it may be modeled as an element of $F(\Omega;PM)$. In other words a control policy is chosen according to some probability law from a set of stochastic controls. See the paper [76] for some related definitions of relaxed stochastic controls.

In the setting here the uncertain system "randomizes" the set of probability distributions representing the input and it is this point of view that is used in the latter portions of section 3.1.

CHAPTER 3

ASYMPTOTIC PROPERTIES OF STOCHASTIC SYSTEMS

3.1 Asymptotic Properties of General Feedback Systems:

The results in this chapter summarize an analysis of the asymptotic properties of feedback systems described by possibly random operators and subjected to stochastic inputs. In this section the properties of general feedback systems are investigated, and a theorem akin to the Small-Gain Theorem (section 2.2) used to establish moment bounds for signals in feedback systems. Under certain conditions on the system operator and the input the distributions of the feedback signals are shown to be asymptotically invariant. These results are reviewed in sections 3.2 and 3.3 for certain feedback systems described by random convolution operators. In section 3.4 a summary of the related existing theory for systems described by differential equations is presented.

Before undertaking the analysis of the asymptotic properties of uncertain systems it is important to define precisely the nature of such a system in feedback form. First the notion of a proper signal space is required. Let $\mathcal{H} \subset R$ be a linearly ordered set, the time set. Let X be a set of R-valued functions on $\mathcal{H}$, assumed to be Borel measureable (i.e. $x^{-1}(\mathcal{B}(R)) \subset \mathcal{B}(\mathcal{H})$ for every $x \in X$. Let $\{\pi_t\}_{t\in \mathcal{H}}$ denote the set of causal truncations introduced earlier, and denote by $\{\xi_t\}_{t\in \mathcal{H}}$ the set of anti-causal truncations. Assume that X is closed under both species of truncation. In that case and $y \in F(\Omega;X)$ by continuity of G. Furthermore, for $x \in S(\Omega;X)$, $G \in \mathcal{L}_c(X)$ and assuming $G0 = 0$, then $y \in S(\Omega;X)$ because G is causal (and $G0 = 0$).

A feedback system will be specified by a set of inputs, a plant, and a feedback controller. It will be assumed here that the system signals have their values in the same space as the inputs.

The input space is defined as follows: Let X, Θ , and $(\Omega_1, \mathcal{F}_1)$ be specified as above and let $\{P_\alpha\}_{\alpha\epsilon A}$ (A an index set) be a set of probability measures on $(\Omega_1, \mathcal{F}_1)$. For each $\alpha \in A$, $f \in S(\Omega_1;X)$ induces a probability distributioanon $(X, \mathcal{B}(X))$ according to the rule

$$\mu_{\alpha,\pi_t f}(B) = P_\alpha\{\omega \in \Omega_1 : (\pi_t f)(\omega) \in B \in \mathcal{B}(X)\} .$$

The input space is defined as an element of $(S(\Omega_1;X),\mu_\alpha)_{\alpha\epsilon A}$. For some choice $\alpha \in A$. The flexibility allowed by specifying a set of distributions $\{P_\alpha\}_{\alpha\epsilon A}$ rather than a single distribution reflects the empirical nature of the analysis of physical systems containing uncertainties. Frequently a number of hypothetical distributions for any uncertainty are proposed and some method of hypothesis testing used to determine the "best" of the candidates. This selection process should be regarded as preliminary to the analysis contained here.

The plant is defined by the following procedure: Let $\mathcal{L}_c(X)$ be specified as above, and let $(\Omega_2, \mathcal{F}_2)$ be a measureable space (possibly distinct from $(\Omega_1, \mathcal{F}_1)$). Let $\{P_\beta\}_{\beta\epsilon B}$ be a set of distributions on $\mathcal{F}_2$. Let $F(\Omega_2;\mathcal{L}_c)$ be the set of $\mathcal{L}_c$-valued random operators on Ω_2 governed by the law μ_β induced on $\mathcal{B}(\mathcal{L}_c)$ according to

$$(\xi_t x)(s) = \begin{cases} x(s) & s > t \\ 0 & s \leq t \end{cases}$$

or symbolically $\xi_t = I - \pi_t$ (I the identity on X).

Giving Θ an appropriate topology (relative to R) X may be topologized and a (least) Borel σ-algebra $\mathcal{B}(X)$ induced by the topology of X. To emphasize the fact that the systems to be studied here are to be considered as control systems, the set of signals is constrained to begin at some finite time. Thus, the set of signals admitted in the system is constrained as

$$S(\Omega;X) \triangleq \{f \in F(\Omega;X) : \xi_t f \equiv 0 \text{ for } \underline{\text{some}} \; t \in \Theta\} .$$

Let $\mathcal{J}(X)$ again denote the set of operators mapping X into itself. Indicate by $\mathcal{J}_c(X)$ the subset of $\mathcal{J}(X)$ consisting of causal, continuous operators. All systems to be studied here will by assumption be constructed from elements of $\mathcal{J}_c(X)$. Note, however, that this does not imply that the overall system will be either causal or continuous, and in general additional conditions will be required to assure preservation of these properties. See Willems [64] for a discussion of this feature of feedback systems which he calls well-posedness. Every element of $\mathcal{J}_c(X)$ induces a natural map on $F(\Omega;X)$ into itself using the continuity assumption and a natural map on $S(\Omega;X)$ by the additional restriction of causality. That is, for $x \in F(\Omega;X)$, $G \in \mathcal{J}_c(X)$ then

$$y(\omega) = G[x(\omega)]$$

$$\mu_{\beta,m}(D) = P_\beta\{\omega \in \Omega_2 : m(\omega) \in D \in \mathcal{B}(\mathcal{S}_c)\}$$

by any element m of $F(\Omega_2; \mathcal{S}_c)$. The collection

$$\{F(\Omega_2; \mathcal{S}_c), \mu_\beta\}_{\beta \epsilon B}$$

is the set of plants "selected" according to the law μ_β chosen as optimaly accounting for the physical observations.

For the purposes of this analysis the feedback operator is also assumed to be uncertain, though in design problems it usually may be freely chosen. Under this assumption the set of feedback controlers is specified in exactly the same manner as was the set of plants. For a given measureable space $\{\Omega_3, \mathcal{F}_3\}$ and a set of hypothetical distributions $\{P_\gamma\}_{\gamma\epsilon\Gamma}$ on $\mathcal{F}_3$ a feedback controller is an element of $F(\Omega_3; \mathcal{S}_c)$ governed by the law P_γ specified as best.

For any element x of $F(\Omega;X)$ let $\mathcal{B}(\pi_t x)$ denote the least Borel algebra generated by $\pi_s x$, $s \leq t$; in symbols

$$\mathcal{B}(\pi_t x) = \bigcup_{\substack{s \leq t \\ s,t \epsilon \theta}} \mathcal{B}(\pi_s x)^{-1}(\mathcal{B}(X))$$

The assumption of measureability of x assures that $\mathcal{B}(\pi_t x) \subset \mathcal{F}$ for every $t \epsilon \theta$.

Definition 1: Given a measureable space $(\Omega, \mathcal{F})$ and a set of probability measures $\{P_\alpha\}_{\alpha\epsilon A}$ on $\mathcal{F}$, a functionall h on $F(\Omega;X)$ into itself is said to be α-non-anticipative if for every $x \epsilon F(\Omega;X)$, $\mathcal{B}(\pi_t[h(x)])$ is independent of $\mathcal{B}(\xi_t x)$ for every $t \epsilon \theta$ with respect to P_α.

See for instance, Gikhman and Skowkhod [29, section 3.3] for a discussion of independence of set algebras. Call h <u>non-anticipative</u> if h is α-non-anticipative for every $\alpha \in A$.

Informally stated the definition says that values of the function $h(x)(t)$ are independent of the future $\xi_t x$ of x, at least with respect to the distribution P_α.

<u>Proposition 2</u>: If h is causal, $h(\pi_t x) = \pi_t h(\pi_t x)$, then h is non-anticipative.

<u>Definition 3</u>: A stochastic dynamical system is a 4-tuple $\{S(\Omega_1;X), \{P_\alpha\}_{\alpha\epsilon A}; F(\Omega_2,\tilde{\mathcal{G}}), \{P_\beta\}_{\beta\epsilon B}\}$ where $S(\Omega_1;X)$ is the set of X-valued signals, $\{P_\alpha\}$ a set of distributions on $(\Omega_1, \mathcal{F}_1)$, $\{P_\beta\}$ a set on $(\Omega_2, \mathcal{F}_2)$ and $F(\Omega_2,\tilde{\mathcal{G}})$ a set of $\tilde{\mathcal{G}}$-valued maps on Ω_2. Here each element G of $\tilde{\mathcal{G}}$ is non-anticipative (with respect to $\{P_\alpha\}$) on $S(\Omega_1;X)$ into itself. Moreover, for each $G \in \tilde{\mathcal{G}}$ assume $G0 = 0$.

<u>Definition 4</u>: A stochastic dynamical system is said to be in <u>feedback form</u> if it may be written as the 6-tuple

$$\{S(\Omega_1;X), \{P_\alpha\}_{\alpha\epsilon A} ; F(\Omega_2;\tilde{\mathcal{G}}), \{P_\beta\}_{\beta\epsilon B}; F(\Omega_3,\tilde{\mathcal{G}}), \{P_\gamma\}_{\gamma\epsilon\Gamma}\}$$

where the components have the meanings and implications established above. Moreover, that the operator H selected on $(\Omega_2 \times \Omega_3, \mathcal{F}_2 \times \mathcal{F}_3)$ according to $\{P_\beta\} \times \{P_\gamma\}$ given by

$$H(\omega_2,\omega_3) = (I+K(\omega_3) \circ G(\omega_2))$$

$(G \in F(\Omega_2;\tilde{\mathcal{G}}), K \in F(\Omega_3;\tilde{\mathcal{G}}))$ is one-to-one and non-anticipative with respect to $\{P_\alpha\}$ on $S(\Omega_1;X)$ into itself. In addition $H0 = 0$.

Clear from this definition is the observation that by identifying

$(\Omega_2 \times \Omega_3, \mathcal{F}_2 \times \mathcal{F}_3, \{P_\beta\} \times \{P_\gamma\})$ with some space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{P}_\delta\})$ the random operator H may be specified on $\tilde{\Omega}$ by

$$H(\tilde{\omega}) = I + \tilde{G}(\tilde{\omega})$$

where $\tilde{G}$ is a $\tilde{\mathcal{B}}$-valued random variable on $\tilde{\Omega}$. Moreover, by combining $(\Omega_1, \mathcal{F}_1, \{P_\alpha\})$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{P}_\delta\})$ in the same way, it is possible to define H and the input signals $\mathcal{S}(\Omega_1;X)$ on the same probability space, governed by the same collection of probability laws. The conclusion of this argument is that, for the purposes of this analysis at least, it suffices to consider the random operator equation

$$x(\omega) + G(\omega)[x] = u(\omega)$$

defined on some probability space $(\Omega, \mathcal{F}, \{P_\alpha\})$ as representative of the feedback system under investigation. Here $u, x \in S(\Omega;X)$, u an admissible input, x to be studied, and G is a random operator on $F(\Omega;X)$ into itself, non-anticipative with respect to $\{P_\alpha\}_{\alpha \in A}$. Moreover, for the purposes of the analysis to follow it is a useful simplification to assume that $G(\omega)$ is an element of $\mathcal{B}_c(X)$ $(\subset \tilde{\mathcal{B}})$, the causal, continuous operators on X. Thus, using Proposition 2 above, the qualifier "non-anticipative with respect to $\{P_\alpha\}$" may be ignored for such operators G. Finally, the assumption is made that by some decision process the "best" distribution has been chosen from among $\{P_\alpha\}_A \times \{P_\beta\}_B \times \{P_\gamma\}_\Gamma$ on the product space $(\Omega_1 \times \Omega_2 \times \Omega_3, \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3)$. Designate this underlying basic space by the customary symbols $(\Omega, \mathcal{F}, P)$.

Recall from section 2.3 the definition of the spaces $\xi_q(\Omega;X;\ell)$

and $\mathcal{E}_{qe}(\ell)$ (from here on the arguments Ω and X will be omitted when not of central concern). Let G be an element of $F(\Omega; \mathcal{J}_c)$ and u any element of $S(\Omega;X)$, then make the following:

Assumption (A1): (Existence of a locally bounded solution) For the equation

(1) $u(t,\omega) = x(t,\omega) + (G(\omega)[x(\omega)])(t), \quad t \in \Theta, \ \omega \in \Omega$

assume that $u \in \mathcal{E}_{qe}(\ell)$ implies that $x \in \mathcal{E}_{qe}(\ell)$.

That is, that $\pi_t u \in \mathcal{E}_q(\ell)$ for any $t \in \Theta$ implies $\pi_t x \in \mathcal{E}_q(\ell)$. As remarked ~~above~~ the assumptions of causality and continuity of G on the function space X (and $G0 = 0$) establishes that $x \in S(\Omega;X)$. What is assumed here is roughly (dependent on ℓ) the additional property that x has a "locally" bounded qth absolute moment.

The following result is the analog of Theorem 2.2.3 (Small Gain Theorem) in this setting.

Theorem 5: For the equation (1) above subject to the assumptions introduced with $G \in F(\Omega; \mathcal{J}_c)$ and $u \in \mathcal{E}_q(\ell) \cap S(\Omega;X)$, a sufficient condition that $x \in \mathcal{E}_q(\ell) \cap S(\Omega;X)$ is that

$$||G||_{q,\ell} \leq \alpha(\ell) < 1$$

for some $\alpha(\ell)$ independent of u.

Proof: By the assumption (A1) x exists and by virtue of the causality of $(I+G)^{-1}$ on X, x is an element of $\mathcal{E}_{qe}(\ell) \cap S(\Omega;X)$. Moreover, using the causality of G

$$\pi_t x(\omega) = \pi_t u(\omega) - \pi_t G(\omega)[\pi_t x(\omega)], \quad t \in \Theta, \ \omega \in \Omega$$

so x does not anticipate u and is a well-defined solution. Next using the triangle inequality property of $||\cdot||_{q,\ell}$ as a norm, it follows that

$$||\pi_t x||_{q,\ell} \leq ||\pi_t u||_{q,\ell} + ||\pi_t G[\pi_t x]||_{q,\ell}$$

The assumption on G permits

$$||\pi_t x||_{q,\ell} \leq ||\pi_t u||_{q,\ell} + \alpha(\ell)||\pi_t x||_{q,\ell}$$

The restriction on $\alpha(\ell)$ and the assumption $u \in \mathcal{E}_q$ lead to

$$||\pi_t x||_{q,\ell} \leq [1-\alpha(\ell)]^{-1}||u||_{q,\ell} \quad .$$

Observing the right hand side to be independent of $t \in \theta$ it follows that

$$||x||_{q,\ell} = \sup_{t \in H} ||\pi_t x|| \leq [1-\alpha(\ell)]^{-1}||u||_{q,\ell}$$

and hence, that the conclusion of the theorem is valid.

QED

Note that the inequality $||x||_{q,\ell} \leq K||u||_{q,\ell}$ for some $K < \infty$ is a "bonus" not required in the theorem. In deterministic stability theory this property ($||x|| \leq K||u||$) is sometimes called "finite-gain stability" and is frequently included as a condition in the definition of stability to preclude certain uniform boundedness arguments. See Willems [65] for a discussion of this point. Though not explicitly required above the finite gain property will be decisive below, where certain assumptions on u are used to deduce properties of x other than boundedness (see the proof of Theorem 3.2.5).

By the assumptions preceding the theorem I+G has a causal inverse on X or more generally on $S(\Omega;X)$, and that inverse is locally bounded (maps $\mathcal{E}_{qe} \to \mathcal{E}_{qe}$), the theorem guarantees that the inverse is globally bounded ($\mathcal{E}_q \to \mathcal{E}_q$). An important corollary to the theorem preceeds from the definition

$$\tilde{\alpha}(\ell) = \sup_{\substack{x_1, x_2 \varepsilon \mathcal{E}_q \\ ||x_1 - x_2||_{q,\ell} \neq 0}} \frac{||Gx_1 - Gx_2||_{q,\ell}}{||x_1 - x_2||_{q,\ell}}$$

of the <u>incremental gain</u> of $G \varepsilon \mathcal{B}_c(X)$.

<u>Corollary 6</u>: For the equation

$$u_1(\omega) - u_2(\omega) = x_1(\omega) - x_2(\omega) + G(\omega)[x_1(\omega)] - G(\omega)[x_2(\omega)]$$

with u_1, $u_2 \ \varepsilon \ \mathcal{E}_{qe}(\ell) \cap S(\Omega;X)$ and $G \ \varepsilon \ F(\Omega; \mathcal{B}_c)$ subject to the additional constraint

$$u_1 - u_2 \ \varepsilon \ \mathcal{E}_{q,\ell} \cap S(\Omega;X)$$

a sufficient condition that $x_1 - x_2 \ \varepsilon \ \mathcal{E}_{q,\ell} \cap S(\Omega;X)$ is that $\tilde{\alpha}(\ell) < 1$.

<u>Proof</u>: By assumption (A1) above $x_1 - x_2 \ \varepsilon \ \mathcal{E}_{qe}(\ell)$ and by the causality of the inverse of I+G on X, $x_1 - x_2 \ \varepsilon \ \mathcal{E}_{qe}(\ell) \cap S(\Omega;X)$. Moreover, causality of G assures that $x_1 - x_2$ does not anticipate $u_1 - u_2$ and so that $x_1 - x_2$ is well-defined as a solution of the equation. The remainder of the theorem follows directly from the definition of $\tilde{\alpha}(\ell)$ and the equation

$$\pi_t x_1(\omega)-\pi_t x_2(\omega)=\pi_t u_1(\omega)-\pi_t u_2(\omega)-\pi_t G(\omega)[\pi_t x_1(\omega)]+\pi_t G(\omega)[\pi_t x_2(\omega)]$$

along the line of the proof of Theorem 5.

QED

Remark 7: That Theorem 6 is a more stringent requirement for a system that Theorem 5 follows immediately from the observation $\alpha(\ell) \leq \tilde{\alpha}(\ell)$ for every $G \in F(\Omega; \mathcal{B}_c)$ (choose $x_2 \equiv 0$ in the definition of $\tilde{\alpha}(\ell)$). Thus, Theorem 5 may hold and Theorem 6 not. When valid, Theorem 6 guarantees that not only does I+G have a causal, bounded inverse on $\mathcal{E}_{q,\ell}$, but also that the inverse is continuous. This property is essential in the sequel.

Let Θ be the fixed set $R^+ = [0,\infty)$ (another choice is $R_{t_o} = [t_o,\infty)$ for some $t_o \in R$). Let (X,d) be a complete, separable metric space of functions mapping R^+ into R. Then with this choice of Θ it is possible to identify $F(\Omega;X)$ and $S(\Omega;X)$ (that is, all elements of $F(\Omega;X)$ are for each ω elements of $S(\Omega;X)$; the opposite inclusion holds by definition). Moreover, for the two functionals mentioned earlier

$$\ell_1(f) = \int_0^\infty |f(t)|dt, \qquad f \in X$$

$$\ell_2(f) = \operatorname*{ess\,sup}_{t \in R^+} |f(t)|$$

the spaces $\mathcal{E}_q(\Omega;X;\ell_{1,2})$ are Banach spaces.

In the next two sections below specific choices of the space X (as the set of continuous functions, and as the set of piecewise continuous functions) permit the use of bounds on the space $\mathcal{E}_q(\ell_2)$ to make Prohorov's Theorem applicable to certain feedback systems. Theorem 8 here is intermediate in this process.

Recall from section 2.4 the definitions of the Prohorov topology and the definition of totally L-bounded sets of random variables. Assume that $(X,||\cdot||)$ is a Banach space. For $H \in \mathcal{J}_c(X)$ define the norm of H on X (distinct from the norm of H as an operator on $\mathcal{E}_q$) as

$$\rho(H) = \sup_{0 \neq x \in X} \left|\left|\frac{Hx}{x}\right|\right|$$

and let $X_e = \{f: R^+ \to R : \pi_t f \in X\}$ be the extended space associated with X.

<u>Theorem 8</u>: (Deterministic plant) Consider the equation on $S(\Omega;X)$

$$u(\omega) = x(\omega) + G[x(\omega)]$$

where

$$u \in S(\Omega;X),\ G \in \mathcal{J}_c(X)$$

and the existence of a solution $x \in X_e$ such that $\pi_t x \in S(\Omega;X)$ is assumed. Moreover, assume that the set of distributions $\{\mu_{\pi_t u}\}_{t \in R^+}$ induced by u on $\mathcal{B}(X)$ is relatively compact, then a sufficient condition that the set $\{\mu_{\pi_t x}\}_{t\in R^+}$ be relatively compact is that

(i) $\rho(G) < 1$

(ii) $(I+G)^{-1} \in \mathcal{J}_c(X)$

<u>Proof</u>: Condition (i) assures that the solution $x(\omega)$ is an element of X for every $\omega \in \Omega$. The argument is familiar

$$\pi_t x(\omega) = \pi_t u(\omega) - \pi_t G[\pi_t x(\omega)]$$

$$\begin{aligned} ||\pi_t x(\omega)|| &\leq ||\pi_t u(\omega)|| + ||\pi_t G[\pi_t x(\omega)]|| \\ &\leq ||u(\omega)|| + \rho(G)||\pi_t x(\omega)|| \end{aligned}$$

Thus,

$$||\pi_t x(\omega)|| \leq [1-\rho(G)]^{-1}||u(\omega)|| \qquad \text{for every } t \in R^+, \ \omega \in \Omega$$

and the conclusion is immediate. That $x \in S(\Omega;X)$ is a consequence of the facts that $\pi_t x \in S(\Omega;X)$ for every t, and $x = \lim_{t\to\infty} \pi_t x$.

Again using the causality of G, for every $t \in R^+$, $\omega \in \Omega$

$$\pi_t x(\omega) + \pi_t G[\pi_t x(\omega)] = \pi_t u(\omega) \ .$$

Hence, for any $A \in \mathcal{B}(X)$

$$\begin{aligned} &P\{\omega : \pi_t x(\omega) \in A\} \\ &\quad = P\{\omega : \pi_t (I+G)^{-1}[\pi_t u(\omega)] \in A\} \\ &\quad = P\{\omega : \pi_t u(\omega) \in (I+G)^{-1}A\} \end{aligned}$$

Where $(I+G)^{-1}A \in \mathcal{B}(X)$, since $(I+G)^{-1}$ is continuous on X. Thus, the formula

$$\mu_{\pi_t x}(A) = \mu_{\pi_t u}[(I+G)^{-1}A]$$

follows from the above equalities and the definition of induced distributions.

Let $\{t_n\}_{n=1}^{\infty}$ be an increasing (unbounded) sequence of elements of R^+ and consider the sequence $\{\mu_{\pi_{t_n} x}\}_{n=1}^{\infty}$. Let f be any element of BC(X), the bounded, continuous functionals on X, then

$$\int_X f(y)\mu_{\pi_{t_n} x}(dy) = \int_X f(y)\mu_{\pi_{t_n} u}\,[(I+G)^{-1}dy]$$

$$= \int_X f[(I+G)y]\mu_{\pi_{t_n} u}(dy) \quad .$$

Since I+G is an element of $\mathcal{B}_c(X)$, the function $f[(I+G)(\cdot)] : X \to X$ is an element of BC(X). Moreover, since the set $\{\mu_{\pi_t u}\}_{t\varepsilon R^+}$ is assumed to be relatively compact in the weak topology, there exists a subsequence (unbounded) $\{t_{n'}\}_{n'=1}^{\infty} \subset \{t_n\}_{n=1}^{\infty}$ such that the subsequence

$$\left\{ \int_X f[(I+G)y]\mu_{\pi_{t_{n'}} u}(dy) \right\}_{n'=1}^{\infty}$$

converges. Hence, the original sequence $\{\mu_{\pi_{t_n} x}\}_{n=1}^{\infty}$ has a convergent subsequence. The arbitrary nature of the set $\{t_n\}_{n=1}^{\infty}$ leads to the desired conclusion that $\{\mu_{\pi_t x}\}_{t\varepsilon R^+}$ is relatively compact.

<u>QED</u>

In other words the theorem says that on the function space X, totally L-bounded (stochastic) inputs give rise to totally L-bounded outputs if the (deterministic) system operator I+G possesses a bounded, continuous, causal inverse on X. Boundedness of the signals is not the usual notion of boundedness in norm, but a more refined concept defined in terms of the distributions induced on X by the signals.

Although it may be considered as a rather direct consequence of Topsφe's Theorem (section 2.4) Theorem 8 serves a number of purposes. First it unites in a simple way the Prohorov theory of convergence and the deterministic operator stability theory to give interesting results for stochastic systems. And it executes this union in such a way as to make directly applicable the deterministic stability criteria (at least in their incremental form) to problems in this setting. Secondly it again establishes the invertibility of operators as a key tool in the class of problems being considered here. In this way Theorem 8 is the analog of Willems' result (Theorem 2.2.2). Corollary 9 below makes the Small-Gain Theorem applicable in this general setting and provides the link to explicit criteria based on this result.

<u>Corollary 9</u>: Let G be an element of $\mathscr{G}_c(X)$ and

$$\tilde{\rho}(G) \triangleq \sup_{\substack{x_1, x_2 \in X \\ x_1 - x_2 \neq 0}} \frac{||Gx_1 - Gx_2||}{||x_1 - x_2||} ,$$

then the system operator I+G under consideration maps totally L-bounded inputs ($u \in S(\Omega;X)$) into totally L-bounded outputs ($x \in S(\Omega;X)$) if $\tilde{\rho}(G) < 1$.

<u>Proof</u>: Clearly $\rho(G) < \tilde{\rho}(G)$ and so (i) of Theorem 8 is satisfied. An easy calculation suggested in the proof of Corollary 6 shows that

$(I+G)^{-1}$ is Lipschitz on x with Lipschitz constant $[1-\tilde{\rho}(G)]^{-1}$ and, hence, is bounded and continuous. Causality of $(I+G)^{-1}$ is assumed, thus (ii) of Theorem 8 holds.

QED

Examples illustrating the last four theorems are postponed until the sections following this one. In the remainder of this section the operator G defining the feedback system will be permitted to be random and the results from the latter paragraphs of section 2.4 used to investigate the system properties. Thus, let G be an element of $F(\Omega; \mathcal{J}_c)$ and let $u \in S(\Omega;X)$. Assume that G and u are independent under P. The properties of x defined by

$$(2) \qquad x(\omega) + G(\omega)[x(\omega)] = u(\omega)$$

are at issue here. Referring to section 2.3 for comments on the existence and measureability of solutions to (2), the assumption of locally bounded solutions will as usual be made.

Assumption (A2): For the equation (2) it is assumed that $\pi_t u \in S(\Omega;X)$ implies that $\pi_t x \in S(\Omega;X)$. That is, that bounded, measureable inputs $(\pi_t u)$ give rise to bounded, measureable outputs, at least on finite intervals $[0,t]$. Bounded means in $||\cdot||$ on X.

This assumption implies that for every $\omega \in \Omega$, $I+G(\omega)$ has a locally bounded inverse on X_e, and moreever, that this inverse maps measureable signals (elements of $S(\Omega;X)$) into measureable signals.

Now let $\{\mu_{\pi_t u}\}_{t\in R^+}$ denote the distributions induced by $\pi_t u$

on $\mathcal{B}(X)$. Put

$$\{\nu_{\pi_t x}(\omega)(\cdot)\} = \{\mu_{\pi_t u}[(I+G(\omega))(\cdot)]\}$$

The Borel measureability of $G \in F(\Omega; \mathcal{G}_c)$ assures that $\nu_{\pi_t x} \in F(\Omega;PM(X))$ (recall this notation from section 2.4). Then from these remarks and Theorem 2.4.11 the following result gives a partial description of x.

Theorem 10: For the equation (2) defined on the function space X, subject to the above assumptions on G, let $u \in S(\Omega;X)$, then by (A2) $\pi_t x \in S(\Omega;X)$ for every $t \in R^+$. Moreover, if $\tilde{\rho}_1(G) < 1$ where

$$\tilde{\rho}_1(G) = \operatorname*{ess\,sup}_{\omega \in \Omega} \sup_{\substack{x_1,x_2 \in X \\ x_1-x_2 \neq 0}} \frac{||G(\omega)[x_1]-G(\omega)[x_2]||}{||x_1-x_2||}$$

then $x \in S(\Omega;X)$; and if the set $\{\mu_{\pi_t u}\}_{t\in R^+}$ is relatively compact, (as a subset of the metric space $(PM(X),L)$), then so is $\{\nu_{\pi_t x}(\omega)\}$ in the (w,L_∞) topology on $F(\Omega;PM(X))$.

Proof: Using the causality of $G(\omega)$ for every ω

$$\pi_t x(\omega) = \pi_t u(\omega) - \pi_t G(\omega)[\pi_t x(\omega)]$$

Thus,

$$\begin{aligned} ||\pi_t x\,|| &\leq ||\pi_t u|| + ||\pi_t G(\omega)[\pi_t x(\omega)]|| \\ &\leq ||u|| + \tilde{\rho}_1(G) \cdot ||\pi_t x(\omega)|| \end{aligned}$$

which proves that $x \in S(\Omega;X)$ when combined with (A2) (establishing measureability of the truncated signals), using a simple limiting argument.

By a modification of the usual argument the condition $\tilde{\rho}_1(G) < 1$ implies that $[I+G(\omega)]^{-1}$ maps X into itself X (and is Lipschitz continuous) for each $\omega \in \Omega$. Let $\{\mu_{\pi_t u}\}_{t\in R^+}$ be the distributions induced on $\mathcal{B}(X)$ by $\pi_t u$. Then using Theorem 2.4.11 (i) the conclusion of the theorem follows.

QED

Corollary 11: If $\tilde{\rho}_2(G) < 1$ where

$$\tilde{\rho}_2(G) = \sup_{\substack{x_1,x_2\in X \\ x_1-x_2\neq 0}} E\left\{\frac{||G(\omega)[x_1]-G(\omega)[x_2]||}{||x_1-x_2||}\right\}$$

then $\{\mu_{\pi_t u}\}$ L-relatively compact implies that $\{\nu_{\pi_t x}(\omega)\}$ is relatively compact in the (w,L_1) topology on $F(\Omega;PM(X))$.

Proof: In Theorem 2.4.11 put

$$K = [1-\tilde{\rho}_2(G)]^{-1} < \infty \quad .$$

QED

The lack of symmetry in these results renders them provisional in nature. In the next two sections this deficiency is avoided by specializing the random operator G to be a nonlinear convolution in a special form. The space X is also restricted to be the continuous functions or the piecewise continuous functions. In the general case, however, this problem remains open.

3.2 Convolution versus a Wiener Process:

In this section the general results of the last section are reconsidered for a special class of random operators formed by convolution versus a Wiener process. Three particular problems are analyzed here: For a general convolution versus a Wiener process sample properties are discussed and moment inequalities derived. For a nonlinear convolution equation moment bounds are obtained for the solution and a condition similar to the Circle Theorem (section 2.2) is used to guarantee the existence of an invariant solution distribution. Finally, as a corollary to the analysis of the nonlinear case a linear convolution is considered and a condition like the Nyquist Criterion given to guarantee the asymptotic invariance of the solution distributions.

Let w denote the usual real-valued Wiener process on R^+, normalized so that $w(0) = 0$. The Wiener measure w is a probability measure on $(C(R^+;R), \mathcal{B}(C))$ satisfying two properties. For each $t, s \in R^+$ the random variable $w(t)-w(s)$ is normally distributed (on R) with mean

$$E\{w(t)-w(s)\} = m(t-s)$$

and variance

$$E\{[w(t)-w(s)-m(t-s)]^2\} = \sigma^2|t-s| \quad .$$

And for any finite collection of elements $\{t_i\}_{i=1}^n \subset R^+$ such that $t_1 \leq t_2 \leq \ldots \leq t_n < \infty$, the random variables $w(t_2)-w(t_1), w(t_3)-w(t_2), \ldots, w(t_n)-w(t_{n-1})$ are independent under (the measure) w.

For any C-valued random variable x on $(\Omega, \mathcal{F}, P)$ let $\mathcal{B}_{st}(x) \subset \mathcal{F}$

denote the minimal Borel σ-algebra over which $x(r)$ is measureable for $r \in [s,t]$. Symbolically,

$$\mathcal{B}_{st}(x) = \bigcup_{r\in[s,t]} \pi_r x^{-1}(\mathcal{B}(C)).$$

In particular let $\mathcal{B}_{st}(dw)$ denote the least Borel σ-algebra over which $w(r)-w(q)$ is measureable for $s \leq r \leq q \leq t$.

Endow $C(R^+)$ with the metric (see section 2.4)

$$d(x,y) = \sum_{n=1}^{\infty} 2^{-n} \frac{||x-y||_n}{1+||x-y||_n}, \quad ||z||_n = \sup_{t\in[0,n]} |z(t)|.$$

Let f be a continuous functional on (C,d), and assume that the measureable function $g : R^+ \times R^+ \to R$ is a causal convolution kernel, i.e., $g(t,s) \equiv 0$ for $s > t$. Then the operator

$$(Gx)(t,\omega) = \int_0^t g(t,s)f(s,(\pi_s x)(\omega))dw(s,\omega)$$

is well defined as an Ito integral [40] on non-anticipating random functions $x \in F(\Omega;C)$, i.e., those for which $\mathcal{B}_{ot}(x) \vee \mathcal{B}_{ot}(dw)$ is independent of $\mathcal{B}_{t\infty}(dw)$ for every $t \in R^+$. (Here $\mathcal{B}_1 \vee \mathcal{B}_2$ denotes the least Borel algebra containing both $\mathcal{B}_1$ and $\mathcal{B}_2$).

Let $u \in F(\Omega;C)$ be a non-anticipating random function in the above sense. As a special case of the general feedback equations of the last section consider the following equation.

$$(1) \quad x(t,\omega) = u(t,\omega) - \int_0^t g(t,s)f(s,x(s,\omega))dw(s,\omega)$$

<u>Theorem 1</u>: Conditions sufficient for the existence of a solution $x \in F(\Omega;C)$ (with locally bounded second moments) such that $\mathcal{B}_{ot}(x) \vee \mathcal{B}_{ot}(u) \vee \mathcal{B}_{ot}(dw)$ is independent of $\mathcal{B}_{t\infty}(dw)$ are that

$$\text{(i)} \quad |f(s,z)|^2 \leq a^2(s)|z|^2, \qquad z \in R$$

$$\text{(ii)} \quad \int_0^t g^2(t,s)a^2(s)ds < \infty ,$$

See [55, Chapter 3] where a much more general existence theorem is proved using the usual Picard approximations.

The properties of the moments of x are of fundamental importance in establishing the ultimate invariance properties of the distribution of x. The existence theorem above guarantees that the first and second moments of x are locally bounded (finite on any bounded interval [0,T]). The next theorem gives a bound on the entire half-line. Assume that $f : R^+ \times R \to R$ is continuous and that

$$|f(s,z)| \leq |a(s)| \; |z| , \qquad z \in R,$$

for some real-valued continuous function a. Assuming the hypothesis of Theorem 1, the mean of x the solution of (1) evolves according to

$$E\{x(t)\} = E\{u(t)\} - \int_0^t g(t,s)E\{f(s,x(s))\}mds \quad .$$

<u>Theorem 2</u>: (i) If

$$\sup_{t\in R^+} \int_0^t |g(t,s)| \; |a(s)| \; |m|ds \leq \alpha < 1$$

then

$$\sup_{t\in R^+} E\{|x(t)|\} \leq (1-\alpha)^{-1} \sup_{t\in R^+} E\{|u(t)|\} \quad .$$

(ii) Let $E\{u(t)\} = 0$, $m=0$, then $E\{x(t)\} = 0$, and if

$$\sup_{t\in R^+} \sigma^2 \int_0^t g^2(t,s)a^2(s)ds \leq \alpha_1 < 1$$

then

$$\sup_{t\in R^+} (E\{x^2(t)\})^{1/2} \leq (1-\sqrt{\alpha_1})^{-1} \sup_{t\in R^+} (E\{u^2(t)\})^{1/2}$$

Proof: (i) This part of the theorem follows easily from the inequalities (assume $m > 0$):

$$E\{|x(t)|\} \leq E\{|u(t)|\} + \int_0^t |g(t,s)| E\{|f(s,x(s))|\} mds$$

$$\leq E\{|u(t)|\} + \int_0^t |g(t,s)| \; |a(s)| E\{|x(s)|\} mds$$

(ii) The first statement of this part follows from the Theorem 1 and the properties of the Ito integral [50, p.24]. The remainder of (ii) follows from

$$(E\{x^2(t)\})^{1/2} \leq (E\{u^2(t)\})^{1/2}$$

$$+ (E\{(\int_0^t g(t,s)f(s,x(s))dw(s))^2\})^{1/2} \leq (E\{u^2(t)\})^{1/2}$$

$$+ \sigma(\int_0^t g^2(t,s)a^2(s)ds)^{1/2} \sup_{0\leq s\leq t} (E\{x^2(s)\})^{1/2}$$

QED

It is via bounds on the second moment that Corollary 2.4.8 is used to establish the existence of an invariant limit (in distribution) for x. The remainder of this section will be devoted to a statement and proof of this property for two special cases of (1) corresponding to certain restrictions on the functional f in (1). The first result below gives an improved moment bound for the nonlinear case under these restrictions.

Let (1) be replaced by

$$\text{(2)} \qquad x(t,\omega) = u(t,\omega) - \int_0^{t=} g(t-s)f(s,x(s,\omega))dw(s,\omega)$$

where g is now a time-invariant kernel and Theorem 1 is assumed to be in force. Assume moreover, that

$$0 < a \leq \frac{f(s,z)}{z} \leq b < \infty , \qquad s \in R^+, \quad z \in R \quad .$$

<u>Theorem 3</u>: Under the additional assumptions that $E\{u(t)\} = 0$, $E\{dw(t)\} = 0$, for every $t \in R^+$ the conditions:

(i) $\int_0^\infty e^{-r_o t} g^2(t)dt < \infty$ for some $r_o < 0$;

(ii) for $H(r+j\nu) = \int_0^\infty e^{-rt}e^{-j\nu t}g^2(t)dt$, and some $r \in (r_o,0)$, the exclusion below holds

$$(-2\sigma^{-2}(a^2+b^2)^{-1},j0) \notin \bigcup_{\substack{\nu\in R \\ r\geq r_o}} H(r+j\nu) \qquad ; \text{ and}$$

(iii) $\sup\limits_{\nu\in R} |H^{-1}(r+j\nu)+\frac{\sigma^2}{2}(a^2+b^2)| > \frac{\sigma^2}{2}(b^2-a^2)$

for some $r \in (r_o,0)$.

imply that $\sup\limits_{t\in R^+} E\{x(t)^2\} \leq \beta \sup\limits_{t\in R^+} E\{u(t)^2\}$ for some finite $\beta > 0$.

Proof: An easy calculation gives

$$E\{x^2(t)\} = E\{u^2(t)\} + \sigma^2 \int_0^t g^2(t-s)E\{f^2(s,x(s))\}ds$$

$$= E\{u^2(t)\} + \frac{1}{2}\ \sigma^2(a^2+b^2) \int_0^t g^2(t-s)E\{x^2(s)\}ds$$

$$+ \ \sigma^2 \int_0^t g^2(t-s)E\{\tilde{f}^2(s,x(s))\}ds$$

where $\tilde{f}^2(s,z) = f^2(s,z) - \frac{1}{2}(a^2+b^2)z^2$.

By (ii) $\frac{1}{2}\sigma^2(a^2+b^2)H(r+j\nu) \neq -1$, thus by two lemmas of Benes [5, Lemmas 4,5, p. 32] the operator $I + \frac{1}{2}\sigma^2(a^2+b^2)H$ (H defined by (ii)) has a continuous inverse represented by the identity minus a convolution. Hence,

$$E\{x^2(t)\} \leq (I + \frac{\sigma^2}{2}(a^2+b^2)H)(Eu^2)(t)$$

$$+ \ \sigma^2 \int_0^t \tilde{h}(t-s)E\{\tilde{f}^2(s,x(s))\}ds$$

where $\tilde{h}$ is the function whose Fourier transform is $\tilde{H}(j\nu) = H(j\nu)[1+\frac{\sigma^2}{2}(a^2+b^2)H(j\nu)]^{-1}$. An easy calculation verifies

$$|\tilde{f}^2(s,z)| \leq \frac{1}{2}(b^2-a^2)z^2$$

Thus,

$$E\{x^2(t)\} \leq (I + \frac{\sigma^2}{2}(a^2+b^2)H)(Eu^2)(t)$$

$$+ \frac{1}{2}\sigma^2(b^2-a^2) \int_0^t \tilde{h}(t-\tau)E\{x^2(s)\}ds \quad .$$

Condition (iii) establishes

$$\sup_{\nu} \frac{1}{2} \sigma^2 (b^2-a^2) \tilde{H}(r+j\nu) < 1$$

and the conclusion of the theorem follows from the L_∞-version of the deterministic Circle Theorem in Zames [75] (given in section 2.2).

QED

Remark: Note that

$$H(r+j\nu) = \int_{-\infty}^{\infty} G(r+j(\nu-\nu_o))G(r+j\nu_o)d\nu_o$$

where

$$G(r+j\nu) = \int_0^{\infty} e^{-rt}e^{-j\nu t}g(t)dt \quad .$$

And so, the criteria could have easily been stated in terms of the r-shifted Fourier transform of g.

The sufficiency of the following theorem is easily established using the techniques of the last proof.

Theorem 4: Consider the linear integral equation

$$x(t,\omega) = u(t,\omega) = \int_0^t g(t-s)x(s,\omega)dw(s,\omega),$$

then subject to $E\{u(t)\} = 0$, $E\{dw(t)\} = 0$ and Theorem 1, the condition

$$\sigma^2 \int_{-\infty}^{\infty} |G(j\nu)|^2 d\nu < 2\pi$$

is necessary and sufficient to guarantee

$$\sup_{t\in R^+} E\{x^2(t)\} \leq \gamma \sup_{t\in R^+} E\{u^2(t)\} \qquad \text{for some } \gamma.$$

Proof: (Necessity) Using the properties of the stochastic integral, the following equation is easily derived

$$E\{x^2(t)\} = E\{u^2(t)\} + \sigma^2 \int_0^t g^2(t-s)E\{x^2(s)\}ds.$$

Rewriting this equation as

$$y(t) = z(t) + \int_0^t h(t-s)y(s)ds$$

where the L_∞-boundedness of y is at question, the conclusion (both parts) of the theorem follows from a result of Davis [11] and the observation that y is a continuous function on the half-line R^+ which follows from Theorem 1.

QED

By further specializing the input process u it is possible to use the criteria of Theorems 2 and 3 to establish the asymptotic invariance of the solution distribution.

Theorem 5: Consider the integral equation

$$(3) \qquad x(t,\omega) = u(t,\omega) - \int_0^t g(t-s)f(s,x(s,\omega))dw(s,\omega)$$

subject to the existence condition of Theorem 1. Assume that u and w are independent, $E\{u(t)\} \equiv 0$, $E\{dw(t)\} = 0$ and, moreover, that the process u satisfies the Lipschitz condition

$$|u(t,\omega)-u(s,\omega)|^2 \leq \gamma|t-s| \ , \quad \gamma > 0, \ t,s \in R^+$$

almost surely (ω), and the moment bound $E\{u^2(t)\} \leq \beta^2 < \infty$. Then a necessary and sufficient condition that the solution x of (3) be totally L-bounded in $(S(\Omega;C),L)$ is that

$$E\{x^2(t)\} \leq \alpha^2 < \infty, \qquad t \in R^+ \quad ,$$

for some constant $\alpha > 0$. Moreover, if u is stationary then x is asymptotically stationary with respect to u and the increments of w.

Remark: Clearly then Theorem 3 gives a sufficient condition for the distributions of x to be bounded (or ultimately invariant) for nonlinear, conic functions f. Theorem 4 gives a necessary and sufficient condition in the special case of linear, constant functions f. Both criteria are stated in terms of the Fourier transform of g, and are thus subject to the usual design interpretations used for feedback systems including a linear, time-invariant, convolution operation.

Proof of the theorem: The proof is based on a lemma of Ito and Nisia [41] stated as Corollary 2.4.8 above. It follows the pattern of a similar proof in [41]. The verification of the hypothesis of that lemma proceeds in three steps, the first showing that the solution x of (3) is totally L-bounded.

Lemma 6: Let the kernel g be locally L_2, that is $\int_s^t |g(\tau)|^2 d\tau < \infty$ for $t, s \in R^+$; then there exists a constant $\eta = \eta(\varepsilon,T)$ such that for any $\varepsilon > 0$, $T \geq 0$,

$$P\{\omega : \sup_{s \leq t \leq s+T} |x(t,\omega)| > \eta\} < \varepsilon \ , \quad \text{for every } s \in R^+$$

Proof: From the definition of a solution

$$x(t) = x(s)+u(t)-u(s) - \int_s^t g(t-\tau)f(\tau,x(\tau))dw(\tau)$$

$$- \int_0^s [g(t-\tau)-g(s-\tau)]f(\tau,x(\tau))dw(\tau)$$

And so, setting

$$S = \sup_{s\le t\le s+T} |x(t)| \quad ,$$

the inequality

$$S \le |x(s)| + |u(s)| + \sup_{s\le t\le s+T} |u(t)|$$

$$+ \sup_{s\le t\le s+T} \left|\int_s^t g(t-\tau)f(\tau,x(\tau))dw(\tau)\right|$$

$$+ \sup_{s\le t\le s+T} \left|\int_0^s [g(t-\tau)-g(s-\tau)]f(\tau,x(\tau))dw(\tau)\right|$$

$$= V + W + X + Y + Z$$

follows. Thus

$$P(S > \eta) \le P(V > \eta/5) + P(W > \eta/5) + P(X > \eta/5) + P(Y > \eta/5) + P(Z > \eta/5).$$

Now

$$P(V > \eta/5) \le \frac{5}{\eta} (E\{x^2(s)\})^{1/2} \le \frac{5\alpha}{\eta} \quad ,$$

and $P(W > \eta/5) \le \frac{5\beta}{\eta}$ in the same way. The analysis of the next three terms is somewhat more delicate. From the Lipschitz assumption on u

$$|u(t,\omega)| \le \gamma|t-s| + |u(s,\omega)|$$

Hence,

$$\sup_{S\leq t\leq S+T} |u(t,\omega)| \leq \gamma T + |u(s,\omega)|$$

and so, for $\eta > 5\gamma T$

$$P\{x > \frac{\eta}{5}\} \leq P\{|u(s,\omega)| > \frac{\eta}{5} - \gamma T\}$$

$$\leq \frac{5\beta}{\eta - 5\gamma T} \quad .$$

For Y and Z consider the following

$$P\{Y > \eta/5\} \leq \sigma^2 (\frac{5}{\eta})^2 b^2 \sup_{s\leq t\leq S+T} \int_S^{S+T} g^2(\tau) E\{x^2(t-\tau)\} d\tau$$

$$\leq \sigma^2 \alpha^2 (\frac{5}{\eta})^2 b^2 \int_S^{S+T} g^2(\tau) d\tau \quad .$$

Similarly, for Z

$$P\{Z > \eta/5\} \leq \sigma^2 (\frac{5}{\eta})^2 b^2 \int_0^S [g(t-\tau) - g(s-\tau)]^2 E\{x^2(\tau)\} d\tau$$

$$\leq 4\sigma^2 \alpha^2 b^2 (\frac{5}{\eta})^2 \int_0^{S+T} g^2(\tau) d\tau \quad .$$

Therefore, the bound for $\eta > 5\gamma T$

$$P\{S > \eta\} \leq \frac{5}{\eta}(\alpha+\beta) + \frac{5\beta}{\eta - 5\gamma T} + 5(\frac{5}{\eta})^2 \sigma^2 \alpha^2 b^2 \int_0^{S+T} g^2(\tau) d\tau$$

holds, and clearly for any ε, $T > 0$ an η may be chosen sufficiently large enough to imply

$$P\{S > \eta\} < \varepsilon$$

<u>QED</u> (Lemma 6)

The second step in the proof requires verification of the following lemma.

Lemma 7: There exists a constant $\xi = \xi(m,T)$ such that for every $t, v \in [s,s+T]$ the following inequality holds (almost surely ω)

$$E\{|x(t)-x(v)|^4 \Big| \sup_{s\leq r\leq s+T} |x(r)| \leq m\} \leq \xi|t-v|^2$$

if for every $s, T \in R^+$

$$\delta_1(T) = \sup_{0\leq t\leq T} \left(\frac{1}{t}\int_0^t g^2(\tau)d\tau\right)^2 < \infty$$

$$\delta_2(s,T) = \sup_{s\leq v\leq t\leq s+T} \left(\frac{1}{t-v}\int_0^v [g(t-v-\tau)-g(\tau)]^2 d\ \right)^2 < \infty$$

Proof: Again express the solution to (3) as

$$x(t)-x(v) = u(t)-u(v) - \int_v^t g(t-\tau)f(x(\tau))dw(\tau)$$

$$- \int_0^v [g(t-\tau)-g(v-\tau)]f(x(\tau))dw(\tau)$$

where the arbitrary assumption $t \geq v$ has been made. Using $(c+d)^4 \leq 8c^4+8d^4$ and the pointwise assumption on u, the following obtains

$$E\{|x(t)-x(v)|^4 \Big| \sup_{s\leq r\leq s+T} |x(r)| \leq m\}$$

$$\leq 8\gamma^2|t-v|^2$$

$$+ 64E\{\left(\int_0^t g(t-\tau)f(x(\tau))dw(\tau)\right)^4 \Big| \sup_{s\leq r\leq s+T} |x(r)| \leq m\}$$

$$+ 64E\{(\int_0^v [g(t-\tau)-g(v-\tau)]f(x(\tau))dw(\tau))^4 \Big| \sup_{s\leq r\leq s+T} |x(r)| \leq m\}$$

$$= X + Y + Z.$$

Now

$$Y = 64 \int_v^t \int_v^t g^2(t-\tau)g^2(t-\mu)E_m\{f^2(x(\tau))f^2(x(\mu))dw^2(\tau)dw^2(\mu)\}$$

where E_m includes the conditioning $\sup_{s\leq r\leq s+T} |x(r)| \leq m$. Thus,

$$Y \leq 64\ \sigma^4 b^4 m^4 (\int_v^t g^2(\tau)d\tau)^2 \leq 64\ \sigma^4 b^4 m^4 \delta_1(s+T)|t-v|^2$$

By similar arguments

$$Z \leq 64\sigma^4 m^4 b^4 (\int_0^v [g(t-\tau)-g(v-\tau)]^2 d\tau)^2$$

$$\leq 64\sigma^4 m^4 b^4 \delta_2(s,T)|t-v|^2$$

where δ_1 and δ_2 are given in the hypothesis of the Lemma. Choosing

$$\xi = 8\gamma^2 + 64\sigma^4 b^4 m^4 [\delta_1(s+T)+\delta_2(s,T)]$$

satisfies the assertion of the lemma.

QED (Lemma 7)

Next the assertion that the solution x of (3) is totally L-bounded is verified.

Lemma 8: The conditions of Lemmas 6 and 7 imply that x is totally L-bounded.

Proof: Denote by θ_s the shift operator

$$(\theta_s x)(t) = x(s+t),$$

and by $(\cdot)_+$ the function $(r)_+ = \max(r,0)$. Using Lemma 6, define the constants $\eta_k = \eta(\varepsilon(k),T(k)) = \eta(2^{-k},2k+\tau)$, then

$$P\{\sup_{-k-\tau\leq t\leq k} |\theta_s(x(t))| \leq \eta_k\}$$

$$= P\{\sup_{(s-k-\tau)_+\leq t\leq s+k} |x(t)| \leq \eta_k\} \leq 1-2^{-k}$$

Let the function ξ in Lemma 7 define the constants $\xi_k = \xi(\eta_k, 2k+\tau)$, then from Lemma 7 for $t,v \in [(s-k)_+,s+k]$

$$E\{|x(t)-x(v)|^4 \Big| \sup_{(s-k-\tau)_+\leq t\leq s+k} |x(t)| \leq \eta_k\} \leq \xi_k|t-v|^2 \quad .$$

Define $A_k \subset C(R)$ as $\{h \in C: \sup_{-n-\tau\leq t\leq n} |h(t)| \leq \eta_k\}$, then

$$E\{|(\theta_s x)(t)-(\theta_s x)(v)|^4 \Big| \theta_s x \in A_k\} \leq \eta_k|t-v|^2$$

and the conclusion of this Lemma follows from Corollary 2.4.8.

QED (Lemma 8)

The remainder of the proof of the theorem follows from the last lemma. Let $(PM(C),L)$ be the set of probability measures on $C(R^+;R)$ equipped with the Prohorov metric. Then from Lemma 8 the induced distributions $\{\mu_{\theta_s x}\}_{s\in R^+}$ on $(C(R^+), \mathcal{B}(C))$ is relatively compact, By the Lipschitz assumption on u the set $\{\mu_{\theta_s u}\}_{s\in R^+}$ is relatively compact, and setting $(\tilde{\theta}_s w)(t) = w(t+s)-w(s)$ it is easily shown (using Corollary 2.4.8) that $\{\mu_{\tilde{\theta}_s w}\}$ is relatively compact. Recalling the fact that the direct product of (relatively) compact sets is (relatively) compact, then the set of distributions $\{\mu_s\}_{s\in R^+}$ induced on $(C\times C\times C, \mathcal{B}(C\times C\times C))$ by $(\theta_s x, \theta_s u, \tilde{\theta}_s dw)$ is relatively compact. This establishes the first assertion of the Theorem.

Now using the fact that $\theta_s h$ is continuous in (s,h) on $R \times C(R^+)$ (for the metric d), the function $\mu_{(\cdot)}(A)$ is measureable on R^+ for any set $A \in \mathcal{B}(C\times C\times C)$. Hence, the function (of t)

$$\nu_t(A) = \frac{1}{t}\int_0^t \mu_s(A)\, ds$$

is continuous on R^+ for any A as above.

Since the set $\{\mu_s\}$ is relatively compact, by Prohorov's Theorem (Theorem 2.4.2 here) for any $\varepsilon > 0$ there exists a compact subset $K(\varepsilon) \subset (C\times C\times C)(R^+)$ independent of $s \in R^+$ such that $\mu_s(K) \geq 1-\varepsilon$ and therefore such that $\nu_t(K) > 1-\varepsilon$ for every $t \in R^+$. Thus, the set $\{\nu_t\}_{t\in R^+}$ is relatively compact, and there exists a measure $\nu_\infty \in PM(C\times C\times C)$ and an increasing sequence $\{t_n\}_{n=1}^{\infty}$ such that $\nu_{t_n} \underset{w}{\to} \nu_\infty$, or equivalently in the L-topology.

Let $(\tilde{x},\tilde{u},\tilde{w})$ be the $(C\times C\times C)(R^+)$-valued random variable whose probability law is ν_∞. It remains to show that

(i) $(\tilde{u},\tilde{w}) = (u,w)$

(ii) $\tilde{x}$ is stationarily correlated with respect to $(\tilde{u},\tilde{w})$,

and (iii) $\tilde{x} = \tilde{u} - G\tilde{x}$

Point (i) follows from the stationarity of u and of the increments of w.

To show (ii) consider continuous, bounded functionals ψ_1, ψ_2, ψ_3

on R^k, R^m, R^n respectively and the series of equalities

$$E\{\psi_1(\tilde{x}_{t_1+t},\ldots,\tilde{x}_{t_k+t})\psi_2(\tilde{u}_{t_1'+t},\ldots,\tilde{u}_{t_m'+t})\psi_3(\tilde{w}_{t_1''+t},\ldots,\tilde{w}_{t_n''+t})\}$$

$$= \lim_{r\to\infty} \frac{1}{T_r}\int_0^{T_r} ds\ E\{\psi_1(\tilde{x}_{t_i+t+s})_{i=1}^{k}\ \psi_2(\tilde{u}_{t_j'+t+s})_{j=1}^{m}\ \psi_3(\tilde{w}_{t_\ell''+t+s})_{\ell=1}^{n}\}$$

$$= \lim_{r\to\infty} \frac{1}{T_r}\int_t^{T_r+t} ds\ E\{\psi_1(\tilde{x}_{t_i+s})\ \psi_2(\tilde{u}_{t_j'+s})\ \psi_3(\tilde{w}_{t_\ell''+s})\ \}$$

$$= \lim_{r\to\infty} \frac{1}{T_r}\int_0^{T_r} \text{———}\ ds$$

$$= E\{\psi_1(\tilde{x}_{t_1},\ldots,\tilde{x}_{t_k})\ \psi_2(\tilde{u}_{t_1'},\ldots,\tilde{u}_{t_m'})\ \psi_3(\tilde{w}_{t_1''},\ldots,\tilde{w}_{t_n''})\}\ .$$

Here the third equality follows from the symbolic decomposition

$$\int_t^{T_r+t} = \int_0^{T_r} + \int_{T_r}^{T_r+t} - \int_0^t$$

and the boundedness properties of $(\tilde{x},\tilde{u},\tilde{w})$ over finite intervals. That this series of equalities for all ψ_1,ψ_2,ψ_3 determines the properties of the finite dimensional distributions of $(\tilde{x},\tilde{u},\tilde{w})$ is fundamental, see Gikhman and Skorokhod [29, Chapter 3].

To show (iii) it suffices to show that for every $s \in [0,t]$

(iii)' $\tilde{x}(t) = \tilde{x}(s) + \tilde{u}(t) - \tilde{u}(s) - (G\tilde{x})(t) + (G\tilde{x})(s)$.

An argument used in Ito and Nisio [41] may be applied directly at this point to yield the desired conclusion.

<u>QED</u> (Theorem 5)

In the event that the function f is linear ($f(z) = az$, $a > 0$) Theorem 5 may be sharpened using Theorem 4 to prove:

Corollary 9: For the linear integral equation

$$x(t) = u(t) - a \int_0^t g(t-s)x(s)dw(s)$$

subject to the assumptions on u, w, and g expressed in the hypothesis of Theorem 5, a necessary and sufficient condition that the distribution of x be ultimately stationary is that

$$a^2\sigma^2 \int_{-\infty}^{\infty} |G(j\nu)|^2 \, d\nu < 2\pi \quad . \tag{4}$$

3.3 Convolution Versus a Levy Process:

The most immediate modification of the integral equation investigated in the last section is to consider the convolution operator with the Wiener measure replaced by a Levy measure, representative of the most general process with independent increments. As is well-known [29] the Levy process has sample paths with at most countable jump discontinuities in any finite interval. Moreover, it may be decomposed into a linear combination of a Wiener process and a general Poisson process. In a feedback system jump process may be considered as models of random shock phenomena and Levy process models as descriptive of combinations of continuous and shock random signals. It is therefore appropriate to review the properties of such processes, whose sample paths are quite different from those of the Wiener process and its transformations.

Let $\{\xi_n\}_{n=1}^{\infty}$ be a set of independent, identically distributed random variables on some probability space $(\Omega, \mathcal{F}, P)$. Assume that the distribution function of the ξ_n is

$$F_\xi(\omega) = P\{\omega:\ \xi_n(\omega) \leq \alpha\} = \begin{cases} 1 - e^{-\lambda\alpha} & ; \quad \alpha \geq 0 \\ 0 & ; \quad \alpha < 0 \end{cases}$$

where $\lambda > 0$. Note $E\{\xi_n\} = \lambda^{-1}$. Let $S_n(\omega) = \sum_{i=1}^{n} \xi_i(\omega)$

then the distribution of S_n is

$$F_n(\alpha) = \begin{cases} 1 - e^{-\lambda\alpha} \sum_{k=0}^{n-1} \frac{(\lambda\alpha)^k}{k!} & ; \quad \alpha \geq 0 \\ 0 & ; \quad \alpha < 0 \ . \end{cases}$$

A Poisson process $x(t,\omega)$, $t \in R^+$, $\omega \in \Omega$, may be defined via

$$x(t,\omega) = \begin{cases} \max\{k:\ S_k(\omega) \leq t\} & , \quad S_0(\omega) = 0 \\ \infty, \text{ if } S_k(\omega) < t & \text{for all } k \ . \end{cases}$$

Note that $x(t,\omega) = n$ if and only if $S_n(\omega) \leq t$ and $S_{n+1}(\omega) > t$. Thus, the induced distribution of x is

$$P\{\omega:\ x(t,\omega) = n\} = \begin{cases} \frac{(\lambda t)^n}{n!} e^{-\lambda t} & ; \quad n = 0, 1, 2, \ldots \\ 0 & ; \quad n = \infty \ . \end{cases}$$

From this expression $E\{x(t)\} = \lambda t$, $E\{(x(t) - \lambda t)^2\} = \lambda t$. Intuitively, the Poisson process represents a quantity increasing by unit jumps occuring at random instants of time.

A somewhat more general process which accounts for random jump amplitudes is defined as follows. Let $\{\eta_k\}_{k=1}^{\infty}$ be a set of independent, identically distributed random variables with common distribution function $F_\eta(\alpha) = P\{\omega:\ \eta(\omega) \leq \alpha \in R\}$. Let x be a Poisson process defined as above, independent of the η_k, and governed by parameter $\lambda > 0$. A compound

Poisson process y may be defined by the expression

$$y(t,\omega) = \begin{cases} \sum_{k=1}^{x(t,\omega)} \eta_k(\omega) & ; \quad x(t,\omega) \geq 1 \\ 0 & ; \quad x(t,\omega) = 0. \end{cases}$$

In words, $y(t,\omega)$ jumps by $\eta_k(\omega)$ at the instant that $x(t,\omega)$ changes from k-1 to k. The distribution function of y(t) is determined as [36]

$$F_{y(t)}(\alpha) = P\{\omega: \ y(t,\omega) \leq \alpha\} = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \tilde{F}^{(n)}(\alpha)$$

where $\tilde{F}^{(n)} = \tilde{F}^{(n-1)} * F_\eta$ and $\tilde{F}^{(1)} = F_\eta$ (* denotes convolution).

Continuing the reasoning of the previous sections, the paragraphs that follow define an operator capable of describing the presence of "random shocks" in a feedback system. The asymptotic properties of such systems are then analysed using this operator.

Let $(X, \mathcal{B}(X))$ be a measureable space and consider the random measure on $\mathcal{B}(R^+) \times \mathcal{B}(X)$ denoted by $\nu([s,t],A)$, $[s,t] \subset R^+$, $A \in \mathcal{B}(X)$, as expressing the number of events in the set A during the interval [s,t]. Assume that the random variable ν takes on non-negative values independent on disjoint sets from $\mathcal{B}(R^+) \times \mathcal{B}(X)$. And for each set $[s,t] \times A \in \mathcal{B}(R^+) \times \mathcal{B}(X)$, assume that $\nu([s,t],A)$ is Poisson with parameter $\int_s^t \Pi(\tau,A)\, d\tau$; i.e.,

$$P\{\omega: \nu(\omega,[s,t],A) = n\} = \frac{1}{n!} \left(\int_s^t \Pi(\tau,A)\, d\tau\right)^n \exp\left(- \int_s^t \Pi(\tau,A) d\tau\right).$$

Here $\Pi(t,A)$ is a probability measure for $\mathcal{B}(X)$ for each $t \in R^+$, and a measureable function $R^+ \to R$ for each $A \in \mathcal{B}(X)$.

It follows that the random process ν is a process with independent increments (on R+); so the stochastic integral

$$\int_0^t \int_X \ell(\tau,x) \quad \nu(d\tau,dx)$$

for non-anticipating random functionals ℓ on $R+ \times X$ such that

$$\int_0^t \int_X E|\ell(\tau,x)|^k \, \Pi(\tau,dx) \, d\tau \qquad , \; k = 1,2; \; t \in R^+,$$

is well-defined as the usual limit of Riemann sums, see also Ito [40] and Gikhman and Dorogovcev [28].

Let $\tilde{\nu}(t,A) = \nu([0,t],A) - \int_0^t \Pi(\tau,A)\, d\tau$, then the following hold

(i) $\quad E\{ \int_0^t \int_X \ell(\tau,x) \; \tilde{\nu}(d\tau,dx) \} = 0$

(ii) $\quad E\{(\int_0^t \int_X \ell(\tau,x) \; \tilde{\nu}(d\tau,dx))^2\} = \int_0^t \int_X E|\ell(\tau,x)|^2 \, \Pi(\tau,dx) \, d\tau$.

Now let the process x be defined on $R^+ \times \Omega$ into X as a non-anticipating ($\mathcal{B}_{0t}(x) \vee \mathcal{B}_{0t}(\nu([0,s],\cdot))$ is independent of $\mathcal{B}_{t\infty}(\nu([s,t],\cdot))$) functional of ν. Let H be an operator on X-valued non-anticipating random functions behaving as follows: if the "input" to H at time t is x(t), then H causes a displacement of x by

$$\int_s^t \int_X h(\tau,x(\tau),y) \; \nu(d\tau,dy)$$

over the interval $[s,t] \subset R^+$. Here h is some (continuous) function mapping $R^+ \times X \times X \to X$.

Recalling the definitions of the last section, the remainder of this section is devoted to an analysis of the integral equation (1) below as a model of a stochastic system with unity feedback (here the space X = R).

$$(1)\quad x(t,\omega) = u(t,\omega) - \int_0^t g(t-s)f(s,x(s,\omega)\; dw(s,\omega)$$
$$- \int_0^t g(t-s) \int_R h(s,x(s,\omega),y)\; \nu(\omega,ds,dy)$$

From Skorokhod [55] the following existence theorem gives conditions under which the equation (1) is well-posed.

<u>Theorem 1</u>: [55,Section 3.3] Assume that the functions u,g,f,h satisfy the following conditions:

(i) $u(\omega)$ for each $\omega \in \Omega$ has only finite jump discontinuities (u is real-valued), and $E\{u(t)^2\} < \infty$ for $t \in [0,T]$, T finite.

(ii) There exists a $K < \infty$ such that for all $t \in R^+$

$$\int_0^t |g(t-s)|^2\; |f(s,x) - f(s,y)|^2\, ds$$
$$+ \int_0^t |g(t-s)|^2 \int_R |h(s,x,\alpha) - h(s,y,\alpha)|^2\; \Pi(s,d\alpha)\; ds$$
$$\leqq\; K|x-y|^2 \quad ; \quad x,y \in R\, .$$

(iii) There exists a $K < \infty$ such that for all $t \in R^+$

$$\int_0^t |g(t-s)| \left|\int_R |h(s,x,y)|\; \Pi(s,dy)\; ds < K(1+|x|) \quad x \in R.\right.$$

Then a solution x of the integral equation (1) exists, is locally bounded almost surely, and has only jump discontinuities. Moreover, if $\sup_{0\leqq t\leqq T} E\{u(t)^2\} < \infty$, then $\sup_{0\leqq t\leqq T} E\{x(t)^2\} < \infty$ for any $T \in R^+$. The solution x is unique at all points of continuity.

Before proceeding to the analysis of the nonlinear equation (1) consider the linear case (corresponding to f and h linear)

$$(2)\ x(t) = u(t) - \int_0^t g(t-s)\ x(s)\ dw(s) - \int_0^t g(t-s)\ x(s) \int_R h(y)\ \nu(ds,dy)$$

where

$$E\{dw(t)\} = m\ dt$$

$$E\{(dw(t) - mdt)^2\} = \sigma^2\ dt$$

$$E\{\nu(dt,A)\} = \Pi(A)\ dt$$

$$E\{(\nu(dt,A) - \Pi(A)dt)^2\} = \Pi(A)\ dt.$$

Assume that u, w, and ν are independent processes. Then clearly, assuming Theorem 1 holds,

$$E\{x(t)\} = E\{u(t)\} - \int_0^t g(t-s)\ E\{x(s)\}\ m\ ds - \int_0^t g(t-s)\ E\{x(s)\}\int_R h(y)\Pi(dy)dt$$

Hence,

Theorem 2: Assume that $g \in L_1(R^+)$ and let $G(s)$ denote the Laplace transform of g. Then $E\{|u(t)|\} < \infty$ implies $E\{|x(t)|\} < \infty$ if and only if

$$(-(m+\hat{\pi})^{-1}, j0) \notin \bigcup_{Re(s)\in R^+} G(s)$$

where $\hat{\pi} = \int_R h(y)\Pi(dy)$.

Now consider the problem of bounding the second moment of x. An easy transformation of equation (2) gives

$$(3)\ x(t) = u(t) - \int_0^t g(t-s)x(s)[m+\hat{\pi}]ds - \int_0^t g(t-s)x(s)d\tilde{w}(s)$$

$$- \int_0^t g(t-s)x(s) \int_R h(y)\ \tilde{\nu}(ds,dy)$$

where $d\tilde{w}(s) = dw(s) - m\ ds$ and $\tilde{\nu}(ds,dy) = \nu(ds,dy) - \Pi(dy)ds$.

Assuming now the conditions of Theorems 1 and 2, the following holds

$$x(t) = (G_1u)(t) - \int_0^t \tilde{g}(t-s)x(s)d\tilde{w}(s) - \int_0^t \tilde{g}(t-s)x(s) \int_R h(y)\ \tilde{\nu}(ds,dy)$$

where G_1 is the linear deterministic convolution whose kernel g_1 has Fourier transform $G_1(j\alpha) = [1 + (m+\hat{\pi})G(j\alpha)]^{-1}$ and the kernel $\tilde{g}$ has Fourier transform $\tilde{G}(j\alpha) = G(j\alpha)G_1(j\alpha)$. In this case

$$E\{x(t)^2\} = \int_0^t \int_0^t g_1(t-s)g_1(t-r)E\{u(s)u(r)\}dsdt + \int_0^t \tilde{g}^2(t-s)E\{x(s)^2\}(\sigma^2+\hat{\pi})ds$$

from which the following is clear.

<u>Theorem 3</u>: Let $g \in L_1(R^+)$, and assume that Theorems 1 and 2 apply, then $\sup_{t\in R^+} E\{u(t)^2\} < \infty$ implies $\sup_{t\in R^+} E\{x(t)^2\} < \infty$ if and only if

(i) $(-(m + \hat{\pi})^{-1},j0) \notin \bigcup_{Re(s)\in R^+} G(s)$

and (ii) $||\tilde{g}||_2 < (\hat{\pi} + \sigma^2)^{-1/2}$

or equivalently,

(ii)' $$\int_{-\infty}^{\infty} \left| \frac{G(j\alpha)}{1 + (\hat{\pi} + m)G(j\alpha)} \right|^2 d\alpha < 2\pi(\hat{\pi} + \sigma^2) .$$

As an illustrative example, consider the linear convolution represented by $G(s) = k/(s+p)$, then

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{G(j\alpha)}{1 + (m + \hat{\pi})G(j\alpha)} \right|^2 d\alpha = \frac{k^2}{2(p + m + \hat{\pi})} .$$

Hence, $\sup_{t\in R^+} E\{x(t)^2\} \leq \beta \sup_{t\in R^+} E\{u(t)^2\}$ for some $\beta \in R^+$ if and only if

$$\frac{k^2 - (\sigma^2 + \hat{\pi})(m + \hat{\pi})}{2(\hat{\pi} + \sigma^2)} < p.$$

Two sufficient conditions were proved in [66] for a special case of (2) (corresponding to $\nu \equiv 0$); these may be modified to apply in this case, and they yield conditions more easily checked for a given kernel g than the criteria (ii) or (ii)' of Theorem 3.

Corollary 4: Assume that $g \in L_1(R^+)$. Then for equation (3) $\sup_{t\in R^+} E\{x(t)^2\} \leq \beta \sup_{t\in R^+} E\{u(t)^2\}$ for some $\beta \in R^+$ if there exists a $\gamma \in R$ such that

(i) $\dfrac{\sigma^2 + \hat{\pi}}{m+\hat{\pi}+\gamma}\ \tilde{g}(0) < 1$,

(ii) and either of the following conditions is satisfied

(a) $(m + \hat{\pi})/\gamma > 0$, and the Nyquist locus $\bigcup_{\alpha\in R} G(j\alpha)$ lies inside the circle centered on the real axis of the complex plane at $(\frac{1}{2}\gamma^{-1}, j0)$ and passing through the origin.

(b) $-1 < (m + \hat{\pi})/\gamma < 0$, and the Nyquist locus $\bigcup_{\alpha\in R} G(j\alpha)$ lies inside the disc centered on the real axis at $(\frac{1}{2}\gamma^{-1}, j0)$ and passing through the origin.

(c) $(m + \hat{\pi})/\gamma < -1$, and the Nyquist locus $\bigcup_{\alpha\in R} G(j\alpha)$ does not intersect or encircle the disc centered at $(\frac{1}{2}\gamma^{-1}, j0)$ passing through the origin.

Proof: By Theorem 3 it suffices to show that

$$(\sigma^2 + \hat{\pi})\frac{1}{2\pi} \int_{-\infty}^{\infty} |G(j\alpha)/(1 + (m + \hat{\pi})G(j\alpha))|^2\, d\alpha < 1.$$

Using the restrictions on the graph of $G(j\alpha)$, it follows that

$$\left| \frac{(m + \hat{\pi})G(j\alpha)}{1 + (m + \hat{\pi})G(j\alpha)} \right|^2 \leq [1 + \gamma(m + \hat{\pi})^{-1}]^{\pm 1}\ \mathrm{Re}\frac{(m + \hat{\pi})G(j\alpha)}{1 + (m + \hat{\pi})G(j\alpha)} .$$

Thus,

$$\frac{(\sigma^2 + \hat{\pi})}{2\pi} \int_{-\infty}^{\infty} \left| \frac{G(j\alpha)}{1 + (m + \hat{\pi})G(j\alpha)} \right|^2 d\alpha \leq \frac{\sigma^2 + \hat{\pi}}{m + \hat{\pi} + \gamma} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{G(j\alpha)}{1 + (m + \hat{\pi})G(j\alpha)} d\alpha$$

$$= \frac{\sigma^2 + \hat{\pi}}{m + \hat{\pi} + \gamma} \tilde{g}(0)$$

The last step using $g \varepsilon L_1(R^+)$, and the assumption that zero is a Lebesque point of $\tilde{g}$ [1, p.5].

QED

The next result is a special case of Corollary 4 as $\gamma \to 0$.

Corollary 5: Assume that $g \varepsilon L_1(R^+)$, then for equation (3), $\sup_{t \varepsilon R^+} E\{x(t)^2\}$ $\leq \beta \sup_{t \varepsilon R^+} E\{u(t)^2\}$ for some $\beta \varepsilon R^+$ if

(i) $m + \hat{\pi} > (\sigma^2 + \hat{\pi})\, \tilde{g}(0)$

and (ii) $\mathrm{Re}G(j\alpha) \geq 0$ for all $\alpha \varepsilon R$.

While Corollary 5 involves a "passivity" property of the operator G, Corollary 4 is reminiscient of the various "circle criteria" introduced above (sections 2.2, 3.2, and Theorem 6 below), and its primary use is to provide easily verified conditions for moment bounds in the equations being considered. That is, for any of the integral conditions given above (Theorem 3, Theorems 3.2.3 and 3.2.4) sufficient conditions may be derived directly in terms of restraints on the kernel g rather than the quantity $||\tilde{g}||_2$ appearing in the results mentioned by using arguments similar to those in the proof of Corollary 4.

Returning then to the analysis of the nonlinear equation (1), assume that

$$E\{dw(t)\} = 0$$

$$E\{[dw(t)]^2\} = \sigma^2 dt$$

$$E\{\nu(dt,dy)\} = \Pi(dy)dt$$

$$E\{[\nu(dt,dy) - \Pi(dy)dt]^2\} = \Pi(dy)dt$$

and that there exist constants a,b,c,d such that

$$0 < a \leqq f(t,x)/x \leqq b < \infty; \quad t\varepsilon\ R^+, \quad x \in R ;$$

$$0 < c \leqq h(t,x,y)/x \leqq d < \infty; \quad t \in R^+, \quad x,y \in R .$$

Moreover, assume that $E\{u\} = 0$ and that u, w, and ν are independent processes.

Theorem 6: For equation (1) under the assumptions of the last paragraph

$$\sup_{t\varepsilon R^+} E\{x(t)^2\} \leqq \beta \sup_{t\varepsilon R^+} E\{u(t)^2\} \qquad \text{for some } \beta\ \varepsilon R^+ \text{ if}$$

(i) There exists an $r_o \geqslant 0$ such that

$$\int_0^\infty \exp(r_o t)|\, g(t)|\, dt < \infty .$$

(ii) $\tilde{\pi} = \int_R \Pi(dy) < \infty$.

(iii) $\{[-\tilde{\pi}(c+d)/2]^{-1}, j0\} \notin \bigcup_{Re(s) \geqq -r_o} G(s)$

(iv) For $\tilde{G}(s) = G(s)[1 + \frac{1}{2}\tilde{\pi}(c+d)G(s)]^{-1}$ (see (iii)) and $\tilde{G}_2 = \tilde{G} * \tilde{G}$, then

$$\left\{(\tfrac{1}{2}[\sigma^2(a^2+b^2) + \tilde{\pi}(c^2+d^2)])^{-1}, j0\right\} \notin \bigcup_{Re(s) > -r_o} \tilde{G}_2(s)$$

(v) For some $\alpha\ \varepsilon(0,1)$ and $r \in (0,r_o)$

$$\sup_{Re(s) \geqq -r} |\tilde{G}_2^{-1}(s) + \tfrac{1}{2}[\sigma^2(a^2+b^2) + \tilde{\pi}(c^2+d^2)]| > \frac{\alpha}{2}[\sigma^2(b^2-a^2) + \tilde{\pi}(d^2-c^2)]$$

and (vi) For some $r \in (0,r_o)$ and

$$H(r+j\xi) = \int_{-\infty}^{\infty} \frac{\tilde{G}(r + j(\xi - \xi_o))\tilde{G}(r + j\xi_o)}{(r + j(\xi - \xi_o))(r + j\xi_o)}\, d\xi_o$$

then

$$\sup_{\xi\in R}\left| \frac{\frac{1}{2} H(r + j\xi)\ \tilde{\pi}^2(d^2-c^2)}{1 + \frac{1}{2}[\sigma^2(a^2+b^2) + \tilde{\pi}(c^2+d^2)\tilde{G}_2(r+j\xi)]} \right| < 1 - \alpha$$

<u>Proof</u>: A transformation of (1) gives

$$x(t) = u(t) - \int_0^t g(t-\tau) \int_{-\infty}^{\infty} \tilde{h}(\tau,x(\tau),y)\ \Pi(dy)d\tau$$

$$- \frac{1}{2}\tilde{\pi}(c+d) \int_0^t g(t-\tau)x(\tau)d\tau - \int_0^t g(t-\tau)f(\tau,x(\tau))dw(\tau)$$

$$- \int_0^t g(t-\tau) \int_{-\infty}^{\infty} \tilde{h}(t,x(\tau),y)\ \tilde{\nu}(d\tau,dy)$$

where $\tilde{\nu}(d\tau,dy) = \nu(d\tau,dy) - \Pi(dy)d\tau$ and $\tilde{h}(t,x,y) = h(t,x,y) - \frac{1}{2}(c+d)x$ and $\tilde{\pi}$ is defined in condition (ii). Let $W(s) = [1 + \frac{1}{2}\tilde{\pi}(c+d)G(s)]$, then by (ii) and (iii) and from, for example, [12], W^{-1} exists on $L_\infty(R^+)$ functions. Hence,

$$x(t) = (W^{-1}u)(t) - \int_0^t \tilde{g}(t-\tau) \int_{-\infty}^{\infty} \tilde{h}(\tau,x(\tau),y)\ \Pi(dy)d\tau$$

$$- \int_0^t \tilde{g}(t-\tau)f(\tau,x(\tau))\ dw(\tau) - \int_0^t \tilde{g}(t-\tau) \int_{-\infty}^{\infty} h(\tau,x(\tau),y)\tilde{\nu}(d\tau,dy)$$

where G the Fourier transform of g is defined above. Then taking into account the assumptions on u, w, and ν:

$$E\{x(t)^2\} = E\{(W^{-1}u)^2(t)\} + E\{(\int_0^t \tilde{g}(t-\tau) \int_{-\infty}^{\infty} \tilde{h}(\tau,x(\tau),y)\ \Pi(dy)d\tau)^2\ \}$$

$$+ \int_0^t \tilde{g}^2(t-\tau)\ E\{f^2(\tau,x(\tau))\}\sigma^2\ d\tau$$

$$+ \int_0^t \tilde{g}^2(t-\tau) \int_{-\infty}^{\infty} E\{h^2(\tau,x(\tau),y)\}\Pi(dy)\ d\tau \quad .$$

Again adding and subtracting the terms

$$\frac{1}{2}\sigma^2(a^2+b^2) \int_0^t \tilde{g}^2(t-\tau)\ E\{x(\tau)^2\}\ d\tau$$

$$\frac{1}{2}\tilde{\pi}(c^2+d^2) \int_0^t \tilde{g}^2(t-\tau)\ E\{x(\tau)^2\}\ d\tau$$

the result is

$$E\{x(t)^2\} + \frac{1}{2}[\sigma^2(a^2+b^2) + \tilde{\pi}(c^2+d^2)] \int_0^t \tilde{g}^2(t-\tau)\ E\{x(\tau)^2\}\ d\tau$$

$$= E\{(W^{-1}u)^2(t)\} + E\{(\int_0^t \tilde{g}(t-\tau) \int_{-\infty}^{\infty} \tilde{h}(\tau,x(\tau)y)\ \Pi(dy)d\tau)^2\}$$

$$\pm\ \sigma^2 \int_0^t \tilde{g}^2(t-\tau)\ E\{\hat{f}^2(\tau,x(\tau))\}\ d\tau$$

$$+ \int_0^t \tilde{g}^2(t-\tau) \int_{-\infty}^{\infty} E\{\hat{h}^2(\tau,x(\tau),y)\ \Pi(dy)d\tau$$

where $\hat{f}^2(t,x) = f^2(t,x) - \frac{1}{2}(a^2+b^2)x^2$ and $\hat{h}^2(t,x,y) = h^2(t,x,y) - \frac{1}{2}(c^2+d^2)x^2$. Setting $\hat{K}$ to be the linear convolution operator whose Fourier transform is $\hat{K}(s) = [1 + \frac{1}{2}[\sigma^2(a^2+b^2) + \tilde{\pi}(c^2+d^2)]\tilde{G}_2(s)]^{-1}$, where $\tilde{G}_2(s)$ is defined in the theorem statement, and using (iv)

$$(4)\quad E\{x(t)^2\} = \hat{K}(E\{(W^{-1}u)^2\})(t)$$
$$+ \int_0^t \hat{k}(t-\tau)E\{(\int_0^\tau \tilde{g}(\tau-s)\int_{-\infty}^{\infty}\tilde{h}(s,x(s),y)\ \Pi(dy)ds)^2\}$$
$$+ \sigma^2 \int_0^t \hat{g}(t-s)\ E\{\hat{f}^2(s,x(s))\}\ ds$$
$$+ \int_0^t \hat{g}(t-s)\int_{-\infty}^{\infty} E\{\hat{h}^2(s,x(s),y)\}\ \Pi(dy)\ ds$$

where $\hat{g}$ is the kernel whose transform is $\hat{G}(s) = \tilde{G}(s)\hat{K}(s)$ and $\hat{K} \sim \hat{k}$.

Using the bounds,

$$|\hat{f}^2(s,x)| \leq \frac{1}{2}(b^2-a^2)x^2 \quad, \text{ for every } s \in R^+,$$
$$|\hat{h}^2(s,x,y)| \leq \frac{1}{2}(d^2-c^2)x^2 \quad, \text{ for every } s \in R^+,\ y \in R,$$

and condition (v) it is clear that the last two terms in equation (4) are bounded by $\alpha \sup_{0\leq s\leq t} E\{x(s)^2\}$. Closer consideration of the decisive term (T2) second on the right of (4) will yield the desired conclusion.

Expanding the square

$$\int_0^\tau\int_0^\tau \tilde{g}(\tau-s)\tilde{g}(\tau-\mu)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\{\tilde{h}(s,x(s),y)\tilde{h}(\mu,x(\mu),z)\ \Pi(dy)\Pi(dz)\ ds\ d\mu$$

$$\leq \int_0^\tau\int_0^\tau \tilde{g}(\tau-s)\tilde{g}(\tau-\mu)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [E\{\tilde{h}^2(s,y)\}]^{1/2}\ [E\{\tilde{h}^2(\mu,z)\}]^{1/2}\Pi(dy)\Pi(dz)dsd\mu$$

$$\leq \frac{1}{2}\tilde{\pi}^2(d^2-c^2)(\int_0^\tau \tilde{g}(\tau-s)\ ds)^2 \sup_{0\leq s\leq\tau} E\{x(s)^2\}\ .$$

Hence,

$$T2 \leq \frac{1}{2}\tilde{\pi}^2(d^2-c^2)\int_0^t |\hat{k}(t-\tau)|\ (\int_0^\tau \tilde{g}(s)ds\ )^2 \sup_{0\leq s\leq\tau} E\{x(s)^2\}d\tau\ ,$$

And so, condition (vi) implies that

$$|T2| < (1-\alpha) \sup_{0 \leq s \leq t} E\{x(s)^2\}$$

and that the combined operator composed of T2 and the sum of the last two terms is a contraction on the Banach space defined by the norm $||x|| = [\sup_{s \in R^+} E\{x(s)^2\}]^{1/2}$. The conclusion of the Theorem follows easily from this point using familiar arguments from the earlier sections.

QED

While Theorem 6 may be regarded as a direct generalization of Theorem 3.2.3 (nonlinear convolution versus a Wiener process), the comparatively more complicated conditions (i)-(vi) of Theorem 6 would seem to preclude the graphical interpertation possible for the conditions of the earlier theorem. No attempt will be made here to weaken Theorem 6 to permit such an interpertation, though the promise of such a procedure is acknowledged.

In order to complete the extension begun in this section it is necessary to prove the analog of Theorem 3.2.5 using Theorem 6 to prove asymptotic invariance of the solution of equation (1) under appropriate assumptions on u, w, and ν. While conceptually no more difficult, the statement and proof of the analog is technically more complex because of the nature of the solution sample paths of equation (1). Recall that the basic existence theorem for this situation (Theorem 1 here) adapted from [55] guarantees only that the solution trajectories will be piecewise continuous. It is therefore necessary to discuss weak convergence of distributions on spaces of piecewise continuous functions. Recall that in section 2.4, it was rather easy to determine conditions for a set of distributions on the space of continuous functions to be compact by using a modification of the Ascoli Theorem [16] to characterize compact sets of continuous functions and Prohorov's Theorem (2.4.5) .

Needed thus, are a topology on the set of piecewise continuous functions rendering them separable and complete (so that Theorem 2.4.2 will be necessary and sufficient in this case) and a characterization of the compact subsets in this topology. Combining the work of Skorokhod [56], Billingsly [7], and Stone [58] the necessary framework is available. Rather than state this technical structure and then prove the theorem, the result will be stated, and the appropriate elements of the theory of weak convergence of measures on piecewise continuous functions used in the proof stated as lemmas.

Theorem 7: Consider the equation (1) under the assumptions

(i) f and h satisfy the sector conditions with the parameters (a,b) and (c,d) respectively.

(ii) $E\{dw(t)\} = 0$ and $E\{[dw(t)]^2\} = \sigma^2\,dt$.

(iii) $E\{\nu(dt,dy)\} = \Pi(dy)dt$ and $E\{[\nu(dt,dy) - \Pi(dy)dt]^2\} = \Pi(dy)dt$,

(iv) u, w, ν are independent, u is piecewise continous (from the right) almost everywhere (P), and $E\{u(t)\} = 0$; $E\{u(t)^2\}\varepsilon L_\infty(R^+)$. For s,t points of continuity (almost sure) of u and $\tau\ \varepsilon[s,t]$

$$E\{|u(t) - u(\tau)|^{1/2}|u(\tau) - u(s)|^{1/2}\} \leq \gamma|t-s|^2$$

(v) The kernel $g\ \varepsilon\ L_1(R^+) \cap L_\infty(R^+)$. (Much less restrictive conditions are possible here.)

Then the criterion of Theorem 6 is sufficient to guarantee the asymptotic invariance of the solution process x.

Outline of Proof:

Definition 8: Let $D(R^+;R)$ denote the space of real-valued fuctions on R^+, which have a limit from the right and are continuous from the left.

Elements of $D(R^+;R)$ are bounded on compact intervals, and for any $\varepsilon > 0$ have at most a finite number of jumps of amplitude greater than ε in any bounded interval [7].

Lemma 9: [55,section 3.3] The existence Theorem 1 implies that $x \in F(\Omega;D)$, the set of D-valued random variables on Ω, if $u \in F(\Omega;D)$.

Lemma 10: [55],[58],[7, p.115] A metric d_o exists on $D(R^+;R)$ such that (D,d_o) is a complete, separable metric space.

This lemma assures that Theorem 2.4.2 applies in its full power on (D,d_o).

Lemma 11: [7] For a subset J of $D(R^+;R)$ to be relatively compact (with respect to d_o) it is necessary and sufficient that for every $T \in R^+$, and partition $\{t_i\}_{i=1}^{r}$ of $[0,T]$

$$\sup_{f\in J}\ \sup_{t\in[0,T]} |f(t)| < \infty ,$$

$$\lim_{\delta\to 0}\ \sup_{f\in J}\ \inf_{\{t_i\}}\ \max_{0\leq i\leq r} \{|f(t)-f(s)| \ ; \ t,s \in [t_i,t_{i+1})\} = 0,$$

where $\delta = \max_i \{t_i - t_{i-1}\}$ is the size of the partition.

This result is the counterpart of the Ascoli Theorem defining compact sets of continuous functions. The necessary convergence criterion (compare Corollaries 2.4.7 and 2.4.8) is provided by:

Lemma 12: A subset $\Lambda \subset F(\Omega;D)$ is totally L-bounded if the following conditions are satisfied for any sequence $\{x_n\} \subset \Lambda$:

(i) The sequence $\{x_n(0)\}$ is tight

(ii) For s,t continuity points of x_n and any $\tau \in [s,t]$

$$E\{|x_n(t)-x_n(\tau)|^{\alpha}|x_n(\tau)-x_n(s)|^{\alpha}\} \leq \rho|t-s|^{2} ,$$

for $\alpha \geq 0$, $\beta > 1/2$, and some $\rho > 0$, all independent of n.

The proof of the Theorem 7 then proceeds to verify the inequality of Lemma 12 (ii) along the lines of the proof of Theorem 3.2.5, the particular values of α and β used are 2 and 1 respectively. The proof is, however, tedious and somewhat removed from the main focus of this work and will be omitted.

In the next section the properties of the solutions of differential eqaations subject to totally L-bounded inputs is examined. Conditions on the cofficients of the equations are found to guarantee that the solution is totally L-bounded when the driving function has this property.

3.4 Differential Equations with Totally Bounded Inputs:

In order to illuminate the results of the earlier sections of this chapter it is worthwhile to consider them in the usual setting provided by stochastic differential equations. This section consists of two distinct parts. First a general class of nonlinear functional differential equations is considered and conditions for L-total boundedness of the solution given. By assuming the functional coefficients in this equation to be memoryless functions the solution becomes a diffusion (strong Markov process), and the latter portion of this section contains a few remarks on this case.

Following Fleming and Nisio [26] (see also Ito and Nisio [41]), consider the functional stochastic differential equation

(1) $dx(t) = (a(\pi_t x))(t)\ du(t) + (b(\pi_t x)(t))\ dw(t)$

where a and b are continuous functionals on $C(R^-;R)$ (the R-valued continuous functions on the negative real line R^-, with the metric d introduced in section 3.2); w is a Wiener process on $(\Omega, \mathcal{F}, P)$; u is a control to be specified later; and π_t is the truncation operator. An initial function $x\_$ such that $x(t) = x\_(t)$, $t \in R^-$, completes the specification of the equation. Assume the initial function $x\_$ is an element of $F(\Omega;C(R^-;R))$.

Let $U_r \subset C(R^+;R)$ be the subset of the continuous functions satisfying the Lipschitz condition below: for $f \in U_r$

$$|f(t) - f(s)| \leq \gamma|t-s| \ ; \quad t,s \in R^+, \quad f(0) = 0,$$

for some constant γ independent of f. Let U_r have the relative topology induced as a closed subset of $(C(R^+),d)$. Let $S(\Omega;U_r)$ be the set of U_r-valued random variables (signals because the half-line R^+ is the time set).

<u>Proposition 1</u>: (i) Let $PM(U_r)$ have the Prohorov topology, then $PM(U_r)$ is relatively compact. (ii) As a subset of $F(\Omega;C)$, the $C(R^+)$-valued random variables, $S(\Omega;U_r)$ is totally L-bounded.

<u>Proof</u>: (i) It is easy to verify that (U_r,d) is a compact (hence complete and separable) subset of $(C(R^+),d)$. Part (i) follows from this observation and Prohorov's Theorem (2.4.5 here). Part (ii) is immediate from (i) or from Billingsly's result (Theorem 2.4.6).

<u>QED</u>

Thus, the set of stochastic processes permitted as inputs is, in the terminology of section 2.4, totally L-bounded. The Lipschitz condition, though severe, is not altogether uncommon in the literature dealing with stochastic control; see for instance, Fleming and Nisio [26], and Fleming [24] for some remarks on this assumption. It is a natural constraint in the framework of the studies here.

A few assumptions are in order on the coefficients in (1) and on the past condition x_-. Assume

(i) a,b are continuous on $C(R^+),d)$.

(ii) For $f \in C(R^+)$, $t \in R$,

$$|a(f)(t)| + |b(f)(t)| \leq \int_{-\infty}^{0} |f(s)|dK(s)$$

for some measure dK, $\int_{-\infty}^{0} dK(s) < \infty$.

(iii) $E\{x_-(t)^4\} \leq c$, $t \leq 0$, for some $c<\infty$.

(iv) $\mathcal{B}_{0t}(u) \vee \mathcal{B}_{-\infty 0}(x_-) \vee \mathcal{B}_{0t}(dw)$ is independent of $\mathcal{B}_{t\infty}(dw)$ for every $t \in R^+$.

<u>Theorem 2</u>: [26,p. 783] Under assumptions (i) through (iv) above, equation (1) has a unique solution x with locally bounded second moments, such that $x \in F(\Omega;C(R^+))$ and $\mathcal{B}_{0t}(x) \subset \mathcal{B}_{-\infty 0}(x_-) \vee \mathcal{B}_{0t}(u) \vee \mathcal{B}_{0t}(dw)$ for every $u \in S(\Omega;U_r)$.

Let $\Xi = \{(\tilde{x}_-,u,w)\}$ the collection of triples such that $\tilde{x}_-$ has the same probability law as x_- on $C(R^+)$, $u \in S(\Omega;U_r)$, and w is a standard Wiener process. Following Fleming and Nisio [26], let s denote the generic element of Ξ and let $\Psi = \{x_s : s \in \Xi\}$ denote the set of solutions generated by elements of Ξ.

<u>Theorem 3:</u> [26,p. 787] The set Ψ is a sequentially compact subset of $(S(\Omega;C(R^+)),L)$ where L denotes the Prohorov metric.

C-2

Thus, for all admissible inputs, the state x is confined to a compact subset of the state space, in this case $S(\Omega;C(R^+))$. Hence, using the same techniques employed in section 3.2, for any element s of Ξ, it is possible to show that the distributions of x_s, the corresponding solution, are convergent in the Cesaro sense used there to an invariant distribution on $C(R^+)$. In Ito and Nisio [41] a rather more detailed treatment of equation (1) is presented for the case when $du(t) = dt$, corresponding in a sense to an autonomous system.

Though giving the desired analogy to the results of sections 3.2 and 3.3, Theorem 3 was used for quite a different purpose in [26]. Consider the problem of selecting a control u from $S(\Omega;U_r)$ to minimize the functional $E\{\Phi(x,u)\}$ where Φ is some positive (values in R^+), continuous functional on $C(R^+) \times U_r$ (jointly).

Theorem 4: [26,p.792] Let $\Xi_1 \subset \Xi$ be closed under L-sequential limits, then there exists an element $s_1 \in \Xi$ such that $E\{\Phi(x_1,u_1)\} \leq E\{\Phi(x,u)\}$ for any other $s \in \Xi_1$. Here x_1 (x) is the solution of (1) corresponding to s_1 (s).

As a theorem in stochastic control theory the above result has a proper place as a preliminary existence theorem; however, it suffers from being non-constructive and from requiring "total knowledge" of the state x. The existence problem in optimal stochastic control theory is in any case very difficult, and attempts to proceed beyond theorems of this nature have not been altogether successful. Some recent work holding the promise of a solution to the problem is contained in the papers of Beneš [3], [4] and Duncan and Variaya [15], and the comprehensive survey of Fleming [24].

In the present context Theorem 4 serves to illustrate the earlier sections of this chapter by providing an alternative application for the mathematical techniques involved. Note that in the equation (1), the solution x need not be a Markov process. The same observation holds in the work in the references [3],[4], [15]. If the solution x is a Markov process, the additional mathematical structure available has compelling consequences. In the remaining paragraphs of this section some of the important aspects of this case will be summarized. As most of the analysis of stochastic systems has been done in this setting only a few of those results related to this research will be presented.

Consider the stochastic equation (all elements are real valued)

$$dx(t,\omega) = a(t,x(t,\omega))dt + b(t,x(t,\omega))dw(t) \;;\quad x(0,\omega) = x_0(\omega),\ t \in R^+, \omega\in\Omega \ .$$

As usual this equation is but a shorthand for the integral equation

$$(2)\quad x(t)-x(s) = \int_s^t a(s,x(s))ds + \int_s^t b(s,x(s))dw(s) \;;\quad t,s \in R^+ \ .$$

Here subject to the assumption of Lipschitz continuity on the coefficient functions a and b, and the assumption that $\mathcal{B}(x_0)$ is independent of $\mathcal{B}_{0\infty}(dw)$ a solution of (2) may be shown to exist as an element of $S(\Omega;C(R^+))$. Moreover, from the form of (2) it is easy to see that the solution x is a Markov process. In fact x is a strong Markov process (begins afresh at random times, see Ito [39] or McKean [50]), and so is a diffusion.

As a markov process the solution x is characterized by a transition operator (Dynkin [19, Chapter 3]) $P: R^+ \times R^+ \times R \times \mathcal{B}(R) \to R$. Here $P(t,s,x_0,E)$ expresses the probability that at time $t \in R^+$, starting in state x_0 at time $s \in R^+$ ($s \leq t$), the solution $x(t)$ is an element of $E \in \mathcal{B}(R)$. The following properties characterize P:

(i) For every $(t,s,x) \in R^+ \times R^+ \times R$, $P(t,s,x,\cdot)$ is a probability measure on $\mathcal{B}(R)$.

(ii) For every $E \in \mathcal{B}(R)$, $P(t,s,x,E)$ is a measureable function of t,s,x jointly on the appropriate domain.

(iii) For every (t,s,x,E) and $r \in [s,t]$

$$P(t,s,x,E) = \int_R P(t,r,y,E)\, P(r,s,x,dy) \; .$$

(iv) $P(s,s,x,E) = 1_E(x)$ for every $s \in R^+$.

Condition (iii) is the familiar Chapman-Kolmogorov equation for Markov processes [19]. This key property of P defines a two-parameter family of operators on $L_\infty(R)$, the bounded measureable functions mapping R into itself, according to the rule

$$T_{t,s} f(x) = \int_R f(y)\, P(t,s,x,dy) \quad .$$

For those fuctions f for which it exists, the limit

$$L_\infty - \lim_{t \downarrow s} \frac{T_{t,s} f - f}{t - s}$$

defines the operator

$$(A_s f)(x) = a(s,x) \frac{\partial f}{\partial x}(x) + \frac{1}{2} b^2(s,x) \frac{\partial^2 f}{\partial x^2}(x)$$

whose domain $\mathcal{D}(A_s)$ includes at least $C_0^2(R;R)$, the space of functions $R \to R$, having compact support and two continuous derivatives. See for instance [19] for more details.

On the set of probability measures on R (PM(R)) the transition operator P defines a second two parameter-family of operators, for $\mu \in PM(R)$

$$(U_{t,s}\mu)(E) = \int_R P(t,s,x,E)\ \mu(dx)$$

In a sense (which may be made precise [19]) U may be regarded as the adjoint of T. Observe that U defines the evolution of the probability distribution of x, the solution of (2). That is, if μ_s is the distribution of the initial state x(s), then $U_{t,s}\ \mu_s$ is that of x(t), t > s, on $\mathcal{B}(R)$.

If P as a measure on $\mathcal{B}(R)$ has a "derivative" (Radon-Nikodym [30]) with respect to Lebseque measure dy, then denoting this function by p:

$$P(t,s,x,E) = \int_R p(t,s,x,y)\ dy \quad .$$

Moreover, the function p (which may be a generalized function if need be) satisfies the equations, for $0 \leq s \leq t$

(3a) $$A_s p = a(s,x)\ \frac{\partial p}{\partial x}(t,s,x,y) + \frac{1}{2}\ b^2(s,x)\ \frac{\partial^2 p}{\partial x^2}(t,s,x,y) = -\frac{\partial p}{\partial s}(t,s,x,y),$$

(3b) $$A_t^* p = -\frac{\partial[a(t,y)p(t,s,x,y)]}{\partial y} + \frac{1}{2}\,\frac{\partial^2[b^2(t,y)p(t,s,x,y)]}{\partial y^2} = \frac{\partial p(t,s,x,y)}{\partial t}$$

Here A_s is the "generator" of $T_{t,s}$ and A_t^* is formally its adjoint. Of course (3a) and (3b) are the well-known Kolmogorov backward and forward equations. The latter is also frequently called the Fokker-Planck equation. For the (3b) the <u>fundamental solution</u> is generated by the initial condition (δ = the Dirac function) $p(s,s,x,y) = \delta(x-y)$. And for (3a) $p(t,t,x,y) = 1_\Gamma(x)$ defines $P(t,s,x,\Gamma)$ for $0 \leq s \leq t$.

Note that (3b) makes little sense unless thecoefficients a and b are sufficiently smooth. Equation (3a) has the obvious advantage that it applies even if the coefficients are not well-behaved. Moreover, it is known [19] that if a and b are bounded and Lipschitz (Holder) continuous*, and b^2 is everywhere positive-definite, then (3a) has a smooth, unique, fundamental solution. This solution precisely defines the distribution of the process x corresponding to A_s, starting from any initial distribution, according to

$$(U_{t,s}\mu)(E) = \int_R p(t,s,x,y)\ \mu(dy)$$

where μ is the distribution (on $\mathcal{B}(R)$) of x(s).

A modification of this concept yields a means of solving arbitrary equations of the form

$$(4) \qquad A_s u = \frac{\partial u}{\partial t}, \ u(s,x) = f(x).$$

That is, since $E_{t,x} f(x(s)) = \int_R p(t,s,x,y)\ f(y)\ dy$ ($E_{t,x}\xi$ is the expectation of ξ conditioned on $x(t) = x$), then clearly $u(t,x) = E_{t,x} f(x(s))$ "solves" (4). See [19, Chapter 13] for more details. Taking into account the interpertations afforded by the stochastic differential equation for x, this solution method is more than a tautology.

The problem corresponding to the analysis of the past three sections in this setting is to study the behavior of the function p(t,s,x,y) as a solution of (3a) as t-s approaches infinity. In other than specific instances this analysis uses certain auxillary functions with properties similar to Lyapunov functions. For the case of time-varying coefficients (a and b) under consideration here the best result is due to Il'in and Khas'minskii [38]:

---

*See [59] for an analysis of equations with less restricted coefficients.

Theorem 5: [38,p.248] Let $p(t,s,x,y)$ be the fundamental solution of (3a) for $t \geq s$. Let $V(t,r)$ be a positive function, monotonically decreasing with respect to t, and nondecreasing with respect to $r \geq 0$, such that for all s

(i) $$a(t-s,x) \frac{\partial V(t,|x|)}{\partial x} + \frac{1}{2} b^2(t-s,x) \frac{\partial^2 V(t,|x|)}{\partial x^2} < 0$$

(ii) $V(0,r) \geq 1$; $r \geq 0$.

(iii) $\int_{R^+} V(t,r)\, dt < \infty$; $r \geq 0$.

Then for every measureable function f, $\int_R p(t,s,x,y)\, f(y)\, dy \to \alpha$,

as $t-s \to \infty$, where α is a constant and $\alpha > 0$ if $f(x) \geq 0$.

Proof: Put

$$u(t,x) = \int_R p(t,s,x,y)\, f(y)\, dy$$

in Theorem 3 of [38] and the result follows.

Corollary 6: [38,p. 255] Let $a(t,x)$ be bounded for all $x \in R$, $t \geq s$, and $a(t,x) + xb(t,x) < -\beta < 0$, then the conclusion of Theorem 5 holds.

If the coefficients are time-invariant in (2), that is,

(5) $dx(t) = a(x(t))\, dt + b(x(t))dw(t)$

then the Markov process x may be described by a transition operator $P: R^+ \times R \times \mathcal{B}(R) \to R^+$. In this case $P(t,x,E)$ gives the probability that $x(t) \in E$ given that $x(0) = x$. The sets of operators $\{T_t\}_{t \in R^+}$ and $\{U_t\}_{t \in R^+}$

$$(T_t f)(x) = \int_R f(y) P(t,x,dy)$$

$$(U_t \mu)(E) = \int_R P(t,x,E) \mu(dx)$$

are in this case semigroups, $T_t \circ T_s = T_{t+s}$ and $U_t \circ U_s = U_{t+s}$, as a consequence of the Chapman-Komogorov relations. The infinitesimal generator A of T is defined as the limit

$$L_\infty - \lim_{t \downarrow s} \frac{T_t f - f}{t} = Af .$$

Here $Af = a(x)\dfrac{\partial f}{\partial x} + \dfrac{1}{2} b^2(x) \dfrac{\partial^2 f}{\partial x^2}$, and the equations (3a) and (3b) for the density function p of P are

$$\text{(6a)} \quad \frac{\partial p(t,x,y)}{\partial t} = a(x)\frac{\partial p(t,x,y)}{\partial x} + \frac{1}{2} b^2(x) \frac{\partial^2 p(t,x,y)}{\partial x^2}$$

$$\text{(6b)} \quad \frac{\partial p(t,x,y)}{\partial t} = -\frac{\partial[a(y)p(t,x,y)]}{\partial y} + \frac{1}{2}\frac{\partial^2[b^2(y)p(t,x,y)]}{\partial y^2}$$

Or concisely,

$$\text{(6a)}' \quad \partial p/\partial t = Ap \quad , \quad p(0,x,y) = \delta(x-y)$$

$$\text{(6b)}' \quad \partial p/\partial t = A^*p \quad , \quad p(0,x,y) = 1_E(x) \quad .$$

The problem corresponding to Theorem 5 above is to establish the existence of an invariant distribution for x. Such a distribution is an element μ of PM(R) such that $\mu = U_t\mu$ for every $t \geq 0$. It is an equivalent problem to look for solutions to $A^*u = 0$. For let u be the density of the invariant measure μ with sespect to Lebesque measure, $\mu(E) = \int_E u(x)\,dx$, and let p(t,x,y) be the density of P(t,x,E). Then again the definition of an invariant distribution is $\mu(E) = (U_t\mu)(E)$ or

$$\int_E u(z)dz = \int_R P(t,x,E)\,\mu(dx)$$

$$= \int_R \int_E p(t,x,y)\,dy\, u(x)\, dx$$

$$= \int_E \left[\int_R p(t,x,y)\, u(x)\, dx \right] dy$$

or since $E \in \mathcal{B}(R)$ was arbitrary

$$u(y) = \int_R p(t,x,y)\, u(x)\, dx \quad .$$

Assuming the right-hand side to be twice differentiable in x and once in t under the integral sign, and assuming p as a function of t and y satisfies $A^*p = \partial p/\partial t$, then

$$(\partial/\partial t - A^*)u(y) = \int_R (\partial/\partial t - A^*)p(t,x,y)\, u(x)\, dy$$

$$= 0.$$

And so, $A^*u = 0$ justifying the claim.

Before considering the invariant measure problem from this point of view, it is appropriate to return to the transition operator and examine it more closely. The next paragraphs follow Khas'minskii [43]. Assume the following:

(i) The process x as a solution of (5) has continuous sample paths.

(ii) The operators T_t: $C(R) \to C(R)$, or that x is a <u>Feller Process</u> [19].

(iii) The process x is <u>non-degenerate</u>, or equivalently, $P(t,x,U) > 0$ holds for any open set of positive Lebesque measure.

(iv) The process x is a strong Markov process.

(v) The process z is <u>recurrent</u>; i.e., there exists a compact subset K of R such that for every $x \in R$, $P(t,x,K) = 1$ for <u>some</u> $t \in R^+$.

<u>Proposition 7</u>: [43,p. 180] The trajectories of the process x are every-dense in R.

The relevant result derived from these assumptions is given in the

Theorem 8: [43,p.182] For the recurrent, diffusion process x as a solution of (5) there exists a non-trivial, unique σ-finite invariant measure μ. If $\mu(R) < \infty$, then

$$\frac{1}{T}\int_0^T P(t,x,E)\,dt \rightarrow \mu(E)/\mu(R) \quad .$$

Compare the second assertion of this theorem with the arguments in the proof of Theorem 3.2.5. See Doob [14] for related remarks on this convergence. Using Doob [14, Theorem 5], Khas'minskii is actually able to conclude that if $\mu(R)$ is finite, then $P(t,x,E) \rightarrow \mu(E)$ for every $x \in R$. Finiteness of μ may be shown under minor additional restrictions on the process x.

Returning to the density equations, the precise conditions for x to have an invariant measure are given in

Theorem 9: [43,p.190] In order that x have a finite invariant measure, it is necessary and sufficient that $Au = -1$ have a positive solution in $R \sim D$ for some bounded domain D with smooth boundary ∂D. Moreover, in this case, for any measureable function f

$$\lim_{t\to\infty} \int_R p(t,s,y)\, f(y)\, dy = \int_R f(y)\, \mu(dy)$$

where μ is the invariant measure, and p the fundamental solution of $A^*p = \partial p/\partial t$.

Proved by arguments involving the first entrance times into the domain D, Theorem 9 depends critically on the smoothness properties of ∂D. This is of course a significant condition and in most instances a handicap. Based on the paper [43], Wonham's paper [71] contains some important sufficient conditions guaranteeing the hypothesis of Theorem 9, and thus recurrence and invariance of the solution process x. His conditions use Lyapunov functionals of the state.

Let S_r denote the open ball of radius $r > 0$ in Euclidean space and assume that the function v on R satisfies the following

(i) v is twice continuously differentiable.

(ii) $v(x) \geqq 0$ for $x \in S_r$, $v(x) \to \infty$ as $x \to \infty$.

Theorem 10: [71,p.200] If there exists a function v, satisfying (i) and (ii) above and such that $Av \leqq -1$, the x, the solution of (5), has a unique invariant distribution.

Although the analysis of Wonham and Khas'minskii relies almost exclusively on the analytical structure of Markov processes, it is more illuminating to outline the proofs of Theorems 9 and 10 in the framework used earlier in this chapter. The idea is simple: from any initial distribution μ_o, the distributions of x(t) for $t \in R^+$ evolve according to $U_t\mu_o = \mu_t$, where U_t is the semigroup defined above:

$$\mu_t(E) = U_t\mu_o(E) = \int_R P(t,x,E) \quad \mu_o(dx) \quad .$$

Clearly, U_t is linear and continuous on PM(R) with the topology of weak convergence; continuity following from the Feller property. Thus, on a compact set contained in PM(R), U_t is closed and has a fixed point [16,p.456]. Thus, it remains to show that the distributions of x form a compact subset of PM(R). It is at this point that the Lyapunov functional is used, see Elliot [22, section 4.3].

Let v be a functional on R satisfying the assumptions (i) and (ii) above; further assume that $v(0) = 0$. Then Elliot [22, p.39] shows that for $\gamma > 0$ the set

$$\Phi(v;\gamma) = \{\mu \in PM(R): \int_R v(x)\ \mu(dx) \leq \gamma\}$$

is compact (in the weak topology on PM(R)). Elliot's result is the following:

<u>Theorem 11</u>: [22, p. 35] Let v satisfying (i) and (ii) above be such that for positive c_1, c_2, c_3

$$\left.\begin{array}{ll} \text{(i)}' & |(Av)(x)| \leq c_1(1 + |x|^2) \\ \text{(ii)}' & (Av)(x) < c_2 - c_3 v(x) \end{array}\right\} \quad ; \quad x \in R,$$

then there exists an invariant distribution for x.

<u>Proof</u>: Consider $T_t v(x)$, then

$$T_t v(x) = v(x) + \int_0^t T_s Av(x)\ ds$$

from Dynkin's Formula ([48,p.10] or [22]). So

$$c_3 \int_0^t T_s v(x)\ ds + T_t v(x) \leq v(x) + \int_0^t c_2\ dt$$

from which it follows that

$$T_t v(x) \leq v(x) \exp(-c_3 t) + \frac{c_2}{c_3}(1 - \exp(-c_3 t)) .$$

For any $\gamma > c_2/c_3$ and $\mu \in \Phi(v;\gamma)$ the equality

$$\int_R T_t v(x)\ \mu(dx) = \int_R v(x)\ (U_t\mu)(dx)$$

$$\leq c_2/c_3 < \gamma .$$

The compactness of $\Phi(v;\gamma)$ and the continuity of U_t yield the result.

<u>QED</u>

It is appropriate to remark that this technique of proving the existence of distributions invariant under time-shifts is commonly used in the ergodic theory of Markov processes per se. See for example Foguel [27] for an interesting introduction to this subject. While less constructive than the use of the steady state Fokker-Planck equation, the technique is quite similar to that used in the earlier sections of this chapter.

CHAPTER 4

APPLICATIONS, CONCLUSIONS, AND FURTHER RESEARCH

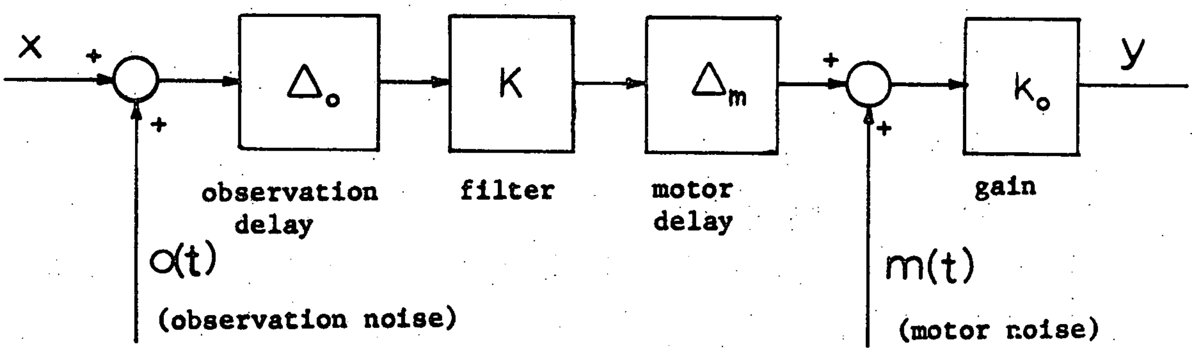
4.1 A Few Remarks on Applications:

In this section two feedback systems profitably modelled as random will be considered. The purpose here is not to give a complete investigation of these examples but rather to indicate treatments within the framework established in the last chapters.

A. The human operator:

As a first example consider the human as a feedback controller. Feedback systems containing humans arise naturally in many settings [2], perhaps the most familiar one in an engineering context is as a pilot. In the design of control mechanisms and instrument displays for aircraft it is important to have some model of the pilot as the "actuator link" between the instruments and the control mechanism. Because of the highly individual techniques of pilots [45] and the possibility of a large number of pilots flying any particular aircraft, it is appropriate to model the human as containing some random parameters when operating in this situation.

In controlling an aircraft about some nominal trajectroy, the human may be modelled as an essentially linear element subject to random perturbations in the following manner. In reading the instruments errors are made, and these errors, being characteristic of individuals, are usually modelled as the effect of additive noise. Attempting to deduce the state of the aircraft from these imperfect observations, the human performs a kind of filtering operation in some optimal manner. This step is usually modelled as operating on the noisy observation signal with an optimal linear (Kalman) filter. The next step in the control process is operating the control mechanism so as to correct for any perceived errors from the nominal trajectory. At this stage a delay is introduced as a consequence of the neuro-motor delays of the human. Moreover, noise is usually added here to account for the errors in manipulating the controls. This model of the human controller in a steady-state control task reduces to the cascade of elements shown in Figure 1.



<u>Figure 1</u>: A model of the human controller.

Defining the Kalman filter by its impulse response k, the input-output equation of the model is

$$y(t) = \begin{cases} k_o m(t) + k_o \displaystyle\int_{\Delta_o}^{t-\Delta_m} k(t-\Delta_m-s)\,[x(s-\Delta_o) + o(s-\Delta_o)]\,ds\,; & t \geq \Delta_m+\Delta_o \\ k_o m(t)\,; \qquad t < \Delta_m+\Delta_o\,. & \end{cases}$$

For any model of the aircraft (about the nominal operating point) analysis of feedback systems including the human operator model above is straight-forward from this point (except for the presence of delays) by familiar methods.

Frequently, however, a less detailed model of the human as a white noise gain is used to obtain worst case results in experiments involving a wide range of operating conditions [2]. In this case the model of Figure 2 applies.

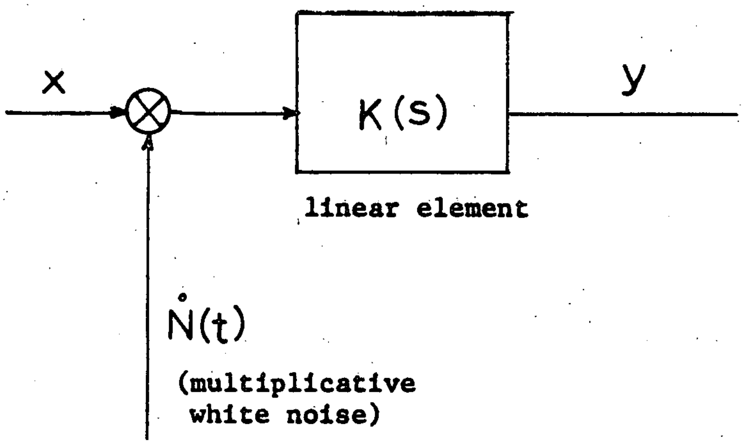


Figure 2: A crude model of the human operator.

Here K represents the combined effects of the human's filtering action and (Padè) approximations to the delays. Thus,

$$y(t) = \int_{0}^{t} k(t-s)\ x(s)\ dN(s)$$

as an Ito integral, describes the transfer of observation (x) into control action (y) by the human. Here k is the impulse response of the linear element K. Again for an appropriate linear model of the aircraft in steady state operation, analysis of the human as a controller is straight-forward using results like Theorems 3.2.4, 3.3.4, and 3.3.5. The latter give easy sufficient conditions in terms of the frequency response of the linear elements for boundedness of the signals in the control loop.

In the event that the human model includes a nonlinearity satisfying sector conditions as used in Chapter 3, perhaps reflecting thresholds of no response [45], then analysis using Theorem 3.2.5, etc., is no more difficult than in the linear case.

B. Analysis of round-off errors in numerical computations:

In the first chapter the point was raised that the accumulation of round-off errors in a numerical computation could be considered as a stochastic process. Though in actuality a deterministic phenomena, the randomization of the error evolution is warranted by the extreme complexity of any nontrivial computation on a large machine. The development of a statistical model takes the following foute (this analysis is drawn from Henrici [34],[35]).

Most numerical algorithms consist of generating a sequence of numbers $x_o, x_1, \ldots$, defined by the relations

$$x_n = F_n(x_o, \ldots, x_{n-1}) \quad ; \quad n = 1,2,\ldots \quad .$$

In actual machine computations, however, the algorithm is only approximately realized and machine numbers $\tilde{x}_n$ (of finite length) are generated by the approximate realizations $\tilde{F}_n$ by

$$\tilde{x}_n = \tilde{F}_n(\tilde{x}_o, \ldots, \tilde{x}_{n-1}) \quad ; \quad n = 1,2,\ldots \quad .$$

Write

$$\tilde{x}_n = F_n(\tilde{x}_o, \ldots, \tilde{x}_{n-1}) + e_n$$

and consider this as the definition of the local rounding error e_n. Thus,

$$e_n = \tilde{F}_n(\tilde{x}_o, \ldots, \tilde{x}_{n-1}) - F_n(\tilde{x}_o, \ldots, \tilde{x}_{n-1}) \quad .$$

Each local rounding error is propagated through the remainder of the computation (from n on), and in this process its effect on the final accumulated error may be amplified or diminished. The accumulated rounding error r_n at any stage is defined as the difference between the numerical result and the correct theoretical result; here $r_n = \tilde{x}_n - x_n$.

Clearly, knowledge of the machine approximations $\tilde{F}_n$ would permit one to determine worst case bounds for the error evolution under the unlikely hypothesis that each local rounding error has the maximum bad effect on the accumulated error. Such a systematic reinforcement of errors is unlikely in any typical computation, and the bounds obtained under this assumption are usually uninformative. It is the need to have some appraisal of the "average" growth of round-off errors that motivates the statistical assumptions.

Therefore, assume that each e_n is for each n a random variable on some probability space $(\Omega, \mathcal{F}, P)$. The accumulated error evolves according to

$$\begin{aligned} r_n &= e_n + F_n(\tilde{x}_o, \ldots, \tilde{x}_{n-1}) + F_n(\tilde{x}_o, \ldots, \tilde{x}_{n-1}) \\ &= e_n + H_n(r_o, \ldots, r_{n-1}) \quad . \end{aligned}$$

The stability problem becomes the following: given the statistics of the stochastic process $\{e_n\}_{n\varepsilon Z^+}$ describe those of the process $\{r_n\}_{n\varepsilon Z^+}$ as $n \to \infty$. Of particular interest are bounds on the mean and variance of the process $\{r\}$, as these are easily determined and indicate the average rate of growth of the errors. The general conditions of section 3.1 enable one to constrain the operator H so as to assure compactness of the distributions of $\{\pi_n r\}_{n\varepsilon Z^+}$ (the truncations of r) on some sequence space and guarantee asymptotic invariance (with n) of the statistics of r if e is stationary. Moreover, the moment bounds determine the asymptotic limit distribution approximately. In certain linear integration schemes (the operator H becomes a linear convolution) the results of section 3.2 apply immediately. In relation to this point see [66], where the analog of Theorem 3.2.4 is given for random sequences.

4.2 Conclusions and Suggestions for Future Research:

In order to place the present work in perspective it is necessary to place the study of stochastic systems within the theory of dynamical systems. Although it is too early for the latter task, some points are clear. First the study of dynamical systems has proven to be one of the most fruitful branches of engineering and mathematics, and for this reason any extensions and generalizations should be pursuited for additional insight. The admission of stochastic variables in optimization problems has led to a much better understanding of the role of information patterns in control systems as may be judged from the several papers on this subject in the Bibliography. Secondly the application of stochastic systems as models for complex physical systems would seem to be promising; the demonstrated success of a few definitive case studies would strengthen this assertion.

Of a more technical nature is the observation that the properties of causal, dynamical systems are deeply related to those of Markov processes. A general examination of the relationship between causality and the Markov property beyond the obvious would seem valuable. Certainly the description of systems by stochastic differential equations interperted in the analytical theory of Markov process has provided a rich class of systems described by partial differential equations. Viewed from the field of what has come to be called distributed parameter systems, this aspect of stochastic systems permits an easy interpertation of the properties of the distributed solution as a probability density function. Moreover, the additional interpertation provided by the differential equation for the sample trajectories of the process cannot be but an asset in the analysis of the partial differential equation. This relationship between distributed systems and Markovian systems is largely unexploited as such.

As the remarks above reflect some of the tentative aspects of the stochastic systems theory, so must the present work be regarded as preliminary in nature. For as an investigation of the problem of determining the transformations of probability distributions by dynamical (feedback) systems, its provisional aspects are apparent. Perhaps the most significant drawback is the non-constructive nature of the analysis. It would be an important extension of this work to render the process of analysis constructive, though this is likely to be equivalent to solving the implicit feedback equations and hence impossible in general.

However, as an alternative approach to the analysis of the asymptotic properties of stochastic systems, this work has succeeded in making the Prohorov theory directly applicable to this kind of analysis. In this context the work is antedated by that of Ito and Nisio [41] and Fleming and Nisio [26], though the explicit connection of deterministic operator stability theory and the Prohorov theory, using the results of Topsøe, appears to be novel. Finally, the specific results of sections 3.2 and 3.3 are interesting as generalizations of deterministic counterparts-the

Nyquist and Circle Theorems. Examined along with papers like [71], [47], [73], [46], and [66], these theorems should increase the understanding of stochastic systems containing linear elements.

Further comments on this work may be usefully made by suggesting a few extensions and modifications. In addition to the general statements above, consider then the following precise problems.

A. Stability conditions based on empirical distributions:

Of course one of the primary objections to this work is its a priori assumption of given distributions for the perturbation inputs and random parameters. In any practical experiment these are seldom given and usually difficult to determine experimentally, though appropriate statistical methods are available. About the most complete characterization one could reasonably hope for is a number of empirical distributions for the uncertainties derived from samples of the processes. It would be, therefore, very useful to determine conditions based on empirical distributions of the inputs and outputs that assure the asymptotic regularity of the outputs in the sense used previously. These conditions would have to apply for a class of distributions which could give rise to those observed empirically. The Prohorov theory has potential applications here especially on the space D of piecewise continuous functions, see some comments to this effect in [7]. The definitions of stochastic systems given in section 3.1 are designed to permit a number of possible distributions for the uncertainties present, and may prove useful in the early stages of work on this problem.

B. Stochastic systems with nonlinear state spaces:

Consider the problem of designing a feedback control law to accurately orient a rigid body (satellite) in orbit. The perturbations are essentially stochastic in nature, arising from position sensor errors and natural phenomena. As is well known the attitude of a rigid body in a fixed coordinate system is described by a set of 3×3 orthogonal matrices, a set not closed under addition. Hence, the control problem must be analyzed in a setting where the state space of the system is a nonlinear manifold.

Note that the analysis of round-off errors may be considered in this framework, as the local errors are confined to a fixed interval and may be considered as random variables on a circle, see [23,p.61].

One of the reasons for seeking problems with nonlinear state spaces is the good possibility of obtaining explicit analytical solutions to the diffusion equations (for the probability density functions of the state). There are rather few diffusion equations, aside from the Guass-Markov case, that admit an explicit solution in the usual vector space setting. For certain special manifolds explicit solutions to Laplace's equation are well known and may be used to describe Brownian motions on these manifolds [18]. Other references are Elliot [22], McKean [50] and the references therein. Research on this problem should provide interesting enhancements of the work in Brockett [9].

C. Passive stochastic systems:

Of a rather more technical nature is the problem of describing the analog of passivity in a stochastic setting. Recall that a deterministic operator G on the Hilbert space $(H,\langle\cdot,\cdot\rangle)$ is said to be passive if

$$\mathrm{Re}\langle x,Gx\rangle \geq 0 \quad \text{for every } x \in H.$$

This is equivalent to the physical notion of a system which always disspates energy [64].

Let $L_2(R^+)$ denote the space of square integrable, real-valued functions on R^+, and $F(\Omega;L_2)$ be the set of L_2-valued random variables. Then clearly the set $(\mathcal{L}(\Omega;L_2),\langle\cdot,\cdot\rangle)$, where for $x,y \in \mathcal{L}(\Omega;L_2) \subset F(\Omega;L_2)$

$$\langle x,y\rangle = E\{\int_0^\infty x(t,\omega)y(t,\omega)\,dt\}$$

and $\langle x,x\rangle < \infty$ for every x, is a Hilbert space. Moreover, the inequality $\langle x,Gx\rangle \geq 0$ makes perfect mathematical sense for some (random) endomorphism G on $\mathcal{L}(\Omega;L_2)$, and it is easy to give the Positive Operator stability theorem [74,p.235] of the deterministic theory in this setting. Physical interpertations of the result are less easy, however, and apparently some notion of random spectra must be developed. Useful ideas are likely to be found in statistical mechanics [62].

BIBLIOGRAPHY

1. Bachman, G., Elements of Abstract Harmonic Analysis, Academic Press, New York, 1964.

2. Bekey, G.A., "Description of the Human Operator in Control Systems," in C.T. Leondes, Modern Control Systems Theory, McGraw-Hill, New York, 1965, Chap. 12, pp. 431-462.

3. Benes, V., "Existence of Optimal Stochastic Control Laws," to appear, SIAM J. on Control, 1971.

4. Benes, V., "Existence of Optimal Strategies Based on Specified Information," SIAM J. on Control, 8, 1970, pp. 179-188.

5. Benes, V., "A Nonlinear Integral Equation in the Marcinkiewicz Space M_2," J. of Math. and Phys., 44, 1965, pp. 24-35.

6. Bharucha-Reid, A., "On the Theory of Random Equations," Proc. of Sym. in Applied Math., Amer. Math. Soc., Providence, R.I., 16, 1964, p.40-69.

7. Billingsly, P., Convergence of Probability Measures, J. Wiley, New York, 1968.

8. Birkhoff, G.D., "Dynamical Systems," Amer. Math. Soc. Colloq. Publ., IX, Providence, 1927.

9. Brockett, R., "System Theory on Group Manifolds and Coset Spaces," submitted for publication, SIAM J. on Control, 1971.

10. Bucy, R.S., "Stability and Positive Supermartingales," Jour. of Diff. Eqs., 1, 1965, pp. 151-155.

11. Davis, J., "Stability Conditions Derived from Spectral Theory: Discrete Systems with Periodic Feedback," to appear SIAM J. on Control, 1971.

12. Davis, J., "Stability of Linear Feedback Systems," Ph.D. Thesis, Dept. of Mathematics, M.I.T., Cambridge, Mass., June, 1970.

13. Doob, J.L., "Semimartingales and Subharmonic Functions," Trans. Amer. Math. Soc., 77, 1954, pp. 86-121.

14. Doob, J.L., "Asymptotic Properties of Markov Probabilities," Trans. Amer. Math. Soc., 63, 1948, pp. 393-421.

15. Duncan, T., Varaiya, P., "On the Solutions of a Stochastic Control System," Research Rep. Elec. Res. Lab, ERL-M273, Univ. of Calif., Berkeley, 1970.

16. Dunford, N., Schwartz, J.T., Linear Operators, Part I : General Theory, Interscience, New York, 1957.

17. Dym, H., "Stationary Measure for the Flow of a Linear Differential Equation Driven by White Noise," Trans. Amer. Math. Soc., 123, 1966, pp. 130-164.

18. Dynkin, E., "Brownian Motion in Symmetric Spaces and Nonnegative Eigenfunctions of the Laplace-Beltrami Operator," Amer. Math. Soc. Translations, ser. 2, 72, 1968, p. 203-228.

19. Dynkin, E.B., Markov Processes, I and II, Springer-Verlag, Berlin, 1965.

20. Dynkin, E., "Markov Processes and Problems in Analysis," Amer. Math. Soc. Trans., Ser. 2, 31, 1963, p. 1-24.

21. Dynkin, E., Theory of Markov Processes, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1961.

22. Elliott, D.L., Controllable Nonlinear Systems Driven by White Noise, Ph.D. Thesis, Dept. Elec. Engrg., Univ. of Calif., Los Angeles, 1969.

23. Feller, W., An Introduction to Probability Theory and its Applications, I and II, J. Wiley, New York, 1966.

24. Fleming, W., "Optimal Continuous-Parameter Stochastic Control," SIAM Review, 11, 1969, pp. 470-509.

25. Fleming, W., "Optimal Control of Partially Observable Diffusions," SIAM J. on Control, 6, 1968, pp. 194-214.

26. Fleming, W., Nisio, M., "On the Existence of Optimal Stochastic Controls," J. of Math. and Mech., 15, 1966, pp. 777-794.

27. Foguel, S., The Ergodic Theory of Markov Processes, Van Nostrand, New York, 1969.

28. Gikhman, I., Dorogovcev, Ja., "On the Stability of Solutions of Stochastic Differential Equations," Amer. Math. Soc. Trans., Ser. 2, 72, 1968, pp. 229-250.

29. Gikhman, I.I., Skorokhod, A.V., Introduction to the Theory of Random Processes, Saunders, Philadelphia (English ed.), 1969.

30. Halmos, P.R., Measure Theory, D. Van Nostrand, Princeton, 1950.

31. Hans, O., "Random Operator Equations," Proc. Fourth Berkeley Sym. on Math. Statistics and Prob., Vol. 2, Univ. of Calif. Press, Berkeley, 1961, pp. 185-202.

32. Hans, O., "Generalized Random Variables," Trans. First Prague Conf., 1957, pp. 61-104.

33. Hans, O., "Random Fixed Point Theorems," Trans. First Prague Conf., 1957, pp. 105-126.

34. Henrici, P., Elements of Numerical Analysis, J. Wiley, New York, 1964, pp. 302-320.

35. Henrici, P., Error Propagation for Difference Methods, J. Wiley, New York, 1963.

36. Hida, T., Stationary Stochastic Processes, Princeton Univ. Press, Princeton, N.J., 1970.

37. Hunt, G.A., "Markov Processes and Potentials, I, II," Illinois Math. Jour. 1, 1957, pp. 44-93, 316-362.

38. Il'in, A., Khas'minskii, R., "Asymptotic Behavior of Solutions of Parabolic Equations and an Ergodic Property of Nonhomogeneous Diffusion Processes," Amer. Math. Soc. Trans., ser. 2, 49, 1966, p. 241-268.

39. Itô, K., Lectures on Stochastic Processes, Tata Institute, Bombay, 1961.

40. Itô, K., "On Stochastic Differential Equations," Memoirs of AMS, No. 4, 1951, Providence, R.I.

41. Ito, K., Nisio, M., "On Stationary Solutions of a Stochastic Differential Equation," J. Math. Kyoto, 4, 1964, pp. 1-76.

42. Kantorovich, L.V., Akilov, G.P., Functional Analysis in Normed Spaces, MacMillan Co., (English edition), New York, 1964.

43. Khas'minskii, R.Z., "Ergodic Properties of Recurrent Diffusion Processes and Stabilization of the Solution to the Cauchy Problem for Parabolic Equations," Th. of Prob. and Appl., 5, 1960, pp. 179-196.

44. Khas'minskii, R., "Diffusion Processes and Elliptic Differential Equations Degenerating on the Boundary of a Domain," Th. Prob. and Appl., 3, 1958, p. 400-419.

45. Kleinman, D., "The Human Operator as a Controller," Seminar at M.I.T., 1971

46. Kleinman, D., "On the Stability of Linear Stochastic Systems," IEEE Trans. on Aut. Con., AC-14, Aug. 1969, pp. 429-430.

47. Kushner, H.J., "The Cauchy Problem for a Class of Degenerate Parabolic Equations and Asymptotic Properties of the Related Diffusion Processes," J. Diff. Equs., 6, 1969, pp. 209-231.

48. Kushner, H.J., Stochastic Stability and Control, Academic Press, New York, 1967.

49. Kushner, H., "On the Stability of Stochastic Dynamical Systems," Proc. Natl. Acad. Sci., 53, 1965, p. 8-12.

50. McKean, H.P., Stochastic Integrals, Academic Press, New York, 1969.

51. Meyer, P.A., Probability and Potentials, Blaisdell, Waltham, Mass., 1966.

52. Parthasarathy, K.R., Probability Measures on Metric Spaces, Academic Press, New York, 1967.

53. Prohorov, Yu., "Convergence of Random Processes and Limit Theorems in Probability Theory," Th. Prob. Appl. 1, 1956, pp. 157-214.

54. Sandberg, I.W., "Some Results on the Theory of Physical Systems Governed by Nonlinear Functional Equations," Bell Sys. Tech. J., 44, 1965, pp. 871-898.

55. Skorokhod, A.V., Studies in the Theory of Random Processes, Addison-Wesley, Reading, Mass., 1965.

56. Skorokhod, A., "Limit Theorems for Stochastic Processes," Th. Prob. Appl., 1, 1956, pp. 261-290.

57. Spacek, A., "Random Metric Spaces," Trans. Second Prague Conf., 1960, pp. 627-638.

58. Stone, C., "Weak Convergence of Stochastic Processes Defined on Semi-infinite Time Intervals," Proc. Amer. Math. Soc., 14, 1963, p. 694-696.

59. Stroock, D.W., Varadhan, S.R.S., "Diffusion Processes with Continuous Coefficients, I and II," Comm. on Pure and Appl. Math., 22, 1969, pp. 345-400, 479-530.

60. Topsφe, F., Topology and Measure, Lecture Notes in Mathematics No. 133 , Springer-Verlag, New York, 1970.

61. Topsφe, F., "Preservation of Weak Convergence Under Mappings," Ann. Math. Stat., 38, 1967, pp. 1661-1665.

62. Willems, J.C., Private communication, April, 1971.

63. Willems, J.C., "Input-Output Stability Versus Global Stability: The Generation of Lyapunov Functions for Input-Output Stable Systems," submitted for publication, Jan., 1970.

64. Willems, J.C., The Analysis of Feedback Systems, M.I.T. Press, Cambridge, Mass., 1971.

65. Willems, J.C., "Stability, Instability, Invertibility and Causality," SIAM Journal on Control, 7, 1969, pp. 645-671.

66. Willems, J.C., Blankenship, G.L., "Frequency Domain Stability Criteria for Stochastic Systems," to be published in IEEE Tran. on Automatic Control, August, 1971.

67. Witsenhausen, H.S., "On Performance Bounds for Uncertain Systems," SIAM J. Control, 8, 1970, pp. 55-89.

68. Witsenhausen, H.S., "Inequalities for the Performance of Suboptimal Uncertain Systems," Automatica, 5, 1969, pp. 507-512.

69. Witsenhausen, H.S., "Minimax Control of Uncertain Systems," Rep. ESL-R-269, Mass. Inst. of Tech., Cambridge, Mass., 1966.

70. Wonham, W.M., "On the Separation Theorem of Stochastic Control," SIAM J. on Control, 6, 1968, pp. 312-326.

71. Wonham, W.M., "Liapunov Criteria for Weak Stochastic Stability," Jour. of Diff. Equns., 2, 1966, pp. 195-207.

72. Zaanen, A.C., Integration, John Wiley, New York, 1967.

73. Zakai, M. , Snyders, J., "Stationary probability measures for linear differential equations drives by white noise," J. Diff. Equs., 8, 1970, pp. 27-33.

74. Zames, G., "On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems - Part I : Conditions Derived Using Concepts of Loop Gain, Conicity, and Positivity; Part II: Conditions Involving Circles in the Frequency Plane and Sector Nonlinearities," IEEE Trans. on Aut. Con., AC-11, 1966, pp. 228-239, 465-476.

75. Zames, G., "Nonlinear Time Varying Feedback Systems - Conditions for L_∞-Boundedness Derived Using Conic Operators on Exponentially Weighted Spaces," Proc. Third Allerton Conf. on Circuit and System Theory, 1965, p. 460-471.

76. Becker, H., Mandrekar, V., "On the Existence of Optimal Random Controls," J. Math. and Mech., 18, 1969, pp.1151-1166.