N73 18606

Department of Mathematics
UNIVERSITY OF GEORGIA

N.A.S.A. Research Grant NGR 11-003-020

Technical Report

STATIONARITY CONDITIONS FOR STOCHASTIC DIFFERENTIAL EQUATIONS

CASE FILE COPY

Stochastic Systems, Methods of Physics, Tensor Analysis, Probability Algebras and Stochastic Spaces, Mathematics of Quantum Theory, Functional Analysis, Variational Methods, Control Theory, Analysis of Large Systems, Probability, Stochastic Processes, Computer Science, Green's Function Methods, Stochastic Differential Equations, Stochastic Operator Theory, Stochastic Generalized Functions

SIGILLVM · VNIVERSITATIS · GEORGIAE · ET DOCERE ET RERVM EXQVIRERE CAVSAS · 1785

N.A.S.A. Research Grant NGR 11-003-020

Technical Report

STATIONARITY CONDITIONS FOR STOCHASTIC DIFFERENTIAL EQUATIONS

by

G. Adomian and W.W. Walker

STATIONARITY CONDITIONS FOR STOCHASTIC DIFFERENTIAL EQUATIONS

This is a preliminary study of possible necessary and sufficient conditions to insure stationarity in the solution process for a stochastic differential equation. It indirectly sheds some light on ergodicity properties and shows that the spectral density is generally inadequate as a statistical measure of the solution. Further work is proceeding on a more general theorem which gives necessary and sufficient conditions in a form useful for applications.

TABLE OF CONTENTS

		Page
Section 1.	Introduction	1
Section 2.	Analytic Random Functions, Stationarity and Applications to Stochastic Differential Equations	8
Section 3.	Random Harmonic Analysis and Applications to Stochastic Differential Equations	30
Section 4.	Stochastic Transformations	44
Bibliography	. . . . .	52

SECTION 1.

The following definitions and notation will be used consistently throughout this thesis.

$(\Omega, \mathcal{J}, \mu)$ will be a fixed (but otherwise arbitrary) probability space with points $\omega \in \Omega$, a σ-algebra $\mathcal{J}$ of subsets (probabilizable events), and a complete measure μ such that $\mu(\Omega) = 1$.

A <u>random variable</u> (r.v.) $x = x(\omega)$ will be a finite real-valued (or complex valued) measurable function defined on Ω, <u>i.e.</u>, we require that $\{\omega : x(\omega) < \lambda\} \in \mathcal{J}$ for all real numbers λ. If $x(\omega)$ is complex-valued, we require that $x(\omega) = u(\omega) + iv(\omega)$ where u and v are real-valued r.v.'s.

Given a r.v. $x(\omega)$, $E(x)$ or $\langle x \rangle$ denotes the integral $\int_\Omega x(\omega)\, d\mu(\omega)$ if this integral is defined.

Given a parameter set T (which we usually take to be the real numbers), a <u>stochastic process</u> (s.p.) or <u>random function</u> (r.f.) on T is a real or complex-valued function $X: T \times \Omega \longrightarrow R(C)$ such that for each fixed $t \in T$ the function $x_t(\omega) = X(t, \omega)$ is a r.v. Very often in our notation we will suppress the variable ω and write $X(t)$ for the r.f. Notice that a s.p. has two convenient interpretations. First of all, a s.p. is a family of r.v.'s indexed by T, <u>i.e.</u>, $X = \{x_t(\omega) : \Omega \longrightarrow R\}_{t \in T}$. On the other hand, if we emphasize the variable t and let $T = R$, then a s.p. X is a

collection of real-valued functions of a real variable (indexed by Ω). These functions are called the sample paths (realizations, trajectories) of the process. The measurability of X with respect to the variable ω says nothing about measurability with respect to the t, and in general the sample paths may be very badly behaved. However, we will consider only <u>measurable</u> processes, <u>i.e.</u>, functions $X(t,\omega)$ which are measurable with respect to the σ-algebra $\mathcal{J} \times \mathcal{J}$ where $\mathcal{J}$ is the family of Lebesgue measurable subsets of the real line. Then all the sample paths will be measurable.

We let $L^2(\Omega)$ denote the Hilbert space of all square integrable r.v.'s on Ω making the usual identification of r.v.'s which are equal almost everywhere with respect to the measure μ.

We say that a r.f. $X(t,\omega)$ is <u>second order</u> if each r.v. x_t, $t \in T$, is a member of $L^2(\Omega)$. Thus X is second order if and only if $\int_\Omega |X(t,\omega)|^2 d\mu(\omega) < \infty$ for all $t \in T$. Note that a second order r.f. X induces a map into a space of r.v.'s $Y: T \longrightarrow L^2(\Omega)$ defined by $(Y(t))(\omega) = X(t,\omega)$. The <u>covariance function</u> associated with the r.f. X is defined by $\Gamma_x(s,t) = E(X(s)\overline{X(t)})$.

Once again let $T = R$ and let $X(t,\omega)$ be a second order s.p. Let Y be defined as in the preceeding paragraph. We say that X is <u>continuous in the mean square sense</u> at t_0 if $\underset{t \to t_0}{\text{l.i.m.}}\ X(t) = X(t_0)$. This is equivalent to saying that the function $Y(t)$ is continuous at t_0 relative to the

standard topology on the reals and the norm topology on $L^2(\Omega)$. Similarly, we say that $X(t)$ is <u>differentiable in mean square</u> at t_0 if there is a r.v. (second order) r such that $\underset{t \to t_0}{\text{l.i.m.}} \frac{X(t) - X(t_0)}{t - t_0} = r$. Thus X is differentiable in mean square at t_0 if and only if $Y(t)$ is differentiable at t_0, and moreover $r = \frac{dY}{dt}(t_0)$. This same analogy carries over to integration (Riemann, Riemann-Stieltjes, Lebesgue-type) in mean square of X and the corresponding integration of Y. Hence the study of the mean square analytic properties of a (second-order) s.p. X is equivalent to the study of the corresponding properties of a function $Y: R \longrightarrow L^2(\Omega)$.

Throughout this thesis we will deal with the concept of <u>wide-sense stationarity</u>. Moreover, without loss of generality we consider only zero-mean processes, and consequently we take as the defining characteristic of a <u>stationary</u> process X the existence of a <u>correlation function</u> f such that $\Gamma_x(s,t) = f(t-s)$. We define the spectral density function of the process by $\varphi(u) = \int e^{2\pi i t u} f(t)dt$.[1]

---

[1]We use this definition of correlation function in accordance with Adomian [1]. Often in the literature the correlation function is defined $g(s-t) = E(X(s)\overline{X(t)})$ which is the complex conjugate of our definition. We let $\varphi(u) = \int e^{2\pi i t u} f(t)dt$ be the spectral density function of the process whereas some authors may have $\psi(u) = \int e^{-2\pi i t u} f(t)dt$ as the spectral density function. This of course will be the complex conjugate of our spectral density function.

Usually we consider only real processes, however if some result takes on a much cleaner form in the complex case we will note it.

The physical interpretation of stationarity is well known (see Yaglom [11]). The following geometrical interpretation may provide some insight however. Note that a second-order process with the real line as the parameter set is a map $X: R \longrightarrow L^2(\Omega)$ from the reals into a particular Hilbert space. Hence the relationship $E(X(s)\overline{X(t)}) = f(t-s)$ is merely a restriction on the behavior of the inner products of points in $L^2(\Omega)$ which lie on the curve associated with X. In particular, $\|X(t)\|^2 = E(X(t)\overline{X(t)}) = f(t-t) = f(0)$ for all t, and so the curve X must lie on a sphere of radius f(0) centered at the origin. For stationary X, $E(X(s+\tau)\overline{X(t+\tau)}) = E(X(s)\overline{X(t)})$ for all s, t, and τ, and so if we think of the inner product as determining an angle between say the vectors X(s) and X(t), then this angle is invariant under translations of the parameter set, e.g. the angle between X(s) and X(t) is the same as the one between X(0) and X(t-s). A circle in R^2 centered at the origin is an example of such a curve if the standard parameterization is taken: thus, consider the curve $x(t) = e_1 \cos t + e_2 \sin t$ where e_1 and e_2 are the standard basis vectors for R^2. In general, we can replace the unit vectors by arbitrary orthogonal vectors in $L^2(\Omega)$ of equal norm, call two such (distinct) vectors A and B. Then the process defined by $X(t) = A \cos t + B \sin t$ is really just

a circle in $L^2(\Omega)$, and a simple calculation shows it is stationary. One of the most powerful results in the general theory of stationary processes is that every (continuous) stationary process is the limit of sums of processes of this special type.

One is naturally interested in determining what sort of transformations of stochastic processes preserve stationarity. For the moment, we interpret the term "stochastic transformation" in the loosest sense, namely we call any rule which associates one or more processes with another process a <u>stochastic transformation</u> (s.t.).[2] In this sense there are many s.t.'s which carry stationary processes into stationary processes, and we list here just a few:

i) Let $U: L^2(\Omega) \longrightarrow L^2(\Omega)$ be any isometry. Let $F(t)$ be a stationary process. Then the process $G(t) = U(F(t))$ is stationary since $E(G(s)\overline{G(t)}) = E(F(s)\overline{F(t)}) = f(t-s)$. Let α be any complex number. Then the process $H(t) = \alpha U(F(t))$ is stationary since $E(H(s)\overline{H(t)}) = |\alpha|^2 f(t-s)$.

ii) Let $F(t)$ and $G(t)$ be stationary processes such that the smallest closed linear manifolds containing F and G respectively are orthogonal. Then $F(t) + G(t)$ is a

---

[2]For a more complete discussion of this term, see Section 4. One also may wish to consider the possibility of mapping a random function into a random sequence. In his dissertation Adomian presented and discussed the important example of a randomly sampled random function. He also sets up conditions under which a stationary random function is mapped in this manner into a stationary random sequence.

stationary process since $E([F(s)+G(s)][\overline{F(t)+G(t)}])$ $= f(t-s) + g(t-s) = u(t-s)$ where $u = f+g$. Note that in general the sum of two stationary processes is not stationary. A necessary and sufficient condition that $F(t) + G(t)$ be stationary is that $E(G(s)\overline{F(t)}) + E(F(s)\overline{G(t)})$ be a function of $t-s$, and we see that this is a fairly strong restriction. The fact that stationary processes do not form a linear manifold (in the space of all processes) causes a certain amount of difficulty in determining what s.t.'s preserve stationarity.

iii) Let the stationary process $F(t)$ be n times continuously differentiable and let $c_0, \ldots, c_n$ be constants. Then the s.p. $G(t) = \sum_{k=0}^{n} c_k F^{(k)}(t)$ is stationary and

$$E(G(s)\overline{G(t)}) = \sum_{k,j=0}^{n} c_k \bar{c}_j \, (-1)^j \, f^{(k+j)}(t-s).$$

We note that limits (in mean square) of stationary processes need not be stationary, and so it is unusual that linear combinations of derivatives of stationary processes are stationary. As a rule, the function $G(s) = X_0 + \int_0^s F(t)dt$ is not stationary even though F is, and so integral operators do not in general preserve stationarity. Intuitively the solution to a stochastic differential equation is representable in the form of applying a stochastic integral operator to the forcing function of the differential equation; thus we see this operation will not often yield as a stationary solution. We now seek to determine conditions under which stochastic

differential equations do possess stationary solutions.

SECTION 2

This section is devoted to the study of analytic random functions and their application to stochastic differential equations. We recall that if $X(t)$ is a <u>second order</u> random function, then $X(t)$ can be thought of as a map $X:T \longrightarrow L^2(\Omega)$. Usually T will be the set of real numbers or some subset of the reals; more generally, T will be a subset of a Euclidean space. There is already a general theory of analytic maps from finite dimensional Euclidean spaces into Banach spaces, so we list here only the most relevant parts of this theory.

<u>Definition</u>: Let B be a real Banach space. Let $\{c_n\}_{n=0}^{\infty}$ be a sequence of elements of B. Suppose there is a positive real number r such that the series $\Sigma \|c_n\| t^n$ converges for all real numbers t satisfying $|t| < r$. Then the series $\Sigma c_n t^n$ is called a <u>power series centered at 0 with coefficients in B</u>.

Notice that since a power series converges absolutely (by definition) in the space B and since B is complete, the series does indeed converge to an element of B for each appropriate t. The absolute convergence of power series allows us to rearrange the series however we like, and the

rearranged series will still converge to the same limit.

Theorem 1. Suppose $\Sigma a_n t^n$ and $\Sigma b_n t^n$ are two power series for $|t| < r$ with coefficients in a Banach space B. If $\Sigma A_n t^n = \Sigma b_n t^n$ for all $t \in (-r,r)$, then $a_n = b_n$ for each n. We will find this theorem on uniqueness of coefficients especially useful. For a proof, see Dieudonne [3].

Definition. Suppose f is a function from the reals into a Banach space B. Suppose there are elements $\{C_n\}_{n=0}^{\infty}$ in B such that $f(t) = \Sigma(C_n/n!)t^n$ for $|t| < r$. Then f is said to be analytic at 0.

In accordance with this definition, a second order random function X(t) is analytic at 0 (in the mean square sense) when there are second order random variables $X_0, X_1, \ldots$ such that $X(t) = \Sigma(X_n/n!)t^n$ for $|t| < r$. If we include the dependence on ω in our notation, we see that $X(t,\omega) = \Sigma(X_n(\omega)t^n)/n!$ and so an analytic random function is one which has this special sort of separation of variables.

We have the following theorem which relates analyticity of a random function to the analyticity of its covariance function.

Theorem 2. A second order random function X(t) is analytic if and only if its covariance function $\Gamma_X(s,t)$ is analytic at every diagonal point (t,t). If this condition is

satisfied, then $\Gamma_X(s,t)$ *is analytic at each* (s,t).

More simply, a random function is analytic if and only if its covariance function is analytic. See Loeve [7] for a proof of Theorem 2.

Returning to more general Banach space considerations, we have the following very important theorem, again from Dieudonne.

Theorem 3. Suppose $f: R \longrightarrow B$, B *a Banach space, is analytic at* 0. *Let* $f(t) = \Sigma(C_n/n!)t^n$. *Then* f *is infinitely differentiable. Moreover,* $f^{(k)}(t) = \sum_{n=0}^{\infty} (C_{n+k}/n!)$, $f^{(k)}(t)$ *is analytic, and* $f^{(n)}(0) = C_n$.

Hence we see that the random variables occurring in a power series expansion of a random function are related in a simple way to the mean-square derivatives of the random function.

Let us now turn our attention to the question of forming a product of two elements each from a (perhaps different) Banach space. We are motivated by ordinary differential equations of the form $x'(t) + a(t)x(t) = f(t)$ but we would like to replace the functions involved in the equation by second order stochastic processes. Then we would have an equation $X'(t) + A(t)X(t) = F(t)$ where X, A, and F are maps from the reals into the Banach space

$L^2(\Omega)$. There is a natural way of attaching meaning to the formal product $A(t)X(t)$, we can form the pointwise product $[A(t)X(t)](\omega) = A(t,\omega)X(t,\omega)$. In general, this expression no longer defines a function into the space $L^2(\Omega)$ since $A(t,\omega)X(t,\omega)$ may not be square integrable (with respect to ω) for each t. Consequently, some care is needed in handing these products.

Consider now the general case of forming products in Banach spaces.

<u>Definition</u>. Let E and F be two (real) Banach spaces. A map $P: E \times F \longrightarrow E$ is called a <u>product</u> on the spaces E and F if P is bilinear and satisfies the inequality $\|P(e,f)\| \leq \|e\| \, \|f\|$ for every $e \in E$ and $f \in F$.

We usually write ef for the product $P(e,f)$. There are many examples of products, and the one which we will find useful is the following.

Let E be a Banach space, let $F = L(E,E)$ be the space of bounded linear operators on E. Note that F is a Banach space. Define the product $P: E \times F \longrightarrow E$ by $P(x,f) = f(x)$. It is elementary to verify that P defines a product in our sense.

We prove now a few generalizations of well-known theorems and see how they fit into our more general framework.

<u>Theorem 4</u>. <u>Suppose</u> $P: E \times F \longrightarrow E$ <u>is a product</u>. <u>Let</u> $A = \Sigma a_n \in E$ <u>and</u> $B = \Sigma b_n \in F$ <u>be two absolutely convergent series</u>. <u>Define</u> $C_n = \sum_{j=0}^{n} a_{n-j} b_j = \sum_{j=0}^{n} P(a_{n-j}, b_j)$. <u>Then</u> $\Sigma C_n = AB = P(A,B)$.

We can prove this theorem by slightly modifying the proof for the case $E = F = R$ found in Rudin [9] so we omit the proof here. However, as an important corollary we have: Let $A: I \longrightarrow L(E,E)$ and $x: I \longrightarrow E$ be two functions defined on I, an open interval containing 0. If A and X are both analytic at 0, then the map $f: I \longrightarrow E$ defined by $f(t) = A(t)(x(t))$ is analytic at 0.

<u>Proof</u>. Let $A(t) = \Sigma A_n t^n$ and $x(t) = \Sigma X_n t^n$. Let $t_0 > 0$ be such that both $A(t_0)$ and $X(t_0)$ converge absolutely. Then

$$f(t) = A(t)(x(t)) = (\sum_n A_n t^n)(\sum_m X_m t^m) = \sum_{m,n} A_n x_m t^{n+m} =$$

$\sum_{k=0}^{\infty} t^k \sum_{j=0}^{k} A_{k-j} X_j$. Then by Theorem 4, $f(to)$ converges, and so $f(t)$ converges absolutely for $|t| < t_0$. Hence $f(t)$ is analytic.

Keeping this concept of products in mind, we turn now to the question of differential equations involving functions from the reals into Banach spaces. Let I be an open interval containing 0 and let U be an open set in the Banach space B. Then a function $f: I \times U \longrightarrow B$ is said to be a <u>time-dependent</u>

vector field on U. A map $\alpha: I \longrightarrow U$ is an integral curve for f if α is differentiable and satisfies the equation $\alpha'(t) = f(t,\alpha(t))$. f is said to be Lipschitz at $t \in I$ if there is a constant $K > 0$ such that $\|f(t,x) - f(t,g)\| \leq K\|x-y\|$ for all x,y in U. f is said to be uniformly Lipschitz on I if there is a single constant $K > 0$ such that $\|f(t,x) - f(t,y)\| \leq K\|x-y\|$ for all $x,y \in U$ and all $t \in I$. We let $C^p(I \times U)$ denote the set of all functions from $I \times U$ into B which are p times continuously differentiable. Let $B_a(X_0) = \{y \in B: \|y-X_0\| < a\}$. Now we can state an existence theorem for certain differential equations in Banach spaces.

Theorem 5. Let I, U, and B be as above. Let $X_0 \in U$. Let $a \in (0,1)$ be a number such that $B_{2a}(X_0) \subset U$. Let $f: I \times U \longrightarrow B$ be continuous, bounded by C, and satisfy a Lipschitz condition (with constant K) uniformly with respect to I. If $b < a/C$ and $b < 1/K$, then there is a unique integral curve $\alpha: (-b,b) \times B_a(X_0) \longrightarrow U$ such that $\alpha(0) = X_0$. If $f \in C^p(I \times U)$, so is α.

In particular, we note that if f is continuously differentiable, it is continuous and satisfies a uniform Lipschitz condition. For a proof of Theorem 5, see Lang [6].

Let us see how to apply this theorem to stochastic differential equations. In particular, consider the equation $X'(t) + A(t)X(t) = F(t)$, $X(0) = X_0$, where $A(t)$ and $F(t)$

are second order random functions and X_0 is a given second order random variable. Solving the above equation for $X'(t)$, we have $X'(t) = F(t) - A(t)X(t)$, and so the vector field (on $L^2(\Omega)$) associated with this equation is given by $f(t,x) = F(t) - A(t)x$. To insure that the formal product $A(t)x$ is well defined for every $t \in I$ and every $x \in L^2(\Omega)$, we assume that there is a constant $K > 0$ such that $\operatorname{ess\,sup}_{\omega \in \Omega} |A(t,\omega| < K$ for all $t \in I$. We will show that for each t, $A(t)$ can be thought of as a bounded linear operator on $L^2(\Omega)$, whereupon the formal product $A(t)x$ will be a product as defined earlier; in particular, $A(t)x \in L^2(\Omega)$ for every $t \in I$ and every $x \in L^2(\Omega)$, and so the function $f(t,x)$ is a well defined vector field on $L^2(\Omega)$.

Suppose then that $\operatorname{ess\,sup}_{\omega} |A(t,\omega)| < K$. Consider the map $\overline{A}: I \longrightarrow L(L^2(\Omega), L^2(\Omega))$ defined by $(\overline{A(t)}x)(\omega) = A(t,\omega)x(\omega)$. We show first that for a fixed $t \in I$, $\overline{A(t)} \in L(L^2(\Omega), L^2(\Omega))$. Thus we must show that if $x \in L^2(\Omega)$, then $\overline{A(t)}x$ must be a square integrable random variable on Ω. We have

$$\int_\Omega [\overline{A(t)}x]^2(\omega)d\mu(\omega) = \int_\Omega A^2(t,\omega)x^2(\omega)d\mu(\omega) \leq$$

$$\operatorname{ess}_\omega\sup |A^2(t,\omega)| \int_\Omega x^2(\omega)d\mu(\omega) \leq K^2\|x\|^2 < \infty.$$ Thus

$\overline{A(t)}x \in L^2(\Omega)$. For a fixed t, $\overline{A(t)}$ is clearly linear, and moreover

$$\|\overline{A(t)}x\| = \left(\int_\Omega (\overline{A(t)}x)^2(\omega)d\mu(\omega)\right)^{1/2} \leq (K^2\|x\|^2)^{1/2} = K\|x\|.$$

Hence $\overline{A(t)}$ is a bounded operator and $\|\overline{A(t)}\| \leq K$. This

$X'(t) + A(t)X(t) = \sum_{k=0}^{\infty} (t^k/k!)[X_{k+1} + \sum_{j=0}^{k} \binom{k}{j} A_j X_{k-j}]$. But this equals $F(t)$, so by equating coefficients (Theorem 1), we have $F_k = X_{k+1} + \sum_{j=0}^{k} \binom{k}{j} A_j X_{k-j}$, or

$$X_{k+1} = F_k - \sum_{j=0}^{k} \binom{k}{j} A_j X_{k-j}.$$

We are given X_0, so this formula allows us to determine each X_n by induction. Now we have to show that the power series for $X(t)$ with these coefficients converges (absolutely) in some neighborhood of the origin.

We observe that the expression for X_n can be put in a more convenient form. We claim that

$$X_n = x_n X_0 + \sum_{k=0}^{n-1} {}_k f_n \, F_k,$$

where x_n is the coefficient of X_0 in the original expression for X_n (after successively substituting the previously calculated X_j's, $j \leq n-1$) and ${}_k f_n$ is the coefficient of F_k in the original expression for X_n. x_n and ${}_k f_n$ are defined inductively by the relations

$$x_0 = 1, \qquad x_{n+1} = -\sum_{j=0}^{n} \binom{n}{j} A_j x_{n-j}$$

$${}_k f_n = 0 \qquad \text{for } n \leq k$$

$${}_k f_{k+1} = 1, \quad {}_k f_{n+1} = -\sum_{j=0}^{n} \binom{n}{j} A_j \, {}_k f_{n-j} \quad \text{for } n > k.$$

Notice that the expressions for x_n and ${}_kf_n$ are combinations of the A_j's only; the initial condition X_0 and the forcing function $F(t)$ are not involved in these coefficients.

Proof of claim: We use induction on n. The claim clearly holds for $n = 1$. Suppose the claim holds for $k \leq n$ and examine the case $k = n+1$.

$$X_{n+1} = F_n - \sum_{j=0}^{n} \binom{n}{j} A_j X_{n-j} = F_n - \sum_{j=0}^{n} \binom{n}{j} \left[\sum_{k=0}^{n-j-1} {}_kf_{n-j} F_k + x_{n-j} X_0 \right] =$$

$$= F_n - \sum_{j=0}^{n} \sum_{k=0}^{n-j-1} \binom{n}{j} A_j \, {}_kf_{n-j} F_k - \sum_{j=0}^{n} \binom{n}{j} A_j \, x_{n-j} \, X_0 =$$

$$\left(-\sum_{j=0}^{n} \binom{n}{j} A_j x_{n-j}\right) X_0 + F_n - \sum_{j=0}^{n} \sum_{k=0}^{n-j-1} \binom{n}{j} A_j \, {}_kf_{n-j} F_k .$$

But $-\sum_{j=0}^{n} \binom{n}{j} A_j x_{n-j} = x_{n+1}$, so all that remains is to show

$$F_n - \sum_{j=0}^{n} \sum_{k=0}^{n-j-1} \binom{n}{j} A_j \, {}_kf_{n-j} F_k = \sum_{k=0}^{n} {}_kf_{n+1} F_k .$$

Fix an integer p such that $0 \leq p < n$. What is the coefficient of F_p in the left hand side of the above equation? Notice that $k = p$ only when j satisfies $n-j-1 \geq p$, i.e., $j \leq n-p-1$. Thus we get an F_p for $j = 0,1,\ldots,n-p-1$ and $k = p$.

Hence the complete contribution involving F_p is $-F_p \sum_{j=0}^{n-p-1} \binom{n}{j} A_j \, {}_pf_{n-j}$. But ${}_pf_k = 0$ for $p \geq k$, so if $p \geq n-j$ we have ${}_pf_{n-j} = 0$. Since $p \geq n-j$ for $j \geq n-p$, we have

$$-\sum_{j=0}^{n-p-1}\binom{n}{j}A_j\ {}_pf_{n-j} = -\sum_{j=0}^{n-p-1}\binom{n}{j}A_j\ {}_pf_{n-j} - \sum_{j=n-p}^{n}\binom{n}{j}A_j\ {}_pf_{n-j} =$$

$$= -\sum_{j=0}^{n}\binom{n}{j}A_j\ {}_pf_{n-j} = {}_pf_{n+1}.$$ Hence $F_n - \sum_{j=0}^{n}\sum_{k=0}^{n-j-1}\binom{n}{j}A_j\ {}_kf_{n-j}\ F_k =$

$= \sum_{k=0}^{n} {}_kf_{n+1}\ F_k$ and the induction is complete.

As a consequence of the above relationships, we have

$$X(t) = \Sigma(t^n/n!)X_n = \sum_n(t^n/n!)[x_nX_0 + \sum_{k=0}^{n-1}{}_kf_n\ F_k] =$$

$$= X_0\ \sum_n(x_n/n!)t^n + \sum_n(t^n/n!)\ (\sum_{k=0}^{n-1}{}_kf_n\ F_k).$$ We will now

place additional restrictions on $A(t)$ so that we can prove the above power series converge.

Let $A(t) = \Sigma(A_n/n!)t^n$, and suppose there is some constant K such that $\operatorname{ess\,sup}_\omega |A_n(\omega)| \leq K^n$ for each n. Then each $A_n(\omega)$ can be thought of as a bounded linear operator on $L^2(\Omega)$, i.e., $A_n \in L(L^2(\Omega), L^2(\Omega))$. Also $\operatorname{ess\,sup}_\omega |A(s,\omega)| < K' < \infty$ for some K' and for all s in some neighborhood of 0, and so our original restriction on $A(s,\omega)$ (allowing us to form products) is satisfied. We wish to show that $A(s) = \Sigma(A_n/n!)s^n$ is an analytic map from I into the Banach space $L(L^2(\Omega), L^2(\Omega))$. First we need to calculate the norm of A_n considered as an element of $L(L^2(\Omega), L^2(\Omega))$, and we have the following theorem.

<u>Theorem 6</u>. <u>Let</u> $A: \Omega \longrightarrow R$ <u>be a random variable such that ess sup</u> $|A(\omega)| < \infty$. <u>Then the norm of A considered as an element of</u> $L(L^2(\Omega), L^2(\Omega))$ <u>is ess sup</u> $|A(\omega)|$.

<u>Proof</u>: Let $\|A\|$ denote the norm of the linear operator generated by A. Then $\|A\| = \sup_{\|x\|=1} \|Ax\|$. Let $x: \Omega \longrightarrow R$ be such that $\|x\| = 1$. Then

$$\|Ax\| = \left(\int_\Omega (Ax)^2(\omega)d\mu(\omega)\right)^{1/2} = \left(\int_\Omega A^2(\omega)x^2(\omega)d\mu(\omega)\right)^{1/2} \leq (\text{ess}_\omega \sup A^2(\omega)$$

$$\int_\Omega x^2(\omega)d\mu(\omega))^{1/2} = \text{ess}_\omega \sup |A(\omega)| \; \|x\| = \text{ess}_\omega \sup |A(\omega)|.$$

Hence $\|A\| \leq \text{ess}_\omega \sup |A(\omega)|$. We show now that $\|A\| \geq \text{ess sup } |A(\omega)|$. Let $\epsilon > 0$ be given. Let $K = \text{ess sup } |A(\omega)|$ and define the set $D = \{\omega : |A(\omega)| \geq K - \epsilon\}$. We may suppose without loss of generality that $\mu(D) = \delta > 0$. Define $x(\omega) = \chi(\omega)\delta^{-1/2}$ where $\chi(\omega) = 1$ for $\omega \in D$ and $\chi(\omega) = 0$ otherwise. Then

$$\|x\| = \left(\int_\Omega x^2(\omega)d\mu(\omega)\right)^{1/2} = \left(\int_D \delta^{-1}d\mu(\omega)\right)^{1/2} = (\mu(D)\delta^{-1})^{1/2} = 1,$$

and $\|Ax\| = \left(\int_\Omega A^2(\omega)x^2(\omega)d\mu(\omega)\right)^{1/2} = \left(\delta^{-1} \int_D A^2(\omega)d\mu(\omega)\right)^{1/2} \geq$

$\geq (\delta^{-1}(K-\epsilon)^2\mu(D))^{1/2} = K-\epsilon$. Hence $\|A\| \geq K-\epsilon$, and since ϵ is arbitrary, we see that $\|A\| \geq K$. Thus $\|A\| = \text{ess}_\omega \sup |A(\omega)|$.

QED

We now see that $\Sigma(A_n/n!)t^n$ is a power series in $L(L^2(\Omega), L^2(\Omega))$ since $\Sigma(\|A_n\| \; |t^n|/n!) =$

$$\Sigma\ (\mathrm{ess}_{\omega}\sup\ |A_n(\omega)|\ \ |t|^n/n!) \leq \Sigma\ (K^n\ \ |t|^n/n!) = e^{K|t|} < \infty.$$

Hence $A(t) = \Sigma(A_n t^n/n!)$ is analytic in $L(L^2(\Omega), L^2(\Omega))$. Form the function $y(t) = -\int_0^t A(s)ds$ where integration takes place in $L(L^2(\Omega), L^2(\Omega))$, and then consider the bounded linear operator (for each t) exp y(t). This is an analytic map, and simple algebra and an inductive proof shows that its power series has the coefficients x_n defined previously. Hence applying the corollary to Theorem 4, the function $X_0 \exp\ (-\int_0^t A(s)ds)$ is an analytic map from I into $L^2(\Omega)$.

Now consider the expression:

$\exp\ (-\int_0^t A(s)ds)\ \int_0^t \exp\ (\int_0^y A(s)ds)\ F(y)dy$ where the indicated integrations involving the exponentials take place in $L(L^2(\Omega), L^2(\Omega))$ and the remaining integration takes place in $L^2(\Omega)$. Since F(y) is analytic (in $L^2(\Omega)$), $\exp\ (\int_0^y A(s)ds)\ F(y)$ is analytic (in $L^2(\Omega)$) as before, and consequently so is its integral $\int_0^t$; we apply once more the operator $\exp\ (-\int_0^t A(s)ds)$, so the whole expression defines an analytic function in $L^2(\Omega)$. Again an inductive proof shows that the coefficients of this analytic map are $\sum_{k=0}^{n-1} {}_k f_n\ F_k$ as previously defined. Hence our power series for X(t) converges (absolutely) in some neighborhood of 0, and analyticity is established. We summarize our results with the following theorem.

<u>Theorem 7. Let</u> $F(t)$ <u>and</u> $A(t) = \Sigma(A_n(\omega)/n!)t^n$ <u>be analytic second order random functions and let</u> X_0 <u>be a square integrable random variable. Suppose there is a constant</u> K <u>such that</u> $\operatorname{ess\,sup}_\omega |A_n(\omega)| \leq K^n$ <u>for every</u> n. <u>Then the stochastic differential equation</u> $X'(t) + A(t)X(t) = F(t)$, $X(0) = X_0$ <u>has a unique analytic solution.</u>

Note that the extension of this theorem to higher order equations is trivial. If we have the equation $X^{(n)}(t) + \alpha_{n-1}(t)X^{(n-1)}(t) + \cdots + \alpha_0(t)X(t) = F(t)$ and the coefficients are analytic and satisfy $\operatorname{ess\,sup}_\omega |\alpha_j^{(k)}(0)| \leq K_j^k$ for some set $\{K_j\}_{j=0}^{n-1}$ and every k, then we write the equation in a vector form

$$X'(t) = A(t)X(t) + G(t)$$

where

$$A(t) = \begin{bmatrix} 0 & 1 & 0 & & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & & \cdots & 1 \\ -\alpha_0(t) & -\alpha_1(t) & & & \cdots & -\alpha_{n-1}(t) \end{bmatrix}$$

and $G(t) = \begin{bmatrix} 0 \\ \vdots \\ F(t) \end{bmatrix}$.

Note that analyticity of the $\alpha_j(t)$'s implies that of $A(t)$ (as a bounded linear operator on $(L^2(\Omega))^n$) and $G(t)$ is obviously analytic. Hence the same techniques of our

theorem extend to this case and to the state space equations of stochastic control theory, in particular to the recent work of Leon H. Sibul (dissertation).[1]

Turning to the question of stationarity, we are interested in finding necessary and sufficient conditions that an analytic random function be stationary. We will assume all our random functions satisfy $< X(t) > = 0$. Suppose then that $X(t) = \Sigma(X_n/n!)t^n$ is a real analytic random function with the (analytic) covariance function $\Gamma(s,t)$. Suppose also that $X(t)$ is stationary. Then there is some function $f: R \longrightarrow R$ such that $\Gamma(s,t) = f(t-s)$. Note that $f(u) = f(-u)$. Since $\Gamma(s,0) = f(-s)$, we see that f is analytic (at 0), so there are real numbers c_n such that $f(s) = \Sigma(c_n/n!)s^n$. Let us see how these constants are related to $X(t)$. We have

$$\Gamma(s,t) = < X(s)X(t) > = < \sum_m (s^m X_m/m!) \sum_n (t^n X_n/n! > =$$

$$= \sum_{n=0}^{\infty} \sum_{j=0}^{n} (s^j t^{n-j}/j!(n-j)!) < X_j X_{n-j} > =$$

$$= \sum_{n=0}^{\infty} (1/n!) \sum_{j=0}^{n} \binom{n}{j} s^j t^{n-j} < X_{n-j} X_j >. \quad \text{But}$$

$$f(s-t) = \sum_{n=0}^{\infty} (c_n (s-t)^n/n!) = \sum_{n=0}^{\infty} (1/n!) \sum_{j=0}^{n} (-1)^{n-j} c_n \binom{n}{j} s^j t^{n-j}.$$

Since $f(s-t) = \Gamma(s,t)$, we have these two power series (in two variables) representing the same function, hence their coefficients must be equal, thus $(-1)^{n-j} c_n = < X_{n-j} X_j >$. Rewriting this last equation, we have $< X_n X_m > = (-1)^n c_{n+m}$.

---

[1] Sibul, L. H., Application of Linear Stochastic Operator Theory, Pennsylvania State University dissertation, 1968.

On the other hand, suppose we have a zero mean analytic random function $X(t) = \Sigma(X_n t^n/n!)$ such that there exist constants $\{c_n\}$ satisfying $< X_n X_m > = (-1)^n c_{n+m}$. Then the above equations show that $X(t)$ is stationary. Thus we have the following theorem.

<u>Theorem 8</u>. <u>Let</u> $X(t) = \Sigma(X_n t^n/n!)$ <u>be a (zero mean) real analytic random function. Then</u> $X(t)$ <u>is stationary if and only if there are constants</u> $\{c_n\}$ <u>such that</u> $< X_n X_m > = (-1)^n c_{n+m}$.

It is clear that a set of constants $\{c_n\}$ satisfying the condition of Theorem 8 cannot be completely arbitrary. In fact, we must have

i) $c_{2k+1} = 0$, ii) $c_{4k} \geq 0$, and iii) $c_{4k+2} \leq 0$.

To see i), note that $< X_{2k+1} X_0 > = (-1)^{2k+1} c_{2k+1} = -c_{2k+1}$ and $< X_0 X_{2k+1} > = (-1)^0 c_{2k+1}$. Since $X(t)$ is real-valued, we have $-c_{2k+1} = c_{2k+1}$, <u>i.e.</u>, $c_{2k+1} = 0$. This is to be expected since $f(t)$ must be an even function. Also $0 \leq < X_{2k}^2 > = < X_{2k} X_{2k} > = (-1)^{2k} c_{4k} = c_{4k}$ and $0 \leq < X_{2k+1}^2 > = < X_{2k+1} X_{2k+1} > = (-1)^{2k+1} c_{4k+2}$, hence ii) and iii) are proved. By defining $d_n = |c_{2n}|$, we can write $f(t)$ in the form $f(t) = \sum_{n=0}^{\infty} (-1)^n (d_n/(2n)!) t^{2n}$ where $d_n \geq 0$ for all n, hence $f(t)$ is representable by an even, alternating power series. (Note that $d_n = < X_n^2 >$).

Theorem 8 can be reformulated in the following way.

<u>Theorem 9</u>. <u>Let</u> $X(t) = \Sigma(X_n t^n/n!)$ <u>be a (zero mean) real analytic random function</u>. <u>Let</u> $A = \{X_0, X_2, \ldots, X_{2n}, \ldots\}$ <u>and</u> $B = \{X_1, X_3, \ldots, X_{2n+1}, \ldots\}$. <u>Then</u> $X(t)$ <u>is stationary if and only if</u> A <u>is orthogonal to</u> B <u>and</u>

i) $\langle X_{2n}X_0\rangle = \langle X_{2(n-k)}X_{2k}\rangle \qquad k = 0, \ldots, n$

ii) $\langle X_{2n+1}X_1\rangle = \langle X_{2(n-k)+1}X_{2k+1}\rangle \qquad k = 0, \ldots, n$

<u>and</u> iii) $\langle X_{2n}X_0\rangle = -\langle X_{2n-1}X_1\rangle$ for all $n \geq 1$.

<u>Proof</u>: Suppose $X(t)$ is stationary. We show that $A \perp B$ first: $\langle X_{2n}X_{2k+1}\rangle = (-1)^{2n}c_{2n+2k+1} = c_{2(n+k)+1} = 0$. Hence $A \perp B$. Also,

1) $\langle X_{2(n-k)}X_{2k}\rangle = (-1)^{2(n-k)}c_{2n-2k+2k} = c_{2n} = \langle X_{2n}X_0\rangle$

2) $\langle X_{2(n-k)+1}X_{2k+1}\rangle = (-1)^{2(n-k)+1}c_{2n-2k+1+2k+1} =$

$= -c_{2n+2} = (-1)^{2n+1}c_{2n+1+1} = \langle X_{2n+1}X_1\rangle$

3) $\langle X_{2n-1}X_1\rangle = (-1)^{2n-1}c_{2n} = -c_{2n} = -\langle X_{2n}X_0\rangle$,

so the first half of the theorem is proved.

Suppose now the second half of the theorem holds. Define $\{c_n\}$ by the equations $c_{2m+1} = 0 \quad m = 0, 1, \ldots,$ and $c_{2m} = \langle X_{2m}X_0\rangle \quad m = 0, 1, \ldots$. We will show that $\langle X_{n-j}X_j\rangle = (-1)^{n-j}c_n$ for $0 \leq j \leq n$ and all n, whereupon Theorem 8 tells us that $X(t)$ is stationary. Suppose first that n is odd. If j is even, $n-j$ is odd and $X_{n-j} \in B$, hence

$\langle X_{n-j}X_j \rangle = 0$ since $X_j \in A$. Thus $\langle X_{n-j}X_j \rangle = 0 = (-1)^{n-j}c_n$ since $n = 2m+1$ for some m. If j is odd, $n-j$ is even, so $X_{n-j} \in A$ and $X_j \in B$, hence $\langle X_{n-j}X_j \rangle = 0 = (-1)^{n-j}c_n$ as before. Now we show Theorem 8 is satisfied for even n. Let $n = 2p$. First consider the case where j is odd, let $j = 2q+1$. Then

$$\langle X_{n-j}X_j \rangle = \langle X_{2(p-q)-1}X_{2q+1} \rangle = \langle X_{2(p-1-q)+1}X_{2q+1} \rangle =$$

$$= \langle X_{2(p-1)+1}X_1 \rangle = \langle X_{2p-1}X_1 \rangle = -\langle X_{2p}X_0 \rangle =$$

$$= (-1)^{2(p-q)-1}c_{2p} = (-1)^{n-j}c_n.$$

Now suppose j is even, $j = 2q$. Then

$$\langle X_{n-j}X_j \rangle = \langle X_{2(p-q)}X_{2q} \rangle = \langle X_{2p}X_0 \rangle c_{2p} =$$

$$= (-1)^{2(p-q)}c_{2p} = (-1)^{n-j}c_n$$

as was to be shown. Hence $X(t)$ is stationary. QED

We now have developed a technique for finding the power series coefficients of solutions to stochastic differential equations and we also have theorems which tell us when a given analytic stochastic process is stationary, so in principle we have the machinery to determine conditions under which a given equation will have stationary solutions. We present some examples to show how this may be done.

<u>Example 1</u>. We know that the derivative of a stationary random function is stationary. When will the integral of

a stationary (analytic) random function be stationary? This question is equivalent to determining stationary solutions to the differential equation $X'(t) = F(t)$, $X(0) = X_0$ where $F(t)$ is the given stationary (analytic) random function and the random variable X_0 is yet to be specified. The soltuion to this equation is

$$X(t) = X_0 + \int_0^t F(s)ds = X_0 + F_0 t + F_1 t^2/2! + \cdots = \sum_{n=0}^{\infty} (X_n t^n/n!)$$

where $X_n = F_{n-1}$ for $n \geq 1$.

We now apply Theorems 8 and 9 to this random function to see what additional conditions we need to place on X_0 to guarantee stationarity. Theorem 9 tells us that in order for $X(t)$ to be stationary, it is necessary that $\langle X_{2n}X_0 \rangle = (-1)^n \langle X_n^2 \rangle$, i.e., we must have $\langle F_{2n+1}X_0 \rangle = \langle X_{2(n+1)}X_0 \rangle = (-1)^{n+1} \langle X_{n+1}^2 \rangle = (-1)^{n+1} \langle F_n^2 \rangle$. Hence our first restriction on X_0 is that it must satisfy the relations $\langle F_{2n+1}X_0 \rangle = (-1)^{n+1} \langle F_n^2 \rangle$. Moreover, Theorem 9 requires that $\langle X_0 X_{2n+1} \rangle = 0$ for all n, hence our second requirement is that $\langle X_0 F_{2n} \rangle = 0$ for all n. Since $F(t)$ is stationary, there are constants $\{f_n\}$ such that $\langle F_{n-j}F_j \rangle = (-1)^{n-j} f_n$. Define constants $\{c_n\}$ by the relations $c_{2n+1} = 0$, $c_0 = \langle X_0^2 \rangle$, and $c_{n+2} = -f_{2n}$. We will show that $\langle X_{n-j}X_j \rangle = (-1)^{n-j} c_n$.

Case 1. Let n be odd, $n = 2k+1$. We must show $\langle X_{n-j}X_j \rangle = 0$ for all appropriate j. For $j = 0$, $\langle X_{n-j}X_j \rangle = \langle F_{n-1}X_0 \rangle = \langle F_{2k}X_0 \rangle = 0$. For $j \geq 1$, $\langle X_{n-j}X_j \rangle =$

$= \langle F_{n-j-1}F_{j-1}\rangle = (-1)^{n-j-1}f_{n-2} = (-1)^{n-j-1}f_{2k-1} = 0$ since the f_n's with odd subscripts must be zero.

Case 2. Let n be even, $n = 2k$. For $j = 0$, $\langle X_{n-j}X_j\rangle = \langle X_nX_0\rangle = \langle F_{n-1}X_0\rangle = \langle F_{2k-1}X_0\rangle = \langle F_{2(k-1)+1}X_0\rangle =$

$= (-1)^k \langle F_{k-1}^2\rangle = (-1)^k f_{2(k-1)}(-1)^{k-1} = (-1)^{2k-1}f_{2(k-1)} =$

$(-1)^{2k}c_{2k} = (-1)^n c_n$ as required. For $j \geq 1$, $\langle X_{n-j}X_j\rangle =$

$= \langle F_{n-j-1}F_{j-1}\rangle = (-1)^{n-j-1}f_{n-2} = (-1)^{2k-j-1}f_{2k-2} =$

$= (-1)^{2k-j-1}f_{2(k-1)} = (-1)^{2k-j}c_{2(k-1)+2} = (-1)^{2k-j}c_{2k} =$

$= (-1)^{n-j}c_n$ as required. Hence we see that necessary and sufficient conditions for $X(t)$ to be stationary are that $\langle F(t)\rangle = 0$, $\langle F_{2n+1}X_0\rangle = (-1)^{n+1}\langle F_n^2\rangle$, and $\langle F_{2n}X_0\rangle = 0$.

These last two conditions may be combined by requiring that

$$\langle X_0F(t)\rangle = \langle X_0\ \Sigma(F_nt^n/n!)\rangle = \Sigma(t^n/n!)\langle X_0F_n\rangle =$$

$$\Sigma(t^{2n}/(2n)!)\langle X_0F_{2n}\rangle + \Sigma(t^{2n+1}/(2n+1)!)\langle X_0F_{2n+1}\rangle =$$

$$= \Sigma(t^{2n+1}/(2n+1)!)\ (-1)^{n+1}\langle F_n^2\rangle = -\Sigma(t^{2n+1}/(2n+1)!)\ \times$$

$$\times\ (-1)^n\langle F_n^2\rangle,$$

<u>i.e.</u>, we require that the correlation function of $F(t)$ and X_0 be given by

$$\langle F(t)X_0\rangle = -\Sigma((-1)^n t^{2n+1}/(2n+1)!)\ \langle F_n^2\rangle.$$

Example 2. We present here an example of a differential equation in which the coefficient function as well as the forcing function is random. Consider the equation

$$X'(t) - Ae^{-tA}X(t) = AB(e^{tA} - 1)$$
$$X(0) = B$$

where A is an essentially bounded r.v. Writing the equation in the form $X'(t) + A(t)X(t) = F(t)$, we calculate that $A(t) = -Ae^{-tA} = \Sigma(-A)^{n+1}t^n/n!$ i.e., $A_n = (-A)^{n+1}$. Also we have $F(t) = AB(e^{tA} - 1) = \sum_{n=1}^{\infty} A^{n+1}Bt^n/n!$ so we have $F_0 = 0$ and $F_n = A^{n+1}B$ for $n \geq 1$. We have a solution $X(t) = \Sigma X_n t^n/n!$ where $X_n = x_n X_0 + \sum_{k=0}^{n-1} {}_k f_n F_k =$

$$= x_n B + \sum_{k=1}^{n-1} {}_k f_n (A^{k+1}B) = B(x_n + \sum_{k=1}^{n-1} {}_k f_n A^{k+1}) = BC_n$$

where we define C_n to be the expression in the brackets. Remember that both x_n and ${}_k f_n$ are polynomials in the variables $A_0, \ldots, A_{n-1}$, and since the variables A_k are polynomials in A ($A_k = (-A)^{k+1}$), we conclude that C_n is a polynomial in the r.v. A. Now if X(t) is to be stationary, we must have $\langle X_0X_2\rangle = -\langle X_1^2\rangle$, and this condition becomes $\langle A^2B^2\rangle = 0$ and hence we know that any polynomial in A is orthogonal to any polynomial in B. Thus we have $\langle X_nX_m\rangle =$ $= \langle BC_nBC_m\rangle = \langle B^2C_nC_m\rangle = 0$ for $n + m \geq 1$ and consequently Theorem 8 is satisfied. Thus the solution is stationary if and only if $\langle A^2B^2\rangle = 0$. Note that we did not actually have to calculate the solution to make this conclusion.

Example 3. Consider the equation $X''(t) + BX(t) = 0$, $X(0) = X_0$, $X'(0) = X_1$, where B is a random variable which is positive a.s. This equation is a stochastic analog of the harmonic oscillator equation. Suppose $X(t) = \Sigma(X_n/n!)t^n$, then $X''(t) = \Sigma(X_{n+2}/n!)t^n$, so $0 = X''(t) + BX(t) =$

$= \Sigma(X_{n+2}/n!)t^n + B\ \Sigma(X_n/n!)t^n = \Sigma((X_{n+2}+BX_n)/n!)t^n$, i.e., $X_{n+2} = -Bx_n$ for all n. It is easy to see that the coefficients X_n are given by the relations $X_{2n} = (-B)^n X_0$ and $X_{2n+1} = (-B)^n X_1$. The closed form expression with these coefficients is $X(t) = X_0 \cos\sqrt{B}\ t + (X_1/\sqrt{B})\sin\sqrt{B}\ t$.

To see when this random function will be stationary, let us assume that the random variables B, X_0, and X_1 are all independent and $\langle X_0\rangle = \langle X_1\rangle = 0$. We apply Theorem 9, so define $A = \{X_0, X_2, X_4, \ldots\}$ and $C = \{X_1, X_3, X_5, \ldots\}$. Now A and C are orthogonal since $\langle X_{2n}X_{2p+1}\rangle = \langle (-B)^n X_0 (-B)^p X_1\rangle =$ $= \langle (-B)^{n+p}\rangle \langle X_0\rangle\langle X_1\rangle = 0$. We check the last three conditions of the theorem:

i) $\langle X_{2(n-k)}X_{2k}\rangle = \langle (-B)^{n-k}X_0(-B)^k X_0\rangle =$

$\langle (-B)^n X_0 X_0\rangle = \langle X_{2n}X_0\rangle$

ii) $\langle X_{2(n-k)+1}X_{2k+1}\rangle = \langle (-B)^{n-k}X_1(-B)^k X_1\rangle =$

$\langle (-B)^n X_1 X_1\rangle = \langle X_{2n+1}X_1\rangle$

iii) $\langle X_{2n}X_0\rangle = \langle (-B)^n X_0^2\rangle = \langle (-B)^n\rangle\langle X_0^2\rangle =$

$(-1)^n\langle B^n\rangle\langle X_0^2\rangle$.

Also $\langle X_{2n-1}X_1\rangle = \langle (-B)^{n-1}X_1^2\rangle = (-1)^{n-1}\langle B^{n-1}\rangle\langle X_1^2\rangle$.

But these expressions must be negatives of one an[illegible]

and so $\langle B^n\rangle = \langle B^{n-1}\rangle\langle X_1^2\rangle/\langle X_0^2\rangle$. Let $c = \langle X_1^2\rangle/$[illegible]

then $\langle B^n\rangle = c\langle B^{n-1}\rangle$, i.e., $\langle B^n\rangle = c^n$ for every n. The characteristic function for B is $\langle \exp itB\rangle =$

$\Sigma\, i^n t^n \langle B^n\rangle/n! = \Sigma(itc)^n/n! = \exp itc$, so $B = c$ a.s. since characteristic functions are unique. Thus if we assume B, X_0, and X_1, are independent and $\langle X_0\rangle = \langle X_1\rangle = 0$, the solution is stationary if and only if $B(\omega) = \langle X_1^2\rangle/\langle X_0^2\rangle$ a.s..

These three examples serve to indicate a fairly wide range of questions which the techniques of this section can answer. The first example provides us with a criterion which we will use in section 3 to characterize the general form of stationary solutions to the equation. Example 2 shows that we may characterize conditions for the existence of stationary solutions without having to find the solution itself. Example 3 says that a more general canonical form of simple stationary processes (i.e. adding randomness in the time functions of A cos t + B sin t) is not needed.

SECTION 3

In this section we use some of the tools of random harmonic analysis to develop further conditions guaranteeing stationarity of solutions to certain stochastic differential equations. We make extensive use of the harmonic decomposition theorems concerning stationary processes to find sufficient conditions to be placed on the initial values of the equations. We then calculate the correlation function of the resulting solutions. Moreover, we show that under fairly general conditions we can get asymptotic stationarity independent of the initial values. We begin with the notion of an orthogonal random measure.

Let $\mathcal{B}$ denote the family of Borel subsets of the real line and let $\mathcal{A}$ denote the subfamily of bounded Borel subsets. Then a function $\xi:\mathcal{A} \times \Omega \longrightarrow C$ is called an <u>orthogonal random measure</u> if

i) $\xi(A) \in L^2(\Omega)$ for each $A \in \mathcal{A}$

ii) $E(\xi(A)) = 0$ for each $A \in \mathcal{A}$

iii) $E(\xi(A)\overline{\xi(B)}) = 0$ if $A \cap B = \emptyset$, $A,B \in \mathcal{A}$

iv) the relation $M(A) = E(|\xi(A)|^2)$ defines a measure on $\mathcal{A}$.

The measure M is called the <u>absolute measure associated with ξ</u>. If $f:R \longrightarrow C$ is a complex valued measurable function such that $\int_a^b |f(\lambda)|^2 dM(\lambda) < \infty$, we can define the integral

$\int_a^b f(\lambda)\xi(d\lambda)$ in a natural way (see Rozanov [8]).

We say that a second order r.f. $\eta(t)$ is <u>separable</u> if the smallest closed linear manifold in $L^2(\mathfrak{H})$ containing the r.v.'s $\{\eta(t)\}_{t\in R}$ is separable, <u>i.e.</u>, it contains a countable dense subset. We remark that a continuous r.f. $\eta(t)$ is necessarily separable, for let $\mathcal{H}$ be a basis for $L^2(\mathfrak{H})$ and let $\{t_n\}_{n=1}^{\infty}$ denote the set of rational numbers. Then for each n, there is a countable subset of $\mathcal{H}$ (call it $H_n = \{h_{nm}\}_{m=1}^{\infty}$) such that $\eta(t_n) = \sum_{m=1}^{\infty} E(\eta(t_n)\overline{h_{nm}})h_{nm}$. Let $H = \{a_n\}_{n=1}^{\infty}$ be an orthonormal basis for the smallest closed linear manifold containing $\bigcup_{n=1}^{\infty} H_n$. We claim that for every t, $\eta(t) = \sum_{n=1}^{\infty} E(\eta(t)\overline{a_n})a_n$. This equation obviously holds if t is rational, so we suppose that t is irrational and $\epsilon > 0$. Since η is continuous, there is a t_k such that $\|\eta(t) - \eta(t_k)\| < \epsilon/3$. There is an N such that if $n \geq N$, then $\|\eta(t_k) - \sum_{m=1}^{n} E(\eta(t_k)\overline{a_m})a_m\| < \epsilon/3$. Hence, for $n \geq N$

$$\|\eta(t) - \sum_{m=1}^{n} E(\eta(t)\overline{a_m})a_m\| \leq \|\eta(t) - \eta(t_k)\| +$$

$$+ \|\eta(t_k) - \sum_{m=1}^{n} E(\eta(t_k)\overline{a_m})a_m\| + \|\sum_{m=1}^{n} E(\eta(t_k)\overline{a_m})a_m - \sum_{m=1}^{n} E(\eta(t)\overline{a_m})a_m\| <$$

$$< \epsilon/3 + \epsilon/3 + \|\sum_{m=1}^{n} [E(\eta(t_k)\overline{a_m}) - E(\eta(t)\overline{a_m})]a_m\| \leq$$

$$\leq 2\,\epsilon/3 + \|\eta(t_k) - \eta(t)\| < \epsilon.$$

Hence the curve $\eta(t)$ is contained in a separable manifold and so $\eta(t)$ is separable.

Rozanov proves that every separable stationary process (with measurable correlation function) is representable in the form

$$\eta(t) = \int_{-\infty}^{\infty} e^{2\pi i t\lambda} \xi(d\lambda)$$

where ξ is an orthogonal random measure. (ξ is called the <u>spectral random measure</u> associated with η). This is one of the most powerful results in random harmonic analysis and we will make extensive use of this theorem throughout the rest of the paper. Relevant discussions of these integrals and decompositions are presented in Rozanov's text [8] and in Irzhina's paper [5].

In particular, we are especially interested in the equation

$$(1) \qquad X'(t) + aX(t) = F(t)$$
$$X(0) = X_0$$

where a is an essentially bounded r.v., $X_0 \in L^2(\mathcal{B})$, and F(t) is a stationary continuous r.f. We will also suppose that a and F(t) are independent. As is well known, the solution to the above equation can be expressed as

$$X(t) = X_0 e^{-ta} + e^{-ta} \int_0^t e^{ya} F(y)dy.$$

We write $F(y) = \int_{-\infty}^{\infty} e^{2\pi i y\lambda} \xi(d\lambda)$ where ξ is the spectral random measure associated with the process F(y). Assume for the moment that the following calculations are valid and calculate $\int_0^t e^{ya} F(y)dy = \int_0^t e^{ya} \int_{-\infty}^{\infty} e^{2\pi i y\lambda} \xi(d\lambda)dy =$

$$= \int_{-\infty}^{\infty} \int_0^t e^{y(2\pi i\lambda+a)} dy\, \xi(d\lambda) = \int_{-\infty}^{\infty} ((e^{t(2\pi i\lambda+a)}-1)/2\pi i\lambda+a)\xi(d\lambda).$$

Thus

$$X(t) = X_0 e^{-ta} + e^{-ta}\int_{-\infty}^{\infty} ((e^{t(2\pi i\lambda+a)}-1)/2\pi i\lambda+a)\xi(d\lambda)$$

$$= X_0 e^{-ta} + \int_{-\infty}^{\infty} ((e^{2\pi i t\lambda}-e^{-ta})/2\pi i\lambda+a)\xi(d\lambda)$$

(2)

$$= e^{-ta}[X_0 - \int_{-\infty}^{\infty} (\xi(d\lambda)/2\pi i\lambda+a)] + \int_{-\infty}^{\infty} (e^{2\pi i t\lambda}/2\pi i\lambda+a)\xi(d\lambda)$$

$$= K(t,a,X_0,F) + Y(t)$$

where K and Y are the respective summands from above. We show now that the random function Y(t) is a stationary process.

$$E(Y(s)\overline{Y(t)}) = E(\int(e^{2\pi i s\lambda}/2\pi i\lambda+a)\xi(d\lambda) \cdot$$

$$\cdot \overline{\int(e^{2\pi i t\mu}/2\pi i\mu+a)\xi(d\mu))} =$$

$$= E(\int\int(e^{2\pi i(s\lambda-t\mu)}/(2\pi i\lambda+a)(\overline{2\pi i\mu+a}))\xi(d\lambda)\overline{\xi(d\mu)})$$

$$= \int\int E(1/(2\pi i\lambda+a)(\overline{2\pi i\mu+a}))e^{2\pi i(s\lambda-t\mu)}E(\xi(d\lambda)\overline{\xi(d\mu)})$$

$$= \int E(1/|2\pi i\lambda+a|^2)e^{2\pi i\lambda(s-t)}G(d\lambda)$$

where G is the absolute spectral measure of F. As we see $E(Y(s)\overline{Y(t)})$ is a function of t-s only, hence Y is stationary and its correlation function is

$$f(\tau) = \int E(1/|2\pi i\lambda+a|^2)e^{-2\pi i\lambda\tau}G(d\lambda).$$

If $F(t)$ is a real process and a is a real valued r.v., the above equations reduce to the representation $X(t) = K(t) + Y(t)$ where

$$Y(t) = \int_0^\infty (a \cos 2\pi\lambda t + 2\pi\lambda \sin 2\pi\lambda t / a^2 + 4\pi^2\lambda^2) u(d\lambda) +$$

$$+ \int_0^\infty (a \sin 2\pi\lambda t - 2\pi\lambda \cos 2\pi\lambda t / a^2 + 4\pi^2\lambda^2) v(d\lambda)$$

where $F(t) = \int_0^\infty \cos 2\pi\lambda t \, u(d\lambda) + \int_0^\infty \sin 2\pi\lambda t \, v(d\lambda)$ is the spectral representation of F.[1] Also then $E(Y(s)Y(t)) = \int_0^\infty E(1/4\pi^2\lambda^2 + a^2) \cos 2\pi\lambda(s-t) G(d\lambda)$ and $Y(t)$ is again stationary with correlation function $f(\tau) = \int_0^\infty E(1/4\pi^2\lambda^2 + a^2) \cos 2\pi\lambda\tau G(d\lambda)$.

We get

$$K(t) = e^{-ta}[X_0 - a \int (u(d\lambda)/a^2 + 4\pi^2\lambda^2) + \int (2\pi\lambda \, v(d\lambda)/ a^2 + 4\pi^2\lambda^2)$$

and so if $a(\omega) \geq \varkappa > 0$ for almost all ω, we see that $\lim_{t\to\infty} K(t,\omega) = 0$ a.e. and then $X(t)$ is asymptotically stationary (independent of the initial conditions). See also Adomian's discussion of a similar problem in his dissertation.

We now need to investigate carefully the various integrations involved in the above calculations. In particular, the problem may be phrased as follows: let I and J be intervals, $f: R^2 \times \Omega \longrightarrow C$ measurable on the product space $R \times R \times \Omega$, ξ a random orthogonal measure; 1) can we define

---

[1] See Doob [4] also.

the integral $\int_J f(x,y,\omega)\xi(dx)$ in a consistent way?, 2) if so, when can we say that the interated integrals $\int_I\int_J f(x,y,\omega)\xi(dx)dy$ and $\int_J\int_I f(x,y,\omega)dy\xi(dx)$ are equal? Let us answer the first question. Let $g: R \times \Omega \longrightarrow C$ be a function with the property that there exists a mutually disjoint sequence of bounded Borel sets $\{A_k\}_{k=1}^n$ and a sequence $\{g_k\}_{k=1}^n$, $g_k \in L^2(\Omega)$, such that $g(x,\omega) = \sum_{k=1}^n C(A_k,x)g_k(\omega)$ where $C(A_k,x)$ is the characteristic function of the set A_k. Moreover we will require that the families $\{g_k\}$ and $\{\xi(A_k)\}$ be independent. Then define $\int g(x,\omega)\xi(dx) = \sum_{k=1}^n g_k(\omega)\xi(A_k,\omega)$. We get:

$$E(|\int g(x,\omega)\xi(dx)|^2) = E(\sum_{j,k=1}^n g_j\overline{g_k}\ \xi(A_j)\ \overline{\xi(A_k)}) =$$

$$(3) \quad = \sum_{j,k=1}^n E(g_j\overline{g_k})E(\xi(A_j)\overline{\xi(A_k)}) = \sum_{k=1}^n E(|g_k|^2)M(A_k)$$

$$= \int E(|g(x)|^2)M(dx) = \int\|g(x)\|^2M(dx),$$

where M is the absolute spectral measure associated with ξ. Now consider the set $\mathcal{S}$ of all functions $g(x,\omega) = \sum C(A_k,x)g_k(\omega)$ where $\{g_k\}$ is independent of the family $\{\xi(A)\}_{A\in\mathcal{A}}$, $\mathcal{A}$ = all bounded Borel sets. $\mathcal{S}$ is clearly a linear space, and if $g \in \mathcal{S}$ $\int g(x)\xi(dx)$ is defined. Define $\|g\|_{\mathcal{S}} = (\int\|g(x)\|^2M(dx))^{1/2}$, then $\|\cdot\|_{\mathcal{S}}$ is a norm on $\mathcal{S}$. Complete $\mathcal{S}$ with respect to this

norm and denote the completed space S. If $g \in S$ and $\|g-g_n\|_{\mathcal{S}} \longrightarrow 0$, define $\int g(x)\xi(dx)$ as the limit in the mean of the random variables $\int g_n(x)\xi(dx)$. Note that equation (3) guarantees that this limit exists. Just as in the case of ordinary stochastic integrals with respect to orthogonal random measures, we have that g(x) is integrable whenever $\int\|g(x)\|^2 M(dx) < \infty$ and g is the limit of elements of $\mathcal{S}$. Note that if $g, h \in S$, then $E(\int g(x)\xi(dx) \overline{\int h(y)\xi(dy)}) =$

$= \int \mathrm{cov}(g(x), h(x))M(dx)$ since the corresponding relation holds for elements of $\mathcal{S}$. Note also that if g(x) is continuous in mean square and independent of $\{\xi(A)\}_{A\in\mathcal{B}}$, then g is integrable if and only if $\int\|g(x)\|^2 M(dx) < \infty$.

Consider now the question of interchanging the order of iterated integrals. Suppose we have a function $f: R^2 \times \Omega \longrightarrow C$ which is measurable, $f(x,y,\cdot) \in L^2(\Omega)$ for each pair (x,y), and the natural map from R^2 into $L^2(\Omega)$ induced by f is continuous. Suppose also that the family $\{f(x,y)\}_{(x,y)\in R^2}$ of r.v.'s is independent of $\{\xi(A)\}_{A\in\mathcal{A}}$. Let I and J be intervals. Then $g(x) = \int_I f(x,y)\xi(dy)$ exists if and only if $\int_I \|f(x,y)\|^2 M(dy) < \infty$. We want to integrate g(x) over the interval J, and $\|g(x)\| = (\int_I E(|f(x,y)|^2)M(dy))^{1/2}$ so let us require that f satisfy the condition

$$(4) \qquad \int_J (\int_I E(|f(x,y)|^2)M(dy))^{1/2} dx < \infty.$$

Then the integral $\int_J\int_I f(x,y)\xi(dy)dx$ exists. Now we want to

insure that $\int_I\int_J f(x,y)dx\xi(dy)$ exists, so we need to know that $\int_J (E(|f(x,y)|^2))^{1/2}dx < \infty$. Note that independence is preserved under this integration, so all we need to check is that $\int_I E(|h(y)|^2)M(dy) < \infty$ where $h(y) = \int_J f(x,y)dx$. But $E(|h(y)|^2) = E\int_J\int_J f(s,y)\overline{f(t,y)}ds\ dt = \int_J\int_J E(f(s,y)\overline{f(t,y)})ds\ dt$, so we require that f satisfy

$$(5) \qquad \int_I\int_J\int_J E(f(s,y)\overline{f(t,y)})ds\ dt\ M(dy) < \infty.$$

Now if f satisfies (4) and (5), both the iterated integrals exist. Consider now a subclass of function integrable with respect to ξ. We say that $g \in U$ if $g \in \mathcal{J}$ and there is a subset $H = H(g) \subset L^2(\Omega)$ whose finite linear combinations are dense in $L^2(\Omega)$ (call such a set <u>linearly dense)</u> with the property that $E(g(x)\xi(A)\overline{h}) = E(g(x))E(\xi(A)\overline{h})$ for $x \in A$, $A \in \mathcal{A}$, $h \in H(g)$. We sometimes write this relation as $E(g(x)\xi(dx)\overline{h}) = E(g(x))E(\xi(dx)\overline{h})$. If g is a simple function, then

$$E(\int g(x)\xi(dx)\overline{h}) = E(\Sigma\ g_k\xi(A_k)\overline{h}) = \Sigma\ E(g_k)E(\xi(A_k)\overline{h}) =$$

$$= \int E(g(x))M_h(dx)$$

for $h \in H(g)$ where M_h is the measure defined by $M_h(A) = E(\xi(A)\overline{h})$. Hence if $g \in U$, we have $E(\int g(x)\xi(dx)\overline{h}) = \int E(g(x))M_h(dx)$ for $h \in H(g)$.

Return now to the consideration of our function f. We require now that the range of the function $g: R \longrightarrow \mathcal{J}$ defined

by $[g(x)](y,\omega) = f(x,y,\omega)$ be contained in U. Then there is a linearly dense subset $H \subset L^2(\Omega)$ such that if $h \in H$ then

$$E\left(\int_I\int_J f(x,y)dx\,\xi(dy)\overline{h}\right) = \int_I \left(E\left(\int_J f(x,y)dx\right)\right)M_h(dy) =$$

$$= \int_I\int_J E(f(x,y))dx\,M_h(dy) = \int_J\int_I E(f(x,y))M_h(dy)dx =$$

$$= \int_J E\left(\int_I f(x,y)\xi(dy)\overline{h}\right)dx = E\left(\int_J\int_I f(x,y)\xi(dy)dx\,\overline{h}\right).$$

Since H is linearly dense and the above relation holds for all $h \in H$, we conclude that

$$(6) \qquad \int_I\int_J f(x,y)\xi(dy)dx = \int_J\int_I f(x,y)dx\,\xi(dy)$$

under these conditions.

To summarize these results, we have the following theorem:

<u>Theorem 10. Suppose</u> $f: R \times R \times \Omega \longrightarrow C$ <u>is measurable and independent from the orthogonal random measure</u> ξ. <u>Let</u> M <u>be the absolute measure associated with</u> ξ. <u>Let</u> $f(x,y,\cdot)$ <u>be square integrable and continuous when considered as a mapping from</u> $R \times R$ <u>into</u> $L^2(\Omega)$. <u>Suppose there is a linearly dense subset</u> H <u>of</u> $L^2(\Omega)$ <u>such that</u> $E(f(x,y)\xi(dy)\overline{h}) = E(f(x,y))E(\xi(dy)\overline{h})$ <u>for</u> $h \in H$. <u>If for the intervals</u> I <u>and</u> J <u>we have</u>

$$(4) \qquad \int_J \left(\int_I E(|f(x,y)|^2)M(dy)\right)^{1/2}dx < \infty$$

(5) <u>and</u> $\iiint_{IJJ} E(f(x,y)\overline{f(t,y)})ds\ dt\ M(dy) < \infty$

<u>then</u> $\iint_{IJ} f(x,y)\xi(dy)dx = \iint_{JI} f(x,y)dx\ \xi(dy)$.

In our calculations involving the differential equation (1), the function f is defined by

$$f(x,y,\omega) = e^{2\pi ixy}\ e^{ax} = e^{2\pi ixy}\ e^{a(\omega)x}$$

Suppose that $K = \text{ess}_{\omega}\sup |a(\omega)|$ and a is real valued. Then relation (4) becomes

$$\int_0^t (\int_{-\infty}^{\infty} E(|f(x,y)|^2)M(dy))^{1/2}dx = \int_0^t (\int_{-\infty}^{\infty} E(e^{2ax})M(dy))^{1/2}\ dx =$$

$$(\int_{-\infty}^{\infty} M(dy))^{1/2}\int_0^t E(e^{2ax})dx.$$

But $(\int_{-\infty}^{\infty} M(dy))^{1/2} = \|F(0)\| < \infty$, and $E(e^{2ax}) \leq e^{2Kx}$ so $\int_0^t E(e^{2ax})dx < \infty$ for t finite. Hence (4) holds. Consider now relation (5).

$$|E(f(x,y)\overline{f(t,y)})| \leq E(|e^{2\pi isy}e^{as}e^{-2\pi ity}e^{at}|) \leq$$

$$\leq e^{K(s+t)}$$

hence $|\int_{-\infty}^{\infty}\int_0^t\int_0^t E(f(s,y)\overline{f(t,y)})ds\ dt\ M(dy)| \leq$

$\leq \|F(0)\|^2 \int_0^t\int_0^t e^{K(u+v)}dudv < \infty$ for $t < \infty$. Since a is independent of F, f is independent of ξ (being a Borel measurable function). Hence the only additional requirement we place on a and F is that there exist a linearly dense set H for which

$$E(e^{ya}\xi(dy)\overline{h}) = E(e^{ya})E(\xi(dy)\overline{h}),$$

for then we have

$$E(f(x,y)\xi(dy)\overline{h}) = E(e^{2\pi ixy}e^{ya}\xi(dy)\overline{h}) =$$

$$= e^{2\pi ixy}E(e^{ya}\xi(dy)\overline{h}) = e^{2\pi ixy}E(e^{ya})E(\xi(dy)\overline{h}) =$$

$$= E(e^{2\pi ixy}e^{ya})E(\xi(dy)\overline{h}) = E(f(x,y))E(\xi(dy)\overline{h}).$$

Note that if a is not random, then all these conditions are trivially satisfied.

So now we can decompose the solution $X(t) = K(t) + Y(t)$ into the sum of a stationary $Y(t)$ and (in general) non-stationary $K(t)$. Thus if we set $K(t) = 0$ and solve for X_0, we obtain a sufficient condition for stationarity. In particular, if

$$X_0 = \int (1/2\pi i\lambda + a)\xi(d\lambda) \tag{7}$$

then $X(t)$ is stationary. Note that a new difficulty arises in this expression. Namely, if $a = 0$ (in general, if $\mu\{\omega : a(\omega) = 0\} > 0$) and 0 is in the point spectrum of F, then (7) is not defined. This does not contradict our exchange of integrations however, but it does say that we cannot split up the integral in equation (2).

The condition in (7) may actually be a necessary condition for stationarity. For example, consider the equation $X'(t) + X(t) = f$, $f \in L^2(\mathcal{A})$. $F(t) = f$, so $\xi(S) = f$ if $0 \in S$ and $\xi(S) = 0$ otherwise. Hence the condition is

$X_0 = \int(1/2\pi i\lambda+1)\xi(d\lambda) = f$ and the resulting solution is $X(t) = \int(e^{2\pi i t\lambda}/2\pi i\lambda+1)\xi(d\lambda) = f$. Note that $F(t)$ is analytic, so solving by power series we get $X(t) = f + (X_0-f)e^{-t}$, i.e., $X_n = (-1)^n(X_0 - f)$, $n \geq 1$. Calculating the covariance function of this process, we see that a sufficient condition that $X(t)$ be stationary is that $\langle f^2\rangle = \langle fX_0\rangle = \langle X_0^2\rangle$. We show now that this is necessary. For if $X(t)$ is stationary, $\langle X_0X_2\rangle = -\langle X_1^2\rangle$. But $\langle X_0X_2\rangle = \langle X_0^2\rangle - \langle fX_0\rangle$ and $\langle X_1^2\rangle = \langle X_0^2\rangle - 2\langle fX_0\rangle + \langle f^2\rangle$. Also $0 = \langle X_0X_1\rangle = \langle X_0^2\rangle - \langle fX_0\rangle$, i.e., $\langle X_0^2\rangle = \langle fX_0\rangle$ and so we have

$$0 = \langle X_0X_2\rangle + \langle X_1^2\rangle = 2\langle X_0^2\rangle - 3\langle fX_0\rangle + \langle f^2\rangle =$$
$$= 2\langle X_0^2\rangle - 3\langle X_0^2\rangle + \langle f^2\rangle, \text{ i.e., } \langle f^2\rangle = \langle X_0^2\rangle.$$

Hence $\langle X_0^2\rangle = \langle fX_0\rangle = \langle f^2\rangle$ is a necessary and sufficient condition that $X(t)$ be stationary. But then $\langle(X_0-f)^2\rangle = \langle X_0^2\rangle - 2\langle fX_0\rangle + \langle f^2\rangle = 0$ and so $f = X_0$. Thus the condition $X_0 = \int(1/2\pi i\lambda+1)\xi(d\lambda)$ is actually necessary. (Note that we have incidentally proved that if a and b are r.v.'s, then $a + be^{-t}$ is stationary if and only if $b = 0$).

Of course the condition on X_0 expressed by (7) is not always necessary for stationarity. Consider the following interesting example. Let $X'(t) = F(t)$ and suppose that $F(t)$ is real valued and analytic and 0 is not in the spectrum of $F(t)$. Then the sufficient condition we get on X_0 is

$$X_0 = -\int_0^\infty (1/2\pi\lambda)v(d\lambda)$$

where $F(t) = \int_0^\infty \cos 2\pi t\lambda\, u(d\lambda) + \int_0^\infty \sin 2\pi t\lambda\, v(d\lambda)$. Note that $E(X_0F(t)) = -\int_0^\infty (\sin 2\pi t\lambda/2\pi\lambda)M(d\lambda)$ in this case.

If we calculate the coefficients in the power series expansion for $F(t)$, we get

$$F_{2n} = (-1)^n\int_0^\infty (2\pi\lambda)^{2n}u(d\lambda)$$

$$F_{2n+1} = (-1)^n\int_0^\infty (2\pi\lambda)^{2n+1}v(d\lambda).$$

Hence in general we have $\langle F_n^2\rangle = \int_0^\infty (2\pi\lambda)^{2n}M(d\lambda)$. Thus applying the result of example 1, we have that a necessary and sufficient condition for $X(t)$ to be stationary is that

$$\langle F(t)X_0\rangle = -\Sigma((-1)^n t^{2n+1}/(2n+1)!)\ \langle F_n^2\rangle =$$

$$= -\Sigma((-1)^n t^{2n+1}/(2n+1)!)\int(2\pi\lambda)^{2n}M(d\lambda)$$

$$= -\Sigma\int(-1)^n((2\pi t\lambda)^{2n+1}/(2n+1)!2\pi\lambda)M(d\lambda)$$

$$= -\int(\sin 2\pi t\lambda/2\pi\lambda)M(d\lambda).$$

Thus we see that if X_0 is any initial condition for which $X(t)$ is stationary, then the projection of the r.v. X_0 onto the smallest closed linear manifold in $L^2(\Omega)$ containing the process $F(t)$ (call it $\mathcal{J}$) is the r.v. $-\int_0^\infty (v(d\lambda)/2\pi\lambda)$.

Hence in the case of analyticity, a necessary and sufficient condition that $X(t)$ be stationary is that

$$X_0 = Y - \int_0^\infty (v(d\lambda)/2\pi\lambda)$$

where Y is any r.v. orthogonal to $\mathcal{F}$.

In this section we have restricted ourselves to the case of an equation with a constant (r.v.) coefficient. This restriction was made purely to facilitate the various computations made in the interchange of order of integrations. Examining the more general case (time-varying coefficients) the author has been unable to extract a likely candidate for the stationary part of the general solution. Nevertheless, the idea of using random harmonic analysis especially in conjunction with the analytic method of section 2 appears valid, and progress in this direction seems likely in the future.

SECTION 4

In section 2 we presented a general method of determining when stochastic differential equations with random (analytic) coefficient processes, random (analytic) forcing function, and random initial conditions have stationary solutions. In section 3 we presented a method of applying random harmonic analysis to a simpler first order equation and we also presented an extension of integration techniques allowing us to use these more powerful tools. In this section we wish to relate our work to that done by others and to also make a few remarks concerning the abstract notion of stochastic transformations. We also indicate some future work.

First let us relate our results to Adomian's [1] results concerning stochastic Green's functions transforming a given statistical measure of an input process to the corresponding measure of the output process. Suppose we have a stationary r.f. F(t) with spectral representation

$$F(t) = \int e^{2\pi i t\lambda} \xi(d\lambda).$$

We have defined the correlation function

$$f(t) = \langle F(s)\overline{F(s+t)} \rangle \int e^{-2\pi i t\lambda} M(d\lambda)$$

where M is the absolute measure associated with ξ. Now if

M is absolutely continuous with respect to Lebesgue measure on the line, we call the Radon-Nikodym derivative

$$m(x) = dM/d\lambda$$

the <u>spectral density</u> of the process F, <u>i.e.</u>, we have

$$f(t) = \int e^{-2\pi i t\lambda} m(\lambda)d\lambda$$

and
$$m(\lambda) = \int e^{2\pi i t\lambda} f(t)dt$$

Now in the equation discussed in section 3

$$X'(t) + aX(t) = F(t), \quad X(0) = X_0$$

with the r.v. coefficient a, we call F(t) the input process and X(t) the output process. Moreover, applying the results of section 2 we know that we can have a stationary F(t) resulting in a stationary X(t). Now let

$$F(t) = \int e^{2\pi i t\lambda} \xi(d\lambda)$$

and
$$X(t) = \int e^{2\pi i t\lambda} \nu(d\lambda)$$

be the spectral representations of F and X respectively. Suppose also that F has the correlation function f and spectral density function m. We calculated in section 2 the correlation function x for X given by

$$x(t) = \int E(1/|2\pi i\lambda+a|^2)e^{-2\pi i t\lambda} m(\lambda)d\lambda.$$

Hence we immediately recognize that the spectral density of X must be given by

$$n(\lambda) = E(1/|2\pi i\lambda+a|^2)m(\lambda).$$

Now the Green's function transforming the spectral density of an input into the spectral density of the output is a (perhaps generalized) function $H(t,u)$ such that

$$n(t) = \int H(t,u)m(u)du$$

whereupon we see that H is given by

$$(1) \qquad H(t,u) = \delta(t-u)E(1/|2\pi iu+a|^2)$$

Note that this form of the Green's function is the same as Adomian's [1] if the r.v. a is a constant, for then we get

$$H(t,u) = \delta(t-u)|Y(u)|^2$$

where $\quad Y(u) = 1/(2\pi iu+a)$.

Similarly, the stochastic Green's function $G(t,u)$ transforming the correlation function f into x by the relation

$$x(t) = \int G(t,u)f(u)du$$

can be expressed in terms of H by the relation

$$G(t,u) = \iint e^{2\pi i(\tau u-\sigma t)}H(\sigma,\tau)d\sigma d\tau$$

and so applying (1) we get

$$G(t,u) = \int e^{2\pi it(u-t)}E(1/|2\pi i\tau+a|^2)d\tau.$$

In both cases we see that these Green's functions are expressed in terms of the statistics of the coefficient in the stochastic differential operator.

Let us make a few remarks concerning the idea of a stochastic transformation (s.t.), a concept due to Adomian [1]. If $X(t,\omega)$, $t \in T$, $\omega \in \Omega$, is a s.p., a stochastic trans-

formation or stochastic operator T on X carries X into another process $Z = T[X]$, and T in general depends on t and $\omega' \in \Omega'$ (where Ω' may not be identical to Ω). Thus T includes all deterministic transformations as a special case and we usually deal with integral operators, differential operators, partial differential operators, etc.

We wish now to indicate a general framework in which these ideas can be precisely expressed. In general there is a natural desire to distinguish between an operation of the form

$$(2) \qquad X(t,\omega) = \int H(t,u)Y(u,\omega)du$$

where the kernel $H(t,u)$ is a complex valued function and X and Y are r.f.'s and an operation of the form

$$(3) \qquad X(t,\omega) = \int H(t,u,\omega)Y(u,\omega)du$$

where the kernel H also depends on the stochastic variable ω. The transformation expressed by (2) in which the process Y is mapped into the process X is often called a deterministic transformation whereas (3) expresses an operation which includes (2) and conforms more closely with our intuitive notion of a stochastic transformation. Indeed (3) is the general form of a stochastic integral operator. Similarly an equation of the form

$$(4) \qquad X(t,\omega) = aY(t,\omega) + bZ(t,\omega)$$

where a and b are complex constants and X, Y, and Z are r.f.'s is rightfully considered a deterministic mapping of the

pair (Y,Z) into the process X. On the other hand, an equation

$$X(t,\omega) = A(t,\omega)Y(t,\omega) + B(t,\omega)Z(t,\omega) \tag{5}$$

mapping the pair (Y,Z) into X is a "truly stochastic" operation. Also we need to consider maps of the form $Y \longrightarrow X$ defined by

$$X(t,\omega) = Y(f(t),\omega) \tag{6}$$

where $f:T \longrightarrow T$ indicates a re-parameterization of the time variable. One wants to think of (6) as expressing a deterministic relationship whereas

$$X(t,\omega) = Y(f(t),\varphi(\omega)) \tag{7}$$

where $f:T \longrightarrow T$ and $\varphi:\Omega \longrightarrow \Omega$ would again be "truly stochastic". Bharucha-Reid's [2] "random transformation" refers to a map

$$T:\Omega \times R \longrightarrow R \tag{8}$$

with the property that the function $T(\cdot,x)$ is a r.v. for each $x \in R$ and such a random transformation may induce a stochastic transformation defined by

$$X(t,\omega) = T(\omega,Y(t,\omega)) \tag{9}$$

carrying Y into X. One would call such a transformation deterministic if the function T did not depend on the first coordinate of its argument. The examples of trans-

formations in the introduction to this thesis indicate more different forms that a s.t. can take.

We propose to examine a function space approach of representing a r.f. as a measure on a fixed collection of functions; then a change of measures on this function space will be identified as a stochastic transformation. In particular, let $X(t)$ be a real-valued s.p. with parameter set $T = (a,b)$. Following the construction outlined in Skorokhod [10], we let Φ be the space of all functions $x: (a,b) \longrightarrow R$. If A is a Borel set in R and $t_0 \in (a,b)$, we let $C_{t_0}(A) = \{x \in \Phi: x(t_0) \in A\}$. A set which is the intersection of a finite number of sets of the form $C_{t_0}(A)$ is a <u>cylindrical set</u>. We let F be the minimal σ-algebra of subsets of Φ generated by all cylindrical sets. Now the measure μ determined on F by the relations

$$(10) \qquad \mu\left(\bigcap_{n=1}^{k} C_{t_n}(A_n)\right) = P\{X(t_i, \omega) \in A_i,\ i = 1,\ldots,k\}$$

for all k, $t_1,\ldots,t_k$ in (a,b) and all Borel sets $A_1,\ldots,A_k$ is called the measure <u>in the function space corresponding to the process $X(t)$</u>. Kolmogorov's theorem guarantees us that (10) defines a unique measure on F.

Conversely, suppose that we have a measure μ defined on F such that $\mu(\Phi) = 1$. Then we have a probability space $(\Omega, \mathcal{J}, P)$ where $\Omega = \Phi$, $\mathcal{J} = F$, and $P = \mu$ and so we can define a process $X(t)$, $t \in (a,b)$ by the relation

$$X(t,\omega) = X(t,x) = x(t)$$

We remark that $X(t,\cdot)$ is measurable, for let A be a Borel set. Then $\{\omega: X(t,\omega) \in A\} = \{x: x(t) \in A\}$ and this is a cylindrical set and so $X(t,\cdot)$ is measurable, i.e., $X(t)$ is a s.p. Note also that the function space measure corresponding to X is just μ.

Hence every process on (a,b) is associated with a measure on Φ and conversely. Now let M be the set of all measures μ on Φ such that $\mu(\Phi) = 1$. Then any function $f: D \longrightarrow M$ where $D \subset M$ could be called a stochastic transformation. More generally, a function $f: D \longrightarrow M$ where $D \subset M^n$ can be a stochastic transformation.

The above interpretation of a stochastic process as a measure on an appropriate function space does not in itself obviate any computational difficulties associated with the analysis of s.p.'s. However, we can now use the full power of general measure theory to gain new insight. For instance, the author is currently attempting to prove theorems answering the following questions: Let F be a stochastic transformation and let $X(t)$ be a stationary (in some sense) process. Suppose that $Y = F(x)$ is stationary (in the same sense). Let μ and ν be the measures associated with X and Y respectively. Let M_x be the manifold in $L^2(\Omega)$ generated by X. Let $Z = Z(t)$ be the process obtained by projecting Y onto M_x and let η be the measure associated with Z. Is $\nu \ll \mu$? Is $\eta \ll \mu$? If $\nu = \nu_0 + \nu_1$ is the Lebesgue decomposition of ν with respect to μ, does $\eta = \nu_1$? If not, is $\eta \ll \nu_1$ or $\nu_1 \ll \eta$? Theorems along

these lines would give information concerning stochastic transformations which carry stationary processes into stationary processes. Also we naturally ask what effect on a measure associated with a s.p. is induced by a stochastic differential operator on the process. Also using our interpretation of s.t.'s in this way, perhaps we can discover measure-theoretic properties of "deterministic" transformations (equations (2), (4), (6)) that distinguish them from "truly stochastic" transformations ((3), (5), (7)) and thus allow us to give a more useful and precise interpretation of these notions.

BIBLIOGRAPHY

1. Adomian, G., "Stochastic Green's functions," Proc. Sympos. Appl. Math. Vol. 16, Amer. Math. Soc., Providence, R. I., 1964.

2. Bharucha-Reid, A. T., "Theory of Random Equations," Proc. Sympos. Appl. Math. Vol. 16, Amer. Math. Soc., Providence, R. I., 1964.

3. Dieudonne, J., Foundations of Modern Analysis, Academic Press, 1962.

4. Doob, J. L., Stochastic Processes, Wiley, 1953.

5. Irzhina, M., "Ordinary Differential and Difference Equations with Random Coefficients and Random Right-hand Side," Selected Translations in Math. Stat. and Prob., Vol. 7, Amer. Math. Soc., 1968, pp. 139-155.

6. Lange, S., Analysis I, Addison-Wesley Publishing Company, 1968.

7. Loeve, M., Probability Theory, D. Van Nostrand, 1963.

8. Rozanov, Yu. A., Stationary Random Processes, Holden-Day, 1963.

9. Rudin, W., Principles of Math Analysis, McGraw-Hill Book Co., 1964.

10. Skorokhod, A. V., Studies in the Theory of Random Processes, Addison-Wesley, 1965.

11. Yaglom, A. M., Theory of Stationary Random Processes, Prentice-Hall, 1962.