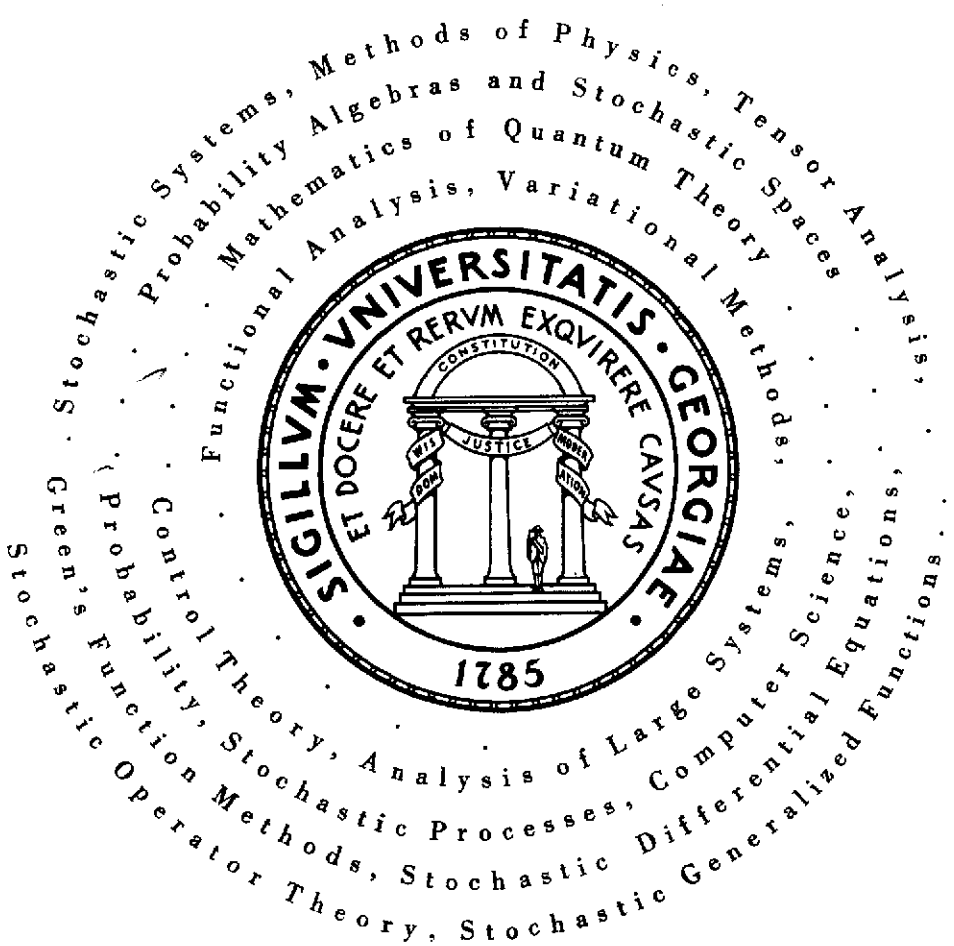
Department of Mathematics
UNIVERSITY OF GEORGIA

Final Report

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

1971-72

G. Adomian



E UNIVERSITY OF GEORGIA
DEPARTMENT OF MATHEMATICS
GRADUATE STUDIES BUILDING
ATHENS, GEORGIA 30602

January 5, 1973

Office of Grants & Research Contracts
National Aeronautics & Space Administration
600 Independence Avenue
Washington, D. C.

(ATTN: CODE SC)

Dear Sir:

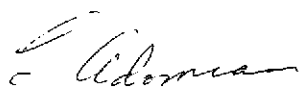
Attached is a final report on work accomplished including reprints of journal publications and some preprints. As a result of two years of N.A.S.A. support, I have 10 published journal papers, 6 published abstracts, 1 published technical report (and doctoral dissertation), 2 journal papers which were submitted but not yet published at the expiration of the grant, and several journal papers and one dissertation almost completed at the expiration of the grant - all of which carry (or will carry) acknowledgement of N.A.S.A. research support. (Listing follows.)

We have developed a useful analytic method and a very versatile and powerful new computer technique for solving stochastic differential equations. We recently computed some 14,000 terms for a series representing the solution process of such an equation for our methods and other existing methods of solution to compare results and errors involved - which show unequivocally that we have achieved extremely important results because of your support. I regret this work was not yet in a form to send to you with this letter and final report although an earlier preliminary discussion was included. The completed work will become a dissertation in June, 1973, and will then be published with appropriate acknowledgement to N.A.S.A.

I, of course, regret the termination of N.A.S.A. support since we clearly have results which will be of value to N.A.S.A. and to other agencies as well. Our work will tie in closely with the work of M.M.R. Williams on kinetics of nuclear reactors, with the work of Klimontovich on plasmas, to stochastic control, weather prediction, and several other areas. In view of results achieved and under way, I hope N.A.S.A. will again be involved with my group in the future. You will also be pleased to know the U.S.S.R. Academy of Sciences new Merguelian Institute of Applied Mathematics in Erevan, U.S.S.R., has asked to send postdoctorate researchers to work with us on the new methods and several professors wish to come to Georgia to work with us.

I would like to express my appreciation to N.A.S.A. for its essential support of this research and for its accomplishments for the U.S. in space. With best wishes for your continued success.

Sincerely,



George Adomian
Professor of Mathematics

Copies to:

Office of University Affairs (Code Y)

Mr. J. Albus (Code RET), Data Systems Branch

Mr. Joseph T. Davis, Grants Officer

Dr. Samuel A. Rosenfeld, former Project Officer for Grant (Please Forward)

Dr. Raymond Wilson, former head of Applied Mathematics (Please Forward)

(N.A.S.A. GRANT NGR 11-003-020)

PUBLICATIONS IN JOURNALS

1. The Closure Approximation in the Hierarchy Equations, G. Adomian, (A72-15257)
Journal of Statistical Physics, 3, No. 2, 127-133, Summer 1971.

2. Random Operator Equations in Mathematical Physics I, G. Adomian, (A70-26550)
Journal of Mathematical Physics, 11, No. 3, 1069-1084, March, 1970.

3. Erratum: Linear Random Operator Equations in Mathematical Physics I,
G. Adomian, Journal of Mathematical Physics, 12, No. 9, p. 2031,
September, 1971.

4. Linear Random Operator Equations in Mathematical Physics II, G. Adomian,
Journal of Mathematical Physics, 12, No. 9, 1944-1948, September, 1971.
(A71-43114)

5. Linear Random Operator Equations in Mathematical Physics III,
G. Adomian, Journal of Mathematical Physics, 12, No. 9, 1948-1953, (A71-43115)
September, 1971.

6. Addendum: Linear Random Operator Equations in Mathematical Physics III,
G. Adomian, Journal of Mathematical Physics, 13, No. 2, p. 272,
February, 1972.

7. Signal Processing in a Randomly Time Varying System, G. Adomian, (A73-23800)
Bulletin of the Academy of Sciences of the Armenian S.S.R., U.S.S.R.
(Also presented at International Symposium on Information Theory,
U.S.S.R., September 2-8, 1971), Vol. VII, No. 1, 14-22, January, 1972.

8. Local Theorems for Non-Identically Distributed Random Variables, (A72-25356)
J. D. Mason, SIAM Journal on Applied Mathematics, 22, No. 2, 259-265,
March, 1972.

9. Local Theorems in Strengthened Form for Lattice Random Variables,(A72-15256)
J. D. Mason, Annals of Mathematical Statistics, 42, 784-788, and 43,
No. 2, p. 696, 1972.

10. Comments on the Continuity of Distribution Functions Obtained by (A72-15260)
Superposition, B. W. Huff, Proc. of the Amer. Mathematical Society,
27, No. 1, 141-146, January, 1971.

(N.A.S.A. GRANT NGR 11-003-020)

PUBLISHED ABSTRACTS

1. Stationarity and Stochastic Differential Equations, G. Adomian and W. W. Walker, Notices of the Amer. Math. Soc., Annual Meeting, San Antonio, Texas, January, 1970.

2. A Theory of Stochastic Systems With Applications to Physics, G. Adomian, Bulletin of the Amer. Physical Society, Series II, Vol. 16, No. 4, Washington, D. C., April 26-29, 1971.

3. On the Theory of Stochastic Systems, G. Adomian, Proc. of the Fourth International Conference on Systems Sciences, January, 1971.

4. On a Theory of Stochastic Control, G. Adomian, Proc. of the Fifth International Conference on Systems Sciences, January, 1972.

5. Stochastic Greens Functions, G. Adomian, Notices of the Amer. Math. Society, Annual Meeting, Las Vegas, January 17-21, 1972.

6. Propagation in Random Media, G. Adomian and L. H. Sibul, Proc. of International G-Ap and U.R.S.I. Conference, UCLA, September, 1971.

(N.A.S.A. GRANT NGR 11-003-020)

TECHNICAL REPORTS

1. N.A.S.A. Technical Report, Stationarity Conditions for Stochastic Differential Equations, G. Adomian and W. W. Walker, 1970.

(N.A.S.A. GRANT NGR 11-003-020)

WORK SUBMITTED FOR PUBLICATION

1. Boundedness of Absolute Values of Generalized Fourier Coefficients, Charles C. Miao and G. Adomian, to be published.

2. A Decomposition for Some Operators, Bernard B. Morrel, to be published.

(N.A.S.A. GRANT NGR 11-003-020)

INCOMPLETE WORK AT EXPIRATION OF GRANT

1. Research on Numerical Methods for Stochastic Differential Equations, G. Adomian and M. Elrod (Note: This will be a Ph.D. dissertation and is expected to be a series of journal papers in view of the significance and magnitude of results.)

2. State Space Methods for Stochastic (Control) Systems, G. Adomian and L. H. Sibul (work still in progress but significant results achieved).

3. Coherence Theory for Randomly Time and Space Varying Media, not yet completed.

BOUNDEDNESS OF ABSOLUTE VALUES OF GENERALIZED FOURIER COEFFICIENTS

by

Charles C. Miao and G. Adomian
University of Georgia
Athens, Georgia 30601

ABSTRACT

For $f(x) \in C[a,b]$ and $\{\varphi_k(x)\}$ an orthogonal set with weight function $w(x)$ on $[a,b]$, an upper bound for $\sum_{k=0}^{n} |(\varphi_k, \varphi_k)^{\frac{1}{2}} a_k|$ where a_k are generalized Fourier coefficients is given for odd n by $\sqrt{n+1}\,(\int_a^b w(x)dx)^{\frac{1}{2}}\|f\|$ where $\|f\|$ is the sup norm of f on $[a,b]$.

For even n, the bound is $[(n+2)(\int_a^b w(x)dx\|f\|^2 - \max_k (\frac{n}{2}+1)\,|\varphi_k, \varphi_k)^{\frac{1}{2}} a_k|^2]^{\frac{1}{2}}$ for $0 \le k \le n$, which is a little cumbersome so the result for odd n would be more useful.

For the trigonometric case $f \in C[-\pi, \pi]$ and $w(x) = 1$, this result gives $\sum_{k=0}^{n} |a_k| + |b_k| \le 2\sqrt{n+1}\,\|f\|$ for any n.

Given $f(x) \in C[a,b]$ and letting $\{\varphi_n\}$ be an orthogonal set of continuous functions with weight function $w(x)$ on $[a,b]$, consider the generalized Fourier series $\Sigma a_k \varphi_k(x)$ converging uniformly to $f(x)$.

The normalized Fourier coefficients are given by

$$a_k = (f,\varphi_k)/(\varphi_k,\varphi_k) = \int_a^b w(x)f(x)\varphi_k(x)dx / \int_a^b w(x)\varphi_k^2(x)dx .$$

By Bessel's inequality,

$$\sum_{k=0}^{n} (\varphi_k,\varphi_k)a_k^2 \le \int_a^b w(x)f^2(x)dx \le (\int_a^b w(x)dx)\|f\|^2$$

where $\|f\|$ is the sup norm or maximum of f in $[a,b]$. Denote $(\varphi_k,\varphi_k)^{\frac{1}{2}}a_k$ by A_k and write

$$\sum_{k=0}^{n} |A_k|^2 \le (\int_a^b w(x)dx)\|f\|^2 . \qquad (1)$$

Suppose n is odd. We separate the left side into pairs thus

$$\sum_{k=0}^{(n-1)/2} |A_{2k}|^2 + |A_{2k+1}|^2 \le (\int_a^b w(x)dx\|f\|^2 .$$

We let the sequence $\{|A'_{2k}| + |A'_{2k+1}|\}_{k=0}^{(n-1)/2}$ be a rearrangement of the sequence $\{|A_{2k}| + |A_{2k+1}|\}_{k=0}^{(n-1)/2}$ such that $|A'_{2k}| + |A'_{2k+1}| \ge |A'_{2(k+1)} + A'_{2(k+1)+1}|$ i.e., each pair is greater or equal to the next pair,

e.g., $A_0' + A_1' \geq A_2' + A_3'$, $A_2' + A_3' \geq A_4' + A_5'$, etc.

Now using Tchbychev's inequality, we write

$$\left(\sum_{k=0}^{(n-1)/2} |A_{2k}'| + |A_{2k+1}'|\right)^2 \leq \left(\frac{n-1}{2} + 1\right) \sum_{k=0}^{(n-1)/2} \left(|A_{2k}'| + |A_{2k+1}'|\right)^2$$

$$\leq 2\left(\frac{n-1}{2} + 1\right) \sum_{k=0}^{(n-1)/2} |A_{2k}'|^2 + |A_{2k+1}'|^2 .$$

Since sums are preserved

$$\left(\sum_{k=0}^{(n-1)/2} |A_{2k}| + |A_{2k+1}|\right)^2 \leq (n+1) \sum_{k=0}^{(n-1)/2} |A_{2k}|^2 + |A_{2k+1}|^2$$

or simply,

$$\left(\sum_{k=0}^{n} |A_k|\right)^2 \leq (n+1) \sum_{k=0}^{n} |A_k|^2 \qquad (2)$$

Combining (1) and (2)

$$\left(\sum_{k=0}^{n} |A_k|\right)^2 \leq (n+1)\left(\int_a^b w(x)dx\right)\|f\|^2$$

hence our result

$$\sum_{k=0}^{n} |(\varphi_k, \varphi_k)^{\frac{1}{2}} a_k| \leq \sqrt{n+1}\left(\int_a^b w(x)dx\right)^{\frac{1}{2}}\|f\| . \qquad (3)$$

If n is even we can let $A_{n+1} = 0$ to divide into pairs as before, or somewhat more generally since zero can be inserted anywhere in the sequence, let $0 \leq j_0 \leq n$ and let $\{A_s'\}_{s=0}^{n+1}$ be a rearrangement of

$\{0, A_o, A_1, \ldots, A_n\}$ such that

(i) $|A'_{2s}| + |A'_{2s+1}| \geq |A'_{2(s+1)}| + |A'_{2(s+1)+1}|$

(ii) The zero is paired with the A_{j_o} .

We now have pairs as before and again using Tchbychev's inequality, we have

$$\left(\sum_{s=0}^{n/2} |A'_{2s}| + |A'_{2s+1}|\right)^2 \leq$$

$$2\left(\frac{n}{2} + 1\right) \sum_{k=0}^{n/2} |A'_{2s}|^2 + |A'_{2s+1}|^2 - \left(\frac{n}{2} + 1\right) A_{j_o}^2 .$$

Since the sums above are preserved,

$$\left(\sum_{k=0}^{n} |A_k|\right)^2 \leq (n+2)\left(\sum_{k=0}^{n} |A_k|^2\right) - \left(\frac{n}{2} + 1\right) A_{j_o}^2 .$$

Combining (1) and (4) (for $0 \leq j_o \leq n$)

$$\left(\sum_{k=0}^{n} A_k\right)^2 \leq (n+2)\left(\int_a^b w(x)dx\right)\|f\|^2 - \left(\frac{n}{2} + 1\right) A_{j_o}^2 .$$

Finally (for even n)

$$\sum_{k=0}^{n} |(\varphi_k, \varphi_k)^{\frac{1}{2}} a_k| \leq \sqrt{(n+2)\left(\int_a^b w(x)dx\|f\|^2\right) - \left(\frac{n}{2} + 1\right) A_{j_o}^2} \qquad (5)$$

where $0 \leq j_o \leq n$.

(If $A_{j_o} = 0$ the right side becomes $\left[(n+2)\left(\int_a^b w(x)dx\right)\right]^{\frac{1}{2}}\|f\|$ which is not as tight a bound.) Thus the lowest bound

(not l.u.b.) is

$$\sum_{k=0}^{n} |(\varphi_k, \varphi_k)^{\frac{1}{2}} a_k| \leq$$

$$[(n+2)(\int_a^b w(x)dx \|f\|^2) - \max_k (\tfrac{n}{2}+1) |(\varphi_k, \varphi_k)^{\frac{1}{2}} a_k|^2]^{\frac{1}{2}} \quad (6)$$

where $0 \leq k \leq n$.

Example: Trigonometric series for $f(x) \in C[-\pi, \pi]$

$$f(x) = \frac{a_o}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx$$

$$w(x) = 1$$

$$(\varphi_k, \varphi_k) = \begin{cases} 2\pi & \text{if } k = 0 \\ \pi & \text{if } k \neq 0 \end{cases} .$$

By (1),

$$(2\pi)(\frac{a_o}{2})^2 + \sum_{k=1}^{n} a_k^2 + b_k^2 \leq 2\pi \|f\|^2$$

or

$$\sum_{k=0}^{n} a_k^2 + b_k^2 - \frac{a_o^2}{2} \leq 2\|f\|^2$$

a result also apparent from Parseval's formula

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = \frac{a_o^2}{2} + \sum_{k=1}^{\infty} a_k^2 + b_k^2$$

or

$$\sum_{k=0}^{n} a_k^2 + b_k^2 - \frac{a_o^2}{2} \leq \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = 2\|f\|^2 .$$

By (2), we have

$$\left(\sum_{k=0}^{n} |a_k| + |b_k|\right)^2 \le 2(n+1)\sum_{k=0}^{n} |a_k|^2 + |b_k|^2 - (n+1)a_o^2 \ .$$

Combining, we have

$$\sum_{k=0}^{n} |a_k| + |b_k| \le 2\sqrt{n+1}\ \|f\| \ .$$

Cheney[1] gives the bound $\pi\sqrt{2n-1}\ \|f\|$. It's easy to see $2\sqrt{n-1}\ \|f\| < \pi\sqrt{2n-1}\ \|f\|$ for all $n \ge 1$.

As an elementary example let $f(x) = \pi$ in $[-\pi, \pi]$. We get immediately $f(x) = 2(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \ldots)$. For the sum $\sum_{k=0}^{3} |a_k| + |b_k|$, the bound stated here is 4π , which is quite close to the actual sum, while Cheney's bound is $\sqrt{5}\,\pi^2$.

Acknowledgement: This work has been supported by the National Aeronautics and Space Administration (NGR 11-003-020).

REFERENCES

1. E.W. Cheney, Introduction to Approximation Theory, International Series in Pure and Applied Mathematics, McGraw-Hill, page 125.

2 D

A Decomposition for Some Operators

by

Bernard B. Morrel

Let H be a complex Hilbert space and let $\mathcal{B}(H)$ denote the algebra of all bounded linear operators on H. Then $T \in \mathcal{B}(H)$ is abnormal (sometimes, completely non-normal) if there is no non-trivial subspace $M \subseteq H$ which reduces T and such that the restriction of T to M is normal. Every $T \in \mathcal{B}(H)$ may be written uniquely as the direct sum of a normal operator T_0 with an abnormal operator T_1. We shall refer to T_0 and T_1 as the normal and abnormal parts of T, respectively.

A theorem of von Neumann ([7], p. 96) asserts that every isometry V on a Hilbert space H is unitarily equivalent to the direct sum of a unitary operator and a pure isometry of multiplicity $d = \dim [(VH)^{\perp}]$ (cf. [3], problem 118). It develops that the scalar d is a complete set of unitary invariants for the abnormal part of the isometry V. An operator T is quasinormal if T commutes with T^*T. In particular, every isometry is quasinormal. In [1], Brown obtains both a canonical form and a complete set of unitary invariants for the abnormal part of a quasinormal operator. In the isometric case, Brown's results specialize to those of von Neumann.

In section 1 of this paper we obtain a decomposition for operators, which, as is shown in section 3, is a generalization of Brown's work on quasinormal operators. We associate with each $T \in \mathcal{B}(H)$ a (not necessarily proper) subspace $H_1(T)$ of H which is invariant under T^* and reduces $[T] = T^*T - TT^*$. If $V \in \mathcal{B}(H)$ is isometric, for instance, one has $H_1(V) = (VH)^\perp$. We establish that the structure of the abnormal part of T is completely determined up to unitary equivalence by the restrictions of T^* and $[T]$ to $H_1(T)$. In case $d = \dim (H_1(T)) < \infty$, the structure of the abnormal part of T is determined by two d-by-d matrices.

The results of section 1 are of little interest if $H_1(T)$ is too large. In section 2 we study conditions under which $H_1(T) = H$. We show that if T is abnormal and nearly a finite-dimensional operator (in some appropriate sense), then $H_1(T) = H$. This suggests that the results of section 1 will be of most interest if the operator being studied is far from being finite-dimensional.

The main result of section 3 is that if T is subnormal, then $H_1(T)$ is the closure of the range of $[T]$. This means, for example, that the structure results given in section 1 may be easily applied to subnormal operators whose self-commutator is of finite rank. It also enables us to deduce the results of Brown and von Neumann mentioned above from our results in section 1.

In section 4 we give an application of our results to the study of quasitrangular operators.

The author would like to express his thanks to P.R. Halmos, T.L. Kriete, III, and to Marvin Rosenblum for a number of helpful conversations concerning the results in this paper.

§1. We begin with several lemmas which may be of interest independent of their application here.

LEMMA 1.1: Let H be a Hilbert space and let $A,B \in \mathcal{B}(H)$. Then $M = \bigcap_{s=1}^{\infty} \ker(AB^s - B^sA)$ is the largest subspace of H for which $BM \subseteq M$ and $ABv = BAv$ for every $v \in M$.

PROOF: It is clear that M is a subspace of H. Pick $v \in M$ and let $w = Bv$. Then for all integers $s \geq 1$ we have $AB^sw = AB^{s+1}v = B^{s+1}Av = B^s(BAv) = B^s(ABv) = B^sAw$, since $v \in M$. Hence $BM \subseteq M$. The relation $M \subseteq \ker(AB - BA)$ implies that $ABv = BAv$ for all $v \in M$.

Next, let Y be a subspace of H such that $BY \subseteq Y$ and $ABy = BAy$ for all $y \in Y$. Then $B^sY \subseteq Y$ for all $s \geq 1$. If $y \in Y$, then $AB^2y = AB(By) = BA(By) = B(ABy) = B^2Ay$. By induction, $B^sAy = AB^sy$ for every $y \in Y$ and all $s \geq 1$. Hence, $Y \subseteq \bigcap_{s=1}^{\infty} \ker(AB^s - B^sA) = M$.

An easy modification of the proof of Lemma 1.1 yields a proof of the following result:

LEMMA 1.2: Let $A,B \in \mathcal{B}(H)$. Then the largest subspace $M \subseteq H$ such that $AM \subseteq M$, $BM \subseteq M$ and $ABv = BAv$ for every $v \in M$ is

$$M = \bigcap_{r=1}^{\infty} \bigcap_{s=1}^{\infty} \ker(A^r B^s - B^s A^r) \ .$$

Although we shall not use the results in this generality, we note that Lemmas 1.1 and 1.2 both hold in case the underlying space is a Banach space.

An immediate consequence of Lemma 1.2 and the definition of reducing subspace is the following:

COROLLARY 1.3: Let $T \in \mathcal{B}(H)$. Then the largest subspace H_o of H which reduces T and such that $T|H_o$ is normal is

$$H_o = \bigcap_{r=1}^{\infty} \bigcap_{s=1}^{\infty} \ker((T^*)^r T^s - T^s (T^*)^r) \ .$$

Furthermore, $T|H_o^{\perp}$ is abnormal.

Our next lemma is the basis for the subsequent decomposition theorems.

LEMMA 1.4: Let $T \in \mathcal{B}(H)$. Put $M_o = H$, and, for all $k \geq 1$, define

$$M_k = \bigcap_{r=1}^{k} \bigcap_{s=1}^{\infty} \ker((T^*)^r T^s - T^s (T^*)^r) \ .$$

Then

i.) $M_k \supseteq M_{k+1}$ for all $k \geq 0$,

ii.) $TM_k \subseteq M_k$ for all $k \geq 0$,

iii.) $T^*M_k \subseteq M_{k-1}$ for all $k \geq 1$,

iv.) $T^*M_k^\perp \subseteq M_k^\perp$ for all $k \geq 0$,

v.) $TM_k^\perp \subseteq M_{k+1}^\perp$ for all $k \geq 0$.

PROOF: Parts iv.) and v.) of the assertion follow immediately from parts ii.) and iii.) upon taking orthogonal complements, and part i.) follows directly from the definition of the subspaces M_k .

Part ii.) is trivially true if $k = 0$. For $r \geq 1$, Lemma 1.1 implies that $\bigcap_{s=1}^{\infty} \ker[(T^*)^r T^s - T^s(T^*)^r]$ is an invariant subspace for T , and hence, so also is M_k for every $k \geq 1$.

Part iii.) is trivial if $k = 1$, so suppose that $k \geq 2$. Let $x \in M_k$ and put $z = T^*x$. Then since $x \in M_k \subseteq M_1$, we have $(T^*)^r T^s z = (T^*)^r (T^s T^* x) = (T^*)^{r+1} T^s x$. But if $r + 1 \leq k$, then, since $x \in M_p$ for every $p \leq k$, we have $(T^*)^r T^s z = T^s (T^*)^{r+1} x = T^s (T^*)^r z$. Thus $z \in M_{k-1}$ and iii.) holds.

If we let $\{e_k \mid k \geq 0\}$ be the standard orthonormal basis for ℓ^2 and if we let T be the unilateral shift

on ℓ^2 , then it is instructive to note that

$$M_k = sp\{e_j \mid j \geq k\} \quad \text{for} \quad k \geq 0 .$$

THEOREM 1.5: Let $T \in \mathcal{B}(H)$. Then there exists a (finite or infinite) sequence $\{H_j \mid j \geq 0\}$ of pairwise orthogonal subspaces of H such that

i.) $H = H_0 \oplus H_1 \oplus \ldots \oplus H_k \oplus \ldots$.

ii.) H_0 reduces T , $T|H_0$ is normal, and $T|H_0^{\perp}$ is abnormal.

iii.) $T^*H_1 \subseteq H_1$.

iv.) $T^*H_k \subseteq H_{k-1} \oplus H_k$ for all $k \geq 2$.

v.) $TH_k \subseteq H_k \oplus H_{k+1}$ for all $k \geq 1$.

vi.) $H_k \oplus H_{k+1} = \bigvee\{H_k, TH_k\}$ for all $k \geq 1$.

vii.) $\dim H_k \geq \dim H_{k+1}$ for all $k \geq 1$.

PROOF: We associate with T the subspaces M_k as was done in Lemma 1.4. Put $H_0 = \bigcap_{k=1}^{\infty} M_k$, or, equivalently,

$$H_0 = \bigcap_{r=1}^{\infty} \bigcap_{s=1}^{\infty} \ker((T^*)^r T^s - T^s (T^*)^r) .$$

From Corollary 1.3, H_0 reduces T , $T|H_0$ is normal, and $T|H_0^{\perp}$ is abnormal.

Next, define $H_k = M_k^\perp \cap M_{k-1}$ for all $k \geq 1$. Since $H_o = \bigcap_{k=1}^{\infty} M_k$, we have $H_o \perp H_j$ for every $j > 0$. Noting that $H_j \subseteq M_{j-1}$ for every $j \geq 1$ and that $H_i \subseteq M_i^\perp \subseteq M_{j-1}^\perp$ whenever $i \leq j-1$, we conclude that $H_i \perp H_j$ if $i < j$, or equivalently, $H_i \perp H_j$ if $i \neq j$. By induction,

$$M_k^\perp = H_1 \oplus H_2 \oplus \ldots \oplus H_k , \quad k \geq 1 .$$

It follows immediately that

$$H = H_o \oplus H_1 \oplus \ldots \oplus H_k \oplus \ldots .$$

Thus, both i.) and ii.) hold.

Part iii.) follows from Lemma 1.4, since $T^*H_1 = T^*M_1^\perp \subseteq M_1^\perp = H_1$.

We shall prove iv.) and v.) simultaneously. Note first that $TH_1 = TM_1^\perp \subseteq M_2^\perp = H_1 \oplus H_2$. Using Lemma 1.4 again, we have $T^*H_2 \subseteq T^*M_2^\perp = H_1 \oplus H_2$. Also, $TH_2 \subseteq TM_2^\perp \subseteq M_3^\perp = H_1 \oplus H_2 \oplus H_3$. Note that if $x_1 \in H_1$ and $x_2 \in H_2$, then

$$\langle Tx_2, x_1 \rangle = \langle x_2, T^*x_1 \rangle = 0 ,$$

since $T^*x_1 \in H_1$. Hence $TH_2 \subseteq H_2 \oplus H_3$ and iv.) and v.) hold in case $k = 2$. Suppose that it has been shown that $TH_j \subseteq H_j \oplus H_{j+1}$ and that $T^*H_j \subseteq H_{j-1} \oplus H_j$ for all $j \leq m$. Then $T^*H_{m+1} \subseteq T^*M_{m+1}^\perp \subseteq M_{m+1}^\perp = H_1 \oplus H_2 \oplus \ldots \oplus H_{m+1}$. But if $y \in H_1 \oplus \ldots \oplus H_{m+1}$,

then $Ty \in H_1 \oplus \ldots \oplus H_m$, and so, if $x \in H_{m+1}$, we have $\langle y, T^*x \rangle = \langle Ty, x \rangle = 0$. Thus, $T^*H_{m+1} \subseteq H_m \oplus H_{m+1}$. We also have $TH_{m+1} \subseteq TM_{m+1}^{\perp} \subseteq M_{m+2}^{\perp} = H_1 \oplus \ldots \oplus H_{m+2}$. If $x \in H_{m+1}$ and $y \in H_1 \oplus \ldots \oplus H_m$, then $T^*y \in H_1 \oplus \ldots \oplus H_m$ and $\langle Tx, y \rangle = \langle x, T^*y \rangle = 0$, so that $TH_{m+1} \subseteq H_{m+1} \oplus H_{m+2}$. By induction, both iv.) and v.) hold.

Since $TH_k \subseteq H_k \oplus H_{k+1}$ for every $k \geq 1$, we have $\bigvee\{H_k, TH_k\} \subseteq H_k \oplus H_{k+1}$ for every $k \geq 1$. Assume that $v \in H_k \oplus H_{k+1}$ and that v is orthogonal to $\bigvee\{H_k, TH_k\}$, where $k \geq 1$ is fixed. Clearly, $v \in H_{k+1}$. The fact that $\langle v, Tx \rangle = 0$ for every $x \in H_k$ together with v.) implies $\langle v, Tx \rangle = 0$ for every $x \in H_1 \oplus \ldots \oplus H_k$. Thus, T^*v is orthogonal to $H_1 \oplus \ldots \oplus H_k$; that is, $T^*v \in M_k$. But since $v, T^*v \in M_k$, we have

$$\begin{aligned} T^s(T^*)^{k+1}v &= T^s(T^*)^k(T^*v) \\ &= (T^*)^k T^s(T^*v) \\ &= (T^*)^k(T^*T^s v) \\ &= (T^*)^{k+1}T^s v \end{aligned}$$

for every $s \geq 1$. Hence $v \in \bigcap_{r=1}^{k+1} \bigcap_{s=1}^{\infty} \ker((T^*)^r T^s - T^s(T^*)^r)$; i.e., $v \in M_{k+1}$. Then $v \in [H_{k+1} \bigcap M_{k+1}] = \{0\}$ and vi.) holds. Part vii.) follows immediately from vi.)

It is worth noting that if T is the unilateral shift, then $H_0 = \{0\}$, while for $k \geq 1$, H_k is precisely the one-dimensional subspace spanned by e_{k-1}.

If $T \in \mathcal{B}(H)$, then we shall use the notation $H_k(T)$, $k = 0,1,2,\ldots$, to denote the subspaces associated with T as in Theorem 1.5. Note that in case $H = H_0(T) \oplus H_1(T)$, Theorem 1.5 is nothing more than the decomposition of T into a normal and an abnormal part. In case $H = H_0(T) \oplus H_1(T)$, we shall say that T has a trivial decomposition.

If $T \in \mathcal{B}(H)$ is abnormal (so that $H_0(T) = \{0\}$) and if we let $P_k : H \longrightarrow H_k(T)$ denote the orthogonal projection of H onto $H_k(T)$ for $k \geq 1$, and if we define

$$T_{ij} = P_i T | H_j$$

for all $i,j \geq 1$, then T is represented by the matrix of operators $\{T_{ij}\}$ acting on the direct sum of the spaces H_k, $k \geq 1$. Parts iii.), iv.), and v.) of Theorem 1.5 assert that $T_{ij} = 0$ if either $j > 1$ or $j < i-1$. Thus T is represented by a matrix of operators whose non-zero (operator) entries lie on either the main diagonal or the first subdiagonal of the matrix. To simplify the notation, let $D_i = T_{i,i}$ and let $S_i = T_{i+1,i}$ for every $i \geq 1$. From part vi.) of Theorem 1.5, we have ran S_i dense in H_{i+1} for $i \geq 1$, or, equivalently, that $\ker S_i^* = \{0\}$ for $i \geq 1$.

In the next two theorems we exhibit a canonical form for operators which have a non-trivial decomposition. The technique to be used is a modification of the proof of the fact that every weighted shift is unitarily equivalent to a weighted shift with non-negative weights (cf. [3], problem 75). Roughly speaking, we wish to show that the matrix of operators $\{T_{ij}\}$ described above is unitarily equivalent to a matrix of operators of the same form with non-negative operator weights along the first subdiagonal. Some technical difficulties arise from the fact that the spaces $\{H_k\}$ may be of different dimensions.

For simplicity, we break the reduction to canonical form into two parts.

THEOREM 1.6: Let $T \in \mathcal{B}(H)$ be abnormal. Then there exist a (finite or infinite) sequence of Hilbert spaces $J_1 \supseteq J_2 \supseteq \ldots$ and corresponding sequences of operators $D_i' : J_i \longrightarrow J_i$ and $S_i' : J_i \longrightarrow J_{i+1}$ with $\ker(S_i')^* = \{0\}$ and $\ker S_i' = J_i \ominus J_{i+1}$, such that T is unitarily equivalent to the operator T' defined on $J_1 \oplus J_2 \oplus \ldots$ by the matrix $\{T_{ij}'\}$ of operators given by $T_{i,i}' = D_i'$, $T'_{i+1,i} = S_i'$, and $T'_{i,j} = 0$ if $i \neq j$, $j+1$.

Further, $H_k(T') = J_k$ for $k \geq 1$.

PROOF: We shall assume that all of the subspaces $H_k(T)$, $k \geq 1$, are non-zero. Put $J_1 = H_1(T)$ and define

$W_1^* : J_1 \longrightarrow H_1(T)$ by $W_1^* = I$ on J_1. Put $D_1' = D_1$.

Recall that $S_1 : J_1 \longrightarrow H_2(T)$ and that $cl(\text{ran } S_1) = H_2(T)$. Let $J_2 = (\ker S_1)^\perp \subseteq J_1$. Then $\dim J_2 = \dim[(\ker S_1)^\perp] = \dim[cl(\text{ran } S_1)] = \dim H_2(T)$. Pick a unitary operator $W_2^* : H_2(T) \longrightarrow J_2$ (onto J_2) and define $S_1' = W_2^* S_1 W_1 = W_2^* S_1$. Then $S_1' : J_1 \longrightarrow J_2$. Since $\ker S_1^* = \{0\}$, we have $\ker(S_1')^* = \ker(S_1^* W_1) = \{0\}$, and since $J_2 = (\ker S_1)^\perp$, we have $\ker S_1' = \ker S_1 = J_1 \ominus J_2$. Setting $D_2' = W_2^* D_2 W_2$, we see that $D_2' : J_2 \longrightarrow J_2$.

Suppose that we have defined Hilbert spaces $J_1 \supseteq J_2 \supseteq \ldots \supseteq J_m$, that we have picked unitary operators W_i^* mapping $H_i(T)$ onto J_i, $i = 1,2,\ldots,m$, and that we have put $D_i' = W_i^* D_i W_i$ for $i = 1,2,\ldots,m$ and $S_i' = W_{i+1}^* S_i W_i$ for $i = 1,2,\ldots,m-1$. Then define $J_{m+1} = [\ker(S_m W_m)]^\perp$. Since $cl(\text{ran } S_m) = H_{m+1}(T)$, we have $\dim(J_{m+1}) = \dim(H_{m+1}(T))$. Pick a unitary operator W_{m+1}^* mapping $H_{m+1}(T)$ onto J_{m+1} and define $S_m' = W_{m+1}^* S_m W_m$. Then, as above, S_m' maps J_m into J_{m+1}, $\ker(S_m')^* = \{0\}$, and $\ker S_m' = J_m \ominus J_{m+1}$. Next, put $D_{m+1}' = W_{m+1}^* D_{m+1} W_{m+1}$, noting that D_{m+1}' is an operator on J_{m+1}.

Continuing this process, we obtain a sequence $J_1 \supseteq J_2 \supseteq \ldots$ of Hilbert spaces and a sequence of unitary

operators $\{W_i\}$ with W_i mapping J_i onto $H_i(T)$ for all $i \geq 1$. The associated sequences of operators $\{D_i'\}$ and $\{S_i'\}$ are as in the statement of the theorem.

Next put $J = J_1 \oplus J_2 \oplus \ldots$ and define $W = W_1 \oplus W_2 \oplus \ldots$. Then W is a unitary operator mapping J onto H, and, of course, the operator $T' = W^*TW \in \mathcal{B}(J)$ is unitarily equivalent to T. A straightforward computation with the representations of T and W as matrices of operators shows that the matrix of T' relative to the decomposition $J = J_1 \oplus J_2 \oplus \ldots$ is as desired.

Since $T' = W^*TW$, we find that for all $r,s \geq 1$, the equation

$$\{(T')^*\}^r(T')^s - (T')^s\{(T')^*\}^r = W^*\{(T^*)^rT^s - T^s(T^*)^r\}W$$

holds, so that W^* maps $M_k(T)$ onto $M_k(T')$ in a one-to-one fashion for all $k \geq 1$. It follows immediately that $H_k(T') = J_k$ for all $k \geq 1$.

DEFINITION: Let $J_1 \supseteq J_2 \supseteq \ldots$ be a finite or infinite sequence of Hilbert spaces and let $J = J_1 \oplus J_2 \oplus \ldots \oplus J_k \oplus \ldots$. Let E_i denote the orthogonal projection of J_i onto $J_{i+1} \subseteq J_i$ and let π_i denote the orthogonal projection of J onto J_i for all $i > 0$. Let $T \in \mathcal{B}(J)$. Then we say that T is in standard form if $M_k(T) = J_{k+1} \oplus J_{k+2} \oplus \ldots$ for $k \geq 1$ and if there exist operators $D_i'' \in \mathcal{B}(J_i)$ and non-negative operators $P_i \in \mathcal{B}(J_i)$ with $\ker P_i = J_i \ominus J_{i+1}$

such that $\pi_i T|J_k = 0$ if $k > i$ or $k < i-1$, $\pi_i T|J_i = D_i^w$ and $\pi_{i+1} T|J_i = E_i P_i$ for all $i > 0$.

THEOREM 1.7: Let $T \in \mathcal{B}(H)$ be abnormal. Then T is unitarily equivalent to an operator in standard form.

PROOF: We may as well replace T by the operator T' of Theorem 1.6. We also retain the notation of Theorem 1.6.

Define $V_1 = I$ on J_1. Next, write $S_1' = U_1((S_1')^* S_1')^{\frac{1}{2}} = U_1 P_1$, so that P_1 is a non-negative operator on J_1 and U_1 is a partial isometry with initial space $(\ker S_1')^\perp = J_2 \subseteq J_1$ and final space $\mathrm{cl}(\mathrm{ran}\, S_1') = J_2$. Since U_1 is onto J_2, it follows that $U_1^* : J_2 \longrightarrow J_1$ is an isometry with range J_2. Thus, $V_2^* = E_1 U_1^*$ is a unitary operator on J_2 and $V_2^* S_1' V_1 = V_1^* U_1 P_1 = E_1 U_1^* U_1 P_1 = E_1 P_1$.

Suppose that we have defined unitary operators $V_i \in \mathcal{B}(J_i)$ such that $V_i^*(S_{i-1}')^* V_{i-1} = E_{i-1} P_{i-1}$ for $1 \le i \le m$ and non-negative operators $P_i \in \mathcal{B}(J_i)$ by $P_i = (V_i^*(S_i')^* S_i' V_i)^{\frac{1}{2}}$ for $1 \le i \le m$. We consider the polar factorization $S_m' V_m = U_m P_m$ of $S_m' V_m$. Arguing as above, $U_m^* : J_{m+1} \longrightarrow J_m$ is an isometry with final space J_{m+1}, so that $V_{m+1}^* = E_m U_m^*$ is a unitary operator on J_{m+1}. Thus we obtain a (finite or infinite) sequence of unitary operators $V_i \in \mathcal{B}(J_i)$ and a sequence $P_i \in \mathcal{B}(J_i)$ of non-negative operators such that $\ker P_i = J_i \ominus J_{i+1}$

and $V_{i+1}^* S_i' V_i = E_i P_i$ for all $i > 0$. We define $D_i'' = V_i^* D_i' V_i$ for all $i > 0$ and put $V = V_1 \oplus V_2 \oplus \ldots \oplus V_k \oplus \ldots$. Then $V \in \mathcal{B}(J)$ is unitary, so that $T'' = V^* T' V$ is unitarily equivalent to T', and hence, to T. A straightforward calculation shows that $\pi_1 T''|J_k = 0$ if $k > i$ or $k < i-1$, that $\pi_i T''|J_i = V^* D_i' V_i = D_i''$ for $i > 0$, and that $\pi_{i+1} T''|J_i = V_{i+1}^* S_i' V_i = E_i P_i$ for all $i > 0$.

Finally, we note that an argument analogous to that used in Theorem 1.6 shows that $H_i(T'') = J_k$ for all $k > 0$. We note also that $H_k(T'') = VWH_k(T)$ and $H_k(T) = (VW)^* H_k(T'')$ for all $k > 0$.

To simplify our notation, we shall assume in the future that if $T \in \mathcal{B}(H)$ is abnormal and is in standard form, then $H = H_1 \oplus H_2 \oplus \ldots$, where $H_1 \supseteq H_2 \supseteq \ldots$. The diagonal (operator) entries of the matrix representation for T will be denoted by D_k and the subdiagonal entries by $S_k = E_k P_k$ for all $k > 1$.

Next, suppose that $T \in \mathcal{B}(H)$ is abnormal and that T is unitarily equivalent to $T^{(1)} \in \mathcal{B}(H^{(1)})$, and to $T^{(2)} \in \mathcal{B}(H^{(2)})$, where $T^{(i)}$ is in standard form for $i = 1,2$. Then there exists a unitary operator U such that $T^{(1)} = U^* T^{(2)} U$. Arguing as in Theorem 1.6, we see that U maps $M_k(T^{(1)})$ onto $M_k(T^{(2)})$ for all $k \geq 1$. Viewing $T^{(1)}$, $T^{(2)}$ and U as matrices of operators,

this means that U is a diagonal matrix with diagonal (operator) entries U_k, where U_k is a unitary operator mapping $H_k(T^{(1)})$ onto $H_k(T^{(2)})$ for all $k \geq 1$. With some more work, the details of which we omit, one verifies the relations

$$P_k^{(1)} = U_k^* P_k^{(2)} U_k ,$$

$$[E_k^{(1)}]^* E_k^{(1)} = U_k^* (E_k^{(2)})^* E_k^{(2)} U_k ,$$

and

$$E_k^{(1)} [E_k^{(1)}]^* = U_{k+1}^* E_k^{(2)} [E_k^{(2)}]^* U_{k+1}$$

for all $k \geq 1$. In effect, these relations assert that each of the non-zero entries in the matrix representation of $T^{(1)}$ is unitarily equivalent to the corresponding entry of $T^{(2)}$. It follows that the representation of an operator in standard form is essentially unique.

The next theorem deals with the relations that hold among the entries in the standard form for T. It shows that the structure of the abnormal part of an operator is determined by the action of the operator on the subspace $H_1(T)$.

THEOREM 1.8: Let $T \in \mathcal{B}(H)$ be abnormal and in standard form. Put $C = [T]|H_1(T)$. Then

i.) $[D_1] = C - S_1^* S_1$,

ii.) $[D_i] = S_{i-1} S_{i-1}^* - S_i^* S_i$ for $i \geq 2$.

iii.) $S_i^* D_{i+1} = D_i S_i^*$ for $i \geq 1$.

Further, the operators D_i , $i \geq 2$ and the operators P_i , E_i , $i \geq 1$, may be determined explicitly in terms of C and D_1 .

PROOF: Note first that

$$\ker[T] \supseteq M_1(T) = \bigcap_{r=1}^{\infty} \ker(T^* T^r - T^r T^*) .$$

Hence, $\mathrm{cl}(\mathrm{ran}[T]) \subseteq (M_1(T))^{\perp} = H_1(T)$. It follows that $H_1(T)$ reduces $[T]$, so that $C = [T]|H_1(T)$ is well-defined. In fact, $[T] = C \oplus O$. If we represent $[T]$ as a matrix of operators relative to the decomposition

$$H = H_1 \oplus H_2 \oplus \ldots \oplus H_k \oplus \ldots ,$$

then we obtain a matrix $\{C_{i,j}\}$ of operators with $C_{1,1} = C$ and $C_{i,j} = O$ if $i+j > 2$.

Using the matrix representation for T relative to this same decomposition for H , we obtain another expression for $[T]$. Direct comparison of the entries in these two representations for $[T]$ yield equations i.) , ii.) , and iii.).

To complete the proof, note first that

$\ker S_1 = \ker(S_1^*S_1) = \ker(C-[D_1])$, so that $H_2 = (\ker S_1)^\perp = \mathrm{cl}\{\mathrm{ran}(C-[D_1])\}$. Hence, H_2 (and thus, E_1) is determined by C and D_1 . Noting that $S_1^*S_1 = P_1E_1^*E_1P_1$ and that $E_1^*E_1 \in \mathcal{B}(H_1)$ is the orthogonal projection of H_1 onto H_2 , we have, since $H_2 = \mathrm{cl}(\mathrm{ran}\ P_1)$, that $S_1^*S_1 = P_1^2$, or,

$$P_1 = (S_1^*S_1)^{\frac{1}{2}} = (C-[D_1])^{\frac{1}{2}} .$$

We note for future reference that $S_1S_1^* = E_1P_1^2E_1 = P_1^2|H_2$.

From iii.), $S_1^*D_2 = D_1S_1^*$. If $X \in \mathcal{B}(H_2)$ is any operator satisfying $S_1^*X = D_1S_1^*$, then $S_1^*(D_2-X) = 0$ and, since $\ker S_1^* = \{0\}$, $D_2 = X$. Note also that $S_1S_1^*D_2 = S_1D_1S_1^*$ and hence, $D_2 = (S_1S_1^*)^{-1}S_1D_1S_1^*$. The expression on the right in the last equation represents a bounded operator even though $S_1S_1^*$ will not, in general, have a bounded inverse. Substitution from above yields

$$D_2v = (D-[D_1])^{-\frac{1}{2}}D_1(C-[D_1]^{\frac{1}{2}}v$$

for all $v \in H_2$.

A messy but rather easy use of induction completes the proof. We omit the details.

The formulas in Theorem 1.8 are much more manageable in the special case $\dim H_1 = \dim H_k$ for all $k \geq 1$. The operators E_k are unnecessary in this case, so that $S_k = P_k \geq 0$ for all $k \geq 1$. In this case, one obtains

the formulas

$$P_k = (C - \sum_{i=1}^{k} [D_i])^{\frac{1}{2}}$$

and

$$D_{k+1} = P_k^{-1} D_k P_k \quad \text{for all} \quad k \geq 1 .$$

The formulas in Theorem 1.8 are easy to handle only in special cases. An important observation, however, is that the structure of an abnormal operator T is determined by its action on the subspace $H_1(T)$. In case $H_1(T)$ is an infinite-dimensional subspace, then, in the absence of stronger hypotheses on $[T]$ and D_1, nothing has been gained. If $\dim H_1(T) < \infty$, however, Theorem 1.8 asserts that the structure of (the abnormal part of) T is determined by two finite-dimensional operators.

§2. It is easily seen that the decomposition for operators given in Section 1 may be trivial. If T is normal, for instance, then $H = H_0(T)$. Even if T is abnormal, the decomposition will be trivial if $\ker[T] = \{0\}$, since $H = H_1(T)$ in this case. In this section we shall consider other conditions which imply that our decomposition is trivial.

LEMMA 2.1: Let $T \in \mathcal{B}(H)$ and suppose that M is a subspace of H such that $TM \subseteq M$ and $M \subseteq \ker[T]$. Then $T|M$ is hyponormal. If $T|M$ is normal, then M reduces T.

PROOF: Write

$$T = \begin{bmatrix} A & B \\ O & C \end{bmatrix}$$

with respect to the decomposition $H = M \oplus M^{\perp}$. Then

$$[T] = \begin{bmatrix} [A]-BB^{*} & A^{*}B-BC^{*} \\ B^{*}A-CB^{*} & [C]+BB^{*} \end{bmatrix} .$$

The fact that $M \subseteq \ker[T]$ implies that M reduces $[T]$. Let $X = [T]|M^{\perp}$. Then $[T]$ has the representation $[T] = O \oplus X$ relative to $H = M \oplus M^{\perp}$. Equating corresponding entries in the two representations for $[T]$ yields $[A] = BB^{*} \geq O$, so that $A = T|M$ is hyponormal. If A is normal, then $BB^{*} = O$. Thus $B = O$ and M reduces T.

THEOREM 2.2: Let $T \in \mathcal{B}(H)$ have compact real part. Then $H = H_{o}(T) \oplus H_{1}(T)$.

PROOF: We may as well assume that T is abnormal. Assume that $H \neq H_{1}(T)$ and let $A = T|(H_{1}(T))^{\perp}$. From Lemma 2.1, A is hyponormal. Since Re T is compact, so also is Re A.

Putnam has shown ([5], p. 43) that if T is hyponormal and abnormal, then the measure of the spectrum of Re T is positive. Since a compact self-adjoint operator has countable spectrum, it follows that a hyponormal operator

with compact real part is normal. In particular, A is normal. By Lemma 2.1 again, $[H_1(T)]^{\perp}$ reduces T and $T|(H_1(T))^{\perp}$ is normal, a contradiction, since T was assumed to be abnormal.

It is interesting to note what happens in case H is a finite-dimensional Hilbert space. Since every operator on a finite-dimensional space is compact, if follows from Theorem 2.2 that our decomposition is always trivial for finite-dimensional operators. Hence, non-trivial examples of our decomposition, much like non-unitary isometries, are purely infinite-dimensional phenomena.

As a consequence of Theorem 2.2 we obtain a simpler expression for the normal subspace of an operator having compact real part.

COROLLARY 2.3: If $T \in \mathcal{B}(H)$ has compact real part, then

$$H_1(T) = \bigcap_{r=1}^{\infty} \ker(T^*T^r - T^rT^*) .$$

PROOF:
$$H_1(T) = H \ominus H_1(T)$$
$$= M_1(T)$$
$$= \bigcap_{r=1}^{\infty} \ker(T^*T^r - T^rT^*) .$$

§3. Since our decomposition is trivial for operators on a finite-dimensional Hilbert space, we shall assume hereinafter that the underlying Hilbert space H is infinite-dimensional. Note that the easiest way to guarantee that the decomposition of an abnormal operator T is non-trivial is to assume that $\dim H_1(T) < \infty$. This ensures that $H \neq H_1(T)$, of course, but it also means that $H_k(T) \neq \{0\}$ for all $k \geq 1$, since, from Theorems 1.5 and 1.6, we have

$$\dim(H_1(T) \oplus \ldots \oplus H_k(T)) \leq k \dim H_1(T) < \infty .$$

The condition $\dim H_1(T) < \infty$ is difficult to verify in many cases. In this section we will show that this condition is easy to verify in case T is subnormal.

We recall that $T \in \mathcal{B}(H)$ is subnormal if there exists a Hilbert space $K \supseteq H$ and a normal operator $N \in \mathcal{B}(K)$ such that $NH \subseteq H$ and $T = N|H$, in which case N is called a normal extension of T. We say that N is a minimal normal extension of T is the smallest subspace of K which contains H and reduces N is K itself. Halmos has shown that every subnormal operator has a minimal normal extension and that this extension is unique up to unitary equivalence. For proofs of these facts and an excellent discussion of subnormal operators, see [3], Chapter 16.

LEMMA 3.1: Let $T \in \mathcal{B}(H)$ be subnormal. Then

1.) $\ker[T^r] \subseteq \ker((T^*)^r T^s - T^s(T^*)^r)$ for $r,s > 0$.

ii.) $\ker[T^r] = \bigcap_{s=1}^{\infty} \ker((T^*)^r T^s - T^s(T^*)^r)$ for all $r > 0$.

iii.) $\bigcap_{r=1}^{\infty} \ker[T^r] = H_o(T)$.

PROOF: Let $N \in \mathcal{B}(K)$, $K \supseteq H$, be the minimal normal extension of T . Since $NH \subseteq H$ and $T = N|H$, we may write

$$N = \begin{bmatrix} T & X \\ O & Y \end{bmatrix}$$

with respect to the decomposition $K = H \oplus H^{\perp}$. Then

$$N^k = \begin{bmatrix} T^k & X_k \\ O & Y^k \end{bmatrix}$$

for all $k \geq 1$, where $X_1 = X$ and $X_{n+1} = TX_n + XY^n = T^n X_1 + X_n Y$ for all $n \geq 1$. Computing both $(N^*)^r N^s$ and $N^s(N^*)^r$ and equating corresponding entries yields

(1) $(T^*)^r T^s - T^s(T^*)^r = X_s X_r^*$

and

(2) $(T^*)^r X_s = X_1 (Y^*)^r$.

for all $r,s > 0$. Putting $r = s$ in (1) gives

$\ker[T^r] = \ker(X_r X_r^*) = \ker X_r^*$ for all $r \geq 1$, Thus, $\ker((T^*)^r T^s - T^s(T^*)^r) = \ker(X_s X_r^*) \supseteq \ker X_r^* = \ker[T^r]$, and i.) holds.

From i.), $\bigcap_{s=1}^{\infty} \ker((T^*)^r T^s - T^s(T^*)^r)$ contains $\ker[T^r]$ for $r \geq 1$, and since the reverse containment is trivial, the two sets are equal. Part iii.) is an immediate consequence of part ii.) and Lemma 1.3.

Recall that if $T \in \mathcal{B}(H)$, then

$$M_k(T) = \bigcap_{r=1}^{k} \bigcap_{s=1}^{\infty} \ker((T^*)^r T^s - T^s(T^*)^r) .$$

if T is subnormal, then, applying Lemma 3.1, we obtain

$$M_k(T) = \bigcap_{r=1}^{k} \ker[T^r] .$$

In particular, if $T \in \mathcal{B}(H)$ is subnormal, then $H_1(T) = (M_1(T))^{\perp} = \mathrm{cl}(\mathrm{ran}[T])$. This shows that if T is subnormal and abnormal and if $H \neq \mathrm{cl}(\mathrm{ran}[T])$ (in particular, if $[T]$ has finite rank), then our decomposition for T will be non-trivial.

It follows from Theorem 1.5 and the remarks above that if T is subnormal, then $M_k(T) = \bigcap_{r=1}^{k} \ker[T^r]$ is invariant under T. Actually, a stronger statement is at hand. Taking adjoints in equation (2) in the proof of Lemma 3.1 and putting $r = 1$, we get $X_s^* T = Y X_s^*$ for all $s > 0$.

Thus, $\ker[T^r] = \ker X_r^*$ is invariant under T for all $r \geq 1$. In case $r = 1$ this observation is due to Stampfli ([6]).

If T is subnormal and abnormal and if $H \neq \mathrm{cl}(\mathrm{ran}[T])$, or equivalently, if $\ker[T] \neq \{0\}$, then it follows from Theorem 1.8 that the structure of T is determined (up to unitary equivalence) by $[T]$ and $T^*|H_1(T)$. In the special case in which $[T]$ is of finite rank, the structure of the abnormal part of T is determined by two matrices. In case T is abnormal and $[T]$ is of rank one, there are two constants which are a complete set of unitary invariants for T.

PROPOSITION 3.2: Let $T \in \mathcal{B}(H)$ be subnormal with one-dimensional self-commutator. Let U denote the unilateral shift on ℓ^2. Then there exist scalars s_1, $d_1 (s_1 > 0)$ such that T is unitarily equivalent to the direct sum of a normal operator and $s_1 U + d_1 I$.

PROOF: We may as well assume that T is abnormal. We have $\dim(H_1(T)) = \dim(\mathrm{cl}(\mathrm{ran}[T])) = 1$. It follows from Theorem 1.5 that $\dim(H_k(T)) \leq 1$ for all $k \geq 1$. Since H is the direct sum of the spaces $H_k(T)$, $k \geq 1$, and since H is infinite-dimensional, we must have $\dim(H_k(T)) = 1$ for $k \geq 1$. From Theorem 1.8, T is unitarily equivalent to a matrix with scalars d_i on the main diagonal, positive scalars s_i on the first

subdiagonal, and zeros in the other entries. Further, since $s_k^* d_{k+1} = d_k s_k^*$ and since $s_k > 0$, $k \geq 1$, we have $d_k = d_1$ for all $k \geq 1$. From Theorem 1.8 again, we have $0 = [d_k] = s_{k-1}s_{k-1}^* - s_k^* s_k = |s_{k-1}|^2 - |s_k|^2$ for all $k \geq 2$. We then have $s_k = s_1$ for all $k \geq 1$ and we conclude that T is unitarily equivalent to $s_1 U + d_1 I$.

A careful examination of the proof of Proposition 3.2 shows that the result holds if we assume only that $T \in \mathcal{B}(H)$ satisfies $\dim(H_1(T)) = 1$. Since one may conclude from this that T is subnormal, the apparent generalization is really an artificial one. Finally, we note that Proposition 3.2 has been obtained independently by K. Clancey ([2]).

Recall that $T \in \mathcal{B}(H)$ is quasinormal if T commutes with T^*T, or, equivalently, if $T^*[T] = 0 = [T]T$. Thus, if T is quasinormal, then $T^*x = 0$ for every $x \in \text{cl}(\text{ran}[T])$. The following lemma was first proved by A. Brown in [1].

LEMMA 3.3: If $T \in \mathcal{B}(H)$ is quasinormal, then T is subnormal.

PROOF: Write $x \in H$ as $x = x_1 + x_2$, where $x_1 \in \text{cl}(\text{ran}[T])$ and $x_2 \in \ker[T]$. Then

$$\begin{aligned}
\langle [T]x, x\rangle &= \langle [T]x, x_1\rangle + \langle [T]x, x_2\rangle \\
&= \langle x, [T]x_1\rangle + \langle x, [T]x_2\rangle \\
&= \langle x_1, [T]x_2\rangle + \langle x_2, [T]x_2\rangle \\
&= \langle x_2, T^*Tx_2\rangle - \langle x_2, TT^*x_2\rangle \\
&= \|Tx_2\|^2 \geq 0 .
\end{aligned}$$

Thus, T is hyponormal. Note that $cl(ran[T]^{\frac{1}{2}}) = cl(ran[T])$, so that $T^*[T]^{\frac{1}{2}} = 0 = [T]^{\frac{1}{2}}T$. A direct computation shows that the operator X defined on $H \oplus H$ by

$$X = \begin{bmatrix} T & [T]^{\frac{1}{2}} \\ 0 & T^* \end{bmatrix}$$

is normal. Hence, T is subnormal.

THEOREM 3.4: (A. Brown, [1]) Let $T \in \mathcal{B}(H)$ be quasinormal. Put $R = cl(ran[T]^{\frac{1}{2}})$ and $C = [T]|R$. Then T is unitarily equivalent to the direct sum of a normal operator with the operator defined on $R \oplus R \oplus \ldots$ by the matrix of operators $\{T_{i,j}\}$ with $T_{i+1,i} = C^{\frac{1}{2}}$ for $i \geq 1$, $T_{i,j} = 0$ if $i \neq j+1$.

PROOF: We may as well assume that T is both abnormal and in standard form. Since T is subnormal, we have $H_1 = R$, and since $T^*[T] = 0$, we have $T^*|H_1 = 0$, or,

in the notation of section 1, $D_1 = 0$. Since $S_i^* D_{i+1} = D_i S_i^*$ and $\ker S_i^* = \{0\}$ for all $i \geq 1$, we have $D_i = 0$ for all $i \geq 1$.

We next observe that since $D_i = 0$ for $i \geq 1$, the inclusion $\ker S_i \subset \ker T$ holds for all $i \geq 1$. But T is both abnormal and hyponormal, and hence, $\ker T = \{0\}$. Thus $\ker S_i = \{0\}$ for all $i \geq 1$, and, since $H_{i+1} = \mathrm{cl}(S_i H_i)$, we have $\dim H_i = \dim H_1$ for all $i \geq 2$. We have shown that T is a matrix of operators on $R \oplus R \oplus \ldots$ whose only non-zero entries are the non-negative operators $S_i = P_i$, $i \geq 1$, which appear on the first subdiagonal.

From part ii.) of Theorem 1.8, we have $0 = [D_i] = S_{i-1}^2 - S_i^2$ for all $i \geq 1$, and hence, $S_i = S_1$ for all $i \geq 2$. From part i.) of Theorem 1.8, $0 = [D_1] = C - S_1^2$. Hence $S_i = C^{\frac{1}{2}}$ for all $i \geq 1$.

If $V \in \mathcal{B}(H)$ is isometric, then $V^*V = I$ and V is quasinormal. Recall that if V is isometric and if P denotes the orthogonal projection on $(VH)^\perp$, then $VV^* = I - P$, so that $[V] = I - (I-P) = P$. This implies that $H_1(V) = \mathrm{cl}(\mathrm{ran}[V]) = (VH)^\perp$.

COROLLARY 3.5: (von Neumann, [7]) Every isometry $V \in \mathcal{B}(H)$ is unitarily equivalent to the direct sum of a unitary operator with a unilateral shift of multiplicity $\dim(VH)^\perp$.

PROOF: A normal isometry is unitary, since $V^*V = I = VV^*$. Noting that $[V]|H_1(V)$ is the identity operator on $(VH)^\perp$ and applying Theorem 3.4, we see that V is unitarily equivalent to the direct sum of a unitary operator and a matrix of operators on $R \oplus R \oplus \ldots$ with identity operators on the first subdiagonal and zeros elsewhere, i.e., a unilateral shift of multiplicity $\dim R = \dim(VH)^\perp$.

We note that if T is quasinormal, then a complete set of unitary invariants for $[T]$ is a complete set of unitary invariants for the abnormal part of T. In case V is an isometry, the fact that $[V] = I$ on $H_1(V) = \operatorname{ran}[V] = (VH)^\perp$ means that the scalar $\dim(\operatorname{ran}[V]) = \dim(VH)^\perp$ is a complete set of unitary invariants for the abnormal part of V.

§4. We conclude with a simple application of our decomposition theorems to the study of quasitriangular operators. Recall that $T \in \mathcal{B}(H)$ is triangular if there exists an increasing sequence $\{E_k\}$ of projections of finite rank such that $\{E_k\} \longrightarrow I$ strongly as $k \longrightarrow \infty$ and such that $TE_k - E_kTE_k = 0$ for all k. We say that T is quasi-triangular if there exists an increasing sequence $\{E_k\}$ of projections of finite rank such that $\{E_k\} \longrightarrow I$ strongly as $k \longrightarrow \infty$ and $\|TE_k - E_kTE_k\| \longrightarrow 0$ as $k \longrightarrow \infty$. It is clear that every triangular operator is quasitriangular. We note that the study of quasitriangular

operators was initiated by Halmos ([4]). We shall use the facts, first proved in [4], that every normal operator is quasitriangular and that the direct sum of two quasitriangular operators is quasitriangular.

PROPOSITION 4.1: Let $T \in \mathcal{B}(H)$ be abnormal with $\dim(H_1(T)) < \infty$. Then T^* is a triangular operator.

PROOF: Let E_k denote the orthogonal projection of H onto $[M_k(T)]^\perp$ for all $k \geq 1$. From Theorem 1.5, the sequence $\{E_k\}$ is an increasing sequence of projections of finite rank, and since $TM_k(T) \subseteq M_k(T)$ for all $k \geq 1$, we also have $T^*E_k - E_kT^*E_k = 0$ for all $k \geq 1$. The abnormality of T implies that $\bigcap_{k=1}^{\infty} M_k(T) = \{0\}$, or, equivalently, that E_k tends strongly to I as k tends to infinity.

The preceding proposition, together with Lemma 1.3 and Halmos' results, yield the following:

COROLLARY 4.2: Let $T \in \mathcal{B}(H)$ satisfy $\dim(H_1(T)) < \infty$. Then T^* is quasitriangular.

COROLLARY 4.3: Suppose that $T \in \mathcal{B}(H)$ is subnormal and that its self-commutator has finite rank. Then T^* is quasitriangular.

The answer to the following question is apparently unknown.

QUESTION: If T is subnormal and has compact self-commutator, is T^* quasitriangular?

REFERENCES

[1] A. Brown, On a class of operators, Proc. Amer. Math. Soc. 4(1953), 723-728.

[2] K. Clancey, On the subnormality of some singular integral operators, preprint, University of Georgia, Athens, Georgia, 1971.

[3] P. R. Halmos, A Hilbert Space Problem Book, Van Nostrand, Princeton, N.J., 1967.

[4] P. R. Halmos, Quasitriangular operators, Acta Szeged 29(1968), 283-293.

[5] C. R. Putnam, Commutation properties of Hilbert space operators and related topics, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 36, Springer-Verlag, New York, 1967.

[6] J. G. Stampfli, Hyponormal operators and spectral density, Trans. Amer. Math. Soc. 117(1965), 469-476.

[7] J. von Neumann, Allgemeine Eigenwerttheorie Hermitischer Funktionaloperatoren, Math. Ann. 102(1929), 49-131.

Acknowledgement: This work has been supported in part by the National Aeronautics and Space Administration (NGR 11-003-020).

The University of Georgia
Athens, Georgia 30601

TITLE OF PAPER:

Stochastic Green's Functions

TEXT OF ABSTRACT (Not more than twenty-six 50 space lines. Double space. No displayed equations.)

A *stochastic Green's function* (s.g.f.) is derived which allows writing the two-point correlation (statistical measure) of the solution (stochastic) process for the stochastic differential equation $\mathcal{L}y=x$, where $x(t,\omega)$, $t \in T$, $\omega \in (\Omega, \mathcal{F}, \mu)$, a probability space, and $\mathcal{L}(t,\omega')$, $t \in T$, $\omega' \in (\Omega', \mathcal{F}', \mu')$ is a *stochastic operator* (a differential operator involving random function coefficients). The s.g.f. is calculable in terms of statistical measures of the random fluctuations of the coefficient processes and ordinary (deterministic) Green's functions. Where perturbation theory is adequate to deal with the randomness involved, the s.g.f. approach yields the expected results. In more general cases, where perturbation theory is not adequate, this appears still to be a promising approach. The results have been applied to a linear (control) system with n dimensional state vector y(t) with system equation - $\dot{y}=f(y,x,t)=A(t)y+x(t)$, A a n x n matrix with random elements $a_{ij}(t,\omega)$, x(t) a product of a n x r *stochastic matrix* (not a probability transition matrix) B(t) with a r x 1 matrix u(t). The Green's function G becomes a Green's matrix (state transition matrix in control theory), and we can determine expectations of correlations of y. (This work is supported by National Aeronautics and Space Administration Grant NGR 11 003 020.)

theory (finite and infinite)
16 ☐ Associative rings and algebras
17 ☐ Nonassociative rings and algebras
18 ☐ Category theory, homological algebra
20 ☐ Group theory and generalizations
22 ☐ Topological groups, Lie groups
26 ☐ Functions of real variables
28 ☐ Measure and integration
30 ☐ Functions of a complex variable
31 ☐ Potential theory
32 ☐ Several complex variables and analytic spaces
33 ☐ Special functions
34 ☐ Ordinary differential equations
35 ☐ Partial differential equations
39 ☐ Finite differences and functional equations
40 ☐ Sequences, series, summability
41 ☐ Approximations and expansions
42 ☐ Fourier analysis
43 ☐ Abstract harmonic analysis
44 ☐ Integral transforms, operational calculus
45 ☐ Integral equations
46 ☐ Functional analysis
47 ☐ Operator theory
49 ☐ Calculus of variations and optimal control
50 ☐ Geometry
52 ☐ Convex sets and geometric inequalities
53 ☐ Differential geometry
54 ☐ General topology
55 ☐ Algebraic topology
57 ☐ Manifolds and cell complexes
58 ☐ Global analysis, analysis on manifolds
60 ☑ Probability theory and stochastic processes
62 ☐ Statistics
65 ☐ Numerical analysis
68 ☐ Computer science
70 ☐ Mechanics of particles and systems
73 ☐ Mechanics of solids
76 ☐ Fluid mechanics
78 ☐ Optics, electromagnetic theory
80 ☐ Classical thermodynamics, heat transfer
81 ☐ Quantum mechanics
82 ☐ Statistical physics, structure of matter
83 ☐ Relativity
85 ☐ Astronomy and astrophysics
86 ☐ Geophysics
90 ☐ Economics, operations research, programming, games
92 ☐ Biology and behavioral sciences
93 ☐ Systems, control
94 ☐ Information and communication, circuits, automata
96 ☐ Mathematical education, elementary
97 ☐ Mathematical education, secondary
98 ☐ Mathematical education, collegiate
99 ☐ None of the above

For URSI Commission VI and G-AP

PROPAGATION IN RANDOM MEDIA

G. Adomian, University of Georgia

Abstract The problem of wave motion in a stochastic medium is treated as an application of stochastic operator theory to the case of partial differential equations and wave equations. Stochastic Green's functions are found for the two point correlation of the solution process for a scalar wave equation with randomly time-varying index of refraction without monochromaticity assumptions. The results are connected to the theory of partial coherence and can be used to calculate spectral spreading in a "hot" medium.

PROPAGATION IN STOCHASTIC MEDIA*

G. Adomian, University of Georgia
L. H. Sibul, Pennsylvania State Univ. O.R.L.

Consider the scalar wave equation

$$\nabla^2 y(\bar{r},t,\omega) - \frac{\partial^2}{\partial t^2}\left[\frac{1}{c^2} + \alpha(\bar{r},t,\omega)\right] y(\bar{r},t,\omega) = x(\bar{r},t,\omega)$$

where $t \in T$ represents time, $\bar{r} \in R^3$, $\omega \in \Omega$ on a probability space $(\Omega, \mathcal{F}, \mu)$. The quantities x and α, and consequently y, are all stochastic processes (s.p.) dependent on space position and time, i.e., random fields. The deterministic operator L is given by the ordinary d'Alembertian $\nabla^2 - (1/c^2)\partial^2/\partial t^2$ and the random part of the stochastic operator by $R = (\partial^2/\partial t^2)\alpha$.

Letting $L^{-1}x = F$ we write the above as

$$y(\bar{r},t,\omega) = F(\bar{r},t,\omega) + L^{-1}(\partial^2/\partial t^2)\alpha(\bar{r},t,\omega)y(\bar{r},t,\omega)$$

where L^{-1} is the inverse of the operator $\nabla^2 - (1/c^2)(\partial^2/\partial t^2)$. Denoting the Green's function for L^{-1} by $G(t,\tau)$, the last term is rewritten as $\int G(t,\tau)(\partial^2/\partial t^2)\alpha(\tau)y(\tau)d\tau$, i.e., the random operator R is $-(\partial^2/\partial t^2)\alpha(t)$.

After integrating twice by parts we can write

$$y(t) = F(t) + \int [\partial^2 G(t,\tau)/\partial \tau^2]\alpha(\tau)y(\tau)d\tau$$

if quantities $G(t,\tau)\frac{\partial}{\partial \tau}\alpha(\tau)y(\tau)$ and $[\partial G(t,\tau)/\partial \tau]\alpha(\tau)y(\tau)$ vanish as $t \to +\infty$ which we suppose does happen either because of the initial conditions (G and G' zero) or because α is a reducible-to-stationary stochastic process.

We write $\Gamma(t,\tau) = \sum_{m=0}^{\infty} (-1)^m K_{m+1}(t,\tau)$ with $K_1 = K$ as before,

$$K(t,\tau) = \frac{\partial^2 G(t,\tau)}{\partial \tau^2}\alpha(\tau)$$

$$K_2(t,\tau) = K(t,\tau_1)K(\tau_1,\tau)d\tau_1$$

$$= \int \frac{\partial^2 G(t,\tau_1)}{\partial \tau_1^2}\alpha(\tau_1)\frac{\partial^2 G(\tau_1,\tau)}{\partial \tau^2}\alpha(\tau)d\tau_1$$

*This work has been supported by the National Aeronautics and Space Administration (NGR11-003-020).

$$K_3(t,\tau) = \int K(t,\tau_1)K_2(\tau_1,\tau)d\tau_1, \text{ etc.}$$

$$\Gamma(t,\tau) = K(t,\tau) - K_2(t,\tau) + K_3(t,\tau) \ldots$$

$$= K(t,\tau) - \int K(t,\tau_1)K(\tau_1,\tau)d\tau_1 + \iint K(t,\tau_1)K(\tau_1,\tau_2) \cdot$$

$$\cdot K(\tau_2,\tau)d\tau_1 d\tau_2 - \cdots$$

$$= \frac{\partial^2 G(t,\tau)}{\partial\tau^2}\ \alpha(\tau) - \int \frac{\partial^2 G(t,\tau_1)}{\partial\tau_1^2}\ \frac{\partial^2 G(\tau_1,\tau)}{\partial\tau^2}\ \alpha(\tau_1)\alpha(\tau)d\tau_1 +$$

$$+ \iint \frac{\partial^2 G(t,\tau_1)}{\partial\tau_1^2}\ \frac{\partial^2 G(\tau_1,\tau_2)}{\partial\tau_2^2}\ \frac{\partial^2 G(\tau_2,\tau)}{\partial\tau^2}\ \alpha(\tau_1)\alpha(\tau_2)\alpha(\tau)d\tau_1 d\tau_2 - \cdots$$

Thus we can determine the s.g.f. (stochastic Green's function) either for the spectral density s.m. (statistical measure) if it exists, or immediately the more general two point correlation (and mutual coherence functions) thus $R_y(t_1,t_2) = \iint G_H(t_1,t_2,\sigma_1,\sigma_2)R_x(\sigma_1,\sigma_2)d\sigma_1 d\sigma_2$ where G_H is found from $h(t,\tau)$, the random Green's function.

The first term of G_H (which we do not write out) shows the results for waves propagating in a deterministic medium. The other terms of G_H involving statistics of Γ show the effects of spectral spreading due to the stochastic medium. These are the terms lost by a monochromatic assumption. The calculation for a specific case presents considerable difficulty but can be made knowing the statistics (i.e., s.m.) of α (such as correlation if α is gaussian).

In the general nonstationary case, we make the time domain iterative treatment, and if we assume gaussian behavior for the index of refraction, we observe the odd terms vanish in the series (terms involving products of odd numbers of α's) and the even terms are negative. Thus in forming products $y(t_1)y(t_2)$ for correlations, the contribution of the spectral spreading or non-monochromatic terms of G_H (i.e., the last three of the four term expression) are all positive.

Our procedure involves no assumption of statistical independence of the solution s.p. or wave function and the stochastic index of refraction and makes no closure approximations.

The first application of this work was the processing of a signal by a "stochastic filter" which randomly sampled the signal

at intervals of time governed by a probability law. Work on optimization of stochastic systems and numerous other applications is immediately suggested.

References

(1) Adomian, G., Linear Random Operator Equations in Mathematical Physics I, Jour. of Mathematical Physics, Vol. 11, no. 3, March 1970, pp. 1069-1084; Part II, same journal, Sept. 1971; Part III, same journal, Oct. 1971.

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 theory which gives necessary and sufficient conditions in a form useful for applications.

TABLE OF CONTENTS

SECTION 1

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

$(\Omega, \mathcal{F}, \mu)$ will be a fixed (but otherwise arbitrary) probability space with points $\omega \in \Omega$, a σ-algebra $\mathcal{F}$ 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 $\{\omega : x(\omega) < \lambda\} \in \mathcal{F}$ 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>, $X = \{x_t(\omega) : \cdot\cdot \longrightarrow R\}_{t \in T}$. 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 measurable processes, i.e., functions $X(t,\omega)$ which are measurable with respect to the σ-algebra $\mathcal{A} \times \mathcal{T}$ where $\mathcal{A}$ 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 $\mathfrak{A}$ 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 second order if each r.v. x_t, $t \in T$, is a member of $L^2(\mathfrak{A})$. 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 covariance function 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 continuous in the mean square sense 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 stochastic transformation (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 second order 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.

Definition: 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 power series centered at 0 with coefficients in B.

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.

Definition. Let E and F be two (real) Banach spaces. A map $P: E \times F \longrightarrow E$ is called a product 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.

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

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.

Proof. 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 time-dependent

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

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

$$= -\sum_{j=0}^{n}\binom{n}{j}A_j\ p^{f}_{n-j} = p^{f}_{n+1}. \quad \text{Hence } F_n - \sum_{j=0}^{n}\sum_{k=0}^{n-j-1}\binom{n}{j}A_j\ k^{f}_{n-j}\ F_k =$$

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

As a consequence of the above relationships, we have

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

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

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 $\operatorname{ess}_\omega\sup\ |A_n(\omega)| \leq K^n$ 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 $\operatorname{ess}_\omega\sup\ |A(s,\omega)| < K' < \infty$ 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.

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

Proof: 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

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

Hence $\|A\| \leq \text{ess}_\omega \sup |A(\omega)|$. We show now that $\|A\| \geq$ ess sup $|A(\omega)|$. Let $\epsilon > 0$ be given. Let $K =$ 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 $\Sigma(\|A_n\|\,|t^n|/n!) =$

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

Hence $A(t) = \Sigma(A_n t^n/n!)$ is analytic in $L(L^2(\mathfrak{N}), L^2(\Omega))$. Form the function $y(t) = -\int_0^t A(s)ds$ where integration takes place in $L(L^2(\mathfrak{N}), L^2(\mathfrak{N}))$, 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 $X_0 \exp\left(-\int_0^t A(s)ds\right)$ is an analytic map from I into $L^2(\mathfrak{N})$.

Now consider the expression:

$\exp\left(-\int_0^t A(s)ds\right) \int_0^t \exp\left(\int_0^y A(s)ds\right) F(y)dy$ where the indicated integrations involving the exponentials take place in $L(L^2(\mathfrak{N}), L^2(\Omega))$ and the remaining integration takes place in $L^2(\Omega)$. Since F(y) is analytic (in $L^2(\mathfrak{N})$), $\exp\left(\int_0^y A(s)ds\right) F(y)$ is analytic (in $L^2(\mathfrak{N})$) as before, and consequently so is its integral $\int_0^t$; we apply once more the operator $\exp\left(-\int_0^t A(s)ds\right)$, so the whole expression defines an analytic function in $L^2(\mathfrak{N})$. 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.

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

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.

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

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 $-c_{2k+1} = c_{2k+1}$, i.e., $c_{2k+1} = 0$. 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.

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

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

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

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

Proof: 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 $c_{2m+1} = 0$ $m = 0, 1, \ldots$, and $c_{2m} = \langle X_{2m}X_0\rangle$ $m = 0, 1, \ldots$. 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.

Example 1. 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

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

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

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

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

i.e., 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 $X_n = x_n X_0 + \sum_{k=0}^{n-1} {}_kf_n F_k =$

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

where we define C_n to be the expression in the brackets. Remember that both x_n and ${}_kf_n$ 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 $\langle X_0 X_2 \rangle = -\langle X_1^2 \rangle$, and this condition becomes $\langle A^2 B^2 \rangle = 0$ and hence we know that any polynomial in A is orthogonal to any polynomial in B. Thus we have $\langle X_n X_m \rangle =$ $= \langle BC_n BC_m \rangle = \langle B^2 C_n C_m \rangle = 0$ for $n + m \geq 1$ and consequently Theorem 8 is satisfied. Thus the solution is stationary if and only if $\langle A^2 B^2 \rangle = 0$. 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 another, and so $\langle B^n\rangle = \langle B^{n-1}\rangle\langle X_1^2\rangle/\langle X_0^2\rangle$. Let $c = \langle X_1^2\rangle/\langle X_0^2\rangle$, 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 orthogonal random measure 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 absolute measure associated with ξ. 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 $L^2(\Omega)$ 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 $L^2(\Omega)$ 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 spectral random measure 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)
$$X'(t) + aX(t) = F(t)$$
$$X(0) = X_0$$

where a is an essentially bounded r.v., $X_0 \in L^2(\mathbb{P})$, 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

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

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

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

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

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 $\iint_{IJ} f(x,y,\omega)\xi(dx)dy$ and $\iint_{JI} f(x,y,\omega)dy\xi(dx)$ 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

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

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

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

where M is the absolute spectral measure associated with ξ. Now consider the set $\mathcal{S}$ of all functions $g(x,\omega) = \Sigma C(A_k,x)g_k(\omega)$ 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 $\|g\|_{\mathcal{S}} = (\int \|g(x)\|^2 M(dx))^{1/2}$, 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 $\int \|f(x,y)\|^2 M(dy) < \infty$. 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 $\int\int_{JI} f(x,y)\xi(dy)dx$ exists. Now we want to

insure that $\int\int_{IJ} f(x,y)dx\xi(dy)$ 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 $E(|h(y)|^2) = E\int\int_{JJ} f(s,y)\overline{f(t,y)}ds\ dt = \int\int_{JJ} E(f(s,y)\overline{f(t,y)})ds\ dt$, so we require that f satisfy

$$\int\int\int_{IJJ} E(f(s,y)\overline{f(t,y)})ds\ dt\ M(dy) < \infty. \tag{5}$$

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 $g \in \mathcal{S}$ and there is a subset $H = H(g) \subset L^2(\mathfrak{H})$ whose finite linear combinations are dense in $L^2(\mathfrak{H})$ (call such a set linearly dense) 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 $g: R \longrightarrow \mathcal{S}$ defined

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

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

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

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

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:

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

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

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

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

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 $X'(t) + X(t) = f$, $f \in L^2(\mathbb{R})$. $F(t) = f$, 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 $\mathcal{I}$.

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 spectral density of the process F, i.e., 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

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

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

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

$$X(t,\omega) = \int H(t,u)Y(u,\omega)du \tag{2}$$

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

$$X(t,\omega) = \int H(t,u,\omega)Y(u,\omega)du \tag{3}$$

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

$$X(t,\omega) = aY(t,\omega) + bZ(t,\omega) \tag{4}$$

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 cylindrical set. 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

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

for all k, $t_1,\ldots,t_k$ in (a,b) and all Borel sets $A_1,\ldots,A_k$ is called the measure in the function space corresponding to the process X(t). 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.

PROPOSED RESEARCH ON NUMERICAL METHODS FOR STOCHASTIC DIFFERENTIAL EQUATIONS

Introduction

Stochastic differential equations involving stochastic processes in the initial conditions, forcing functions, and even in the coefficients of the differential operator have been extensively studied, (see, e.g., ref. 11), the latter particularly by Adomian (see ref., 2,3,4,5). As is the case with deterministic differential equations, numerical methods often become necessary for the solution of specific problems so it is of interest to develop numerical methods which insure correct statistic to first or second order. As an initial example we will consider the first order* stochastic differential equation

$$y'(t) + \xi(t)\, y(t) = \psi(t) \tag{1}$$

where $\xi(t)$ and $\psi(t)$ are uncorrelated stochastic processes. This equation has been investigated by Adomian and Tikhonov (see ref. 2 in chapter 4) and by Aström and others more recently. In Adomian's work theoretical expressions were derived for the covariance of the solution process $y(t)$. We now consider methods to calculate $y(t)$ and $\mathrm{Cov}_y(t,t')$ numerically. Two methods have been developed and computer programs have been written to implement the various algorithms. The first method uses quadrature to evaluate the appropriate

---

*We emphasize that the method is not for the solution of first order equations but is to be extended to partial differential equations and nonlinear stochastic equations. The preliminary work is for checking against known results.

integral of the stochastic Green's function, (ref. 4,5). The second method simulates realizations of $\xi(t)$ and $\psi(t)$ from which the corresponding realization of $y(t)$ is then computed. In this manner an ensemble of realizations of $y(t)$ is calculated and statistically evaluated to approximate $\overline{y(t)}$ and $Cov_y(t,t')$. The development of both methods has been for Gaussian $\xi(t)$ and $\psi(t)$. There has been no indication that either method would not be useful if the stochastic processes are not taken to be Gaussian.

The Quadrature Method

The solution to equation (1) can be written

$$y(t) = \int_a^t e^{-\int_u^t \xi(v)dv} \psi(u)du + y_0 e^{-\int_a^t \xi(v)dv} \tag{2}$$

where y_0 is the initial condition random variable given by $y(t_0) = y(a)$. The first moment, $\overline{y(t)}$, is found by taking the expected value of (2) resulting in

$$E\{y(t)\} = \int_a^t E\{e^{-\int_u^t \xi(v)dv}\}\{E\{\psi(u)\}du + E\{y_0\}E\{e^{-\int_a^t \xi(v)dv}\} \tag{3}$$

where it is assumed that y_0 and $\xi(t)$ are independent and that $\xi(t)$ and $\psi(t)$ are independent. With the same assumptions the product $y(t)\ y(t')$ can be calculated and correlation functions $E\{y(t)\ y(t')\}$ can be written

$$E\{y(t)\ y(t')\} = \int_a^{t'}\int_a^{t} E\{e^{-\int_u^t \xi(x)dx\ -\int_v^{t'} \xi(x)dx}\}E\{\psi(u)\ \psi(v)\}dudv \quad (4)$$

$$+E\{y_0\}\int_a^{t'} E\{e^{-\int_a^t \xi(x)dx\ -\int_v^{t'} \xi(x)dx}\}E\{\psi(v)\}dv$$

$$+E\{y_0\}\int_a^{t} E\{e^{-\int_u^t \xi(x)dx\ -\int_a^{t'} \xi(x)dx}\}E\{\psi(u)\}du$$

$$+E\{y_0^2\}\ E\{c^{-\int_a^t \xi(x)dx\ -\int_a^{t'} \xi(x)dx}\}$$

For a Gaussian random variable z,

$$E\{e^{-z}\} = e^{-m_z + 1/2\ \sigma_z^2} \quad (5)$$

where m_z is the mean of z and σ_z^2 is the variance of z.
Defining the function

$$\hat{z}(t, t', u, v) = m_z + 1/2\ \sigma_z^2 \quad (6)$$

where $z = \int_u^t \xi(x)dx + \int_v^{t'} \xi(x)dx$ equation, (3) can be expressed

$$E\{y(t)\} = \int_a^t c^{\hat{z}(t,0,u,0)}E\{\psi(u)\}du + \bar{y}_0\ e^{\hat{z}(t,0,a,0)} \quad (7)$$

and (4) can be written

$$E\{y(t)\ y(t')\} = \int_a^{t'} \int_a^{t} e^{\hat{2}(t,t',u,v)} E\{\psi(u)\ \psi(v)\}dudv \qquad (8)$$

$$+\bar{y}_0 \left\{ \int_a^{t'} e^{\hat{2}(t,t',a,v)}\ E\{\psi(v)\}dv + \int_a^{t} e^{\hat{2}(t,t',u,a)} E\{\psi(u)\}du \right\}$$

$$+\bar{y}_0^2\ e^{\hat{2}(t,t',a,a)}$$

The covariances can be computed from (7) and (8)

$$Cov_y(t,t') = E\{y(t)\ y(t')\} - E\{y(t)\}\ E\{y(t')\} \qquad (9)$$

Inspection of (7) and (8) show that double and single integration must be performed if these equations are to be evaluated. Moreover, a vector of values, $\{\overline{y(t_1)}, \overline{y(t_2)}, \ldots, \overline{y(t_n)}\}$ must be evaluated to numerically represent $\overline{y(t)}$ and a matrix of values must be evaluated, $Cov_y(t_i, t_j)$ to represent $Cov_y(t,t')$. The amount of calculation required to evaluate this matrix and vector in any given case, made it necessary to give special attention to one and two dimensional quadrature. Gaussian quadrature, while more powerful in many respects, lacks one feature of Newton-Cotes quadrature that degrades its sophistication. With Newton-Cotes quadrature, but not with Gaussian quadrature, the integrand may be computed at new points while retaining the old values and then all values may be used to re-evaluate the integral. Moreover, this can

be accomplished without large core requirements for the computer. The only limiting factor on the particular Newton-Cotes formula (closed type) that may be used is the word size of the computer being used.

[1] gives the 2-point through the 11-point Newton-Cotes formula (closed type). A program was written for the FORMAC preprocessor for the PL1 language that will derive this type formula. This program has been used to compute the 12-point through the 32-point formula. Higher formula have not been derived because formula higher than the 29-point formula have weights that can not be expressed exactly on any machine available. The 3-point rule (Simpson's rule) through the 29-point formula have been written into a 3 subroutine program that permits the user to choose any formula desired, 3-point through 29-point. Successive refinements of the interval of integration permit a predetermined accuracy to be selected. It must be noted that there is no mathematical guarantee that the selected accuracy will be achieved. Experience has shown that the convergence criterion employed is reliable for a large class of functions to either achieve the accuracy requested or to note failure. The subroutine package, named INTGL, will evaluate integrals of the form

$$I = \int_a^b f(x)\, dx \tag{10}$$

and

$$I = \int_a^b [\Phi_3(y) \int_{\Phi_1(y)}^{\Phi_2(y)} f(x,y)dx]dy \tag{11}$$

INTGL was modified to use particular properties of (8) and (9) to decrease computation further.

Program SDEQUAD was written to evaluate an approximating vector to $\overline{y(t)}$ and an approximating matrix to $Cov_y(t,t')$. Quadrature is achieved by the modified INTGL package. The user of SDEQUAD must provide the interval of integration, initial mean and 2nd moment of y, the mean and covariance functions of ξ and ψ and the function $\hat{z}$.

Advantages of this method include accuracy of the results and computation time requirements as compared to the simulation method. It is anticipated that this method can be applied to stochastic differential equations of higher degree provided that the stochastic Green's function can be calculated for the appropriate statistical measure to be evaluated, (see ref. 2-5). Additional programming is necessary to accomplish this, however, and quadrature may become overly time consuming for given integrals.

Three areas present themselves for further investigation. First, SDEQUAD may be used as a tool in its present form for analysis of specific cases of equation (1). The function $\hat{z}$, may be evaluated for ξ that are not Gaussian. Second, a modification of SDEQUAD should be developed for a suitable class of 2 degree stochastic differential equations. According to references 2 and 5, stochastic Sturm-Liouville systems are subject to this approach. Consequently, a large class of problems of applied mathematics, engineering and sciences can be analyzed by stochastic methods. Third, partial differential

equations may also be approached by transformation into an integral equation. Equations of this type include meteorological models involving a system of partial differential equations from the mechanics of continua (ref. 8). This system, subject to stochastic analysis, should yield significant improvement in weather prediction.

The Simulation Method

Equation (1) can also be written

$$y'(t) = \psi(t) - \xi(t)\, y(t) \tag{12}$$

and this is the form in which Adomian does his iteration. Deterministic equations that are represented in this form may be solved numerically by Runge-Kutta methods [1]. By letting ω represent a realization event in the probability space and writing

$$y'_\omega(t) = \psi_\omega(t) - \xi_\omega(t)\, y_\omega(t) \tag{13}$$

the realization $y_\omega(t)$ can be solved numerically provided that the realizations $\xi_\omega(t)$ and $\psi_\omega(t)$ are known. The first task in developing a simulation of (2) was to simulate realization of a stochastic process χ in terms of its statistical measures. In general, this depends upon the evaluation of the conditional distribution of

$$\chi(t_i) \,|\, \chi(t_1) = x_1, \quad \chi(t_2) = x_2, \ \ldots, \ \chi(t_{i-1}) = x_{i-1} \tag{14}$$

If $X_1 = \chi(t_i)$, $i = 1, 2, \ldots, n$ and $f_\chi(X_1, X_2, \ldots, X_i)$ is the joint distribution density function of $\{X_1, X_2, \ldots, X_i\}$ then the conditional distribution density function can be written

$$f_\chi(X_i | X_1 = x_1, \quad X_2 = x_2, \ldots, \quad X_{i-1} = x_{i-1}) = \tag{15}$$

$$\frac{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_{i-1}, X_i, \ldots, X_n) dX_{i+1}, \ldots, dX_n}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_{i-1}, X_i, \ldots, X_n) dX_i, \ldots, dX_n}$$

If $\chi(t)$ is a Gaussian stochastic process, then

$$f(X, \ldots, X_n) = \frac{e^{-1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(X_i - a_i)(X_j - a_j)}}{(2\Pi)^{\frac{n}{2}} (|M|)^{1/2}} \tag{16}$$

where M is a n x n matrix such that

$$(M)_{ij} = \mathrm{Cov}_\chi(t_i, t_j)$$

$$w_{ij} = (M^{-1})_{ij}$$

and $$a_i = \overline{\chi(t_i)}$$

Let M_i denote the upper left i x i partition of M

Let R_i denote the right-hand column of M_i^{-1}; R_i is the column vector whose elements are $r_1, r_2, \ldots r_{i-1}, r_i$.

Let $P_i^{-1} = r_i$ so that $P_i = 1/r_i$.

Equation (13) can now be written

$$f_\chi(X_i | X_1 = x_1, \ \ldots, \ X_{i-1} = x_{i-1}) = (2\Pi P_1)^{-1/2} \exp\{-(X_1 - b_1)^2/2P_i\} \tag{17}$$

where $b_i = a_i - P_i \sum_{j=1}^{i-1} r_j(x_j - a_j)$

The value b_i is the conditional expectation of X_i and P_i is the conditional variance of X_i. The iteration to simulate a relationship of $\chi(t)$ is defined by $x_i = P_i\mu_i + b_i$ where $\{\mu_1, \mu_2, \ \ldots, \ \mu_n\}$ are independently chosen values of a Gaussian random variable with zero mean and unit variance. The iteration is started by $P_1 = [M]_{11}$ and $b_1 = a_1$. An ensemble of realization of χ generated in this manner approximate χ in that the means and covariances of the realizations approximate $\overline{\chi(t_i)}$ and $\mathrm{Cov}_\chi(t_i, t_j)$, $1 \leq i, j \leq n$.

With the ability to generate realizations of a stochastic process, our ability to solve stochastic differential equations is limited only by our ability to solve deterministic differential equations. Linear equations of higher degree as well as non-linear equations are subject to the simulation approach. The class of stochastic differential equations that can be solved by this method is larger than the class that can be solved by the quadrature method. Two disadvantages of simulation in general also apply to stochastic simulation. Accuracy is limited. It is difficult to achieve more than 2 digits accuracy Simulation is also by nature time consuming. In many cases,

however, either method will yield results or the accuracy that can be reasonably achieved by a simulation is acceptable.

Three areas for further investigation were mentioned with respect to the quadrature method. The same may be said with respect to simulation.

REFERENCES

1. Abramowitz, M. and Stegun, D. A. (eds.). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Washington, D. C.: U. S. Government Printing Office, National Bureau of Standards Applied Mathematics Series, 1964.

2. Adomian, G. "Linear Stochastic Operators," Ph.D. Dissertation. University of California (Los Angeles), 1961.

3. _____, "Linear Stochastic Operators," *Review of Modern Physics*, vol. 35, (1963), pp. 185-207. (Not the same as ref. 2).

4. _____, "Stochastic Green's Functions," in ref. 5, pp. 1-39.

5. _____, "Linear Random Operator Equations in Mathematical Physics," *Journal of Mathematical Physics*, vol. 11, no. 3, (March 1970), pp. 1069-1084.

6. Bellman, R. (ed.). *Stochastic Processes in Mathematical Physics and Engineering. Proceedings of Symposium in Applied Mathematics*, vol. 16. Providence, Rhode Island: American Mathematical Society, 1964.

7. Bharucha-Reid, A. T., "On the Theory of Random Equations," in ref. 5, pp. 40-69.

8. Kibel, D. A. *An Introduction to Hydrodynamical Methods of Short Period Weather Forecasting*. New York, N. Y.: The MacMillan Company, 1963.

9. Miller, Kenneth S. *Multidimensional Gaussian Distributions*. New York, N. Y.: John Wiley and Sons, Inc. 1964.

10. Saatz, Thomas L. (ed.). *Modern Nonlinear Equations*. New York, N. Y.: McGraw-Hill Book Company, 1967.

11. Syski, R., "Stochastic Differential Equations," in ref. 9, pp. 346-456.