LECTURE NOTES 67-2

W. M. WONHAM

LECTURE NOTES ON STOCHASTIC CONTROL

FEBRUARY, 1967

N67-26371
(ACCESSION NUMBER) (THRU)
240 1
(PAGES) (CODE)
CR 84074 10
(NASA CR OR TMX OR AD NUMBER) (CATEGORY)
FACILITY FORM 602



CENTER FOR DYNAMICAL SYSTEMS

VISION OF APPLIED MATHEMATICS BROWN UNIVERSITY

DRF-00765

CENTER FOR DYNAMICAL SYSTEMS, BROWN UNIVERSITY

ERRATA FOR: LECTURE NOTES 67-2

(LECTURE NOTES ON STOCHASTIC CONTROL BY W. M. WONHAM)

(* means count from bottom)

P. 25, *l.* 4: Insert { after $\mu_j(\alpha^0)$.

P. 26, *l.* 6*: For $\min_{\alpha}$ read $\min_{\alpha} w$

P. 49, *l.* 4*: (following $\leq$) For $\mathcal{L}(\alpha)$ read $\mathcal{L}(\alpha^*)$.

P. 63, *l.* 7: For V_n read V_N.

P. 64, *l.* 2*: For $-L_m^*$ read $-L_m(\alpha)$

P. 79, *l.* 9: For "at least" read "at most"

P. 105, *ll.* 2-5: incl. should be deleted

P. 111. Eq.(2): Read $x'M(t)x$

P. 121. Section 7: Add the remarks:

(1) For existence of $\hat{P}(\infty)$ it is enough to assume that (A,C) is stabilizable: i.e. there exists a constant matrix $\hat{K}$ such that $A-C\hat{K}$ is stable. The following criterion is sometimes useful. Let the minimal polynomial $\varphi(\lambda)$ of A be factored in the form

$$\varphi(\lambda) = \varphi_1(\lambda)\varphi_2(\lambda)$$

where all zeros of $\varphi_1(\lambda)$ have nonnegative real parts and all zeros of $\varphi_2(\lambda)$ have negative real parts. Define

$$E_1 = \{x : \varphi_1(A)x = 0\} .$$

Then (A,C) is stabilizable if and only if the subspace E_1 of

"unstable modes" is controllable; i.e.

$$E_1 \subset \text{Range of } [C, AC, \ldots, A^{n-1}C]$$

(2) If in addition $M > 0$, then $P_\infty > 0$, and the limiting system matrix $A-CK_\infty$ is stable. Then (9) implies that P_∞ has the <u>minimum</u> property:

$$P_\infty = \int_0^\infty e^{t(A-CK_\infty)'}(M+K_\infty' N K_\infty)e^{t(A-CK_\infty)}dt$$

$$\leq \int_0^\infty e^{t(A-CK)'}(M+K'NK)e^{t(A-CK)}dt$$

for all K such that $A-CK$ is stable.

P. 123, Eq. (9b): Read $K = N^{-1}C'P$.

PP.136-146: The proof given here contains a gap: a complete proof can be found in the writer's CDS TR 67-2, "On Pole Assignment in Multi-Input Controllable Linear Systems".

P. 160: The introduction of $\tilde{A}$ etc. is unnecessary. It is sufficient to write $f_n(t) = (t_{j+1}-t_j)^{-1}\int_{t_j}^{t_{j+1}} f(s)ds$ $(t_j \leq t < t_{j+1})$, and then use the relation $f_n(t) \to f(t)$ a.e. $(n \to \infty)$ together with dominated convergence.

P. 170: To the hypotheses of the limit theorem, add: $M > 0$.
Add remark: observability of (C,A) can be weakened to stabilizability of (A',C').

PP.171-173: Delete material between sentence "In the following ..." ($\ell.3^*$, p.171) and "Q.E.D." (p.173). The deleted material is replaced by:

It follows that, for every $\epsilon > 0$, there exists $\hat{T}$ such that

$$P(t) \leq \int_0^\infty e^{s(\tilde{A}-\tilde{C}\hat{K})'}(M+\hat{K}'N\hat{K})e^{s(\tilde{A}-\tilde{C}\hat{K})'}ds + \epsilon I \qquad (8)$$

for all $t \geq \hat{T}$. Denote the integral on the right side of (8) by $J(\hat{K})$. Since $M > 0$, we know from §2.7 that there exists K^o $(= K_\infty)$ with the property

$$J(K^o) \leq J(\hat{K})$$

for all $\hat{K}$; in particular setting $\hat{K} = K^o$ in (8) we have

$$P(t) \leq J(K^o) + \epsilon I \qquad (t \geq T^o).$$

On the other hand, if Φ is the fundamental matrix associated with $\tilde{A}-\tilde{C}K(t)$, then by (4) and (5),

$$\begin{aligned} P(t) &= \Phi(t,0)'P_o\Phi(t,0) + \int_0^t \Phi(s,0)'[M+K(s)'NK(s)]\Phi(s,0)ds \\ &\geq \int_0^t \Phi(s,0)'[M+K(s)'NK(s)]\Phi(s,0)ds \\ &\geq P^*(t), \end{aligned}$$

where $P^*(t)$ is the solution of (4) and (6) corresponding to the special initial condition $P^*(0) = 0$. Here the last inequality holds because P^* is the minimal value of the last written integral with respect to all

possible choices of $K(\cdot)$ (and corresponding Φ). Using the fact (§2.7) that $P^*(\cdot)$ is monotone, we conclude that for some T^*,

$$P(t) \geqq P^*(\infty) - \epsilon I \qquad (t \geqq T^*).$$

In the notation of §2.7,

$$P^*(\infty) = P^*_\infty = \int_0^\infty e^{t(\tilde{A}-\tilde{C}K_\infty)}(M+K_\infty' N K_\infty) e^{t(\tilde{A}-\tilde{C}K_\infty)} dt$$

$$= J(K_\infty).$$

Since $K_\infty = K^o$, we have for $t \geqq \max(T^o, T^*)$,

$$J(K^o) - \epsilon I \leqq P(t) \leqq J(K^o) + \epsilon I .$$

Since ϵ is arbitrary, it follows that

$$\lim_{t \to \infty} P(t) = J(K^o) .$$

Q.E.D.

PP. 175-177: Exercise 3 contains an error and should be deleted.

P. 175, *ll.* 1-3: The proof of this inequality was mistakenly omitted from the Appendix, and will be given elsewhere. The proof is accomplished by a reduction to the case $C = I$, A in Jordan form, and then by considering the blocks of e^{tA}.

P. 185, ℓ. 4: For the inequality read

$$\frac{k}{n}(b-a) \leqq t-a < \frac{(k+1)}{n}(b-a) , \qquad k = 0,1,\ldots,n-1.$$

P. 187, ℓ.6: Read $\int_a^b |f(t)dw(t)|$

P. 193, ℓ.5: Read §2.1

P. 195, ℓ.2*: For dxdw read dxdt

P. 205, ℓ.1*: For $0 \leqq t_1$ read $t_o \leqq t_1$

P. 206, ℓ.6*: Read "property 2"

P. 211: For "stochastic" read "statistic"

P. 218, ℓ.6*: Read $\int_t^T \mathcal{E}\{L$

P. 222, ℓ.9: For F read G .

LECTURE NOTES ON STOCHASTIC CONTROL*

PART I

W. M. Wonham

Center for Dynamical Systems
Division of Applied Mathematics

Brown University

*This research was supported in part by National Aeronautics and Space Administration under Grant No. NGR-40-002-015, in part by the Air Force Office of Scientific Research, Office of Aerospace Research United States Air Force, under AFOSR Grant No. AF-AFOSR-693-66 and in part by National Science Foundation under Grant No. GK-967.

LECTURE NOTES ON STOCHASTIC CONTROL

PART I

PREFACE

The present set of notes is a partial compilation of material presented by the writer in graduate courses at M.I.T. (E.E.Dept, Spring, 1966) and Brown University (Applied Mathematics 207, Fall, 1966). Part I is devoted to Markovian optimization problems for Markov chains on a finite state space and for diffusion processes of Gaussian type (that is, linear systems subject to Gaussian white noise perturbations). Part II, to appear later, is concerned with stochastic stability of continuous Markov processes, and with problems of optimization of nonlinear systems. It is hoped that the notes may be useful to students of stochastic control theory.

W. M. W.
January 31, 1967

LECTURE NOTES ON STOCHASTIC CONTROL

PART I

CHAPTER I. CONTROL THEORY FOR FINITE MARKOV CHAINS. 1
1. Finite Markov Chains 3
2. Stationary Control of Ergodic Chains. 15
3. Intercept Problems. 41
4. Optimal Control with Fixed Terminal Time. 61

References 71
Appendix 73

CHAPTER II. INTRODUCTION TO GAUSSIAN DIFFUSIONS 78
1. Preliminaries 78
2. Brownian Motion 86
3. A Class of Stochastic Integrals 95
4. Gaussian Diffusions 101
5. Quadratic Functionals 107
6. Optimal Control with Quadratic Cost 111
7. Asymptotic Properties of the Optimal Control. 121
8. Open Loop Control Versus Feedback Control 124
9. Bibliographical Note and References 132

References 133
Appendix 1 135
Appendix 2 136

CHAPTER III. OPTIMAL LINEAR FILTERING AND PREDICTION. 147
1. Introduction. 147
2. Derivation of the Filter Equations. 149
3. Qualitative Properties of the Filter. 164

References 178
Appendix. 179

CHAPTER IV. ITÔ'S STOCHASTIC INTEGRAL AND APPLICATIONS. 183
1. Stochastic Integral 183
2. Integration Rule. 203
3. Application to the Linear Regulator 206
4. Application to the Linear Regulator with Noisy Data 210
5. Itô's Stochastic Differential Equation. 222

References 236

CHAPTER I

CONTROL THEORY FOR FINITE MARKOV CHAINS

INTRODUCTION

The structure of stochastic control problems of Markovian type is most simply revealed by considering Markov chains on a finite state space. Here the plant or controlled object is represented by a family of transition matrices labelled by a parameter which represents the control law. Under simple regularity conditions, approximation in policy space leads to monotone convergence of the cost 'potentials' to the optimum, and thus to determination of an optimal or ϵ-optimal control.

In this chapter we consider three of the simplest control problems for chains: optimal stationary control, optimal interception and pursuit, and optimal control on a finite interval fixed in advance. In all cases it is assumed that the controller has complete information about the structure and the current state of the chain; and in all but the third problem we assume a priori that the control law is a function only of current state. These assumptions can be (and have been) relaxed; further information will be found in the references listed at the end of the chapter. In particular we note that our basic formulation is due to Bellman (1957, 1961); the treatment of stationary processes, in essence, follows Howard (1960), although the convergence theorems given here may be novel. Sufficient classes of strategies (control laws) are

discussed by Derman (1962-64), and in greater detail by Dynkin (1965). The inclusion of pursuit problems was suggested by the work of Eaton and Zadeh (1962); the present treatment seems to be more straightforward. Further rigorous discussion of ϵ-optimal controls can be found in Krylov (1965) and Romanovskii (1965). Finally, extension in the direction of adaptive decision making (although not control) is the subject of an interesting paper by Raviv (1965).

1. FINITE MARKOV CHAINS.

1.1 Notation.

Ω - abstract set of elements ω.

$\mathfrak{F}$ - σ-algebra of subsets of Ω : i.e.

(i) $\Omega \in \mathfrak{F}$

(ii) $A \in \mathfrak{F}$, $B \in \mathfrak{F} \Longrightarrow A-B \in \mathfrak{F}$

(iii) $A_n \in \mathfrak{F} (n = 1,2,\ldots) \Longrightarrow \overset{\infty}{\underset{1}{\cup}} A_n \in \mathfrak{F}$

Thus

$\phi \in \mathfrak{F}$ (ϕ = null set)

$A \in \mathfrak{F} \Longrightarrow A' = \Omega - A \in \mathfrak{F}$

$A,B \in \mathfrak{F} \Longrightarrow A \triangle B = (A-B) \cup (B-A) \in \mathfrak{F}$

$A,B \in \mathfrak{F} \Longrightarrow A \cap B = (A \cup B)-(A \triangle B) \in \mathfrak{F}$

Thus $\mathfrak{F}$ is closed under countable numbers of set-theoretic operations.

$P : \mathfrak{F} \to [0,1]$ probability measure

(i) $P(\phi) = 0, \quad P(\Omega) = 1$

(ii) $A \in \mathfrak{F} \Longrightarrow 0 \leqq P(A) \leqq 1$

(iii) $A_n \in \mathfrak{F} \Longrightarrow P(\overset{\infty}{\underset{1}{\cup}} A_n) = \overset{\infty}{\underset{1}{\sum}} P(A_n)$ if the A_n are pairwise disjoint

$(\Omega, \mathfrak{F}, P)$ - Probability triple. Elements of $\mathfrak{F}$ are <u>events</u>. Instead of $P(A) = 1$ we sometimes write A a.s. (almost surely).

Definition: $E = \{1,\ldots,J\}$. State space.

$x : \Omega \to E$.

Definition: x is a random variable if

$$(\forall_j)\ j \in E \Longrightarrow \{\omega : x(\omega) = j\} \in \mathfrak{F}.$$

Definition: A stochastic process is a sequence

$$X = \{x_n(\cdot) : n = 0,1,\ldots.\}$$

of random variables.

$x_n(\cdot)$ is the state of process at time n.

Conditional probability.

If $A,B \in \mathfrak{F}$ and $P(B) \neq 0$ we define

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

the conditional probability of A given B.

For X, let A_n be an event of form

$$A_n = \{\omega : x_n(\omega) = j_n\}\ ,\ n = 0,1,\ldots\ ;\ j_n \in E\ .$$

1.2 Chains.

Definition: X is a Markov chain if for all n, A_n,

$$P\{A_{n+1} | A_o \cap A_1 \cap \ldots \cap A_n\} = P\{A_{n+1} | A_n\}$$

whenever $P(A_o \cap \ldots \cap A_n) \neq 0$. (*)

EXERCISE: (*) $\Longrightarrow P(x_n = j | x_m = i, x_r = k) = P(x_n = j | x_r = k)$. $(m \leq r \leq n)$

Transition matrix:

Define

$$P(n) = [p_{ij}(n)]_{i,j = 1,\ldots,J} \quad (n = 0,1,\ldots)$$

where $p_{ij}(n) = P\{x_{n+1} = j | x_n = i\}$ ($\equiv 0$ if $P\{x_n = i\} = 0$).

For $n \geq m$, let $p_{ij}(m,n) = P\{x_n = j | x_m = i\}$.

Chapman-Kolmogorov equation: if $m \leq r \leq n$,

$$p_{ij}(m,n) = \sum_k p_{ik}(m,r) p_{kj}(r,n) \qquad (CK)$$

Proof

$$p_{ij}(m,n) = P(x_n = j | x_m = i)$$

$$= \sum_k P(x_n = j | x_m = i, x_r = k) \; P(x_r = k | x_m = i)$$

and (CK) follows by the Exercise above.

Write

$$P(m,n) = [p_{ij}(m,n)]$$

$$(CK) \Longrightarrow P(m,n) = P(m,r)P(r,n)$$

$$\Longrightarrow P(0,n) = P(0,n-1)P(n-1,n)$$

$$= P(0,n-1)P(n-1)$$

$$\vdots$$

$$= P(0)P(1) \ldots P(n-1) .$$

Let $p_j(n) = P(x_n = j)$

$$p(n) = [p_j(n)] \quad \text{(row vector)} .$$

Since

$$p_j(n) = \sum_i P(x_n = j \mid x_o = i)p(x_o = i)$$

we have

$$p(n) = p(0)P(0,n) .$$

Definition: X is homogeneous if $P(n) = P = \text{constant}$.

Then

$$P(0,n) = P^n$$

$$p(n) = p(0)P^n .$$

Construction of a (homogeneous) chain

Given the matrix

$$P = [p_{ij}] \quad (J \times J)$$

where

$$p_{ij} \geq 0, \quad \sum_j p_{ij} = 1 ,$$

we wish to construct a chain X with P as transition matrix.

Write $p_{ij}^{(n)} = (i,j)^{th}$ element of P^n. Let $p = (p_j)$ be the row vector of initial probabilities:

$$p_j \geq 0$$

$$\sum p_j = 1 .$$

Finally let

$$\Omega = \{\text{all sequences of form } \omega = (j_o, j_1, \ldots)\}\ (j_i \in E)$$

$$x_n(\omega) = j_n = n^{th} \text{ component of } \omega$$

$$A = \{\omega : x_o(\omega) = j_o, \ldots, x_n(\omega) = j_n\}$$

and

$\mathfrak{F}$ = smallest σ-algebra of ω-sets containing all sets of form A.

Define $P(A) = p_{j_o} p_{j_o, j_1} \cdots\cdots p_{j_{n-1}, j_n}$.

P is then determined uniquely on $\mathfrak{F}$. (Kolmogorov)

We assert that, so defined, X is a Markov chain (For the proof see: Chung(1960)§1.2, Theorem 1).

State classification.

Suppose $x_o = j$ (a.s) The state j is recurrent if $P(x_n = j \text{ for some } n \geq 1) = 1$. We define the recurrence time to state j: Let $\tau_j(\omega) = \min\{n : n \geq 1,\ x_n(\omega) = j\}$ if such an n exists

$= +\infty$, otherwise.

$$\{\omega : \tau_j(\omega) = \nu\} = \{\omega : x_1(\omega) \neq j, \ldots, x_{\nu-1}(\omega) \neq j,\ x_\nu(\omega) = j\} \in \mathfrak{F}$$

j is recurrent iff $P(\tau_j < \infty) = \sum_{\nu=1}^{\infty} P(\tau_j = \nu) = 1$. If $P(\tau_j < \infty) = 1$, then expected recurrence time $= \mathcal{E}\{\tau_j\} = \sum_{1}^{\infty} \nu P(\tau_j = \nu)$.

Definition. If $j \in E$,

j is null recurrent iff $\mathcal{E}\{\tau_j\} = \infty$;

j is positive recurrent iff $\mathcal{E}\{\tau_j\} < \infty$;

j is transient iff $P(\tau_j < \infty) < 1$.

Suppose

$$P(\tau_j(\omega) = \nu) > 0 \Longrightarrow \nu = kt \qquad (*)$$

(t an integer) for some $k = 1,2,\ldots$.

Period of j = largest integer t for which (*) is true

= g.c.d. $\{n : p_{ij}^{(n)} > 0\}$

Definition. If $j \in E$,

j is periodic if period ≥ 2

j is aperiodic if period $= 1$.

X is aperiodic if all states in E aperiodic.

j is ergodic if it is aperiodic and positive recurrent

X is ergodic if all states in E are ergodic.

State i leads to j if $p_{ij}^{(n)} > 0$ for some n. Write $(i \rightsquigarrow j)$

States i,j communicate if $i \rightsquigarrow j$ and $j \rightsquigarrow i$; in this case write $i \leftrightsquigarrow j$. (The terminology and notation are Chung's).

EXERCISE: If $i \leftrightsquigarrow i$ for all $i \in E$, then $\leftrightsquigarrow$ is an equivalence relation on E.

Definition: X is irreducible iff $i \leftrightsquigarrow j$ for all i,j.

THEOREM A finite chain is ergodic if and only if for all $i,j \in E$,

i is aperiodic and $i \rightsquigarrow j \Longrightarrow j \rightsquigarrow i$.

PROOF. Necessity is obvious. For sufficiency, note first that for every i there is a j such that $i \rightsquigarrow j$, and so for all $i, i \rightsquigarrow i$. Hence $\langle\sim\rangle$ is an equivalence relation on E. Let $\tilde{E} \subset E$ be an equivalence class under $\langle\sim\rangle$. If P is the transition matrix of X let $\tilde{P}$ be the restriction of P to $\tilde{E}$: i.e., relable states so that $\tilde{E} = (1,\ldots,\tilde{J})$ and let $\tilde{P} = [p_{ij}]$, $i,j \in \tilde{E}$.

Since $i \rightsquigarrow j \Rightarrow i \langle\sim\rangle j$, if $i \in \tilde{E}$, then

$$\tilde{E} \supset \{j : p_{ij} > 0\} ;$$

hence

$$\sum_{j=1}^{\tilde{J}} p_{ij} = 1 \quad , \qquad (i = 1,\ldots,\tilde{J}) ;$$

that is, $\tilde{P}$ is a transition matrix. It is therefore sufficient to prove the theorem for the restriction $\tilde{X}$ of X to $\tilde{E}$; thus, without loss of generality we can assume that X is irreducible.

(i) <u>For some</u> N, $p_{ij}^{(N)} > 0$ <u>for all</u> i,j.

Let $N_{ij} = \{n: p_{ij}^{(n)} > 0\}$ $(N_{ij} \neq \emptyset)$. Consider N_{ii}. Since $p_{ii}^{(n+m)} = \sum_k p_{ik}^{(n)} p_{ki}^{(m)} \geq p_{ii}^{(n)} p_{ii}^{(m)}$, it follows $n,m \in N_{ii} \Rightarrow n+m \in N_{ii}$; that is, N_{ii} is closed under addition.

Since i is aperiodic we have

$$1 = \text{g.c.d. } \{n : n \in N\},$$

and so there exists finite $\tilde{N} \subset N$ with

$$1 = \text{g.c.d.}\ \{n:\ n \in \tilde{N}\}\ .$$

From elementary number theory

$$1 = a_1 n_1 + \ldots + a_k n_k\ ,\quad \text{some}\ \ n's \in \tilde{N}$$

where the $a's$ are integers. That is, there exist $m,n \in N$ such that $m-n = 1$.

For $k \geq n(n-1)$ we can write

$$k = an + b \qquad \begin{matrix} a \geq n-1 \\ 0 \leq b \leq n-1 \end{matrix}$$

$$= (a-b)n + bm\ ;$$

thus $k \in N_{ii}$, and so N_{ii} contains all sufficiently large integers. Now

$$a \in N_{ij}\ ,\ k \in N_{jj} \Longrightarrow a + k \in N_{ij}$$

$$\Longrightarrow k \in N_{ij} \quad \text{for all} \quad k \quad \text{sufficiently large.}$$

This proves (i), since total number of N_{ij} is finite. Hence for some N and $\epsilon > 0$, $p_{ij}^{(N)} > \epsilon$ for all i,j.

(ii) X <u>is recurrent</u>.

Let $P(x_o(\omega) = i) = 1$. Then

$$P(x_n = i \text{ for some } n \geq 1)$$

$$= 1 - P(x_n \neq i \text{ for all } n \geq 1)$$

$$\geq 1 - P(x_n \neq i \text{ for } n = N, 2N, \ldots, kN)$$

$$= 1 - \sum_{j_1 \neq i} \cdots \sum_{j_k \neq i} p_{ij_1}^{(N)} p_{j_1 j_2}^{(N)} \cdots p_{j_{k-1} j_k}^{(N)}$$

$$\geq 1 - (1-\epsilon)^k \longrightarrow 1 \quad (k \to \infty),$$

where we used

$$\sum_{j_k = i} p_{j_{k-1} j_k}^{(N)} = 1 - p_{j_{k-1}, i}^{(N)} \leq 1 - \epsilon$$

(iii) X <u>is positive-recurrent</u>

Let $P\{x_o = i\} = 1$.

Let $\tau(\omega) = \min \{n : n \geq 1, x_n(\omega) = i\}$

$$\leq \min \{n : n = kN \text{ for some } k \geq 1, \text{ and } x_n(\omega) = i\}$$

$$= \hat{\tau}(\omega), \text{ say.}$$

We observe that $P\{\hat{\tau}(\omega) < \infty\} = 1$. In fact if $\hat{X}$ is the process defined by restricting X to the parameter set $\{0, N, 2N, \ldots\}$ then $\hat{X}$ is a Markov chain (with transition matrix P^N) to which our previous argument in (ii) is applicable. Therefore

$$\mathcal{E}\{\hat{\tau}(\omega)\} = \sum_{k=1}^{\infty} (kN) P\{\hat{\tau}(\omega) = kN\}$$

Now

$$P\{\hat{\tau}(\omega) = kN\} =$$

$$= \sum_{j_1 \neq i} \cdots \sum_{j_{k-2} \neq i} \sum_{j_{k-1} \neq i} p_{ij_1}^{(N)} p_{j_1 j_2}^{(N)} \cdots p_{j_{k-2} j_{k-1}}^{(N)} p_{j_{k-1} i}^{(N)} \leq 1$$

$$\leq (1-\epsilon)^{k-1} ,$$

where we used

$$\sum_{j_{k-1} \neq i} p_{j_{k-2} j_{k-1}}^{(N)} \leq 1 - \epsilon$$

Thus $\mathcal{E}\{\hat{\tau}(\omega)\} < \infty$ and so $\mathcal{E}\{\tau(\omega) < \infty\}$

Q.E.D.

Let $\mu = (\mu_1, \ldots, \mu_J)$ be a probability distribution on E.

Definition: μ is a stationary probability distribution iff

$$(\forall i) P(x_o = i) = \mu_i \Longrightarrow (\forall i) P(x_n = i) = \mu_i, \; n \geq 1 ;$$

that is,

$$\mu P = \mu$$

THEOREM If X is irreducible and ergodic there exists a unique stationary probability distribution for X; and

(i) $p_{ij}^{(n)} \to \mu_j \quad (n \to \infty)$ geometrically fast

(ii) Also, $\mu_j = \dfrac{1}{\mathcal{E}\{\tau_j\}}$

PROOF OF (i) [Doob (1953), Ch. 5, §2]

As we have seen, there exists N such that

$$\min_{i,j} p_{ij}^{(N)} = \delta > 0.$$

Define

$$m_j^{(r)} = \min_i p_{ij}^{(r)}$$
$$M_j^{(r)} = \max_i p_{ij}^{(r)}$$

Then

$$m_j^{(r+1)} = \min_i \sum_k p_{ik} p_{kj}^{(r)} \geq \min_i \sum_k p_{ik} m_j^{(r)} = m_j^{(r)}$$

$$M_j^{(r+1)} = \max_i \sum_k p_{ik} p_{kj}^{(r)} \leq \max_i \sum_k p_{ik} M_j^{(r)} = M_j^{(r)}$$

$$m_j^{(1)} \leq m_j^{(2)} \leq \ldots \leq M_j^{(2)} \leq M_j^{(1)}$$

Fix α,β. We can write

$$\sum_k^{+} \underbrace{(p_{\alpha k}^{(N)} - p_{\beta k}^{(N)})}_{\geq 0} + \sum_k^{-} \underbrace{(p_{\alpha k}^{(N)} - p_{\beta k}^{(N)})}_{< 0} = 1-1 = 0 \qquad (1)$$

If the first sum has s terms then also,

$$\sum_k^+ (p_{\alpha k}^{(N)} - p_{\beta k}^{(N)}) \leqq 1-(J-s)\delta - s\delta = 1-J\delta \qquad (2)$$

$$\Longrightarrow \quad M_j^{(N+n)} - m_j^{(N+n)} = \max_{\alpha\,\beta} \sum_k \underbrace{(p_{\alpha k}^{(N)} - p_{\beta k}^{(N)})}_{(*)} p_{kj}^{(n)}$$

$$\leqq \max_{\alpha\,\beta} \{\sum_k^+ (*) M_j^{(n)} + \sum_k^- (*) m_j^{(n)}\}$$

$$= \max_{\alpha,\beta} \sum_k^+ (*)(M_j^{(n)} - m_j^{(n)}) \quad \text{(by (1))}$$

$$\leqq (1-J\delta)(M_j^{(n)} - m_j^{(n)}) \quad \text{(by (2))}$$

Also

$$M_j^{(N)} - m_j^{(N)} = \max_{\alpha\,\beta} (p_{\alpha j}^{(N)} - p_{\beta j}^{(N)})$$

$$\leqq \max_{\alpha\,\beta} \sum_k^+ (p_{\alpha k}^{(N)} - p_{\beta k}^{(N)})$$

$$\leqq 1 - J\delta$$

$$\Longrightarrow \quad M_j^{(kN)} - m_j^{(kN)} \leqq (1-J\delta)^k$$

$$\Longrightarrow \quad M_j^{(n)},\ m_j^{(n)} \to \mu_j \quad \text{(say)} \quad (n \to \infty)$$

Thus

$$|p_{ij}^{(n)} - \mu_j| \leqq M_j^{(n)} - m_j^{(n)}$$

$$\leqq (1-J\delta)^{(n/N)-1} .$$

Now $p_{ij}^{(n+1)} = \sum_k p_{ik}^{(n)} p_{kj}$. Let $n \to \infty$: $\mu_j = \sum_k \mu_k p_{kj}$.

Existence of μ is proved. Uniqueness is obvious -- but

to give a separate proof, observe: if $1 = (1,\ldots,1)'$ then

$$P^n \to 1\mu = \begin{bmatrix} \mu_1 \cdots \mu_J \\ \\ \mu_1 \cdots \mu_J \end{bmatrix}$$

$$\text{i.e. } |P^n - 1\mu| \to 0$$

(we use $|P| = \max_i \sum_j p_{ij}$; $|\mu| = \max_i \mu_i$)

Thus

$$\tilde{\mu} = \tilde{\mu} P$$

$$\Longrightarrow \tilde{\mu} - \mu = \tilde{\mu}(P^n - \mu)$$

(since $$\tilde{\mu} 1\mu = (\tilde{\mu}_1,\ldots,\tilde{\mu}_J)\begin{bmatrix} \mu_1 \cdots \mu_J \\ \cdots\cdots\cdots \\ \mu_1 \cdots \mu_J \end{bmatrix}$$

$$= (\mu_1 \cdots \mu_J) = \mu)$$

$$\Longrightarrow |\tilde{\mu} - \mu| \leq |\tilde{\mu}|\,|P^n - 1\mu| \to 0 .$$

(ii) Proof omitted -- see Chung (1960), §6, Theorem 1.

2. STATIONARY CONTROL OF ERGODIC CHAINS.

2.1 Functionals on an irreducible ergodic chain.

Let $X = \{x_n\}$ be an irreducible ergodic chain on the state space $E = \{j : 1 \leq j \leq J\}$. Notation as in §1. Let

$$L = (L_1,\ldots,L_n)'$$

be a column vector of numbers $L_i \geq 0$. L_i = cost incurred when X is in state i. Also write $L_i = L(i)$, $i \in E$. Consider

$$V_n(i) = \mathcal{E}\{ \sum_{m=0}^{n} L(x_m) \mid x_o = i\} ,$$

the expected cost incurred in the first $n+1$ stages, given that the chain starts in i. Thus

$$V_n(i) = \sum_{m=0}^{n} \sum_{j} p_{ij}^{(m)} L(j) \quad (p_{ij}^{(0)} = \delta_{ij})$$

or

$$V_n = \sum_{m=0}^{n} P^m L \tag{1}$$

where

$$V_n = [V_n(1),\ldots,V_n(J)]'$$

Thus

$$\begin{aligned} V_{n+1}-V_n &= P^{n+1}L \\ &= \sum_{m=0}^{n} (P^{m+1}L-P^mL) + L \\ &= (P-I)V_n + L . \end{aligned}$$

Let

$$\mathcal{L} = P-I, \quad \Delta V_n \equiv V_{n+1}-V_n$$

Then

$$\begin{aligned} \mathcal{L}V_n - \Delta V_n &= -L \qquad n = 0,1,2,\ldots \\ V_o &= L \end{aligned} \tag{2}$$

Consider asymptotic behavior of V_n as $n \to \infty$. Since

$$\lim_{n\to\infty} P^n = 1\mu \qquad\qquad 1 = (1,\ldots,1)' \qquad \mu = (\mu_1,\ldots,\mu_J)$$

we have

$$\begin{aligned} \lim_{n\to\infty} n^{-1} V_n &= \lim_{n\to\infty} n^{-1} \sum_{m=0}^{n} P^m L \\ &= \lim_{n\to\infty} n^{-1} \sum_{m=0}^{n} (1\mu + Q_m)L \end{aligned}$$

(where $Q_m \to 0 \quad (m \to \infty)$)

$$= (\mu L)1 \tag{3}$$

[Here we used the fact that if $|Q_m| = q_m$ is a sequence with $q_m \to 0$ $(m \to \infty)$ then

$$\lim_{n\to\infty} n^{-1} \sum_{o}^{n} q_m = 0 \quad .$$

Proof: exercise].

Thus

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n} P^m L = (\mu L)1 = \Big(\sum_i \mu_i L_i\Big)1$$

or

Expected cost per stage $\rightarrow$ expected cost relative to stationary distribution

and the limit is independent of the initial state.

To examine the effect of initial state on the asymptotic behavior of V_n we introduce a new vector, the potential.

Definition. The potential v corresponding to the sequence V_n is

$$v = \lim_{n \to \infty} [V_n - (n+1)(\mu L)1] \tag{4}$$

v is well-defined, since if $n > m$,

$$\begin{aligned} &|[V_n - (n+1)(\mu L)1] - [V_m - (m+1)(\mu L)1]| \\ &\quad = \Big| \sum_{m+1}^{n} P^k L - (n-m)(\mu L)1 \Big| \\ &\quad \leq \alpha |L| \sum_{m+1}^{n} \beta^k \rightarrow 0 , \quad (m,n \to \infty) \end{aligned}$$

for some constants $\alpha > 0$, $0 < \beta < 1$.

In particular we have

$$V_n \sim n(\mu L)1 + v \qquad (n \to \infty).$$

Write $\mu L = \lambda$. We shall find an equation relating v and λ. Observe that

$$\begin{aligned}
\mathcal{L}v &= \mathcal{L}\,[\lim_{n\to\infty}(V_n-(n+1)\lambda 1)] \\
&= \lim_{n\to\infty}\mathcal{L}[V_n-(n+1)\lambda 1] \\
&= \lim_{n\to\infty}(\Delta V_n - L) \qquad\qquad (\text{since } \mathcal{L}1 = 0) \\
&= \lim_{n\to\infty}(P^{n+1}L - L) \\
&= (\mu L)1 - L
\end{aligned}$$

There follows the basic relation

$$\mathcal{L}v = \lambda 1 - L \quad . \tag{5}$$

Observe that if u is a vector of form $c1$, where c is a scalar, then $\mathcal{L}u = 0$. Hence if v is a solution of (5), so is $v+u$. But the solution of (5) is otherwise unique: for $\mathcal{L}u = 0 \Longrightarrow Pu = u \Longrightarrow P^n u = u \Longrightarrow u = 1(\mu u)$ (letting $n \to \infty$); i.e. u is of the form $c1$.

Observe also that (5), regarded as an equation for v and λ, has a solution iff $\lambda = \mu L$. In fact, for any vector w,

$$\mu(\mathcal{L}w) = 0 \ ,$$

so that

$$\begin{aligned}
0 &= \mu(\lambda 1 - L) \\
&= \lambda - \mu L \ .
\end{aligned}$$

Let us evaluate the potential v on the chain by writing

$$\tilde{v}_n = v(x_n) \ .$$

Then (5) says that

$$(\mathcal{L}\tilde{v}_n)_i = \mathcal{E}\{\tilde{v}_{n+1} | x_n = i\} - v(i) = \lambda - L(i)$$

Thus $\tilde{v}_n$ increases, on the average, at states where $\lambda > L(i)$ and decreases, on the average, at states where $\lambda < L(i)$.

The potential will be of basic importance in control theory.

EXERCISE: Green's matrix.

Let $$\mathcal{G} = (I - P + 1\mu)^{-1} - 1\mu \ ;$$

Then $$\mathcal{L}v = \lambda 1 - L \quad (\lambda = \mu L)$$

iff $$v = \mathcal{G}L + c1$$

where c is a scalar.

[Note:

$$\mathcal{G} = \sum_0^\infty (P^n - 1\mu)$$

and

$$P^n - 1\mu = (P - 1\mu)^n \quad , \quad n \geq 1 \ .$$

Thus

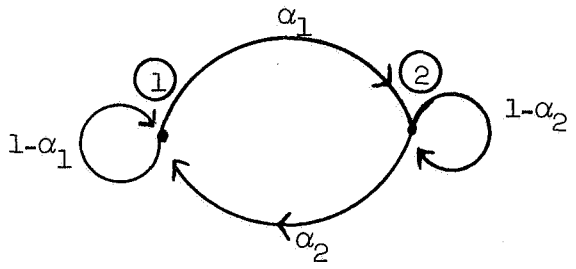
$$\begin{aligned} \mathcal{G} &= I - 1\mu + \sum_1^\infty (P - 1\mu)^n \\ &= \sum_0^\infty (P - 1\mu)^n - 1\mu \\ &= (I - P + 1\mu)^{-1} - 1\mu \ . \quad] \end{aligned}$$

For the chain illustrated below, we have

$$P = \begin{pmatrix} 1-\alpha_1 & \alpha_1 \\ \alpha_2 & 1-\alpha_2 \end{pmatrix}, \qquad \mu = \frac{(\alpha_2, \alpha_1)}{\alpha_1 + \alpha_2}$$

and

$$\mathcal{D} = (\alpha_1 + \alpha_2)^{-2} \begin{pmatrix} \alpha_1 & -\alpha_1 \\ -\alpha_2 & \alpha_2 \end{pmatrix}$$



2.2 Stationary optimal control of an irreducible ergodic chain.

Let A be an abstract set with elements α. Let $\{X(\alpha):\alpha \in A\}$ be a family of irreducible ergodic chains. Write

$$P(\alpha), \quad \mathcal{L}(\alpha), \quad \mu(\alpha), \quad v(\alpha)$$

$$L(\alpha) = [L(1,\alpha),\ldots,L(J,\alpha)]'$$

$$L(i,\alpha) \geqq 0$$

$$\lambda(\alpha) = \mu(\alpha)L(\alpha) \ .$$

Problem: choose (if possible) $\alpha^* \in A$ such that

$$\lambda(\alpha^*) = \inf \{\lambda(\alpha) : \alpha \in A\} .$$

Such an α is said to be optimal. We now obtain a sufficient condition for optimality.

THEOREM 1 Suppose there exist $\alpha^* \in A$ such that

$$\mathcal{L}(\alpha^*)v(\alpha^*) = \lambda(\alpha^*)1 - L(\alpha^*) \qquad (1)$$

and

$$\mathcal{L}(\alpha^*)v(\alpha^*) + L(\alpha^*) \leqq \mathcal{L}(\alpha)v(\alpha^*) + L(\alpha) \quad (\text{all } \alpha \in A) \qquad (2)$$

Then α^* is optimal.

PROOF. Since $\mu(\alpha)\mathcal{L}(\alpha)w = 0$ for $\alpha \in A$ and arbitrary vectors w, and since $\mu(\alpha)1 = 1$ $(\alpha \in A)$,

$$(1) \Longrightarrow \mu(\alpha)[\mathcal{L}(\alpha^*)v(\alpha^*) + L(\alpha^*)] = \lambda(\alpha^*)$$

and then

$$(2) \Longrightarrow \lambda(\alpha^*) \leqq \mu(\alpha)[\mathcal{L}(\alpha)v(\alpha^*) + L(\alpha)] = \lambda(\alpha) \quad (\alpha \in A)$$

Q.E.D.

Note: After a $v(\alpha^*)$ is found, the optimality of α^* can be checked without computing comparison $v(\alpha)$'s and $\lambda(\alpha)$'s explicitly. Furthermore if α^* satisfies the conditions of the theorem, then $v(\alpha^*)$ and $\lambda(\alpha^*)$ are solutions of the equation

$$\inf \{\mathcal{L}(\alpha)v + L(\alpha) : \alpha \in A\} = \lambda \qquad (3)$$

Equation (3) is Bellman's equation (or rather, one of many functional equations of Bellman's type).

Condition (2) is rather special, and certainly not necessary for the problem as formulated thus far. To see under what condition (2) or Bellman's equation (3) may be necessary, we shall be more explicit about the dependence of $P(\alpha)$ on α.

Let

$$A = A_1 \times A_2 \times \ldots \times A_J$$

where the A_i are nonempty abstract sets, and write

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_J), \quad \alpha_i \in A_i .$$

Assume that

$$p_{ij}(\alpha) = p_{ij}(\alpha_i)$$

$$(i,j \in E, \ \alpha \in A)$$

$$L_i(\alpha) = L_i(\alpha_i)$$

Such an α is a control law, assigning to each state in E a set of transition probabilities and a cost value. With this

structure we are free to modify the behavior of the chain "locally", that is, at each state independently.

Write

$$\mathcal{L}_i(\alpha)v = \mathcal{L}_i(\alpha_i)v = \sum_j p_{ij}(\alpha_i)v_j - v_i$$

Under these assumptions we have

THEOREM 2.

If an optimal control $\alpha^* \in A$ exists, then

$$\mathcal{L}_i(\alpha_i^*)v(\alpha^*) + L_i(\alpha_i^*) \leqq \mathcal{L}_i(\alpha_i)v(\alpha^*) + L_i(\alpha_i) \qquad (4)$$

for all $i \in E$, $\alpha \in A$.

PROOF. Suppose α^* is optimal and let $\tilde{E} \subset E$ be the set of states where (4) fails. Assuming $\tilde{E}$ is nonempty, there exists $\tilde{\alpha} \in A$ such that

$$\mathcal{L}_i(\alpha_i^*)v(\alpha^*) + L_i(\alpha_i^*) > \mathcal{L}_i(\tilde{\alpha}_i)v(\alpha^*) + L_i(\tilde{\alpha}_i) \qquad (i \in \tilde{E});$$

but

$$\mathcal{L}_i(\alpha_i^*)v(\alpha^*) + L_i(\alpha_i^*) \leqq \mathcal{L}_i(\tilde{\alpha}_i)v(\alpha^*) + L_i(\tilde{\alpha}_i) \quad (i \in E-\tilde{E}).$$

Define α^0 by setting

$$\alpha_i^{o} = \tilde{\alpha}_i \quad (i \in \tilde{E})$$

$$= \alpha_i^{*} \quad (i \in E-\tilde{E}).$$

Then

$$\lambda(\alpha^{o}) = \sum_j \mu_j(\alpha^{o})\ \mathcal{L}_j(\alpha_j^{o})v(\alpha^{*})+L_j(\alpha_j^{o})$$

$$= \sum_{j\in\tilde{E}} \mu_j(\alpha^{o})\{\mathcal{L}_j(\tilde{\alpha}_j)v(\alpha^{*})+L_j(\tilde{\alpha}_j)\}$$

$$+ \sum_{j\in E-\tilde{E}} \mu_j(\alpha^{o})\{\mathcal{L}_j(\alpha_j^{*})v(\alpha^{*})+L_j(\alpha_j^{*})\}$$

Because each $X(\alpha)$ is assumed to be ergodic and irreducible, $\mu_j(\alpha) > 0$ $(j \in E,\ \alpha \in A)$, and so we obtain the contradiction

$$\lambda(\alpha^{o}) < \sum_j \mu_j(\alpha^{o})\{\mathcal{L}_j(\alpha_j^{*})v(\alpha^{*})+L_j(\alpha_j^{*})\}$$

$$= \sum_j \mu_j(\alpha^{o})\lambda(\alpha^{*})$$

$$= \lambda(\alpha^{*})$$

Q.E.D.

The foregoing argument shows that Bellman's equation is a necessary and sufficient condition for optimality of $\alpha^{*} \in A$, whenever our system has the "product" structure described above.

We now describe a computational method for finding successively better controls α. This is Bellman's <u>approximation in policy space</u>. Namely, choose $\alpha^{(o)} \in A$ arbitrarily and compute $v^{(o)} = v(\alpha^{(o)})$. Choose $\alpha^{(1)} \in A$ to satisfy the inequalities

$$\mathcal{L}_i(\alpha_i^{(1)})v^{(o)} + L_i(\alpha_i^{(1)}) \leqq \mathcal{L}_i(\alpha_i)v^{(o)} + L_i(\alpha_i)$$

$$\alpha \in A_i \ ; \ i \in E \qquad (4)$$

By an argument already used above,

$$(4) \Longrightarrow \lambda(\alpha^{(1)}) \leqq \lambda(\alpha^{(o)})$$

Continuing thus, compute in turn $v^{(1)}$, $\alpha^{(2)}$, $v^{(2)}$, $\alpha^{(3)}$,... .
At the n^{th} stage of computation we have, in simpler notation,

$$\mathcal{L}(\alpha_n)v_n + L(\alpha_n) = \lambda(\alpha_n)1$$

$$\mathcal{L}(\alpha_{n+1})v_n + L(\alpha_{n+1}) = \min\{\mathcal{L}(\alpha)v_n + L(\alpha) : \alpha \in A\} \qquad (5)$$

provided the indicated minima exist (notation: for a vector w, $\min_{\alpha} = (\min_{\alpha} w_i)$). Then

$$\lambda(\alpha_o) \geqq \lambda(\alpha_1) \geqq \ldots$$

and, since $\lambda(\alpha_n) = \mu(\alpha_n)L(\alpha_n) \geqq 0$, it follows that the sequence $\lambda(\alpha_n)$ converges. We shall next investigate whether $\lambda(\alpha_n)$ actually converges to the desired minimum. It will be convient to assume first that α^* exists; this assumption will be dropped later.

THEOREM 3.

If

(i) for every vector v there exists $\beta \in A$ such that

$$\mathcal{L}(\beta)v + L(\beta) = \inf\{\mathcal{L}(\alpha)v+L(\alpha) : \alpha \in A\}$$

(ii) there exists $\alpha^* \in A$ such that

$$\lambda^* \equiv \mu(\alpha^*)L(\alpha^*) = \inf\{\mu(\alpha)L(\alpha) : \alpha \in A\}$$

(iii) $$\eta = \inf\{\mu_i(\alpha) : i \in E, \alpha \in A\} > 0$$

then any sequence λ_n, constructed by approximation in policy space, converges to λ^*.

PROOF. By (i), the sequence λ_n is well-defined. Our result will follow if, for every $\epsilon > 0$, there exists N such that, for all $n \geq N$ and all $\alpha \in A$,

$$\mathcal{L}(\alpha_n)v_n+L(\alpha_n) < \mathcal{L}(\alpha)v_n+L(\alpha)+ \epsilon 1 \qquad (6)$$

In fact, the left side of this inequality is equal to $\lambda_n 1$, so

$$\lambda_n < \mu(\alpha^*)[\mathcal{L}(\alpha^*)v_n+L(\alpha^*)+\epsilon 1] = \lambda^* + \epsilon$$

To verify (6), suppose the contrary: that is, for some $\epsilon > 0$ and every N there exist $m > N$, $\alpha = \alpha_m \in A$, $\beta = \beta_m \in A$ and $i = i_m \in E$ such that

$$\mathcal{L}_{i_m}(\alpha_m)v_m + L_{i_m}(\alpha_m) \geqq \mathcal{L}_{i_m}(\beta_m)v_m + L_{i_m}(\beta_m) + \epsilon .$$

Since, by construction of the α_n,

$$\mathcal{L}_i(\alpha_{m+1})v_m + L_i(\alpha_{m+1}) \leqq \mathcal{L}_i(\beta_m)v_m + L_i(\beta_m), \quad (i \in E)$$

we have

$$\begin{aligned}\mathcal{L}_i(\alpha_{m+1})v_m + L_i(\alpha_{m+1}) &\leqq \mathcal{L}_i(\alpha_m)v_m + L_i(\alpha_m) - \epsilon\,\delta_{i,i_m} \\ &= \lambda_m - \epsilon\,\delta_{i,i_m} \quad (i \in E).\end{aligned}$$

Hence, multiplying by $\mu_i(\alpha_{m+1})$ and summing on i, we obtain in view of (iii),

$$\lambda_{m+1} \leqq \lambda_m - \epsilon\eta .$$

But then for an appropriate subsequence $m = m(\nu)$,

$$\begin{aligned}\lambda_{m(\nu+1)} &\leqq \lambda_{m(\nu)+1} \leqq \lambda_{m(\nu)} - \epsilon\eta \\ &\leqq \lambda_{m(1)} - \nu\epsilon\eta \\ &\to -\infty \quad (\nu \to \infty)\end{aligned}$$

and this contradiction shows that (6) is true.

Q.E.D.

In this proof the only role played by condition (ii), the existence of α^*, was in showing that (6) is a sufficient condition for the desired result. We now show that condition (i), together with another, more fundamental condition, imply both (ii) and (iii). This will be done by showing that the sequence of potentials $\{v_n\}$ is bounded, hence contains a convergent subsequence.

For a vector v let $|v| = \max_i |v_i|$ and for a matrix Q let $|Q| = \max_i \sum_j |q_{ij}|$. Let $\mathcal{G}(\alpha)$ be the Green's matrix

$$\mathcal{G}(\alpha) = [I-P(\alpha) + 1\mu(\alpha)]^{-1} - 1\mu(\alpha)$$

LEMMA. <u>If for some integer</u> K,

$$\delta = \inf_{\alpha\in A} \max_{1\le N\le K} \min_{i,j} p_{ij}^{(N)}(\alpha) > 0$$

<u>then</u> $|\mathcal{G}(\alpha)|$ <u>is bounded above on</u> A, <u>and</u>

$$\mu_i(\alpha) > \delta, \qquad i \in E, \quad \alpha \in A .$$

In the Appendix it is shown that one can take $K = (J-1)^2+1$.

PROOF. Fixing α and writing $\mathcal{G}(\alpha) = \mathcal{G}$ we have

$$\begin{aligned}\mathcal{G} &= (I-P+1\mu)^{-1} - 1\mu \\ &= \sum_{0}^{\infty} (P-1\mu)^n - 1\mu \\ &= I + \sum_{0}^{\infty} (P^n-1\mu) - 1\mu \\ &= \sum_{0}^{\infty} (P^n-1\mu) \ .\end{aligned}$$

Let $\tilde{\delta} = \delta(\alpha) = \max\limits_{1\leq N\leq K} \min\limits_{ij} p_{ij}^{(N)}(\alpha)$.

Then $\tilde{\delta} \geq \delta > 0$; we also have the estimate (see p. 13)

$$|p_{ij}^{(n)} - \mu_j| \leq (1-J\tilde{\delta})^{(n/N)-1}$$

for some N, $1 \leq N \leq K$. Thus

$$\begin{aligned}|P^n - 1\mu| &= \max_i \sum_j |p_{ij}^{(n)} - \mu_j| \\ &\leq J(1-J\tilde{\delta})^{(n/N)-1}\end{aligned}$$

and so

$$\begin{aligned}|\mathcal{G}| &\leq \sum_{0}^{\infty} |P^n-1\mu| \\ &\leq 2N + \sum_{N}^{\infty} J(1-J\tilde{\delta})^{(n/N)-1} \\ &\leq 2N + J[1-(1-J\tilde{\delta})^{1/N}]^{-1} \\ &\leq 2K + J[1-(1-J\tilde{\delta})^{1/K}]^{-1} \ .\end{aligned}$$

Also

$$\mu_j = \sum_j \mu_i p_{ij}^{(N)} \geq \tilde{\delta} \sum_j \mu_i \geq \delta .$$

Q.E.D.

Remark. The parameter δ is a measure of how fast the chains $X(\alpha)$, $\alpha \in A$, approach equilibrium as $n \to \infty$, and we have shown that, in the case where the $X(\alpha)$ are ergodic and irreducible, this parameter also determines a lower bound on the equilibrium state probabilities. It is not true, however, that there is any special relation between "rate of stabilization" and values of equilibrium probabilities: in particular it is not true that $\mathscr{E}(\alpha)$ is necessarily bounded above if the $\mu_i(\alpha)$ are bounded below.

We can now prove a somewhat improved version of Theorem 3.

THEOREM 4.

If (i) for every vector v there exists $\beta \in A$ such that

$$\mathscr{L}(\beta)v+L(\beta) = \inf\{\mathscr{L}(\alpha)v+L(\alpha):\alpha \in A\}$$

(ii) for some integer K, the parameter

$$\delta = \inf_{\alpha\in A} \max_{1\leq N\leq K} \min_{i,j} p_{ij}^{(N)}(\alpha) > 0$$

(iii) the vectors $L(\alpha)$ are bounded on A;

then an optimal control $\alpha^* \in A$ exists. If $\{\alpha^{(n)}\}$ is an arbitrary sequence constructed by approximation

in policy space, then

$$\lambda(\alpha^{(n)}) \downarrow \lambda(\alpha^*) .$$

PROOF. We write α_n for $\alpha^{(n)}$, etc. By (i), an approximation sequence $\{\alpha_n\}$ exists. By the lemma, $|\mathcal{G}(\alpha)|$ is bounded on A, and hence the sequence

$$|v_n| = |\mathcal{G}(\alpha_n)L(\alpha_n)|$$

$$\leq |\mathcal{G}(\alpha_n)||L(\alpha_n)|$$

is bounded. Thus there is a subsequence of integers, which we denote again by $\{n\}$, such that

$$\tilde{v} = \lim v_n \qquad (n \to \infty)$$

exists. Since the sequence $\{\lambda_n\}$ is nonnegative and nonincreasing, we have that

$$\tilde{\lambda} = \lim \lambda_n \qquad (n \to \infty)$$

exists. It will be shown that

$$\tilde{\lambda}1 = \tilde{w} \equiv \min\{\mathcal{L}(\alpha)\tilde{v} + L(\alpha) : \alpha \in A\} . \tag{7}$$

Write

$$w_n = \min\{\mathcal{L}(\alpha)v_n + L(\alpha) : \alpha \in A\} .$$

We first show that $\lim w_n = \tilde{w}$ $(n \to \infty)$. In fact for $\eta > 0$ and n sufficiently large we have

$$|v_n - \tilde{v}| < \eta/2$$

and since

$$|\mathcal{L}(\alpha)| \leqq |P(\alpha)| + |I| \leqq 2$$

there follows

$$\begin{aligned} w_n &= \min \{\mathcal{L}(\alpha)\tilde{v} + L(\alpha) + \mathcal{L}(\alpha)(v_n - \tilde{v}) : \alpha \in A\} \\ &\leqq \min \{\mathcal{L}(\alpha)\tilde{v} + L(\alpha) + |\mathcal{L}(\alpha)(v_n - \tilde{v})|\,1 : \alpha \in A\} \\ &\leqq \tilde{w} + \eta 1 \,. \end{aligned}$$

Similarly $w_n \geqq \tilde{w} - \eta 1$, so that

$$|w_n - \tilde{w}| < \eta \,.$$

Next, by the Lemma, $\mu_i(\alpha) \geqq \delta > 0$ $(i \in E, \alpha \in A)$ and then, by the same argument used in the proof of Theorem 3,

$$\lambda_n 1 = \mathcal{L}(\alpha_n)v_n + L(\alpha_n) < \mathcal{L}(\alpha)v_n + L(\alpha) + \eta 1$$

for $\eta > 0$, n sufficiently large, and all α. Letting $n \to \infty$ we have

$$\tilde{\lambda} 1 = \lim \lambda_n 1 < \mathcal{L}(\alpha)\tilde{v} + L(\alpha) + \eta 1 \,;$$

since $\eta > 0$ is arbitrary

$$\tilde{\lambda}1 \leq \min \{\mathcal{L}(\alpha)\tilde{v} + L(\alpha) : \alpha \in A\} .$$
$$= \tilde{w} .$$

Since

$$\lambda_n 1 = \mathcal{L}(\alpha_n)v_n + L(\alpha_n)$$
$$\geq \min \{\mathcal{L}(\alpha)v_n + L(\alpha) : \alpha \in A\}$$
$$= w_n ,$$

we conclude that $\tilde{\lambda}1 \geq \tilde{w}$, and (7) follows. Finally, by condition (i) and (7) there is $\alpha^* \in A$ such that

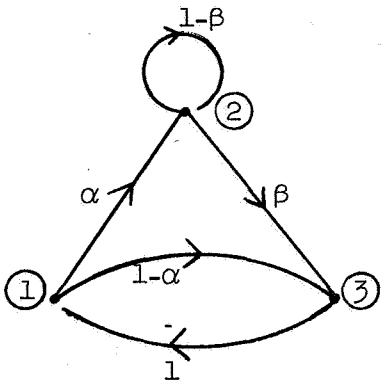
$$\tilde{\lambda}1 = \mathcal{L}(\alpha^*)\tilde{v} + L(\alpha^*) ,$$

and therefore α^* is optimal.

Q.E.D.

Remark. Theorems 3 and 4 can be extended to the situation where E includes a transient subset in addition to a single ergodic subset: the extension is left as an exercise. Thus condition (ii) of Theorem 3 is unnecessarily restrictive for the general case; however, this condition cannot be relaxed altogether, as the example below shows. Roughly speaking, trouble occurs if $\mathcal{L}(\alpha_n)$ is unbounded on an approximation sequence $\{\alpha_n\}$, because this means that as $n \to \infty$ the chain tends to decompose into ergodic subchains. In this case the control problem can be considered separately for each distinct classification of states (into transient and recurrent subsets) possible for the given class of controls A.

EXAMPLE.



$$P = \begin{bmatrix} 0 & \alpha & 1-\alpha \\ 0 & 1-\beta & \beta \\ 1 & 0 & 0 \end{bmatrix}$$

$$\mu = \frac{1}{\alpha+2\beta}\,[\beta,\ \alpha,\ \beta]$$

$$\mathcal{J} = \frac{1}{(\alpha+2\beta)^2} \begin{bmatrix} \alpha+\alpha\beta+\beta^2 & -2\alpha+\alpha^2+\alpha\beta & \alpha-(\alpha+\beta)^2 \\ -2\beta-\beta^2 & 4\beta-\alpha\beta & -2\beta+\alpha\beta+\beta^2 \\ \alpha-\beta^2 & -2\alpha-\alpha\beta & \alpha+\alpha\beta+\beta^2 \end{bmatrix}$$

$$\mathcal{L} = P - I$$

$$\mathcal{L}v = \lambda 1 - L \Longrightarrow v = \mathcal{J} L$$

Suppose $0 \leqq \alpha,\beta \leqq 1$. If $0 < \alpha,\beta < 1$ all states are ergodic; the other possibilities are tabulated:

α	β	Ergodic	Periodic	Transient
0	0	2	1,3	-
0	(0,1]	-	1,3	2
(0,1]	0	2	-	1,3
1	1	-	1,2,3	-
1	(0,1)	1,2,3	-	-
(0,1)	1	1,2,3	-	-

Suppose $L_1 = 2\alpha$, $L_2 = 4\beta$, $L_3 = 1$. It is clear that optimal control is $0 < \alpha \leq 1$, $\beta = 0$, in which case

$$\mu = (0,\ 1,\ 0)$$

and

$$\lambda = \mu L = 0.$$

In an attempt to compute by dynamic programming let

$$\alpha^{(1)} = \beta^{(1)} = \frac{1}{2}$$

Then

$$\mathcal{U}^{(1)} = \frac{1}{9}\begin{bmatrix} 4 & -2 & -2 \\ -5 & 7 & -2 \\ 1 & -5 & 4 \end{bmatrix}$$

$$L^{(1)} = (1,\ 2,\ 1)'$$

$$v^{(1)} = \frac{1}{9}(-2,\ 7,\ -5)$$

$$\mathcal{L}(\alpha,\beta)v^{(1)}+L^{(1)} = \begin{bmatrix} -v_1^{(1)} + \alpha v_2^{(1)} + (1-\alpha)v_3^{(1)} + L_1^{(1)} \\ -\beta(v_2^{(1)} - v_3^{(1)}) + L_2^{(1)} \\ v_1^{(1)} - v_3^{(1)} + L_3^{(1)} \end{bmatrix} = \min$$

$$\Longrightarrow \quad \begin{cases} \frac{3}{9} + \alpha \cdot \frac{12}{9} + 2\alpha = \min \\ -\beta \cdot \frac{12}{9} + 4\beta = \min \\ \frac{3}{9} + L_3 = \min \end{cases}$$

$$\Longrightarrow \quad \alpha^{(2)} = 0, \quad \beta^{(2)} = 0$$

Note that these values are nonoptimal -- and the state space has broken up into two irreducible subsets. Note also that $\mathcal{L}$ is not defined for these values; and becomes unbounded if, for example, we set

$$\alpha = \beta = \gamma \quad \text{and let} \quad \gamma \to 0 \ .$$

In fact if we attempt to solve

$$\mathcal{L}v + L = \lambda$$

with $\alpha = \beta = 0$, we get

$$\left.\begin{aligned} -v_1 + v_3 &= \lambda \\ 0 &= \lambda \\ v_1 - v_3 + 1 &= \lambda \end{aligned}\right\}$$

and no solution exists.

Remark. In Theorems 3 and 4 it was assumed that minimization could actually be achieved at each step of approximation. It is of some practical interest to verify that this condition can be relaxed, provided we are content with a control which is only 'close' to optimal.

Definition. Let $\epsilon \geq 0$. A control $\alpha^* \in A$ is ϵ-optimal if

$$\lambda(\alpha^*) \leq \inf \{\lambda(\alpha) : \alpha \in A\} + \epsilon .$$

Obviously if $\epsilon > 0$ an ϵ-optimal control always exists. We will show that such a control can be computed by approximation in policy space, where we do not assume that the minima can be exactly achieved.

THEOREM 5.

If (i) for some integer K, the parameter

$$\delta = \inf_{\alpha \in A} \max_{1 \leq N \leq K} \min_{ij} p_{ij}^{(N)}(\alpha) > 0$$

(ii) $L(\alpha)$ is bounded on A

then for every $\epsilon > 0$ a sequence

$\{\alpha^{(n)}\}$ can be constructed such that $\alpha^{(n)}$ is ϵ-optimal for all n sufficiently large.

PROOF. Write α_n for $\alpha^{(n)}$ etc. Let α_o be arbitrary and define α_n, v_n, λ_n by:

$$\mathcal{L}(\alpha_n)v_n + L(\alpha_n) = \lambda_n$$

$$\mathcal{L}(\alpha_{n+1})v_n + L(\alpha_{n+1}) \leqq \inf_{\alpha \in A} \{\mathcal{L}(\alpha)v_n + L(\alpha)\} + \epsilon 2^{-n} 1$$

By the Lemma (p.27) we have $\mu_i(\alpha) \geqq \delta$ $(i \in E, \alpha \in A)$. Now we generalize slightly the argument used in the proof of Theorem 3. We wish to show that

$$\mathcal{L}(\alpha_n)v_n + L(\alpha_n) \leqq \inf_{\alpha \in A} \{\mathcal{L}(\alpha)v_n + L(\alpha)\} + \epsilon 1 \qquad (8)$$

for all n sufficiently large. If (8) is false then for every N there exist $m > N$, $\alpha = \alpha_m \in A$, $\beta = \beta_m \in A$ and $i = i_m \in E$ such that

$$\mathcal{L}_{i_m}(\alpha_m)v_m + L_{i_m}(\alpha_m) > \mathcal{L}_{i_m}(\beta_m)v_m + L_{i_m}(\beta_m) + \epsilon .$$

Since, by construction of the α_n,

$$\mathcal{L}_i(\alpha_{m+1})v_m + L_i(\alpha_{m+1}) \leqq \mathcal{L}_i(\beta_m)v_m + L_i(\beta_m) + \epsilon 2^{-m}$$

for $i \in E$, we have

$$\mathcal{L}_i(\alpha_{m+1})v_m + L_i(\alpha_{m+1}) \leqq \mathcal{L}_i(\alpha_m)v_m + L_i(\alpha_m) + \epsilon 2^{-m} - \epsilon\delta_{i,i_m}$$
$$= \lambda_m + \epsilon 2^{-m} - \epsilon\delta_{i,i_m} .$$

Hence, multiplying by $\mu_i(\alpha_{m+1})$ and summing on i, we obtain

$$\lambda_{m+1} \leqq \lambda_m + \epsilon 2^{-m} - \epsilon\delta ;$$

and therefore, for a subsequence $m = m(\nu) \to \infty$

$$\lambda_{m(\nu+1)} \leqq \lambda_{m(\nu)+1} + \epsilon 2^{-m+1}$$
$$\leqq \lambda_{m(\nu)} + \epsilon 2^{-m(\nu)} - \epsilon\delta + \epsilon 2^{-m(\nu)+1}$$
$$\leqq \lambda_{m(1)} + 3\epsilon - \nu\,\epsilon\,\delta$$
$$\to -\infty \quad (\nu \to \infty) .$$

This contradiction verifies (8). But now (8) yields

$$\lambda_m 1 \leqq \mathcal{L}(\alpha)v_n + L(\alpha) + \epsilon 1 \qquad (\alpha \in A)$$

for all n sufficiently large; and therefore

$$\lambda_n \leqq \inf_{\alpha\in A} \{\mu(\alpha)L(\alpha)\} + \epsilon$$

Q.E.D.

3. INTERCEPT PROBLEMS.

We discuss the analogy, for controlled Markov chains, of a typical stochastic intercept problem. With the same notation as before, let X be a homogeneous* Markov chain on state space $E = \{0,1,\ldots,J\}$; the state 0 will be interpreted later as a target state, toward which the chain is to be controlled in an optimal fashion. Later we also consider the extension to moving targets.

On the hypothesis that

$$P\{x_o(\omega) = i \neq 0\} = 1 \qquad (i \in E)$$

let

$$\tau_i(\omega) = \min \{n : x_n(\omega) = 0, \quad n \geq 1\}$$

if $\tau_i(\omega) < \infty$; otherwise put $\tau_i(\omega) = +\infty$. τ_i is thus the first time the chain visits 0, when the chain starts in i.

Write $F = \{1,\ldots,J\}$. Let X have transition matrix P, with

$$p_{o,o} = 1 \quad , \quad p_{o,j} = 0 \quad (j \in F)$$

---

*This simplification is convenient but not essential.

We assume that $i \rightsquigarrow 0$ (i.e., $p_{i,0}^{(n)} > 0$ for some n) for all $i \in F$; thus all states in F are transient. It will now be proved that

$$P\{\tau_i < \infty\} = 1 \tag{1}$$

and even that

$$\mathcal{E}\{\tau_i\} < \infty \tag{2}$$

for $i \in F$. In proving (1) we shall need the

Borel-Cantelli Lemma

If $A_n \in \mathfrak{F}$, $n = 1,2,\ldots,$ *and if*

$$\sum_1^\infty P(A_n) < \infty$$

then

$$P(\overline{\lim}\, A_n) = 0 \quad .$$

PROOF. Recall that $\overline{\lim}\, A_n$ is the set of ω which belong to infinitely many of the A_n, i.e.,

$$\overline{\lim}\, A_n = \bigcap_{n=1}^\infty \bigcup_{m=n}^\infty A_m \quad .$$

Thus

$$\begin{aligned} P(\overline{\lim}\, A_n) &\leq P(\bigcup_{m=n}^\infty A_m) \\ &\leq \sum_{m=n}^\infty P(A_m) \qquad (n = 1,2,\ldots) \\ &\to 0 \qquad n \to \infty \end{aligned}$$

Q.E.D.

Now following Doob (1953), observe that for $i \in F$,

$$\sum_{j\in F} p_{ij}^{(n+1)} = \sum_{j\in F}\sum_{k\in F} p_{ik}^{(n)} p_{kj} \leqq \sum_{k\in F} p_{ik}^{(n)} \quad ;$$

i.e., the probability of staying in F cannot increase with n. Since $p_{i,o}^{(n)} > 0$ for some $n < J$ we have $\sum_{j\in F} p_{ij}^{(n)} < 1$ for some $n < J$ and so we can pick $N < J$ such that

$$\sum_{j\in F} p_{ij}^{(N)} \leqq \eta < 1 \quad , \quad n \geqq N, \; i \in F \; .$$

Then

$$\sum_{j\in F} p_{ij}^{((m+1)N)} = \sum_{j,k\in F} p_{ik}^{(mN)} p_{kj}^{(N)}$$

$$\leqq \eta \sum_{k\in F} p_{ik}^{(mN)} \quad , \quad i \in F$$

so that

$$\sum_{j\in F} p_{ij}^{(mN)} \leqq \eta^{m} \quad , \quad i \in F \; .$$

Since the sum $\sum_{j\in F} p_{ij}^{(n)}$ is monotone in n,

$$\sum_{j\in F} p_{ij}^{(n)} \leqq \eta^{(n/N)-1} \to 0 \quad (n \to \infty), \quad i \in F \; . \tag{3}$$

Since the right side of this inequality is the general term of a convergent series, the Borel-Cantelli Lemma (with $A_n = \{\omega: x_n(\omega) \in F\}$) implies that

$$P\{x_n(\cdot) \in F \text{ for infinitely many } n\} = 0$$

and therefore

$$P\{x_n(\cdot) = 0 \text{ for some } n < \infty\} = 1 .$$

This proves (1). To prove (2) observe that

$$P(\tau_i = n+1) = \sum_{j \in F} p_{ij}^{(n)} p_{j,o} \leqq \sum_{j \in F} p_{ij}^{(n)}$$

and the result follows from (3).

Now let

$$L = (L(0),\ldots,L(J))' \; ; \; L(0) = 0, \; L_i \geqq 0, i \in F \; ;$$

and define the <u>potential</u>

$$v = (v(0),v(1),\ldots,v(J))'$$

where

$$\begin{aligned} v(i) &= \mathcal{E}\{ \sum_{n=0}^{\tau_i} L(x_n) \mid x_o = i\} , \quad i \in F, \\ v(0) &= 0 \end{aligned} \tag{4}$$

Because of (2), $v(i) < \infty$, and since $L(0) = 0$ we can write

$$v(i) = \mathcal{E}\{ \sum_{n=0}^{\infty} L(x_n) \mid x_o = i\} , \quad i \in E .$$

In the following we will write v_i, $v(i)$ etc. interchangeably. We next derive a difference equation for v. [Observe that $v(i)$ is the expected cost of tracking (from time 0 to absorption in the target at time τ_i); our future goal will be to minimize v]. Now

$$\begin{aligned} v(i) &= L(i) + \mathcal{E}\{\sum_{n=1}^{\infty} L(x_n) \mid x_o = i\} \\ &= L(i) + \sum_{j\in E} P(x_1 = j \mid x_o = i)\cdot \mathcal{E}\{\sum_{n=1}^{\infty} L(x_n) \mid x_o=i, x_1=j\} \\ &= L(i) + \sum_{j\in E} P(x_1 = j \mid x_o = i)\mathcal{E}\{\sum_{n=1}^{\infty} L(x_n) \mid x_1= j\} \\ &= L(i) + \sum_{j\in E} p_{ij} v(j) \ . \end{aligned}$$

Here we have used the Markov property, and the fact that the chain X is homogeneous. Defining

$$\mathcal{L} = P - I$$

we have then the difference equation

$$\mathcal{L}v = -L \tag{5}$$

with "boundary condition"

$$v_o = 0 \ . \tag{6}$$

Next we show that (5) and (6) have a unique solution, hence characterize v. For $i,j = 1,\ldots,J$ define

$$\hat{p}_{ij} = p_{ij},\ \hat{P} = [\hat{p}_{ij}],\ \hat{L} = (L_i)',\ \hat{v} = (v_i)',$$

$$\hat{\mathcal{L}} = \hat{P} - I \quad (I = J \times J \text{ identity}).$$

Then (5), (6) are equivalent to

$$\hat{\mathcal{L}}\hat{v} = -\hat{L}$$

$$v_o = 0,$$

and it is sufficient to show that $\hat{\mathcal{L}}^{-1}$ exists. This follows easily from the estimate (3) above. Thus, if $Q = [q_{ij}]$ $(i,j \in F)$ define

$$|Q| = \max_{i\in F} \sum_{j\in F} |q_{ij}| \quad .$$

Then

$$|\hat{P}^n| = \max_{i\in F} \sum_{j\in F} [\hat{P}^n]_{ij}$$

$$= \max_{i} \sum_{j\in F} p_{ij}^{(n)}$$

$$\leq \eta^{(n/N)-1}, \qquad n = 1,2,\ldots$$

using (3). If follows that the series $\sum_{o}^{\infty} \hat{P}^n$ converges; that is,

$$(I - \hat{P})^{-1} = \sum_{o}^{\infty} \hat{P}^n$$

exists.

Our next formula will be useful in the application of dynamic programming: roughly speaking it expresses the potential as the sum of its successive differences.

LEMMA (Summation Rule)

<u>If</u> $v = [v_i]$ $(i \in E)$ <u>is an arbitrary vector with</u> $v(0) = 0$ <u>then</u>

$$v(i) = -\mathcal{E}\{ \sum_{n=0}^{\tau_i} (\mathcal{L}v)(x_n) \mid x_o = i\} \qquad (7)$$

<u>where we interpret</u> $\tau_i = 0$ if $i = 0$.

PROOF. The discussion leading to (5) and (6) did not depend on the fact that $L_i \geq 0$ but only on the restriction $L_o = 0$. For any vector v, $(\mathcal{L}v)(0) = (Pv)(0)-v(0) = v(0)-v(0) = 0$. Now define $\tilde{L} = -\mathcal{L}(v)$, and consider (7) with this replacement. By the discussion leading to (5) and (6), (7) is true if (and only if) $\mathcal{L}v = -\tilde{L}$ and $v_o = 0$. The lemma now follows trivially.

Q.E.D.

Next, let us formulate the control problem in the same way as in §2, namely let A be an arbitrary nonempty set of elements $\alpha = (\alpha_o, \alpha_1, \ldots, \alpha_J)$ where $\alpha_i \in A_i$, and let $X(\alpha)$ be the chain defined by transition matrix

$$P(\alpha) = [p_{ij}(\alpha)] , \quad i,j \in E ,$$

where

$$p_{ij}(\alpha) = p_{ij}(\alpha_i) , \quad i,j \in E .$$

Each $X(\alpha)$ is assumed to have the properties of the X process described above; in particular

$$p_{o,o}(\alpha) = 1, \; p_{o,i}(\alpha) = 0, \quad (i \in F, \; \alpha \in A).$$

Finally, introduce cost vectors $L(\alpha)$, $\alpha \in A$, with

$$L_i(\alpha) = L_i(\alpha_i), \quad L_o(\alpha) = 0 \quad (i \in E, \; \alpha \in A) .$$

We assume that the $L(\alpha)$ are bounded for $\alpha \in A$. For $i \in F$ and $\alpha \in A$ define

$$v_i(\alpha) = \{ \sum_{n=0}^{\tau_i(\alpha)} L[x_n(\alpha),\alpha] \,|\, x_o = i \} \tag{8}$$

where the $x_n(\alpha)$ are random variables of $X(\alpha)$. $v_i(\alpha)$ is thus the expected cost of interception when the chain starts in i and the control is α. Now let

$$u_i^* = \inf \{v_i(\alpha) : \alpha \in A\} \qquad (i \in F)$$

We shall say that $\alpha^* \in A$ is <u>optimal</u> if, for all $i \in F$,

$$v_i(\alpha^*) = u_i^* .$$

Observe that, according to our definition, an optimal control α^* is uniformly optimal with respect to all states $i \in F$.

OPTIMALITY THEOREM. If there exist $\alpha^* \in A$ and a vector v^* $(v_o^* = 0)$ such that

$$0 = \mathcal{L}(\alpha^*)v^* + L(\alpha^*) \leqq \mathcal{L}(\alpha)v^* + L(\alpha) \qquad (9)$$

for all $\alpha \in A$, then α^* is optimal. [Here if w, z are vectors, $w \leqq z$ means $w_i \leqq z_i$ for all i].

PROOF. By (5), (6) and the equality in (9),

$$v^* = v(\alpha^*) ;$$

and by the summation rule, if $\alpha \in A$, $i \in F$,

$$\begin{aligned} v^* &= -\mathcal{E}\{\sum_{n=0}^{\tau_i(\alpha)} \mathcal{L}(\alpha)v^*[x_n(\alpha)] \,|\, x_o = i\} \\ &\leqq -\mathcal{E}\{\sum_{n=0}^{\tau_i(\alpha)} [\mathcal{L}(\alpha)v^*(x_n(\alpha)) + L(x_n(\alpha),\alpha^*) \\ &\qquad - L(x_n(\alpha),\alpha)] \,|\, x_o = i\} \\ &= -\mathcal{E}\{\sum_{n=0}^{\tau_i(\alpha)} [0 - L(x_n(\alpha),\alpha)] \,|\, x_o = i\} \\ &= v_i(\alpha) \end{aligned}$$

Q.E.D.

We observe that (9) can be expressed more concisely as

Bellman's equation

$$\min \{\mathcal{L}(\alpha)v + L(\alpha) : \alpha \in A\} = 0 \qquad (10)$$

Thus if v is a solution of (10) with $v_o = 0$, and if

$$\mathcal{L}(\alpha^*)v + L(\alpha^*) = \min\{\mathcal{L}(\alpha)v + L(\alpha) : \alpha \in A\}$$

then α^* is optimal.

We shall now show that optimal (or ϵ-optimal) controls can be computed by approximation in policy space. We naturally assume A is nonempty: i.e., there is at least one α for which $X(\alpha)$ is such that $\mathcal{E}\{\tau_i(\alpha)\} < \infty$, $i \in F$. To start the approximation let $\alpha^{(0)} \in A$ be arbitrary, let $v^{(0)}$ be defined by

$$\mathcal{L}(\alpha^{(0)})v^{(0)} + L(\alpha^{(0)}) = 0$$
$$v_o^{(0)} = 0$$

and define $\alpha^{(1)}$ by

$$\mathcal{L}(\alpha^{(1)})v^{(0)} + L(\alpha^{(1)}) = \min\{\mathcal{L}(\alpha)v^{(0)} + L(\alpha) : a \in A\}$$

on the assumption that the indicated minimum exists. Given $\alpha^{(n)}$ define $v^{(n)}$ by

$$\mathcal{L}(\alpha^{(n)})v^{(n)} + L(\alpha^{(n)}) = 0 \quad , \quad v_o^{(n)} = 0,$$

and determine $\alpha^{(n+1)}$ by

$$\mathcal{L}(\alpha^{(n+1)})v^{(n)} + L(\alpha^{(n+1)}) = \min\{\mathcal{L}(\alpha)v^{(n)} + L(\alpha) : a \in A\} \quad (11)$$

$(n = 1,2,\ldots)$. By the same kind of argument used in proving the optimality theorem we have

$$v^{(n+1)} \leqq v^{(n)} \quad , \quad n = 0,1,2,\ldots .$$

Indeed (writing α_{n+1} for $\alpha^{(n+1)}$ and using k for the time variable)

$$v_n = -\mathcal{E}\{\sum_{k=0}^{\tau(\alpha_{n+1})} \mathcal{L}(\alpha_{n+1})v_n[x_k(\alpha_{n+1})]\}$$

$$\geqq -\mathcal{E}\{\sum_{k=0}^{\tau(\alpha_{n+1})} [\mathcal{L}(\alpha_n)v_n(x_k(\alpha_{n+1})) + L(\alpha_n, x_k(\alpha_{n+1})) - L(\alpha_{n+1}, x_k(\alpha_{n+1}))]\}$$

$$= +\mathcal{E}\{\sum_{k=0}^{\tau(\alpha_{n+1})} L(\alpha_{n+1}, x_k(\alpha_{n+1}))\}$$

$$= v_{n+1} .$$

Because $v^{(n)} \geqq 0$ it follows that

$$v^* \equiv \lim v^{(n)} \qquad (n \to \infty)$$

exists. Now $\mathcal{L}(\alpha) = P(\alpha)-I$ and $L(\alpha)$ are bounded on A; and so for a suitable subsequence $m = m(\nu)$ $(\nu = 1,2,\ldots)$ the limits

$$\begin{aligned} \mathcal{L}^* &\equiv \lim \mathcal{L}(\alpha^{(m+1)}) \\ L^* &\equiv \lim L(\alpha^{(m+1)}) \end{aligned} \qquad (\nu \to \infty)$$

exist. Therefore, by (11)

$$\mathcal{L}^* v^* + L^* \leqq \mathcal{L}(\alpha) v^* + L(\alpha) \tag{12a}$$

for all $\alpha \in A$. Also, since $v_{m+1} \downarrow v^*$, we have

$$\begin{aligned} \mathcal{L}^* v^* + L^* &= \lim_{\nu \to \infty} [\mathcal{L}(\alpha_{m+1}) v_{m+1} + L(\alpha_{m+1})] \\ &= 0 \ . \end{aligned} \tag{12b}$$

For brevity let us write (7) as

$$v = - \mathcal{E}\{ \sum_{n=0}^{\tau} \mathcal{L} v(x_n)\} \tag{7'}$$

Then if $\alpha \in A$, (12) and (7') imply

$$
\begin{aligned}
v^* &= -\mathcal{E}\{\sum_{n=0}^{\tau(\alpha)} \mathcal{L}(\alpha)v^*(x_n(\alpha))\} \\
&\leqq -\mathcal{E}\{\sum_{n=0}^{\tau(\alpha)} [\mathcal{L}^*v^*(x_n(\alpha)) + L^*(x_n(\alpha)) \\
&\qquad\qquad -L(x_n(\alpha),\alpha)] \\
&= \mathcal{E}\{\sum_{n=0}^{\tau(\alpha)} L(x_n(\alpha),\alpha)\} \\
&= v(\alpha) \ .
\end{aligned}
$$

This shows that for every $\epsilon > 0$, there is an integer N such that

$$v(\alpha^{(n)}) < \inf \{v(\alpha) : \alpha \in A\} + \epsilon 1$$

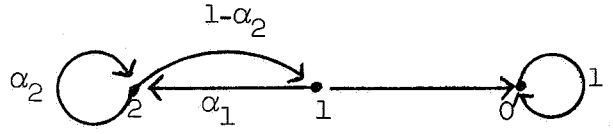
for all $n > N$; that is, an ϵ-optimal control always exists and can be found by approximation in policy space. It is clear that if A can be topologized in such a way that the sequence $\{\alpha^{(n)}\}$ belongs to a (sequentially) compact subset of A, then an optimal control $\alpha^* \in A$ exists which can be found as the limit of an appropriate subsequence of the $\alpha^{(n)}$.

[It is true and somewhat trivial that, by adjoining "ideal" elements to A, we can always obtain a new set $\overline{A} \supset A$ which is sequentially compact in the topology induced on A by the mapping $\alpha \to v(\alpha)$. To do this we merely identify an ideal element α^0 with the equivalence class of all sequences $\{\alpha^{(k)} : \alpha^{(k)} \in A\}$ which are such that the corresponding sequences $\{v(\alpha^{(k)})\}$ converge to the

same (finite) limit v^o: we then write $v^o = v(\alpha^o)$. It is obvious that an optimal control $\alpha^* \in \bar{A}$ always exists. However, this procedure has its dangers, even if $P(\alpha^o)$ can be meaningfully defined, since it may no longer be true that the process $X(\alpha^o)$ retains all "regularity" properties of the $X(\alpha)$, $\alpha \in A$.

EXAMPLE.

$L_1 = 1, \quad L_2 = 0$



If $A_1 = [0,1)$, $A_2 = [\frac{1}{4}, \frac{3}{4}]$, an optimal control $\alpha^* \in A$ exists and is obviously

$$\alpha_1^* = 0, \quad \alpha_2^* : \text{arbitrary}$$

Clearly $\mathcal{E}\{\tau_i^*\} < \infty$. If $A_1 = (0,1)$ and $A_2 = [\frac{1}{4}, \frac{3}{4}]$ an optimal control in $\bar{A}$ is the ideal element $\alpha_1 = 0$, $\alpha_2 = \frac{1}{4}$ and again $\mathcal{E}\{\tau_i^*\} < \infty$. But if $A_1 = (0,1)$, $A_2 = (0,1)$ the optimal control in $\bar{A}$ is α_1^* : arbitrary, $\alpha_2^* = 1$, and now $\mathcal{E}\{\tau_2^*\} = +\infty$.

We remark that this "pathology" cannot arise in the present problem if, for some $\delta > 0$, $L_i(\alpha) \geq \delta$ for $\alpha \in A$, $i \in F$.]

PROBLEM 1 (Time-optimal control).

For a finite controlled chain $X(\alpha)$, formulate the problem

of minimizing the expected time (number of steps) required to reach a designated target state (0). State and prove sufficient conditions for existence of an optimal control. In particular show the following: if

(i) $A_i = [0,1]$, $i \in E-\{0\} = F$

(ii) p_{ij} is a linear function of α_i, $i,j \in F$

(iii) For every $\alpha \in A$, $i \rightsquigarrow 0$ for all $i \in F$

then an optimal control exists and can be chosen in the form $\alpha_i = 0$ or 1. This is a version of the "bang-bang" principle.

PROBLEM 2.

Generalize the intercept problem for the case of a (fixed) target set of states $\theta_1, \theta_2, \ldots, \theta_K$: that is, the state space is now

$$\{\theta_1, \theta_2, \ldots, \theta_K \; ; \; 1,2,\ldots,J\}$$

Assume that with each θ_j there is associated a terminal cost C_j and consider the problem of minimizing the expected sum of pursuit cost plus terminal cost.

PROBLEM 3.

Generalize the intercept problem to the case of pursuit of a moving target: that is, a particle Y wanders on E independently

of but simultaneously with $X(\alpha)$: the pursuer (X) always knows Y's current state (between jumps) and knows also that Y moves in accordance with a Markov process with fixed transition matrix Q. If $x_n = i$ and $y_n = j$ then the cost is $L_{ij}(\alpha)$, with

$$L_{ij}(\alpha) > 0, \quad i \neq j \; ; \; L_{ii}(\alpha) = 0 \quad (i,j \in E; \; \alpha \in A).$$

Formulate the problem of minimizing expected cost of interception and give sufficient conditions for existence of an optimal control. [Suggestion: show that by an appropriate choice of state space the problem can be reduced to Problem 2.]

PROBLEM 4.

In the problem of §3 suppose that

$$\inf \{p_{i,o}^{(J)}(\alpha) : i \in F, \; \alpha \in A\} > 0$$

Show that

$$\sup \{\mathcal{E}\{\tau_i(\alpha)\} : i \in F, \; \alpha \in A\} < \infty. \qquad (*)$$

Defining an ϵ-optimal control in the natural way, show that if (*) is true then an ϵ-optimal control ($\epsilon > 0$) can always be constructed by approximation in policy space, where we do not assume that the

minima can be achieved.

[Suggestion: Let $\sup\{\mathcal{E}\tau_i(\alpha)\colon i \in F, \alpha \in A\} = T$. Pick $\{\alpha_n\}$ such that

$$\mathcal{L}(\alpha_{n+1})v_n + L(\alpha_{n+1}) \leqq \mathcal{L}(\alpha)v_n + L(\alpha) + \epsilon 2^{-n}T^{-1}\,1, \quad \alpha \in A.$$

By the summation rule

$$v_n = -\,\mathcal{E}\{\sum_{o}^{\tau(\alpha_{n+1})} \mathcal{L}(\alpha_{n+1})v_n[x_n(\alpha_{n+1})]\}$$

$$\geqq -\,\mathcal{E}\{\sum_{o}^{\tau(\alpha_{n+1})} [-L(x_k(\alpha_{n+1}),\alpha_{n+1}) + \mathcal{L}(\alpha_n)v_n[x_k(\alpha_{n+1})] + L(x_k(\alpha_{n+1}),\alpha_n) + \epsilon 2^{-n}\,T^{-1}\,1]\}$$

$$= v_{n+1} - \epsilon 2^{-n}\,T^{-1}\,1\,\mathcal{E}\{\tau(\alpha_{n+1})\}$$

$$\Longrightarrow v_{n+1} \leqq v_n + \epsilon\,2^{-n}\,1\,.$$

Thus $0 \leqq v_n \leqq v_o + \epsilon 1$. Now we claim:

$$\mathcal{L}(\alpha_n)v_n + L(\alpha_n) \leqq \mathcal{L}(\alpha)v_n + L(\alpha) + \epsilon T^{-1}\,1, \quad \alpha \in A \qquad (*)$$

for all n sufficiently large: this shows that α_n is ϵ-optimal for all n sufficiently large. To verify (*) suppose, on the contrary, that for some subsequence $m(\nu) \to \infty$, $i_m \in F$, $\beta_m \in A$,

$$\mathcal{L}(\alpha_m)v_m(i_m)+L_{i_m}(\alpha_m) > \mathcal{L}(\beta_m)v_m(i_m)+L_{i_m}(\beta_m)+\epsilon T^{-1}$$

Now

$$\mathcal{L}(\alpha_{m+1})v_m(i)+L_i(\alpha_{m+1}) \leqq \mathcal{L}(\alpha)v_m(i)+L_i(\alpha)+\epsilon T^{-1}\,2^{-m}$$

for all i, α ; and so

$$\begin{aligned}\mathcal{L}(\alpha_{m+1})v_m(i)+L_i(\alpha_{m+1}) &\leqq \mathcal{L}(\alpha_m)v_m(i)+L_i(\alpha_m)+\epsilon T^{-1}\,2^{-m}\\ &\quad - \epsilon T^{-1}\delta_{i,i_m}\\ &= \epsilon T^{-1}(2^{-m}-\delta_{i,i_m}) \quad (i \in E)\end{aligned}$$

Thus by the summation rule,

$$v_m(i) \geqq v_{m+1}(i) - \epsilon T^{-1}(2^{-m}-\delta_{i,i_m})\cdot\mathcal{E}\{\tau_i(\alpha_{m+1})\} \quad (i \in E)$$

or

$$\begin{aligned}v_{m+1}(i_m) &\leqq v_m(i_m)+\epsilon 2^{-m}-\epsilon T^{-1}\mathcal{E}\{\tau_{i_m}(\alpha_{m+1})\}\\ &\leqq v_m(i_m)+\epsilon 2^{-m}-\epsilon T^{-1},\end{aligned}$$

since $\mathcal{E}\{\tau_{i_m}(\alpha_{m+1})\} \geqq 1$. But then

$$v_{m(\nu+1)}(i_m) \leqq v_{m(\nu)+1}(i_m) + \epsilon \sum_{j=m(\nu)+1}^{m(\nu+1)} 2^{-j}$$

and so

$$v_{m(\nu+1)}(i_m) \leq v_{m(\nu)}(i_m) + \epsilon \sum_{j=m(\nu)+1}^{m(\nu+1)} 2^{-j} + \epsilon 2^{-m} - \epsilon T^{-1} .$$

Since the number of states is finite we can assume (by picking a further subsequence if necessary) that $i_m = k \in F$ for all m. Then, iterating the last inequality.

$$v_{m(\nu)}(k) \leq v_{m(o)}(k) + 2\epsilon - \nu \epsilon T^{-1} \quad ;$$

and since $v_n \geq 0$, we obtain a contradiction.]

Remark. Problem 4 can also be solved by the following alternative argument (due to W-U. Haussmann). Observe that

$$0 \leq v_{n+1} \leq v_n + \epsilon 2^{-n} 1$$

$$\Longrightarrow \quad 0 \leq v_n \leq v_1 + \epsilon 1$$

We shall show that the sequence $\{v_n\}$ converges. Indeed let $\{u_n\}$ be a sequence of scalars such that

$$0 \leq u_{n+1} \leq u_n + \epsilon 2^{-n}$$

Then

$$\begin{aligned} u_{n+p} &\leq u_n + \epsilon\{2^{-n} + 2^{-(n+1)} + \ldots 2^{-(n+p-1)}\} \\ &\leq u_n + \epsilon 2^{-(n-1)}\{2^{-1} + 2^{-2} + \ldots\} \\ &\leq u_n + \epsilon 2^{-n+1} . \end{aligned}$$

Also $\{u_n\}$ has a convergent subsequence $\{u_{m(\nu)}\}$; let

$$\overline{m} = \min \{m(\nu) : m(\nu) \geqq n+p\}$$

$$\underline{m} = \max \{m(\nu) : m(\nu) \leqq n\}$$

Then

$$u_n - u_{n+p} \leqq (u_n - u_{\underline{m}}) + (u_{\underline{m}} - u_{\overline{m}}) + (u_{\overline{m}} - u_{n+p})$$

$$\leqq \epsilon(2^{-\underline{m}} + 2^{-(\underline{m}+1)} + \ldots + 2^{-\overline{m}}) +$$

$$+ (u_{\underline{m}} - u_{\overline{m}}) + \epsilon(2^{-\underline{m}} + \ldots + 2^{-\overline{m}})$$

$$\leqq 4\epsilon 2^{-\underline{m}} + (u_{\underline{m}} - u_{\overline{m}}) .$$

Thus given $\delta > 0$ we can choose ν so large that

$$u_{\underline{m}} - u_{\overline{m}} < \delta/2$$

and

$$4\epsilon 2^{-\underline{m}} < \delta/2 .$$

Then for n,p sufficiently large,

$$|u_{n+p} - u_n| < \delta$$

so that $\{u_n\}$ converges. Thus our vector sequence $\{v_n\}$ actually converges, to v^* say.

The argument then follows that of page 51. By (11), we

have

$$\mathcal{L}(\alpha_{n+1})v_n + L(\alpha_{n+1}) \leqq \inf\{\mathcal{L}(\alpha)v_n + L(\alpha)\} + \epsilon 2^{-n}$$

and then

$$\mathcal{L}^*v^* + L^* \leqq \inf\{\mathcal{L}(\alpha)v^* + L(\alpha)\} .$$

But also we can take the limit in (12), since $\{u_n\}$ converges, so that

$$\mathcal{L}^*v^* + b^* = 0 .$$

Thus, for n sufficiently large,

$$v_n \leqq v + \epsilon T^{-1} \leqq v(\alpha) + \epsilon$$

by the same argument as on page 52.

4. OPTIMAL CONTROL WITH FIXED TERMINAL TIME.

Optimal control of a Markov chain with fixed terminal time is in many ways the simplest problem of the type we have been considering. However, in this section it will be studied under broader assumptions on the class of control laws than have been made heretofore.

First, consider the following standard problem. For $\alpha \in A$ let $X(\alpha)$ be an arbitrary Markov chain on $E = \{1,\ldots,J\}$; and let

$$P(n,\alpha) = [p_{ij}(n,\alpha)]$$

$$p_{ij}(n,\alpha) = p_{ij}(n,\alpha_i)$$

$$= P\{x(n+1,\alpha) = j \mid x(n,\alpha) = i\}$$

where as usual,

$$\alpha = (\alpha_1,\ldots,\alpha_J)$$

$$\alpha_i \in A_i$$

$$A = A_1 \times \ldots \times A_J .$$

Let $L(n, i, \alpha) = L(n, i, \alpha_i) \geq 0$ be the cost incurred by $X(\alpha)$ in state i at time n, and let

$$v(n,i,\alpha) = \mathcal{E}\{ \sum_{m=n+1}^{N} L[m, x_m(\alpha), \alpha] \mid x_{n+1}(\alpha) = i\} \qquad (1)$$

In (1), N is fixed, and an empty sum is defined to be zero.

Drop α temporarily and write $V(n,\cdot,\cdot) = V_n(\cdot)$, etc. One easily verifies that

$$V_n(i) - V_{n-1}(i) = - P_n V_n(i) + V_n(i) - L_n(i). \qquad (2)$$

Introducing the column vectors

$$V = [V(i)], \quad L = [L(i)] ;$$

backward difference operator

$$\Delta V_n \equiv V_n - V_{n-1}$$

and generator

$$\mathcal{L}_n V \equiv P_n V - V$$

we can write the basic relation

$$\begin{aligned} \Delta V_n + \mathcal{L}_n V_n &= -L, \quad 1 \leq n \leq N \\ V_n &= 0 \end{aligned} \tag{3}$$

Since (3) has a unique solution, it follows that (1) and (3) are equivalent, if we identify

$$\mathcal{L}_n = \mathcal{L}_n(\alpha), \quad L_n = L_n(\alpha) .$$

Conversely, if V_n $(1 \leq n \leq N)$ is an arbitrary sequence of J-vectors with $V_N = 0$, then we have the summation rule

$$V_n(i) = -\mathcal{E}\{\sum_{m=n+1}^{N} [\Delta V_m(x_m^\alpha) + \mathcal{L}_m^\alpha V_m(x_m^\alpha)] \mid x_{n+1}^\alpha = i\} \tag{4}$$

We now define a <u>control law</u> to be a sequence $\bar{\alpha} = \{\alpha(n) : 1 \leq n \leq N\}$ where each $\alpha(n) \in A$. The foregoing relations remain valid with the replacements

$$V_n(\alpha) = V_n(\bar{\alpha})$$

$$L_n(\alpha) = L_n(\alpha_n)$$

$$P_n(\alpha) = P_n(\alpha_n)$$

$$\mathcal{L}_n(\alpha) = \mathcal{L}_n(\alpha_n)$$

(Here $\alpha_n = \alpha(n)$). We shall use the notation $L_n(\alpha_n) = L_n^{\alpha(n)}$, etc., according to convenience; and write $\bar{A}$ for the class of controls $\bar{\alpha}$.

Now consider the problem of minimizing $V_o(\bar{\alpha})$ with respect to $\bar{\alpha}$. The relevant optimality theorem is obvious from (3) and (4): namely, if $\bar{\alpha}^* \in \bar{A}$ and a vector V^* have the properties

$$\text{(i)} \quad \Delta V_n^* + \mathcal{L}(\alpha_n^*)V_n^* = -L_n(\alpha_n^*)$$

$$\text{(ii)} \quad \mathcal{L}_n(\alpha_n^*)V_n^* + L_n(\alpha_n^*) \leq \mathcal{L}_n(\alpha_n)V_n^* + L_n(\alpha_n) \qquad (5)$$

for all $0 \leq n \leq N$ and all $\bar{\alpha} \in \bar{A}$, then

$$V_n^* \leq V_n \quad , \quad n = 0,\ldots,N \ .$$

In fact, by (4) and (5),

$$V_n^*(i) = -\mathcal{E}\{\sum_{m=n+1}^{N} (\Delta V_m^* + \mathcal{L}_m^{\alpha} V_m^*)(x_m^{\alpha}) \,|\, x_{n+1}^{\alpha} = i\}$$

$$\leq -\mathcal{E}\{\sum_{m=n+1}^{N} (\Delta V_m^* + \mathcal{L}_m^* V_m^* + L_m^* - L_m^*)(x_m^{\alpha}) \,|\, x_{n+1}^{\alpha} = i\}$$

$$= V_n^{\alpha}(i) \ .$$

Thus it is clear that at any stage n, we may seek to determine α_n^* by minimizing

$$\mathcal{L}_n(\alpha_n)V_n^* + L_n(\alpha_n)$$

with respect to $\alpha_n \in A$. Assuming the minimum exists at each stage, an α_n^* and V_n^* are determined by iteration backwards from the terminal time, according to

$$\begin{aligned} V_N^* &= 0 \\ \min \{\Delta V_n^* + \mathcal{L}(\alpha_n)V_n^* + L_n(\alpha_n) &: \alpha_n \in A\} = 0 \end{aligned} \tag{6}$$

EXERCISE: Work out an example for a Markov chain with two states, and carry through the computations explicitly for $N = 3$.

We now look at some qualitative features of problems of this type. It has been assumed throughout that the component of control α_n at time n depends only on the current state x_n of the process. Suppose we broaden the class of control laws to include those which allow the control at the 'present' time n to depend on $x_0,\ldots,x_n$; that is, on the entire past history of the process. To make the extension consider a class of stochastic processes $X = X(\beta) = \{x_0(\beta),\ldots,x_N(\beta)\}$, where:

$$\beta = \{\beta_0,\ldots,\beta_N\}$$

$$\beta_n : \underbrace{E \times \ldots \times E}_{n+1} \to A$$

and if β, γ are such that

$$\beta_\nu = \gamma_\nu , \quad \nu = 0,1,\ldots,n$$

then

$$P\{x_{n+1}^{\beta} = i \mid x_0^{\beta},\ldots,x_n^{\beta}\} = P\{x_{n+1}^{\gamma} = i \mid x_0^{\gamma},\ldots,x_n^{\gamma}\}$$

for $i \in E$.

In plain language the definitions state that a control law is a sequence of mappings $\{\beta_n\}$, such that the n^{th} of these maps the sequence $(x_0^{\beta},\ldots,x_n^{\beta})$ into the set of allowable control vectors A; and that transition probabilities of the X^{β} process at time n (that is, from the n^{th} state to the $(n+1)^{th}$ state) depend only on the values assumed by the 'past' random variables $x_0^{\beta},\ldots,x_n^{\beta}$.

Let

$$\begin{aligned} &P\{x_{n+1}^{\beta} = j \mid x_0^{\beta} = i_0,\ldots,x_n^{\beta} = i_n\} \\ &\qquad = p_{ij}[n,\beta_{n,i}(i_0,\ldots,i_n)] \end{aligned} \tag{7}$$

where $p_{ij}(n,\cdot)$ is the transition probability introduced earlier (p.57), and $\beta_{n,i}(\cdot)$ is the i^{th} component of $\beta_n(\cdot)$. If $\mathcal{B}$ is the class of all possible controls β, then $\mathcal{B} \supset \bar{A}$ because, if $\bar{\alpha} \in \bar{A}$ and

$$\alpha(n) = \alpha' \in A ,$$

we need only define

$$\beta_n(i_o,\ldots,i_n) = \alpha'_{i_n} \quad . \tag{8}$$

With $L = (L_i(\alpha_i))$ as before, we say that $\beta^* \in \mathcal{B}$ is optimal if

$$\mathcal{E}\{ \sum_{m=n+1}^{N} L[m, x_m^\beta, \beta_m(x_o^\beta,\ldots,x_m^\beta)] \,|\, x_o^\beta,\ldots,x_{n+1}^\beta\} \tag{9}$$

is a minimum (with respect to $\beta \in \mathcal{B}$) for every assignment of values to $x_o^\beta,\ldots,x_{n+1}^\beta$ and for all n, $0 \leq n \leq N$. We denote the functional (9), evaluated at $x_\nu^\beta = i_\nu (\nu = 0,\ldots,n+1)$ by

$$V_n^\beta(i_o,\ldots,i_{n+1}) \quad .$$

We shall now show that if $\alpha^* \in \bar{A}$ is optimal in $\bar{A}$ then α^* is optimal in $\mathcal{B}$. This is easily achieved by noting that the summation rule (4) is true even for processes of type X^β. More precisely, let

$$P_n(\beta) = [p_{ij}(n,\beta_{n,i})] \qquad i,j \in E$$

$$\mathcal{L}_n^\beta = \mathcal{L}_n(\beta) = P_n(\beta) - I$$

Then if $\{V_n,\ 0 \leq n \leq N\}$ is an arbitrary sequence of vectors with

$V_N = 0$, we have

$$V_n(i) = -\mathcal{E}\{\sum_{m=n+1}^{N} [\Delta V_m(x_m^\beta) + \mathcal{L}_m^\beta V_m(x_m^\beta)] \mid x_o^\beta = i_o, \ldots, x_n^\beta = i_n, x_{n+1}^\beta = i\} \quad (10)$$

for all $i_o, \ldots, i_n$, $i \in E$. To verify (10) write $x_m^\beta = x_m$ and observe that, for $m \geq n+1$,

$$\mathcal{E}\{\mathcal{L}_m^\beta V_m(x_m) \mid x_o, \ldots, x_n, x_{n+1} = i\}$$

$$= \mathcal{E}\{\mathcal{E}\{V_m(x_{m+1}) \mid x_o, x_1, \ldots, x_{m-1}, x_m\} - V_m(x_m) \mid x_o, \ldots, x_n, x_{n+1} = i\}$$

$$= \mathcal{E}\{V_m(x_{m+1}) - V_m(x_m) \mid x_o, \ldots, x_n, x_{n+1} = i\} .$$

Furthermore

$$\Delta V_m(x_m) = V_m(x_m) - V_{m-1}(x_m) .$$

Thus the right side of (10) becomes (dropping evaluations of $x_o, \ldots, x_n$)

$$-\mathcal{E}\{\sum_{m=n+1}^{N} V_m(x_m) - V_{m-1}(x_m) + V_m(x_{m+1}) - V_m(x_m) \mid x_o, \ldots, x_n, x_{n+1} = i\}$$

$$= -\mathcal{E}\{\sum_{n}^{N-1} -V_m(x_{m+1}) + \sum_{m=n+1}^{N} V_m(x_{m+1}) \mid x_o, \ldots, x_n, x_{n+1} = i\}$$

$$= V_n(i) .$$

[Note: in these computations, a conditional expectation $\mathcal{E}\{f|x\}$, where x is a random variable with values in E, is itself a random variable, defined by forming the function

$$\phi(i) = \mathcal{E}\{f|x = i\}$$

and then setting

$$\mathcal{E}\{f|x\} = \varphi(x) .$$

Thus $\mathcal{E}\{\mathcal{E}\{f|x,y\},x\} = \mathcal{E}\{f|x\}$ since

$$\mathcal{E}\{\mathcal{E}\{f|x,y\},x = i\}$$

$$= \sum_{j\in E} \mathcal{E}\{f|x = i, y = j\} P\{y = i|x = i\}$$

$$= \mathcal{E}\{f|x = i\} .]$$

It is now easy to show that optimality of $\bar{\alpha}^* \in A$ implies optimality in $\mathcal{B}$. In fact if $\bar{\alpha}^* \in \bar{A}$ is optimal then either (5) holds or there exists an optimal $\bar{\alpha}^{**} \in \bar{A}$ for which (5) holds (Exercise), and then the proof that $\bar{\alpha}^*$ (or $\bar{\alpha}^{**}$) is optimal is carried out by applying (10) in the same way that (4) was used in the discussion, p. 64.

In the language of control, this result is expressed by

the statement that, if the controller has perfect information about the current state of the controlled object, then it is sufficient for optimality to generate control decisions as functions of the current state alone. Results of this kind are important because they serve to limit a priori the amount of data-processing required of the control computer.

EXERCISE: Write out the details of the foregoing argument.

Problem.

Let $X(\alpha)$ be a family of homogeneous chains which satisfies the conditions of §2. Let $L(\alpha)$ be bounded for $\alpha \in A$, let α^* be an optimal stationary control, and $\lambda^* = \lambda(\alpha^*)$. Write

$$\bar{\alpha} = \{\alpha_0, \alpha_1, \ldots\}$$

where the $\alpha_n \in A$; let $\{x_n^{\bar{\alpha}}\}$ be the process defined by setting

$$p\{x_{n+1}^{\bar{\alpha}} = j \mid x_n^{\bar{\alpha}} = i\} = p_{ij}(\alpha_n(i)) ;$$

and define

$$V_N(i) = \mathcal{E}\{\sum_{n=0}^{N-1} L(x_n^{\bar{\alpha}}, \alpha_n) \mid x_0 = i\} .$$

As a measure of long-time average cost using $\bar{\alpha}$, let

$$\lambda_i(\bar{\alpha}) = \lim_{n \to \infty} N^{-1} V_N(i) \ .$$

Show that α^* is as good as any $\bar{\alpha}$, in the sense that

$$\lambda^* \leq \lambda_i(\bar{\alpha}) \quad (\text{all } i, \bar{\alpha}) \ .$$

REFERENCES

[1] Bellman, R., "A Markovian decision process", J. Math. Mech., 6(1957), pp. 679-684.

[2] Bellman, R., Adaptive Control Processes, A Guided Tour, Princeton University Press, 1961.

[3] Birkhoff, G. and S. MacLane, A Survey of Modern Algebra, MacMillan, New York, 1953.

[4] Blackwell, D., "On the functional equation of dynamic programming", J. Math. Anal. and Appl., 2(1961), pp. 273-271.

[5] Blackwell, D., "Discontinuous dynamic programming", Ann. Math. Stat., 36(1)1965, pp. 226-235.

[6] Chung, K.L., Markov Chains with Stationary Transition Probabilities, Berlin, Springer-Verlag, 1960.

[7] Cox, D.R. and H.D. Miller, Theory of Stochastic Processes, New York, Wiley, 1965.

[8] Derman, C., "On sequential decisions and Markov chains," Management Science, 9(1)1962, pp. 16-23.

[9] Derman, C., "Stable sequential control rules and Markov chains," J. Math. Anal. and Appl., 6(2)1963, pp. 257-265.

[10] Derman, C., "On sequential control processes", Ann. Math. Stat., 35(1)1964, pp. 341-349.

[11] Doob, J.L., Stochastic Processes, New York, Wiley, 1953.

[12] Dynkin, E.B., "Controlled random sequences," Theory of Prob. and Its Appl., 10(1)1965, pp. 3-18.

[13] Eaton, J.H., and L.A. Zadeh, "Optimal pursuit strategies in discrete-state probabilistic systems," J. Basic Engrg., March, 1962, pp. 23-29.

[14] Howard, R.A., Dynamic Programming and Markov Processes, New York, Wiley, 1960.

[15] Howard, R.A., "Semi-Markovian decision processes", Bull. Inst. Internat. Stat., 40(1963), pp. 625-652.

[16] Kemeny, J.G., and J.L. Snell, Finite Markov Chains, New York Van Nostrand, 1960.

[17] Krylov, N.V., "Construction of an optimal strategy for a finite controlled chain," Th. of Prob. and Its Appl., 10(1)1965, pp. 45-54.

[18] Raviv, J., "Decision making in incompletely known stochastic systems," Int. J. Engrg. Sci. 3(1965), pp. 119-140.

[19] Romanovskii, I.V., "The existence of an optimal stationary policy in a Markovian decision process," Teor. Veroyat. i ee Prim., 10(1)1965, pp. 130-133.

APPENDIX. A Property of Directed Graphs.

We show that if $N = \min \{n : p_{ij}^{(n)} > 0, \text{ all } i,j,\}$ then

$$N \leq (J-1)^2 + 1 .$$

In other words, a strongly connected, aperiodic directed graph is "fully saturated" after at most $(J-1)^2 + 1$ transitions, J being the number of nodes.

For the proof, let N (instead of J) denote number of nodes (states) and let nodes be labelled $\{1,\ldots,N\}$. Denote state space (set of nodes) by E. Let

$$\varphi : 2^E \to 2^E$$

be the "successor function":

$$\left.\begin{aligned} \varphi(i) &= \{j : p_{ij} > 0\} . \\ F \subset E \Longrightarrow \varphi(F) &= \{j : (\exists i)\ i \in F,\ p_{ij} > 0\}. \end{aligned}\right\}$$

LEMMA.

If $F \subset E$ and $\varphi^{(n)}(F) = F$ for some n, then $F = E$.

PROOF. We have

$$i \in F = \sum_{j \in F} p_{ij}^{(n)} = 1 \implies \sum_{j \in F} p_{ij}^{(kn)} = 1, \quad k = 1,2,\ldots .$$

But since all states are aperiodic, $p_{ij}^{(r)} > 0$ for all r sufficiently large. Hence $F = E$.

Q.E.D.

Since the graph is aperiodic there is at least one node (1, say) which has two distinct nodes among its (one-step) successors. Consider the sequence

$$1, \varphi(1), \varphi^2(1), \ldots .$$

Since the graph is strongly connected we have $1 \in \varphi^n(1)$ for some n. Let

$$n_1 = \min \{n : n \geq 1, \ 1 \in \varphi^n(1)\} .$$

We have $n_1 \leq N-1$. In fact, if $1 \in \varphi(1)$ then $n_1 = 1$; otherwise, there are $i,j \in \varphi(1)$ with $1 \neq i \neq j \neq 1$, and loop-free paths of lengths $\leq N-1$ from i to 1 and from j to 1. If a loop-free path from i to 1 is of length $N-1$, then this path contains j, --hence there is a loop-free path from j to 1 of length $\leq N-2$; that is, there is a loop-free path from 1 to 1 of length $\leq N-1$.

Suppose $\varphi^{n_1}(1) \neq E$. By the lemma, there is a node (2, say) in $\varphi^{n_1}(1)$ which is distinct from 1. Let

$$n_2 = \min \{n : n \geq 1, \varphi^{n_1}(1) \subset \varphi^{n_1+n}(1)\} .$$

Again, if $\varphi^{n_1+n_2}(1) \neq E$ there is a node (3, say) in $\varphi^{n_1+n_2}(1) - \varphi^{n_1}(1)$. Also, it is clear that $n_2 \leq n_1$. Continuing thus we obtain sequences

$$N-1 \geq n_1 \geq n_2 \geq \ldots \geq n_{N-2}$$

$$1 \in \varphi^{n_1}(1) \subset \varphi^{n_1+n_2}(1) \subset \ldots \subset \varphi^{n_1+\ldots+n_{N-2}}(1)$$

with

$$\varphi^{n_1+\ldots+n_{N-2}}(1) \supset E - \{N\} \quad , \text{ say.}$$

We observe next that (if $N \geq 3$) then $\varphi^{n_1+1}(1)$ contains at least three distinct nodes: otherwise all nodes in $\varphi^{n_1}(1)$ are mapped into the (two) successors of 1: i.e.,

$$\varphi^{n_1}(\varphi(1)) = \varphi(\varphi^{n_1}(1)) = \varphi(1)$$

and by the lemma, $\varphi(1) = E$. Similarly (if $N \geq 4$) $\varphi^{n_1+n_2+1}(1)$ contains at least four distinct nodes. Otherwise all nodes in $\varphi^{n_1+n_2}(1)$ are mapped into (three) successors of $\varphi^{n_1}(1)$; i.e.,

$$\varphi^{n_2}(\varphi^{n_1+1}(1)) = \varphi(\varphi^{n_1+n_2}(1)) = \varphi^{n_1+1}(1)$$

and by the lemma, $\varphi^{n_1+1}(1) = E$. Continuing by induction we get finally

$$\varphi^{n_1+\ldots+n_{N-2}+1}(1) = E \ .$$

To complete the proof we note that

$$\varphi(1) \cup \varphi^2(1) \cup \ldots \cup \varphi^{N-1}(1) = E$$

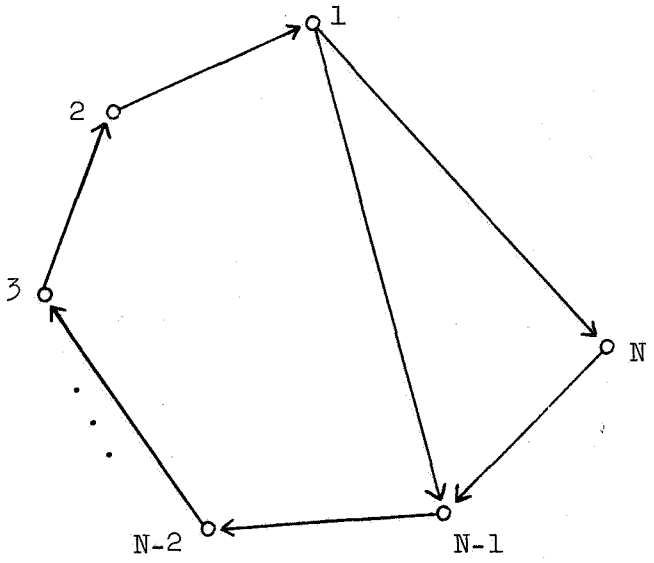
and that $\varphi(E) = E$. Thus we shall have

$$\varphi^n(i) = E \qquad \text{(all } i\text{)}$$

provided

$$\begin{aligned} n &= n_1 + n_2 + \ldots + n_{N-2}+1 + (N-1) \\ &\leqq (N-2)(N-1) + 1 + (N-1) \\ &= (N-1)^2 + 1 \ . \end{aligned} \qquad (*)$$

EXAMPLE:



The upper bound (*) is attained.

CHAPTER II

INTRODUCTION TO GAUSSIAN DIFFUSIONS

1. PRELIMINARIES.

In this chapter we discuss Brownian notion and the control of (Gaussian) diffusion processes determined by linear stochastic differential equations of Itô's type. A more general treatment of Itô's equation is deferred to a later chapter. In fact, nearly all the concrete results of stochastic control theory which are available at present pertain to Gaussian processes only.

In the following, we shall be concerned with stochastic processes $X = \{x(t,\omega);\ 0 \leq t < \infty\}$ defined relative to a probability triple $(\Omega,\mathfrak{F},P)$; the sample functions $x(\cdot,\omega)$ of X are defined on the halfline $[0,\infty)$ and take values $x(t,\omega)$ in Euclidean n-space E. We assume that P-measure is complete: i.e., $\mathfrak{F}$ contains all subsets of sets of P-measure zero.

Certain results of measure theory will be taken for granted; for proofs the reader will be referred to standard textbooks.*

To discuss processes X with continuous time (t) parameter a number of technical difficulties must be dealt with: these arise from the practical necessity of defining our processes in such a way that (a) various sample funtion properties of interest (such as continuity) be events, i.e., have assignable probabilities, and (b) the sample functions of X be sufficiently regular to admit

---

*A bibliographical note and references are given at the end of the chapter.

construction of new random variables by operations on them involving limits, such as integration. Briefly stated, requirements (a) and (b) are met by introducing respectively the concepts of separability and measurability of X.

1.1 Separability.

Let T be a finite or denumerable subset of $I = [0,\infty)$, let $\lambda \in R_1$, and consider the event

$$\{\omega : \sup_{t\in T} |x_t(\omega)| > \lambda\} = \bigcup_{t\in T} \{\omega : |x(t,\omega)| > \lambda\}. \qquad (1)$$

As an (at least) countable union of ω-sets in $\mathfrak{F}$, the ω-set on the right is an event (i.e., is measurable, i.e., belongs to $\mathfrak{F}$). However, if the set T is nondenumerable (e.g., is an interval) the ω-set described is not necessarily measurable, and thus its probability measure need not be defined. This difficulty is avoided if X has the following property.

Definition: The process $X = \{x(t,\omega)\}$ is separable if there exists a countable everywhere dense subset $T \subset I$, and a set $N \in \mathfrak{F}$ with $P(N) = 0$, such that, for every open set $G \subset I$ and every closed set $F \subset E$, the sets

$$\{\omega : x(t_j,\omega) \in F, \quad t_j \in G \cap T\}$$

and

$$\{\omega : x(t,\omega) \in F, \quad t \in G\}$$

differ only by a subset of N.

The set T is a separability set for X.

Of course, X as given may not be separable: nevertheless it is possible under very general conditions to replace X by a new process $\tilde{X}$ such that $\tilde{X}$ is separable and differs only "slightly" from X.

Definition: The processes X, $\tilde{X}$ are (stochastically) equivalent if for each $t \in I$,

$$P\{x(t,\omega) = \tilde{x}(t,\omega)\} = 1 \qquad (2)$$

[Note that (2) is much weaker that the condition

$$P\{x(t,\omega) = \tilde{x}(t,\omega) \ , \quad t \in I\} = 1$$

since the exceptional ω-set of measure zero involved in (2) may depend on t, and a nondenumerable union of null sets may, if measurable, have positive measure.] Thus if X and $\tilde{X}$ are equivalent they have the same finite-dimensional distributions (the converse statement is in general false!).

THEOREM 1 Let $X = \{x(t,\omega), t \in I\}$ be defined on $(\Omega, \mathfrak{F}, P)$ with values $x(t,\omega) \in E$. There is then a stochastic process $\tilde{X} = \{\tilde{x}(t,\omega), t \in I\}$ defined on $(\Omega, \tilde{\mathfrak{F}}, P)$ and with values

in $E \cup \{\infty\}$ such that $\tilde{X}$ is separable and is equivalent to X.

PROOF. See Doob(1953) Ch.2, §2; Loève(1963) 35.2; Gikhman and Skorokhod(1965) Ch.4, §2.

Note. ∞ is an ideal element adjoined to E and identified with the class of all sequences $\{x_n\}$ of points of E such that $|x_n| \to \infty$ $(n \to \infty)$. For the theory of such a "one-point compactification" see Loève(1963)p.71 or Kelley(1955)p.149. In spite of this apparant enlargement of E, it is true (by equivalence) that, for each t, $\tilde{x}(t,\cdot)$ is finite-valued with probability 1.

Theorem 1 gives no information on how a separability sequence can actually be chosen. We will now show that if X has a mild regularity property then any dense sequence $\{t_n\}$ will do (for instance: the rationals).

Definition: The process X is stochastically continuous at $t_o \in I$ if for every $\epsilon > 0$,

$$P\{|x(t)-x(t_o)| > \epsilon\} \to 0$$

as $t \to t_o$, $t \in I$. [i.e., $x(t) \to x(t_o)$ in pr.; $P\lim_{t \to t_o} x(t) = x(t_o)$].

Definition: X is stochastically continuous if it is stochastically continuous for each $t \in I$.

THEOREM 2 Let X be separable and stochastically continuous. Then any sequence $T = \{\tau_j\}$, everywhere dense in I, is a separability set for X.

PROOF. (Gikhman and Skorokhod (1965) p.207).

Let $V = \{S\}$ be the set of all intervals S in I with rational endpoints: thus V is countable.

Let $J = \{t_j\}$ be a separability set for X and let N be the exceptional ω-set of the separability definition. Let $B(S,\omega)$ denote the closure of the set

$$\{x(\tau,\omega) \ : \ \tau \in T \cap S\} \qquad (\subset E)$$

and, for $t_k \in J \cap S$ let

$$N(S,k) = \{\omega : x(t_k,\omega) \notin B(S,\omega)\} .$$

We show that $P\{N(S,k)\} = 0$. In fact, let τ'_r, $r = 1,2,\ldots$ be an arbitrary sequence in $T \cap S$ with $\tau'_r \to t_k$. Then

$$N(S,k) \subset \{\omega : \varlimsup_{r \to \infty} |x(t_k,\omega)-x(\tau'_r,\omega)| > 0\}$$

and if A denotes the ω-set on the right,

$$P(A) = \lim_{n \to \infty} P\{\omega : \varlimsup_{r \to \infty} |x(t_k,\omega)-x(\tau'_r,\omega)| > n^{-1}\}$$

$$\leqq \lim_{n \to \infty} \varlimsup_{r \to \infty} P\{\omega : |x(t_k,\omega)-x(\tau'_r,\omega)| > n^{-1}\}$$

$$= 0, \text{ by stochastic continuity.}$$

Since P-measure is complete, $N(S,k)$ is measurable and $P\{N(S,k)\} = 0$. Let

$$N' = \bigcup_{S \in V} \bigcup_{t_k \in S \cap J} N(S,k) ;$$

then $P(N') = 0$. Let $G \subset I$ be open; let $\omega \notin N \cup N'$, and for some closed set F let $x(\tau,\omega) \in F$ for all $\tau \in T \cap G$. Then for all $t_k \in G \cap J$ and S, such that

$$t_k \in S \subset G ,$$

we have

$$x(t_k,\omega) \in B(S,\omega) \subset F .$$

Since $J = \{t_k\}$ is a separability set, we have $x(t,\omega) \in F$ for all $t \in G$ and $\omega \notin N \cup N'$. Hence T is a separability set for X.

EXERCISE: If $\{A_n\}$ is an arbitrary sequence of measurable sets

show that

$$P\{\underline{\lim}\, A_n\} \leq \underline{\lim}\, P\{A_n\} .$$

If $\{f_n\}$ is an arbitrary sequence of real-valued random variables and $f = \underline{\lim}\, f_n$, show that

$$P\{f > 0\} \leq \underline{\lim}\, P\{f_n > 0\} .$$

State and prove the corresponding facts about $\overline{\lim}$.

EXERCISE: If X is a separable, stochastically continuous process in R_1, and if $0 \leq a < b \leq \infty$, show that

$$\sup\{x(t,\omega): a \leq t \leq b\} = \sup\{x(t,\omega): a \leq t \leq b,\ t \text{ rat.}\}$$

an

$$\inf\{x(t,\omega): a \leq t \leq b\} = \inf\{x(t,\omega): a \leq t \leq b,\ t \text{ rat.}\}$$

with probability one. ("rat." means "rational").

In many applications it is of interest to consider variables such as

$$\int_a^b x(t,\omega)dt$$

and to assign to them probability distributions. For the integral to make sense, it is necessary that $x(\cdot,\omega)$ be a measurable function of t; if the integral itself is to be a random variable we must further require that $x(\cdot,\cdot)$ be measurable in t and ω. Let $\mathcal{L}$ denote the σ-algebra of Lebesque measurable sets of $[0,\infty)$.

Definition: The process $X = \{x(t,\omega): t \geq 0\}$ is measurable if the function $x(t,\omega)$ is $\mathcal{L} \times \mathfrak{F}$-measurable.

Just as in the case of separability, a process X need not be measurable; but under mild restrictions there exists a measurable process $\tilde{X}$ which is stochastically equivalent to X; and if X is separable, $\tilde{X}$ can be chosen to be separable as well.

THEOREM 3 Let $X = \{x(t,\omega): t \geq 0\}$ be a stochastically continuous process in E. Then there exists a separable, measurable process $\tilde{X} = \{x(t,\omega): t \geq 0\}$, defined on the same probability triple as X, and with values in $E \cup \{\infty\}$, such that $\tilde{X}$ is stochastically equivalent to X.

For the proof, which is somewhat lengthy and technical, we refer to Doob (1953), Ch.5,§2, or Gikhman and Skorokhod (1965), Ch.4,§3.

The following result is an immediate consequence of Fubini's theorem on repeated integration [Halmos(1950)p. 148; Berberian (1965), p. 140].

THEOREM 4 Let X be measurable. If $0 \leqq a < b \leqq \infty$ and

$$\int_a^b \mathcal{E}\{|x(t,\omega)|\}dt < \infty$$

then

$$\int_a^b \mathcal{E}\{x(t,\omega)\}dt = \mathcal{E}\{\int_a^b x(t,\omega)dt\} \quad .$$

2. BROWNIAN MOTION.

Let $X = \{x(t,\omega): t \geqq 0\}$. We often write $x(t)$ for $x(t,\cdot)$.

Definition: X is a process with independent increments if, for arbitrary $t_1 < t_2 < \ldots < t_n$ $(n > 3)$, and $t_i \in I$, the random variables

$$x(t_2)-x(t_1), \ \ldots \ , x(t_n)-x(t_{n-1})$$

are independent.

Definition: X is a Brownian motion in E_1 if (i) X is a process with independent increments, (ii) the random variables $x(t)-x(s)$ are real-valued and normally distributed, with

$$\mathcal{E}\{x(t)-x(s)\} = 0$$

$$\mathcal{E}\{[x(t)-x(s)]^2\} = |t-s| \quad .$$

The Brownian motion process is often called a Wiener process.

Note: To define the process initially we assume

$$P\{x(0) = 0\} = 1 .$$

EXERCISE 1: Show that Brownian motion is stochastically continuous.

EXERCISE 2: Show that $\mathcal{E}\{x(t)x(s)\} = t \wedge s$ and compute the joint distribution of $x(t_1),\ldots,x(t_n)$.

We will next show that the sample functions of a separable Brownian motion are almost all continuous.

LEMMA 1 [Doob(1953)p.54]

Let X be separable and stochastically continuous. Let $0 \leq a < b < \infty$ and suppose that, for each n,

$$a \leq s_0^{(n)} < \ldots < s_{j(n)}^{(n)} \leq b$$

and

$$\lim_{n \to \infty} \sup_{a \leq t \leq b} \min_{j \leq j(n)} |t - s_j^{(n)}| = 0$$

Thus, if X is real,

$$P\{ \lim_{n \to \infty} \min_{j} x(s_j^{(n)}, \omega) = \inf_{a \leq t \leq b} x(t,\omega)\} = 1$$

and

$$P\{\lim_{n\to\infty} \max_{j} x(s_j^{(n)},\omega) = \sup_{a\leq t\leq b} x(t,\omega)\} = 1$$

PROOF. Let $t \in [a,b]$ be fixed. For each n, we can choose j so that $s_n \equiv s_j^{(n)} \to t$ $(n\to\infty)$. Since X is stochastically continuous,

$$x(t) = P \lim_{n\to\infty} x(s_n) \quad .$$

Hence [Loève(1963)p.116] there is a subsequence $\{n'\}$ of values of n such that

$$P\{\lim_{n\to\infty} x(s_{n'}) = x(t)\} = 1 \ .$$

We can assume that the sets $\{s^{(n)}\}$ are increasing with n. Then

$$P\{x(t) \leq \varliminf_{n\to\infty} \max_{k\leq j(n)} x(s_k^{(n)})\} = 1 \ .$$

Now let $\{t_r\}$ be a separability sequence; then (with t_r playing the role of t in the inequality above)

$$\sup_{a\leq t\leq b} x(t,\omega) \leq \sup_{t_r\in[a,b]} x(t_r,\omega)$$

$$\leq \varliminf_{n\to\infty} \max_{k\leq j(n)} x(s_j^{(n)},\omega)$$

with probability 1. But the reverse inequality (with $\overline{\lim}$ instead of $\underline{\lim}$) is obvious, and this proves the second equality. The first is proved in the same way.

LEMMA 2 [Doob(1953)p.106]

Let $y_1,\dots,y_n$ be independent real variables and let $x_j = y_1 + \dots + y_j$. Then, if x_n-x_1, $x_n-x_2,\dots$ have symmetric distributions,

$$P\{\max_{j\leqq n} x_j \geqq \lambda\} \leqq 2P\{x_n \geqq \lambda\} .$$

PROOF. Clearly

$$P\{\max_{j\leqq n} x_j \geqq \lambda,\ x_n \geqq \lambda\} = P\{x_n \geqq \lambda\} . \qquad (1)$$

But by symmetry and independence, if x_ν is the first x_j, if any, which is $\geqq\lambda$, then

$$\begin{aligned} &P\{\max_{j\leqq n} x_j \geqq \lambda,\ x_n < \lambda\} \\ &= \sum_{k=1}^{n-1} P\{\nu = k,\ x_n < \lambda\} \\ &\leqq \sum_{k=1}^{n-1} P\{\nu = k,\ x_n-x_k < 0\} \\ &= \sum_{k=1}^{n-1} P\{\nu = k\}P\{x_n-x_k < 0\} \end{aligned}$$

$$= \sum_{k=1}^{n-1} P\{\nu = k,\ x_n - x_k > 0\}$$

$$\leqq \sum_{k=1}^{n-1} P\{\nu = k,\ x_n \geqq \lambda\} \leqq P\{x_n \geqq \lambda\} \qquad (2)$$

Adding (1) and (2), the result follows.

LEMMA 3 [Doob (1953)p.392]

If X is a separable Brownian motion in E_1 then

$$P\{\sup_{0\leqq t\leqq T} [x(t)-x(0)] \geqq \lambda\} \leqq 2P\{x(T)-x(0) \geqq \lambda\} \leqq \lambda^{-1}\sqrt{\frac{2T}{\pi}}\, e^{-\lambda^2/2T} \qquad (3)$$

PROOF. Let $0 = t_0 < \ldots < t_n = T$. Then by Lemma 2

$$P\{\max_{0\leqq n} [x(t_j)-x(0)] \geqq \lambda\}$$

$$\leqq 2P\{x(T)-x(0) \geqq \lambda\}$$

$$= 2\int_{\lambda}^{\infty} \frac{e^{-\xi^2/2T}}{\sqrt{2\pi\, T}}\, d\xi$$

$$\leqq \lambda^{-1}\sqrt{\frac{2T}{\pi}}\, e^{-\lambda^2/2T}$$

By Lemma 1, the left side of (3) is obtained by choosing successively denser sequences $t_0,\ldots,t_n$ and going to the limit. This completes the proof.

THEOREM 1 [Doob (1953) p.393]

Almost all sample functions of a separable Brownian motion are continuous.

PROOF. Let $t \in [0,\infty)$. We show

$$|x(t)-x(j/N)| \leq N^{-1/4} \quad \text{if} \quad |t-j/N| \leq 1/N, \; j \leq N^2 \qquad (4)$$

for sufficiently large N, with probability one. Using Lemma 3,

$$P\{\sup_{\substack{|t-j/N|\leq 1/N \\ 1\leq j\leq N^2}} |x(t)-x(j/N)| \geq N^{-1/4}\}$$

$$\leq N^2 P\{\sup_{|t-1/N|\leq 1/N} |x(t)-x(1/N)| \geq N^{-1/4}\}$$

$$\leq 2N^2 P\{\sup_{0\leq t\leq 1/N} |x(t)-x(0)| \geq N^{-1/4}\}$$

$$\leq 2N^2 \cdot N^{1/4} \cdot \sqrt{\frac{2}{\pi}\cdot\frac{1}{N}} \cdot e^{-N^{-1/2}/2} \cdot \frac{1}{N}$$

$$= 2\sqrt{\frac{2}{\pi}}\, N^{7/4} e^{-\sqrt{N}/2}$$

Since the last expression is the N^{th} term of a convergent series the Borel-Cantelli Lemma implies that the event whose probability is being measured occurs at most a finite number of times, (as $N \to \infty$) with probability one. Thus (4) is true, and states that almost all sample functions are uniformly continuous in every finite t-interval.

Q.E.D.

The sample functions of a Brownian motion are highly irregular in behavior:

THEOREM 2 Almost all Brownian paths are nowhere differentiable.

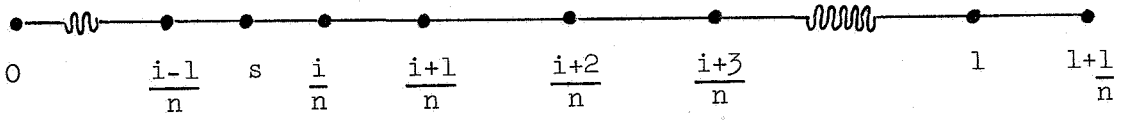
PROOF. (Dvoretski, Erdos, Kakutani)* If a path is differentiable at some point s, $0 \leq s \leq 1$ then

$$|x(t)-x(s)| \leq \ell(t-s) \ , \ s < t < s + \frac{5}{n} \ , \ n \geq m \ ,$$

for some $\ell \geq 1$ and some $m \geq 1$. Claim: this event is a subset of the event

$$\bigcup_{\ell \geq 1} \bigcup_{m \geq 1} \bigcap_{n \geq m} \bigcup_{0 < i \leq n+2} \bigcap_{i < k \leq i+3} \{\omega : |x(\tfrac{k}{n},\omega)-x(\tfrac{k-1}{n},\omega)| < 7\ell/n\}$$

[This is clear from the picture:



$$i < k \leq i+3 \Longrightarrow |x(\tfrac{k}{n})-x(\tfrac{k-1}{n})| \leq |x(\tfrac{k}{n})-x(s)| + |x(\tfrac{k-1}{n})-x(s)|$$

$$\leq \ell(\tfrac{k-i+1}{n}) + \ell(\tfrac{k-1-i+1}{n})$$

$$\leq 4\ell/n + 3\ell/n \qquad]$$

* Quoted by Itô and McKean (1965) p. 15.

But

$$P\{\bigcup_{0<i\leq n+2}\ \bigcap_{i<k\leq i+3}(\omega:|x(\tfrac{k}{n})-x(\tfrac{k-1}{n})|<\tfrac{7\ell}{n})$$

$$\leq (n+2)[P\{|x(\tfrac{1}{n})|<\tfrac{7\ell}{n}\}]^3$$

$$=(n+2)\left[\int_{|\xi|<\frac{7\ell}{\sqrt{n}}}\frac{e^{-\xi^2/2}}{\sqrt{2\pi}}\,d\xi\right]^3$$

$$<\frac{\text{constant}}{\sqrt{n}}\ \downarrow\ 0 \qquad (n\to\infty)\ .$$

Hence

$$P\{\bigcap_{n\geq m}\bigcup_{i}\bigcap_{k}(\ldots)\}=0\ .$$

Q.E.D.

THEOREM 3. Almost no Brownian motion path is of bounded variation in any finite interval.

PROOF. See Doob (1953)p. 395.

EXERCISE: on Brownian motion.

1. $\mathcal{E}\{\frac{|x(t)-x(s)|}{t-s}\,|\,x(s)=x\}\to\infty,\quad t\downarrow s\ .$

2. Let $\{a=t_o^{(n)}<t_2^{(n)}<\ldots<t_{j(n)}^{(n)}=b\}$ be a partition of $[a,b]\subset I$, and suppose

$$\lim_{n \to \infty} \max_{1 \leq j \leq j(n)} | t_j^{(n)} - t_{j-1}^{(n)} | = 0 .$$

Show that

$$\lim_{n \to \infty} \mathcal{E}\{ \sum_{j=0}^{j(n)} | x(t_j^{(n)}) - x(t_{j-1}^{(n)}) | \} = \infty$$

This result may make Theorem 3 plausible.

In later work we shall need Brownian motion in n-dimensions.

Definition: A Brownian motion in Euclidean n-space E_n is a process $X = \{x(t,\omega) : t \geq 0\}$ with sample paths $x(\cdot,\omega) = [x_1(\cdot,\omega),\ldots,x_n(\cdot,\omega)]$ in E_n, where the components

$$\{x_i(t,\omega) : t \geq 0\} \quad i = 1,\ldots m$$

are independent Brownian motion processes in E_1.

EXERCISE: on n-dimensional Brownian motion

For $x \in E_n$ let $|x| = (\sum_1^n x_i^2)^{1/2}$.

1. If B is a Borel set of E_n,

$$P\{x(t) \in B | \ x(0) = 0\} = \int_B p(t,x)dx$$

where

$$p(t,x) = (2\pi t)^{-n/2} e^{-|x|^2/2t}$$

2. Writing $|x| = \rho$, $p(t,\rho)$ satisfies

$$\frac{1}{2}\{\frac{\partial^2 p}{\partial \rho^2} + \frac{n-1}{\rho}\frac{\partial p}{\partial \rho}\} - \frac{\partial p}{\partial t} = 0 \qquad (t > 0,\ \rho > 0)$$

3. X is continuous (i.e., almost all paths of X are continuous), nondifferentiable, and not of bounded variation in any finite t-interval.

3. A CLASS OF STOCHASTIC INTEGRALS.

In applications we need to discuss integrals of the type

$$\int_a^b f(t)\dot{w}(t)dt$$

where $\dot{w}$ is the formal derivative of a Brownian motion (= white noise), and f is a given nonrandom function of t. In future we denote the integral by

$$\int_a^b f(t)dw(t) \qquad (1)$$

and now proceed to give (1) a definite meaning. Observe that since $w(\cdot)$ is not of bounded variation the integral (1) cannot be defined as a Stieltjes integral of conventional type.

For simplicity we assume that w is a (separable) Brownian motion in E_1 and that f is real-valued; the extension to Brownian

motion in E_n and matrix-valued f is straightforward. For (1), the procedure is first to define the integral with f a step function, then extend the class of f's by appropriate passages to the limit. It is convenient (and no restriction) to assume w is a Brownian motion defined for $-\infty < t < \infty$ and in (1) to set also $a = -\infty$, $b = +\infty$.

Now suppose $a_1 < \ldots < a_n$ and

$$\begin{aligned} f(t) &= 0, \quad t < a_1 \\ &= c_j, \ a_{j-1} \leqq t < a_j \ , \ j \leqq n, \\ &= 0, \quad t \geqq a_n \ . \end{aligned}$$

Define

$$\begin{aligned} F &= \int f(t)dw(t) \qquad \left(\int = \int_{-\infty}^{\infty}\right) \\ &= \sum_2^n c_j[w(a_j)-w(a_{j-1})] \end{aligned} \tag{2}$$

Thus if f and g are step functions, then

$$\begin{aligned} \mathcal{E}\{FG\} &= \mathcal{E}\{[\int f(t)dw(t)][\int g(t)dw(t)] \\ &= \int f(t)g(t)dt \end{aligned} \tag{3}$$

[For calculations it is useful to bear in mind the formal relations: $\mathcal{E}\{dw(s)\} = 0$, $\mathcal{E}\{(dw(s))^2\} = ds$, and recall that $dw(s)$, $dw(s')$ are (formally) independent if ds, ds' are nonoverlapping differential

intervals].

Equation (3) suggests that we attempt to extend the class of admissible f's to the class $\mathcal{L}_2(-\infty,\infty) = \mathcal{L}_2(I)$ of functions such that $f(\cdot)$ is Lebesgue measurable and

$$\int f(t)^2 dt < \infty .$$

To do this consider the Hilbert space $\mathcal{L}_2(\Omega)$ of real-valued random variables F such that

$$\mathcal{E}\{F^2\} = \int_\Omega F(\omega)^2 dP(\omega) < \infty .$$

If $f \in \mathcal{L}_2(I)$ there is a sequence $\{f_n\}$ of step functions such that $\|f-f_n\| \to 0$ ($\|\cdot\|$ denotes norm in the appropriate Hilbert space), and thus a corresponding sequence $F_n \in \mathcal{L}_2(\Omega)$, defined by (2), such that

$$\|F_n-F_m\| = \|f_n-f_m\| \to 0 \quad . \tag{4}$$

There is then $F \in \mathcal{L}_2(\Omega)$ such that $F_n \to F$ in $\mathcal{L}_2(\Omega)$, and we <u>define</u>

$$\int f(t)dw(t) = F . \tag{5}$$

F is uniquely defined up to its values on a P-null ω-set; aside from this intrinsic lack of uniqueness F is independent of the particular sequence $\{f_n\}$ chosen, since two approximating sequences can be

combined to form a third mean-convergent sequence, for which the corresponding sequence of random variables is mean convergent.

To define $\int_a^b f(t)dw(t)$ we put $\tilde{f}(t) = f(t)$ if $a \leqq t \leqq b$ and $\tilde{f}(t) = 0$ otherwise, and

$$\int_a^b f(t)dw(t) = \int \tilde{f}(t)dw(t) .$$

Clearly the stochastic integral is additive and homogeneous (with respect to scalar multiplication) in the integrand.

EXERCISE:

1. $\underset{\substack{a \to -\infty \\ b \to +\infty}}{\text{l.i.m.}} \int_a^b f(t)dw(t) = \int_{-\infty}^{\infty} f(t)dw(t)$

2. $\mathcal{E}\{\int f(t)dw(t)\} = 0$

The foregoing definition has been extended to integrand functions $f(t,s)$ where, for instance, $f \in \mathcal{L}_2(I \times I)$; it then turns out that the function

$$g(t,\omega) = \int_I f(t,s)dw(s,\omega)$$

is a measurable random process. For details see Doob (1953)p. 430, etc. One can then also define double integrals of the form

$$\int_a^b \int f(s,t)dsdw(t)$$

in terms of iterated integrals.

Later we will generalize the definition of the integral to include random integrand functions f.

Finally we note that the probability distribution of the stochastic integral is Gaussian. In fact each member of an approximating sequence $\{F_n\}$ is Gaussian and, since

$$F = P \lim_{n \to \infty} F_n$$

we know that the distribution function of F_n tends to that of F, on the continuity set of the distribution of F [Loève(1963)p. 168]. But if F is nondegenerate (i.e., $F \neq$ some constant with probability 1: since $\mathcal{E}\{F\} = 0$, this is true if $\mathcal{E}\{F^2\} > 0$) then, by Khintchine's theorem on convergence of types [Loève(1963)p.203] the distribution of F is also Gaussian.

PROBLEM 1.

Let $\{x_n\}$ be a sequence of random variables in R_k. Show that

$$x_n \to 0$$

with probability one, if for every $\epsilon > 0$

$$\sum_1^\infty P\{|x_n| > \epsilon\} < \infty$$

[Suggestion: observe that

$$\{\omega : x_n(\omega) \to 0\} = \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{\omega : |x_m(\omega)| > \frac{1}{k}\}$$

and use the Borel-Cantelli lemma]

PROBLEM 2.

Let $f(t)$ be measurable and locally bounded. Define a process X :

$$x(t) = \int_0^t f(s)dw(s) \quad , \quad t \geq 0 \ .$$

Show that if X is separable then any countable set dense in $[0,\infty)$ is a separability set for X. Then show that almost every $x(t)$ sample function is a continuous function of t. [Suggestion: to prove the latter, simply indicate in detail how the proofs of Lemma 3 and Theorem 1 should be generalized].

PROBLEM 3.

Let X be Brownian motion in E_1. For a fixed interval $[a,b] \subset I$ let

$$y_n = \sum_{k=1}^{n} |x(t_k^{(n)}) - x(t_{k-1}^{(n)})|$$

$$z_n = \sum_{k=1}^{n} |x(t_k^{(n)}) - x(t_{k-1}^{(n)})|^2$$

where $t_k^{(n)} = a + \frac{k}{n}(b-a)$, $k = 0,1,\ldots,n$.

Using the result of Problem 1 and the Borel-Cantelli Lemma where appropriate, show that

$$y_n \to \infty$$

$$z_n \to b-a$$

as $n \to \infty$, with probability one.

4. GAUSSIAN DIFFUSIONS.

A finite dimensional linear dynamic system with white noise input is formally representable by an equation

$$\dot{x} = A(t)x(t) + B(t)\dot{w}(t) , \qquad (1)$$

w being a Wiener process. To achieve an appearance of respectability we write (1) as

$$dx = A(t)xdt + B(t)dw . \qquad (2)$$

Equation (2) can be interpreted as an Itô equation of special type; deferring a general discussion of Itô equations until later we shall define the solution of (2) in a direct manner, consistent both with Itô's method and with the usual formal rules for manipulating (1).

In the following, all vectors and matrices have real-valued elements.

For $0 \leq t < \infty$ let $A(t)$ be an $n \times n$ matrix with $A(\cdot)$ measurable and locally bounded [i.e., bounded in norm on every finite interval]. Let $B(t)$ be an $n \times m$ matrix with $B(\cdot)$ measurable and locally $\mathcal{L}_2$ (i.e., $\int_0^t |B(s)|^2 ds < \infty$ for every $t < \infty$). Let $\Phi(t,t_o)$ be an $n \times n$ matrix defined for $0 \leq t,t_o < \infty$ such that $\Phi(t_o,t_o) = I$ and

$$\frac{\partial \Phi(t,t_o)}{\partial t} = A(t)\Phi(t,t_o) , \qquad \text{a.e.} \tag{3}$$

It is well known from differential equations that such a <u>fundamental matrix</u> [of the homogeneous system $\dot{x} = A(t)x$] exists and is unique [Coddington and Levinson (1955)Ch.3]. Furthermore, denoting Euclidean norms by $|\cdot|$, we have

$$\Phi(t,t_o) = I + \int_{t_o}^{t} A(s)\Phi(s,t_o)ds$$

$$\Longrightarrow \quad |\Phi(t,t_o)| \leq 1 + \alpha_T \int_{t_o}^{t} |\Phi(s,t_o)| ds$$

$$\Longrightarrow \quad |\Phi(t,t_o| \leq e^{\alpha_T(t-t_o)} , \quad t_o \leq t \leq T , \tag{4}$$

where $\alpha_T = \sup\{|A(s)| : t_o \leq s \leq T\}$ and (4) follows by Gronwall's inequality. Observe also

$$|\Phi(t+h_1,t_o+h_2)-\Phi(t,t_o)|$$

$$\leq |\Phi(t+h_1,t_o+h_2)-\Phi(t,t_o+h_2)|+|\Phi(t,t_o+h_2)-\Phi(t,t_o)|$$

$$= \left|\int_t^{t+h_1} A(s)\Phi(s,t_o+h_2)ds\right| + \left|\int_{t_o}^{t_o+h_2}\{-\Phi(t,s)A(s)\}ds\right|$$

[using the fact that

$$\frac{\partial\Phi(s,t)}{\partial t} = -\Phi(s,t)A(t) \quad \text{a.e. }]$$

$$\leq \alpha_T e^{\alpha T}(|h_1|+|h_2|) \quad , \; 0 \leq t_o, t_o+h_2, t, t+h_1 \leq T \; ;$$

Φ is thus continuous (and hence Lebesgue measurable) jointly in t_o,t. We have then that $\Phi(t,s)B(s)$ is in $\mathcal{L}_2([0,T]\times[0,T])$ for every $0 < T < \infty$.

Let x_o be an arbitrary random variable and consider the formal initial value problem: (2) for $t \geq t_o$, and $x(t_o) = x_o$. It is natural to <u>define</u> the solution to be the process $\{x(t): t\geq t_o\}$ given by

$$x(t) = \Phi(t,t_o)x_o + \int_{t_o}^{t} \Phi(t,s)B(s)dw(s) \qquad (5)$$

As we have seen, the function defined by the integral in (5) is already a separable, measurable and continuous process defined for $t \geq t_o$; hence the same is true of $\{x(t)\}$. Also, if x_o is Gaussian

then the $x(t)$ process is Gaussian.

From now on we assume that the random variable x_o <u>is independent of the family of increments</u> $\{w(t)-w(s): t_o \leq s < t\}$ <u>and also that</u> $\mathcal{E}\{|x_o|^2\} < \infty$.

From (5) we see that

$$\mathcal{E}\{x(t)\} = \Phi(t,t_o)\mathcal{E}\{x_o\}; \tag{6}$$

that is, "the mean value of x satisfies the unperturbed differential equation"; this is a consequence of linearity. The covariance matrix of $x(t)$ is

$$\begin{aligned} Q(t) &= \mathcal{E}\{[x(t)-\mathcal{E}\{x(t)\}][x(t)-\mathcal{E}\{x(t)\}]'\} \\ &= \Phi(t,t_o)Q_o\Phi(t,t_o)' + \\ &\quad + \int_{t_o}^{t} \Phi(t,s)B(s)B(s)'\Phi(t,s)'ds \end{aligned} \tag{7}$$

where Q_o is the covariance matrix of x_o.

Write

$$R(t,t_o) = \int_{t_o}^{t} \Phi(t,s)B(s)B(s)'\Phi(t,s)'ds\ . \tag{8}$$

Evidently R is the contribution to the covariance of x due to the white noise. Using

$$\partial\Phi(t,t_o)/\partial t = A(t)\Phi(t,t_o)$$

one derives that

$$\int_{t_o}^{t}\{A(s)R(s,t_o)+R(s,t_o)A(s)'+B(s)B(s)'\}ds$$

$$= \int_{t_o}^{t}\left[\int_{t_o}^{s}\frac{\partial}{\partial s}\{\Phi(s,u)B(u)B(u)'\Phi(s,u)'\}du+B(s)B(s)'\right]ds$$

$$= \int_{t_o}^{t}[A(s)R(s,t_o)+R(s,t_o)A(s)'+B(s)B(s)']ds \qquad (9)$$

and therefore

$$\frac{\partial R(t,t_o)}{\partial t} = A(t)R(t,t_o)+R(t,t_o)A(t)'+B(t)B(t)' \qquad (10)$$

for almost all t. Combining (7) and (10) yields similarly

$$\frac{\partial Q(t,t_o)}{\partial t} = A(t)Q(t,t_o)+Q(t,t_o)A(t)'+B(t)B(t)' \ . \qquad (11)$$

Since Φ is never singular it is clear from (7) that $Q_o > 0$,* then $Q(t,t_o) > 0$ for all $t \geq t_o$; similarly if $R(s,t_o) > 0$ then $R(t,t_o) > 0$ for all $t \geq s$. Positive definiteness of R reflects qualitatively the property that "all system state variables (i.e., components of $x(\cdot)$) are perturbed by the white noise". In this sense the system (1) is "completely controlled by the white noise".

---

*For symmetric $n \times n$ matrices, $P > Q$ ($P \geq Q$) means $P-Q$ is positive definite (positive semi-definite); see also Appendix 1.

Definition: The dynamic system with differential equation

$$\dot{x} = A(t)x + B(t)u(t)$$

(where $u(\cdot)$ represents an input) is completely controllable at $t = t_o$ if, for each $t > t_o$, the matrix $R(t,t_o)$ defined by (8) is positive definite. [One sometimes says: the pair $\{A(\cdot),B(\cdot)\}$ is controllable at $t = t_o$]. [The definition given here, which is convenient for the sequel, is stronger than the original definition, due to Kalman (1960a). For systems with constant A and B, the two definitions coincide.]

From what was said earlier, controllability at t_o is a local property of A and B, depending only on the behavior of A, B in a neighborhood $[t_o,t_o+\delta]$, $\delta \downarrow 0$.

EXERCISE: If A,B are constant matrices, the pair $\{A,B\}$ is controllable if and only if the matrices

$$B, AB, \ldots, A^{n-1}B$$

contain among their column vectors a basis for E_n (regarded as a linear space over the reals).

EXERCISE: Let A, B be constant matrices, let $t_o = 0$ and write $Q(t,t_o) = Q(t,0) = Q(t)$ etc. A is stable if all eigenvalues of A have negative real parts. If A is stable show that $\lim Q(t)$ $(t \to \infty)$ exists, and that, if $P\{x_o = 0\} = 1$, the limit is approached

monotonically (in the sense of $\geqq$, footnote, p.105). If A is not stable and (A,B) is controllable show that $|Q(t)| \to \infty$ as $t \to \infty$.

EXERCISE: With $\{x(t)\}$ defined by (5) and $t_o \leqq t_1 \leqq t_2$ show that

$$x(t_2) = \Phi(t_2,t_1)x(t_1) + \int_{t_1}^{t_2} \Phi(t_2,s)B(s)dw(s) . \qquad (12)$$

This relation and our assumption on x_o imply that the $x(t)$ process is Markov.

5. QUADRATIC FUNCTIONALS.

Quadratic functionals of Gaussian processes are especially easy to study by dynamic programming methods, and we proceed to a discussion analogous to that of Chapter I,§4. Let $M(t)$ be a $n \times n$ symmetric matrix defined, measurable, and locally bounded for $t \geqq t_o$; and consider the functional

$$V(t,x) = \mathcal{E}\{\int_t^T x(s)'M(s)x(s)ds \,|\, x(t) = x\}. \qquad (1)$$

defined for $t_o \leqq t < T < \infty$ and $x \in E$. The conditional expectation is interpreted here to mean expectation on the paths of the $x(t)$ process determined by setting $t_o = t$, $P\{x_o = x\} = 1$, in (5)§2.4. We write $\mathcal{E}_{t,x}(\cdot)$ for $\mathcal{E}\{\cdot \,|\, x(t) = x\}$. Since

$$\int_t^T \mathcal{E}_{t,x}\{|x(s)|^2\}ds < \infty$$

we have

$$V(t,x) = \int_t^T \mathcal{E}_{t,x}\{x(s)'M(s)x(s)\}ds$$

$$= \int_t^T \overline{x}(s)'M(s)\overline{x}(s)ds + \int_t^T \mathrm{tr}\{M(s)R(s,t)\}ds$$

where

$$\overline{x}(s) = \mathcal{E}_{t,x}\{x(s)\} = \Phi(s,t)x \ ,$$

and $R(t,s)$ is defined by (8)§2.4. Thus

$$V(t,x) = x'P(t)x + p(t) \tag{2}$$

where

$$P(t) = \int_t^T \Phi(s,t)'M(s)\Phi(s,t)ds \tag{3}$$

$$p(t) = \int_t^T \mathrm{tr}\{M(s)R(s,t)\}ds \ . \tag{4}$$

From (3) and (3)§2.4 we get that P satisfies (a.e.) the linear matrix differential equation

$$\frac{dP(t)}{dt} + A(t)'P(t) + P(t)A(t) + M(t) = 0, \quad t \leqq T, \tag{5a}$$

$$P(T) = 0 \ . \tag{5b}$$

Also, using (8)§2.4,

$$p(t) = \int_t^T \mathrm{tr}\{M(s)\int_t^s \Phi(s,u)B(u)B(u)'\Phi(s,u)'du\}ds$$

$$= \int_t^T \mathrm{tr}\{B(u)'[\int_t^T \Phi(s,u)'M(s)\Phi(s,u)ds]B(u)du\}$$

$$= \int_t^T \mathrm{tr}\{B(u)'P(u)B(u)\}du \qquad (6)$$

and thus

$$\frac{dp(t)}{dt} = -\mathrm{tr}\{B(t)'P(t)B(t)\} \qquad \text{(a.e.)}$$

$$p(T) = 0 \qquad (7)$$

Now combining (2), (5) and (7) we have that $V(t,x)$ satisfies the parabolic equation

$$\mathcal{L}V(t,x) + V_t(t,x) = -x'M(t)x \qquad (t \leqq T,\ x \in E) \quad (8a)$$

and terminal condition

$$V(T,x) = 0, \quad x \in E \qquad (8b)$$

where we have put

$$\mathcal{L}u(x) \equiv \frac{1}{2}\,\mathrm{tr}\{B(t)'u_{xx}(x)B(t)\} + \\ +[A(t)x]'u_x(x) \qquad (9)$$

and subscripts denote partial derivatives.

Later we shall see how (8) could have been written down directly by inspection of (1), and (2)§2.4.

Since (5) is a linear equation it is clear that (8) has a unique solution $V(t,x)$ in the class of functions V of the form (2). In this sense the functional (1) is characterized by the relations (8). Thus we have the following integral formula analogous to the summation rule of I§4.

LEMMA (Integration Rule)

If $V(t,x) = p(t) + x'P(t)x$, *where* p *and* P *are absolutely continuous for* $t_o \leqq t \leqq T$, *and* $p(T) = 0$, $P(T) = 0$, *then*

$$V(t,x) = -\mathcal{E}_{t,x}\{\int_t^T (\mathcal{L}V[s,x(s)]+V_t[s,x(s)])ds\} \qquad (10)$$

where the process $\{x(s)\}$ *in the integrand is an arbitrary process of the type defined by* (1)§2.4; *and* $\mathcal{L}$ *is the corresponding operator* (9).

We call $\mathcal{L}$ the *differential generator* of the process X.

EXERCISE: If A,B,M are constant and A is stable show that $\lim P(t)$ $(t \to \infty)$ exists and $\lim |t|^{-1}p(t)$ $(t \to \infty)$ exists.

6. OPTIMAL CONTROL WITH QUADRATIC COST.

Consider the equation

$$dx = A(t)xdt - C(t)u(t)dt + B(t)dw(t) \qquad t \geq t_o \qquad (1)$$

where A and B are as before, and C is an $n \times k$ matrix, defined, measurable and locally bounded for $t \geq t_o$. The k-vector function $u(\cdot)$ represents a control function and will be assumed to be measurable and locally bounded. Thus (1) represents a time-variable linear system perturbed by additive Gaussian white noise.

A variety of interesting control problems arise with (1). Let a quadratic cost function be given, in the form

$$L(t,x,u) = x\,M(t)x + u'N(t)u \qquad (2)$$

where M,N are measurable, locally bounded, positive semidefinite, symmetric matrices. Thus L represents the instantaneous cost when the system is in state x under control u at time t. It is of interest to determine the function $u(\cdot)$ such that, for example, the functional

$$V(t,x;u) = \mathcal{E}_{t,x}\{\int_t^T L[s,x(s),u(s)]ds\}, \quad t_o \leq t \leq T, \qquad (3)$$

is a minimum in the class of all admissible functions $u(\cdot)$ on $[t_o,T]$. Such a control u^* is an optimal open-loop control or (in the language of operations research) an optimal schedule. Observe that an open loop (OL) control is regarded as selected once for all at the start of the control process: i.e., on the basis only of knowledge that the initial time is t and the initial state $x(t)$ is x. With OL control the actual evolution of the process X (or rather the evolution of a particular path) for times $> t$ has no influence on the control function input; and in general process evolution need not and may not be observable.

At the other extreme, suppose that the controller receives instantaneous information about current system state: it is plausible that exploiting this information would lead to superior control. To formalize the idea let $\varphi(t,x)$ be some class of functions from $[t_o,\infty) \times E$ to E_k, such that, if $u(t) = \varphi[t,x(t)]$ in (1), then the corresponding stochastic differential equation makes sense (we know at least that $\varphi(t,x) = K(t)x$, say, is a possibility). Such a φ is a feedback (FB) control. Then the optimization problem could be to minimize (3) for $u = \varphi$ and φ in some class Φ of admissible FB controls.

Clearly,intermediate situations are possible and of practical importance: state information may be available only at a finite number of discrete instants, or if available continuously may be degraded by the presence of observation noise. In some

applications it is difficult and/or expensive to undertake measurements, and the resulting tradeoff between cost of control and cost of information must be determined.

In this section we begin by looking for optimal FB controls in the class of linear controls $\varphi = Kx$. Later the class will be enlarged somewhat, but it will turn out that this enlargement does not change the result. This fact is critically dependent on the assumptions that the right side of (1) is linear in (x,u,dw) and that L is quadratic in x and u.

Let Φ be the class of functions $\varphi(t,x)$ such that

$$\varphi(t,x) = K(t)x$$

and the $k \times n$ matrix $K(t)$ is measurable and bounded for $t_o \leqq t \leqq T$. We say that $\varphi^o \in \Phi$ is <u>optimal</u> if

$$V(t,x;\varphi^o) \leqq V(t,x;\varphi) , \quad \varphi \in \Phi , \tag{4}$$

for all $t \in [t_o,T]$ and $x \in E$. Observe that the optimality condition is at first sight quite strong: that is, we could define φ^o to be "(t,x)-optimal" if (4) holds merely at t,x, and not necessarily at any $(t',x') \neq (t,x)$; thus an optimal control is (t,x)-optimal for all (t,x). It will turn out that an optimal control actually exists so that the more general concept of (t,x)-optimality will not be needed.

THEOREM 1 *Let* $N(t)$ *be positive definite for each* $t \in [t_o,T]$ *and let* $N(t)^{-1}$ *be bounded for* $t \in [t_o,T]$. *Then an optimal control exists.*

Our proof will yield a set of equations by which the optimal "gain" matrix $K^o(t)$ can be computed. The technique of proof is familiar: use the Integration Rule (§2.5) in the same way that the Summation Rule was used in §1.4. To do this let $\mathcal{L}_u$ denote the operator

$$\mathcal{L}_u V(t,x) \equiv \frac{1}{2} \operatorname{tr} \{B(t)' V_{xx}(t,x) B(t)\} + \\ + [A(t)x - C(t)u]' V_x(t,x) \qquad (5)$$

for abritrary k-vectors u. Write $\mathcal{L}_u$ for the operator obtained by substituting $u = \varphi(t,x)$ in (5); and write

$$\mathcal{L}_{\varphi^o} = \mathcal{L}^o \quad ; \quad V(t,x,;\varphi^o) = V^o(t,x) \ .$$

We assert:

OPTIMALITY LEMMA

The control $\varphi^o \in \Phi$ *is optimal if there is a set* $S \subset [t_o,T]$ *of Lebesgue measure zero such that, for all* $t \in [t_o,T]-S$, *all* $x \in E$, *and all k-vectors* u (*with real components*)

$$\mathcal{L}^{o}V^{o}(t,x) + L[t,x,\varphi^{o}(t,x)] \leqq$$

$$\leqq \mathcal{L}_{u}V^{o}(t,x) + L(t,x,u) \ . \qquad (6)$$

PROOF. Let $\varphi \in \Phi$ be arbitrary. We know then that (6) holds with $u = \varphi(t,x)$. Furthermore, φ^{o} being linear, V^{o} is a quadratic function and thus the integration rule applies. Let $\{x^{\varphi}(t): t_{o} \leqq t \leqq T\}$ be the family of Gaussian diffusions determined by setting $u(t) = \varphi(t,x(t))$ in (1). Then, by the integration rule

$$V^{o}(t,x) = -\mathcal{E}_{t,x}\{\int_{t}^{T}(\mathcal{L}_{\varphi}V^{o}[s,x^{\varphi}(s)] +$$

$$+ V_{t}^{o}[s,x^{\varphi}(s)])ds\}$$

$$\leqq -\mathcal{E}_{t,x}\{\int_{t}^{T}(\mathcal{L}^{o}V^{o}[s,x^{\varphi}(s)] + V_{t}^{o}[s,x^{\varphi}(s)]$$

$$+ L[s,x^{\varphi}(s),\varphi^{o}(s,x^{\varphi}(s))]-L[s,x^{\varphi}(s),\varphi(s,x^{\varphi}(s))])ds\}$$

$$= \mathcal{E}_{t,x}\{\int_{t}^{T}L[s,x^{\varphi}(s),\ \varphi(s,x^{\varphi}(s))]ds\}$$

$$= V(t,x;\varphi)$$

for all $t \in [t_{o},T]$ and $x \in E$.

Q.E.D.

On this basis we can now approach the problem by looking for a solution of Bellman's equation

$$\min_u \{\mathcal{L}_u V^o(t,x)+V_t^o(t,x)+L(t,x,u)\} = 0 \qquad (7a)$$

with terminal condition

$$V^o(T,x) = 0 \ , \qquad x \in E \ . \qquad (7b)$$

In principle this is a simple matter: performing the minimization in (7a) we get

$$\frac{1}{2}\,\mathrm{tr}\{B(t)'V_{xx}^o(t,x)B(t)\}+x'[A(t)-C(t)K^o(t)]'V_x^o(t,x) + V_t(t,x)+x'M(t)x+x'K^o(t)'N(t)K^o(t)x = 0 \qquad (8)$$

provided

$$u = K^o(t)x = \frac{1}{2}N(t)^{-1}C(t)'V_x^o(t,x) \ . \qquad (9)$$

Thus the proof of existence amounts to showing that there is a matrix-valued function $K^o(t)$, with the required measurability and boundedness, and a quadratic function $V^o(t,x)$, such that (8) and (9) are true.

It is a pleasant and remarkable fact that a rather explicit solution of (8) and (9) is available. To find it, set

$$V^{o}(t,x) = p(t) + x' P(t)x$$

and substitute. Then, formally

$$K^{o}(t) = N(t)^{-1}C(t)' P(t) \quad ; \tag{10}$$

and p, P satisfy

$$\dot{p} = -tr(B' PB) \tag{11}$$

$$\dot{P} + A' P + PA - PCN^{-1}C' P + M = 0 \tag{12}$$

for $t_o \leqq t \leqq T$, with

$$p(T) = P(T) = 0 .$$

We have reduced the problem to solving the matrix Riccati equation (12). This problem is not quite easy because the quadratic term in P fails (obviously) to satisfy a linear growth condition in $|P|$ and so we cannot, a priori, rule out the possibility that a (bounded) solution of (12) may fail to exist in the whole interval $[t_o, T]$. At this point, approximation in policy space comes to the rescue.* Let $\varphi_1(t,x) = K_1(t)x$ be an arbitrary admissible control; and let $V_1(t,x) = p_1(t) + x' P_1(t)x$ be the corresponding functional (3).

---

* Alternatively, we could exploit the fact that solvability of (12) is equivalent to absence of conjugate points for a certain variational problem. This question can be discussed by transforming (12) to a linear equation. See, for instance, Gelfand and Fomin (1963).

Having defined $\varphi_1,\ldots,\varphi_{n-1}$, $V_1,\ldots,V_{n-1}$, define

$$K_n(t) = N(t)^{-1}C(t)'P_{n-1}(t) \tag{13}$$

and determine p_n, P_n, V_n, φ_n from the equations

$$\begin{aligned} \dot{p}_n &= -\operatorname{tr}(B'P_nB) \\ \dot{P}_n+(A-CK_n)'P_n+P_n(A-CK_n)+M+K_n'NK_n &= 0 \\ p_n(T) = P_n(T) &= 0 \end{aligned} \tag{14}$$

$$V_n = p_n + x'P_nx$$

$$\varphi_n = K_nx \quad .$$

If we observe that (13) is equivalent to the condition

$$\mathcal{L}_{\varphi(n)}V_{n-1}(t,x) + \frac{\partial V_{n-1}}{\partial t}(t,x) + L[t,x,\varphi_n(t,x)]$$

$$\leqq \mathcal{L}_uV_{n-1}(t,x) + \frac{\partial V_{n-1}}{\partial t} + L(t,x,u)$$

(for $t \in [t_o,T]$, $x \in E$, $u \in E_k$) then it follows from the Integration Rule that $0 \leqq p_n(t) \leqq p_{n-1}(t)$ and $0 \leqq P_n(t) \leqq P_{n-1}(t)$, $n = 2,3,\ldots$. Then, by a theorem on monotone convergence of positive operators

[Kantorovich and Akilov (1964)p. 189; see also Appendix 1] we have that the matrix

$$P(t) = \lim P_n(t) \geqq 0$$

exists. Define $K^o(t) = \lim N(t)^{-1}C(t)'P_{n-1}(t)$ $(n \to \infty)$. Write $\hat{A}_n = A\text{-}CK_n$, $\hat{A} = A\text{-}CK$. Then (14) yields

$$P_n(t) = \int_t^T [\hat{A}_n(s)'P_n(s)+P_n(s)\hat{A}_n(s)+M(s)+K_n(s)'N(s)K_n(s)]ds \ . \quad (15)$$

Since the integrand in (15) is uniformly (in n) bounded for $t_o \leqq s \leqq T$, we have by the dominated convergence theorem

$$P(t) = \int_t^T [\hat{A}(s)'P(s)+P(s)\hat{A}(s)+M(s) + K^o(s)'N(s)K^o(s)]ds$$

But then P satisfies a.e. the differential equation (12); $V^o = p+x'Px$ satisfies the conditions of the optimality lemma; and the proof of Theorem 1 is complete.

It is interesting and important that our optimal FB control φ^o does not depend on the noise parameter B, since B does not appear in (12). In other words our control, which is in fact optimal for the nonstochastic problem

$$\int_t^T L[s,y(s),u(s)]ds = \min \qquad \dot{y} = A(s)y-C(s)u$$

remains optimal when noise is present as in (1). The presence of noise does, of course, degrade performance, the additional contribution to quadratic cost being

$$p(t) = \int_t^T \mathrm{tr}\{B(s)'P(s)B(s)\}ds \quad .$$

As a byproduct of this discussion we have that the solution of (12) exists and is essentially unique in the class of measurable (on $[t_o,T]$) matrices P. [Uniqueness follows by the local uniqueness theorem for differential equations (Coddington and Levinson (1955)Ch.1) and the fact that a solution bounded on $[t_o,T]$ is known to exist]. Thus (12) can in principle be solved numerically as a Cauchy problem of conventional type. Furthermore the optimal control φ^o is unique (modulo values of K on a subset of $[t_o,T]$ of Lebesgue measure zero) because, if φ^* is another optimal control in Φ, then

$$V(t,x,\varphi^o) = V(t,x,\varphi^*) \qquad t_o \leqq t \leqq T \ , \quad x \in E$$

so

$$P^o(t) = P^*(t)$$

and then

$$K^o(t) = K^*(t) \qquad \text{a.e.}$$

since the minimum in (6) is unique.

7. ASYMPTOTIC PROPERTIES OF THE OPTIMAL CONTROL.

In the notation of previous sections let the matrices A, B, C, M, N be constant. Let us write $\tilde{t} = T-t$, so that $\tilde{t}$ represents "time-to-go" and let $\tilde{V}(\tilde{t},x) = V(T-\tilde{t},x)$. By substitution in 2.6(8)-(12) we obtain (after dropping the tilde on $\tilde{V}$ and $\tilde{t}$)

$$V(t,x) = p(t) + x'P(t)x \tag{1}$$

with

$$p(t) = \int_0^t \operatorname{tr}\{B'P(s)B\}ds\ , \quad t \geq 0 \tag{2}$$

$$-\dot{P}(t)+[A-CK(t)]'P(t)+P(t)[A-CK(t)] + M+K(t)'NK(t) = 0\ , \quad t \geq 0, \tag{3}$$

with

$$P(0) = 0 \tag{4}$$

and

$$K(t) = N^{-1}C'P(t)\ . \tag{5}$$

In the following, we assume that the pair (A,C) is controllable, and use the following characterization of controllable pairs:

CONTROLLABILITY LEMMA

The pair (A,C) is controllable if and only if, for every set of complex numbers $\Lambda = \{\lambda_1,\ldots,\lambda_n\}$, such that any λ_i with $\operatorname{Im}\lambda_i \neq 0$ occur in conjugate pairs, there exists a matrix K

such that $A-CK$ has the spectrum Λ. [Here the matrices A, C, K have real elements and are of dimension respectively $n \times n$, $n \times k$, $k \times n$].

A proof is given in Appendix 2.

Thus, in particular, if (A,C) is controllable there exists a (constant) matrix $\hat{K}$ such that $A-C\hat{K}$ is stable. Now since $K(t)x$ is an optimal control for the problem of §2.6 we have that

$$\begin{aligned} 0 \leqq P(t) &\leqq \hat{P}(t) \\ &\leqq \int_0^t e^{s(A-C\hat{K})'}(M+\hat{K}'N\hat{K})e^{s(A-C\hat{K})}ds \\ &\leqq \hat{P}(\infty) \end{aligned}$$

and therefore $P(\cdot)$ is bounded on $[0,\infty)$

Next, we observe that $P(\cdot)$ is monotone nondecreasing To see this, let Φ be the class of admissible feedback laws, write $\varphi(s,x(s)) = \varphi_s$, let $0 \leqq t \leqq \tau$, and note:

$$\begin{aligned} V(\tau,x) &= \min_{\varphi\in\Phi} \mathcal{E}_{-\tau,x}\{\int_{-\tau}^{0} L(x_s,\varphi_s)ds\} \\ &= \min_{\varphi\in\Phi} \mathcal{E}_{-\tau,x}\{\int_{-\tau}^{-\tau+t} L(x_s,\varphi_s)ds + \int_{-\tau+t}^{0} L(x_s,\varphi_s)ds\} \\ &\geqq \min_{\varphi\in\Phi} \mathcal{E}_{-\tau,x}\{\int_{-\tau}^{-\tau+t} L(x_s,\varphi_s)ds\} \qquad (*) \end{aligned}$$

$$= \min_{\varphi \in \Phi} \mathcal{E}_{-\tau, x}\{\int_{-\tau}^{0} L(x_s, \varphi_s)ds\}$$

$$= V(t,x) \tag{6}$$

[Here we used the fact that, since the parameter matrices are constant, the quantity (*) is invariant under a shift in origin of time.] Since V has the form (1), our assertion follows.

By the theorem on monotone convergence of positive operators [Appendix 1], we can now conclude that

$$P_\infty = \lim_{t \to \infty} P(t) \tag{7}$$

exists, and therefore

$$K_\infty = \lim_{t \to \infty} N^{-1} C' P(t)$$

$$= N^{-1} C' P_\infty \tag{8}$$

exists as well. Then it follows from (3) that $P_\infty = P$ is given by

$$(A-CK)'P + P(A-CK) + M + K'NK = 0 \tag{9a}$$

$$K = N^{-1}CP \quad . \tag{9b}$$

Finally, $p(t) \sim t.\mathrm{tr}(B'PB)$ $(t \to \infty)$; thus the asymptotic cost rate is $\mathrm{tr}(B'PB)$.

EXERCISE: Show directly that, if A,C is a controllable pair and N^{-1} exists, then (9) has a unique positive semi-definite solution P which can be obtained by successive approximation.

8. OPEN LOOP CONTROL VERSUS FEEDBACK CONTROL.

It was stated earlier that FB control would be expected to be superior to OL control since FB control exploits information about current state. We now verify this fact for the "linear-quadratic" problem just considered. First we define the OL problem for 2.6(1),(2). Let $T > t_o$ be fixed; and let U be the class of bounded, measurable functions $u = u(t)$ defined for $t_o \leqq t \leqq T$ and with values in E_k. We define $V(t,x;u)$ by 2.6(3) and define $u^o \in U$ to be (t,x)-optimal if

$$V(t,x;u^o) \leqq V(t,x;u) \quad , \quad u \in U .$$

Let K^o be defined as in §2.6; and let $\Phi(t,t_o)$ be the fundamental matrix of the system

$$\dot{y} = [A(t)-C(t)K^o(t)]y \quad , \quad t_o \leqq t \leqq T . \qquad (1)$$

In the following, write $K^o = K$.

THEOREM 1 *The open loop control* u^o *given by*

$$u^o(s) = K(s)\Phi(s,t)x \quad , \quad t \leqq s \leqq T \tag{2}$$

is (t,x)*-optimal for all* $t \in [t_o,T]$, $x \in E$.

The proof is by a version of dynamic programming. We have

$$\begin{aligned} dx &= A(s)x\,ds - C(s)u(s)ds + B(s)dw(s), \\ x(t) &= x \qquad\qquad\qquad\qquad t \leqq s \leqq T . \end{aligned} \tag{3}$$

Furthermore, writing $y(s) = \mathcal{E}_{t,x}\{x(s)\}$, $z(s) = x(s)-y(s)$, we have

$$\begin{aligned} \dot{y}(s) &= A(s)y(s) - C(s)u(s) , \\ y(t) &= x ; \end{aligned} \tag{4a}$$

$$\begin{aligned} dz(s) &= A(s)z(s)ds + B(s)dw(s) ; \\ z(t) &= 0 \end{aligned} \tag{4b}$$

and

$$\mathcal{E}_{t,x}\{x(s)x(s)'\} = \mathcal{E}_{t,o}\{z(s)z(s)'\} + y(s)y(s)' .$$

Since the $z(s)$ process does not depend on u, the optimal OL

control is independent of the noise parameter B. We can thus restrict attention to the system (4a), and consider the simpler functional $\bar{V}$ defined by the integral

$$\bar{V}(t,x;u) = \int_t^T L[s,y(s),u(s)]ds \ , \quad y(t) = x \ . \tag{5}$$

It is clear that, if $W(t,x)$ is an arbitrary smooth, scalar-valued function such that

$$W(T,x) = 0 \ , \quad x \in E$$

then

$$W(t,x) = -\int_t^T \{W_t[s,y(s)] + \dot{y}(s)'W_x[s,y(s)]\}ds \tag{6}$$

where $y(s)$ is the solution of (4a). Now let y^o be the solution of (4a) with $u = u^o$ given by (2). Clearly $u^o(s) = K(s)y^o(s)$, and

$$\begin{aligned} \dot{y}^o(s) &= [A(s)-C(s)K(s)]y^o(s) \\ y^o(t) &= x \ . \end{aligned} \tag{7}$$

Thus, writing $\bar{V}^o(t,x) = \bar{V}(t,x,u^o)$, we have

$$\begin{aligned} \bar{V}^o_t(t,y) + [A(t)y-C(t)K(t)y]'V^o_x(t,y) \\ + L[t,y,K(t)y] = 0 \quad t_o \leqq t \leqq T \end{aligned} \tag{8}$$

with the terminal condition

$$\bar{V}^{o}(T,y) = 0 .$$

From (7) and (8) it is easy to check that

$$\bar{V}^{o}(t,y) = y' P(t) y$$

where P is given by the matrix Riccati equation 2.6(12). Because of the minimizing role of $K(t)$, the left side of (8) is no greater than

$$\bar{V}^{o}_{t}(t,y) + [A(t)y - C(t)u]' \bar{V}^{o}_{x}(t,y) + L(t,y,u),$$

for (almost) all $t \in [t_{o},T]$, all $x \in E$, and every k-vector u. Combining this fact with (6) and (8), we conclude that

$$\bar{V}(t,x;u^{o}) \leqq \bar{V}(t,x;u)$$

and consequently

$$V(t,x;u^{o}) \leqq V(t,x;u) ,$$

$t \in [t_{o},T]$, $x \in E$ and $u \in U$. This completes the proof that u^{o} is an optimal OL control.

We are now in a position to compare performance for optimal OL and FB control. To distinguish the two cases let (^) denote quantities pertaining to FB control. Retaining the notation of (4), write

$$\begin{aligned} \hat{y}(s) &= \mathcal{E}_{t,x}\{x(s)\} \\ \hat{z}(s) &= x(s)-y(s) \cdot \end{aligned} \qquad t \leqq s \leqq T \tag{9}$$

Then since

$$dx = (Ax-CKx)ds + B\,dw$$

there follows

$$\hat{y}(s) = y(s) \qquad t \leqq s \leqq T \tag{10}$$

and

$$\begin{aligned} d\hat{z} &= (A-CK)\hat{z}\,ds + B\,dw \\ \hat{z}(t) &= 0 \end{aligned} \qquad t \leqq s \leqq T \tag{11}$$

For optimal FB control $\hat{u}(\cdot)$ and OL control $u(\cdot)$ we have

$$\hat{u}(t) = Kx(t) \qquad t_o \leqq t \leqq T$$

so that

$$\mathcal{E}_{t,x}\{\hat{u}(s)\} = K(s)y(s) = \mathcal{E}_{t,x}\{u(s)\} \quad t \leqq s \leqq T \tag{12}$$

and

$$\hat{u}(s) - \mathcal{E}_{t,x}\{\hat{u}(s)\} = K(s)\hat{z}(s) \ . \tag{13}$$

On the other hand

$$u(s) - \mathcal{E}_{t,x}\{u(s)\} = 0 \ . \tag{14}$$

Now if $L(t,x,u) = x'M(t)x + u'N(t)u$, then identically

$$\mathcal{E}\{L(s,x_s,u_s)\} = L[s,\mathcal{E}\{x_s\},\mathcal{E}\{u_s\}] +$$

$$+\mathcal{E}\{L[s,x_s-\mathcal{E}\{x_s\},u_s-\mathcal{E}\{u_s\}]\} \ .$$

Using this fact together with (10)-(14) we find, after a brief calculation

$$V(t,x) - \hat{V}(t,x) = \mathcal{E}_{t,o}\{\int_t^T z(s)'M(s)z(s)ds\} -$$

$$-\mathcal{E}_{t,o}\{\int_t^T \hat{z}(s)'[M(s)+K(s)'N(s)K(s)]\hat{z}(s)ds\} \ .$$

Denote the integrals on the right by $q(t)$, $p(t)$ respectively. Then by (4b)

$$q(t) = \int_t^T \mathrm{tr}\,\{B(s)'Q(s)B(s)\}ds$$

where

$$\dot{Q} + A'Q + QA + M = 0 \ ; \ Q(T) = 0 \ ; \tag{15}$$

and by (11),

$$p(t) = \int_t^T \mathrm{tr}\{B(s)'P(s)B(s)\}ds$$

where, repeating 2.6(12),

$$\begin{aligned} &\dot{P} + A'P + PA - PCN^{-1}C'P + M = 0 \\ &P(T) = 0 \quad . \end{aligned} \tag{16}$$

Write $R = Q-P$. From (15) and(16)

$$R + A'R + RA + PCN^{-1}C'P = 0 \ .$$

Thus if Φ_A is the fundamental matrix of the system $\dot{\xi} = A\xi$ we have

$$0 \leqq R(s) = \int_s^T \Phi_A(\tau,s)'(PCN^{-1}C'P)(\tau)\Phi_A(\tau,s)d\tau \tag{17}$$

and then

$$q(t)-p(t) = \int_t^T \mathrm{tr}\{B(s)'R(s)B(s)\}ds \geqq 0 \tag{18}$$

Inequality (18) expresses an important, if intuitively evident, result: namely that (optimal) FB control yields superior performance to

(optimal) OL control, the difference in performance increasing with the intensity of noise in the system and with the time-interval of operation. If no noise is present ($B = 0$) then both types of control yield the same result. If the process interval is small ($T-t \to 0$) then $q(t)-p(t) \sim$ constant $(T-t)^2$; on the other hand, if $T-t \to \infty$ then we have, in the case, say, where A is constant and not stable, that $q(t)-p(t) \to \infty$ exponentially fast. That is, "no open loop control can stabilize an inherently unstable system", and feedback is crucial to acceptable operation.

EXERCISE: Show that the discussion of 2.6 and 2.8 can be generalized for a performance index

$$V(t,x) = \mathcal{E}_{t,x}\{\int_t^T L(s,x_s,u_s)ds + x_T'Qx_T\}$$

where $Q \geq 0$ and L is as before. [Discussion of asymptotic behavior is more difficult and is deferred to Chapter 3, where an equivalent problem is treated].

9. BIBLIOGRAPHICAL NOTE AND REFERENCES.

The results of this chapter have been known for some years. Among other problems, that of §2.6 was treated by Florentin (1961) and by Krasovskii and Lidskii (1961). The approach taken here, via the Integration Rule, is based on the Optimality Theorem discussed formally by Wonham (1963), and on Bellman's (1957) method of approximation in policy space. In discussion of asymptotic behavior of the Riccati equation, the important role of controllability was emphasized by Kalman (1960 b).

The linear regulator problem admits a quadratic value functional under more general assumptions than were made here: in particular, to the white noise perturbation terms of Gaussian type can be added those of Poisson type (random impulses), and the parameter matrices themselves can be assumed to jump randomly in value according to a Markov step process. In the latter case, the optimum feedback control law remains linear in the dynamic state, but its coefficients vary with the current state of the step process. Details can be found in Florentin (1961,1962) and Krasovskii and Lidskii (1961). We remark that a value functional quadratic in the state is also obtained if the parameter matrices (A,C) of the dynamic system are additively perturbed by Gaussian white noise: in this case, again, the control law depends parametrically on the intensity of noise; a stationary version of this problem is treated by Wonham (1967).

In Appendix 2 we establish the equivalence between controllability of an open-loop system and the possibility of assigning an arbitrary set of poles to the closed-loop system obtained by linear feedback. This result, for the system considered, is apparently unpublished, but the treatment closely follows an earlier discussion of Langenhop (1964).

REFERENCES

[1] Bellman, R., Dynamic Programming, Princeton University Press, 1957.

[2] Berberian, S.K., Measure and Integration, MacMillan, New York, 1965.

[3] Coddington, E.A., and Levinson, N., Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

[4] Doob, J.L., Stochastic Processes, Wiley, New York, 1953.

[5] Florentin, J.J., Optimal control of continuous-time, Markov stochastic systems, J. Electr. and Control, 10(6), 1961, pp. 473-488.

[6] Florentin, J.J., Optimal control of systems with generalized Poisson inputs, JACC Preprints, 1962.

[7] Gikhman, I.I., and Skorokhod, A.V., Introduction to the Theory of Random Processes, (Russ.), Izd.-vo "Nauka", Moscow, 1965.

[8] Gelfand, I.M., and Fomin, S.V., Calculus of Variations, Prentice-Hall, Englewood Cliffs, New Jersey, 1963.

[9] Halmos, P.R., Measure Theory, Van Nostrand, Princeton, New Jersey, 1950.

[10] Itô, K., and McKean, H.P., Jr., Diffusion Processes and their Sample Paths, Academic Press, New York, 1965.

[11] Kalman, R.E., (1960a), On the general theory of control systems, Proc. of the First Internat. Congress of the IFAC, Moscow, 1960; Butterworths Scientific Publications, London, vol.1, 1961, pp. 481-492.

[12] Kalman, R.E.,(1960b), Contributions to the theory of optimal control, Bol. Soc. Mat. Mexico, 1960, pp. 102-119.

[13] Kantorovich, L.V., and Akilov, G.P., Functional Analysis in Normed Spaces, MacMillan, New York, 1964.

[14] Kelley, J.L., General Topology, Van Nostrand, Princeton, New Jersey, 1955.

[15] Krasovskii, N.N., and Lidskii, E.A., Analytical design of controllers in systems with random attributes, I-III, Auto . and Rem. Control, 22(9), 1961, pp. 1021-1025; 22(10), 1961, pp. 1141-1146; 22(11), 1961, pp. 1289-1294.

[16] Langenhop, C.E., On the stabilization of linear systems, Proc. Amer. Math. Soc., 15(5), 1964, pp. 735-742.

[17] Loève, M., Probability Theory, (Third edition), Van Nostrand, Princeton, New Jersey, 1963.

[18] Wonham, W.M., Stochastic Problems in Optimal Control, RIAS TR 63-14, May, 1963.

[19] Wonham, W.M., Optimal stationary control of a linear system with state-dependent noise, J. Soc. Indust. Appl. Math., Series A, Control, 1967, to appear.

APPENDIX 1. Monotone Convergence of Symmetric Matrices.

Consider the class of symmetric $N \times N$ matrices P with real elements. We write $P \geq 0$ or $0 \leq P$ if P is positive semi-definite and $P \geq Q$ if $P-Q \geq 0$. A set of matrices $\{P_\alpha\}$ is bounded above (below) if there is a real number c such that $P_\alpha \leq c\, I$ ($P_\alpha \geq cI$) for all α. A sequence $\{P_n\}$ is monotone nondecreasing (nonincreasing) if $P_{n+1} \geq P_n$ ($P_{n+1} \leq P_n$). The sequence P_n converges to P if $P_n x \to Px$ for all x; equivalently $P_n \to P$ in matrix norm.

The following proposition is a useful generalization to matrices of a well-known property of real numbers.

PROPOSITION. Let $\{P_n\}$ be a sequence of symmetric $N \times N$ matrices with real elements. If $\{P_n\}$ is monotone nondecreasing (nonincreasing) and bounded above (below), then there exists a matrix P such that $P_n \to P$. P is unique and $P \geq P_n$ ($P \leq P_n$) for all n.

The proposition is a special case of a theorem on monotone convergence of positive operators in Hilbert space: for a proof of the latter see, for instance, Kantorovich and Akilov (1964)p. 189.

APPENDIX 2. On Stabilization of Linear Controllable Systems.

Consider the system

$$\dot{x} = Ax + Bu \tag{1}$$

where $A : n \times n$, $B : n \times m$ are matrices with real elements. We assume that (A,B) is controllable, that is, the $n \times nm$ matrix

$$H = [B,AB,\ldots,A^{n-1}B]$$

has rank n. Let $\{\lambda_1,\ldots,\lambda_n\}$ be an arbitrary set of n complex numbers such that any λ_i with $\operatorname{Im} \lambda_i \neq 0$ occur in conjugate pairs. It will be shown that there exists an $m \times n$ matrix C with real elements such that the matrix $A + BC$ has spectrum $\{\lambda_1,\ldots,\lambda_n\}$. In particular, (1) can always be stabilized by an appropriate choice of control law $u = Cx$. This fact will be proved using the canonical form of A and B derived in [1].* For completeness we re-derive that canonical form here.

Let

$$B = [b_1,\ldots,b_m]$$

where the b_i are n-vectors with real elements; and let S_i $(i = 1,\ldots,m)$ be the A-cyclic subspace generated by b_i. That

---

* See the end of this Appendix.

is, for each $i = 1,\ldots,m$ the vectors

$$b_i,\ Ab_i,\ldots,A^{n_i-1}b_i$$

are linearly independent; S_i is their span; and the minimal polynomial of b_i is of degree n_i. We can assume that $b_i \neq 0$, so that $1 \leqq n_i \leqq n$.

By our assumption of controllability the subspaces $S_1,\ldots,S_m$ span the entire n-space, S say, but in general the S_i are not independent. However, we can write S as a direct sum

$$S = \hat{S}_1 \oplus \hat{S}_2 \oplus \ldots \oplus \hat{S}_r \qquad (2)$$

where the $\hat{S}_j$ are certain subspaces of the S_i. To see this, define

$$\hat{S}_1 = S_1$$

If $b_2 \notin \hat{S}_1$ let ν_2, $1 \leqq \nu_2 \leqq n_2$, be the greatest integer such that the vectors

$$b_1,Ab_1,\ldots,A^{n_1-1}b_1,\ b_2,Ab_2,\ldots,A^{\nu_2-1}b_2$$

are linearly independent, and let $\hat{S}_2$ be the span of the vectors $A^{j-1}b_2$ $(j = 1,\ldots,\nu_2)$. Thus $\hat{S}_2$ is A-invariant and A-cyclic modulo $\hat{S}_1$, and has relative minimal polynomial of degree ν_2. Continuing in

this way, for $i \geq 2$ define ν_i, $1 \leq \nu_i \leq n_i$, to be the greatest integer such that the vectors

$$b_1,\ldots,A^{\nu_{i-1}-1}b_{i-1},\ b_i,Ab_i,\ldots,A^{\nu_i-1}b_i$$

are linearly independent, and let $\hat{S}_i$ be the span of $A^{j-1}b_i$ $(j=1,\ldots,\nu_i)$. Then $\hat{S}_i$ is A-cyclic modulo $\hat{S}_1 \oplus \ldots \oplus \hat{S}_{i-1}$, with relative minimal polynomial of degree ν_i. If at any stage $b_i \in \hat{S}_1 \oplus \ldots \oplus \hat{S}_{i-1}$ $(i \geq 2)$, then b_i is skipped; by re-ordering the columns of B if necessary we can arrange that for $i = 1,\ldots,r$ the vectors b_i generate independent subspaces $\hat{S}_i$, where either $r = m$ or $b_s \in \hat{S}_1 \oplus \ldots \oplus \hat{S}_r$ $(s = r+1,\ldots,m)$. Since $A^{j-1}b_i \in \hat{S}_1 \oplus \ldots \oplus \hat{S}_r$ if $j \geq \nu_i+1$ $(1 \leq i \leq m)$, it follows that (2) is true and thus $\nu_1 + \ldots + \nu_r = n$.

We can now transform A to a convenient canonical form. Let

$$\lambda^{\nu_i} - \sum_{j=1}^{\nu_i} \alpha_{ij}\lambda^{j-1} \qquad (i = 1,\ldots,r) \qquad (3)$$

be the minimal polynomial of b_i modulo $\hat{S}_1 + \ldots + \hat{S}_{i-1}$ (or the absolute minimal polynomial of b_1 in case $i = 1$). Following Langenhop [1] introduce vectors

$$e_{ik} = A^{\nu_i-k}b_i - \sum_{j=k+1}^{\nu_i} \alpha_{ij}A^{j-k-1}b_i \qquad (4a)$$

$(k = 1,2,\ldots,\nu_i\ ;\ i = 1,2,\ldots,r)$, where the summation does not appear in case $k = \nu_i$; that is

$$e_{i\nu_i} = b_i \qquad (i = 1,\ldots,r). \tag{4b}$$

It is clear that for each i the vectors e_{ik} $(k = 1,\ldots,\nu_i)$ are independent linear combinations of the vectors $A^{j-1}b_i$ $(j = 1,\ldots,\nu_i)$, and thus the entire set of e_{ik} is a basis for S. Observe that for $2 \leqq k \leqq \nu_i$,

$$Ae_{ik} = e_{i,k-1} + \alpha_{ik}e_{i\nu_i} \quad (i = 1,\ldots,r). \tag{5a}$$

Furthermore, using the fact that the minimal polynomial of b_i (modulo $\hat{S}_1 \oplus \ldots \oplus \hat{S}_{i-1}$) is of degree ν_i, we have (cf. (3))

$$Ae_{i1} = \alpha_{i1}e_{i\nu_i} + \sum_{k=1}^{i-1} \sum_{j=1}^{\nu_k} \gamma_{ijk}e_{kj} \tag{5b}$$

$(i = 1,\ldots,r)$ for certain scalars γ_{ijk}, and where the double summation does not appear if $i = 1$.

Using (5), we compute A, regarded as a linear transformation in S, relative to the basis $e_{11},\ldots,e_{1\nu_1},e_{21},\ldots,e_{2\nu_2},\ldots,e_{r1},\ldots,e_{r\nu_r}$. That is, let E be the matrix with column vectors e_{ik} in the order just given and let $T = E^{-1}$. Then $TAT^{-1} = \tilde{A}$ has the block form

$$\tilde{A} = \begin{bmatrix} A_1 & A_{12} & \cdots & A_{1r} \\ & A_2 & \cdots & \\ & & \ddots & \vdots \\ & & & A_{r-1,r} \\ \bigcirc & & & A_r \end{bmatrix} \tag{6}$$

In (6) the matrices A_i which occur on the diagonal of $\tilde{A}$ are $\nu_i \times \nu_i$ companion matrices

$$A_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & \cdots & 0 & 1 \\ \alpha_{i1} & \alpha_{i2} & \cdots & & \alpha_{i\nu_i} \end{bmatrix} \quad (i = 1,\ldots,r); \qquad (7)$$

and the matrices A_{ij} have dimension $\nu_i \times \nu_j$.

Finally, $TB = \tilde{B}$, where

$$\tilde{B} = \left[\begin{array}{ccc|c} \tilde{b}_1 & & \bigcirc & \\ & \tilde{b}_2 & & \tilde{B}_o \\ & & \ddots & \\ \bigcirc & & \tilde{b}_r & \end{array}\right] . \qquad (8a)$$

Here the ν_i-vectors $\tilde{b}_i = Tb_i$ have transpose

$$\tilde{b}_i' = (0,\ldots,0,1) , \qquad (8b)$$

and $\tilde{B}_o = T[b_{r+1},\ldots,b_m]$ is an $n \times (m-r)$ matrix (which does not appear if $r = m$).

To verify (6), (7), (8) in detail observe that Tx is the

column of components of x in the basis $\{e_{ik}\}$. Thus the p'th column of $\tilde{A}$ is a list of the components, in this basis, of Ae_{ik}, where e_{ik} is the p'th basis vector. This list can be read off from (5). Similarly $\tilde{b}_i$ is the column of components of b_i, so that (8) follows at once from (4b).

With A,B in canonical form, we next show that the spectrum of $A + BC$ can be prescribed arbitrarily. Put $\tilde{C} = CT^{-1}$ and observe that the matrix $\tilde{A} + \tilde{B}\tilde{C}$ is similar under T to $A + BC$. From now on we drop the tilde on $\tilde{A},\tilde{B},\tilde{C}$. Next, take C in the form

$$C = \begin{bmatrix} c_1' \\ c_2' \\ \\ \\ c_r' \\ \hline \\ \end{bmatrix} \qquad (9)$$

where

$$c_i' = [c_{i1}, c_{i2}, \ldots, c_{i\nu_i}] \qquad (i = 1,\ldots,r)$$

In C the upper block is of dimension $r \times n$ and the lower block of dimension $(m-r) \times n$. From (6) through (9) we see that

$$A + BC - \lambda I = \begin{bmatrix} P_1(\lambda) & A_{12} & \cdots & A_{1r} \\ & P_2(\lambda) & & \\ & & \ddots & A_{r-1,r} \\ & \bigcirc & & P_r(\lambda) \end{bmatrix}$$

where, for $i = 1,\ldots,r$,

$$P_i(\lambda) = \begin{bmatrix} -\lambda & 1 & 0 & \cdots & 1 \\ 0 & -\lambda & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & 0 & -\lambda & 1 \\ \alpha_{i1}+c_{i1} & \alpha_{i2}+c_{i2} & \cdots & \cdots & \alpha_{i\nu_i}+c_{i\nu_i} & -\lambda \end{bmatrix} = A_i + b_i c_i' - \lambda I.$$

Let $p_i(\lambda)$ be the characteristic polynomial of $P_i(\lambda)$; that is,

$$(-1)^{\nu_i} p_i(\lambda) = \lambda^{\nu_i} - \sum_{j=1}^{\nu_i} (\alpha_{ij}+c_{ij})\lambda^{j-1} \qquad (10)$$

From the triangular form of $A + BC - \lambda I$ it follows that

$$\det (A + BC - \lambda I) = p_1(\lambda) \ldots p_r(\lambda) .$$

Also, it is clear from (10) that the zeros of $p_i(\lambda)$ can be assigned

arbitrarily (subject to conjugacy of complex zeros) by proper choice of the real coefficients c_{ij}. This shows that, in particular, (1) can always be stabilized by appropriate choice of C.

To show that any desired spectrum can be assigned, we must take account of the fact that the matrices $P_i(\lambda)$ are of fixed dimension: if for example $r = 2$ and P_1, P_2 are each of dimension 1, it is impossible to assign a complex pair of eigenvalues by independent adjustment of p_1, p_2. To avoid this difficulty let $C = C_1 + C_2$, where C_1, C_2 are of form (9), and C_1 is chosen so that all the zeros of

$$\det |A + BC_1 - \lambda I|$$

are distinct; by the argument above, such a choice is possible. Next observe that $A + BC_1$ is cyclic ([2], p. 187, Theorem 4), and that $(A + BC_1, B)$ is controllable whenever (A,B) is. By the lemma proved below, we conclude that there exists an m-vector g such that the pair

$$(A + BC_1, Bg) \tag{11}$$

is controllable. But then, as we have seen, the pair (11) can be reduced to canonical form (6)-(8) where, this time, $r = 1$ and $\tilde{B}_o = 0$. It is now obvious that, for every set $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$, there exists vector c such that

$$A + BC_1 + Bgc'$$

has the spectrum Λ; then we choose

$$C = C_1 + gc' \ .$$

It remains to prove the lemma. Let $\{B\}$ denote the subspace spanned by the columns of B.

Lemma

Let (A,B) be controllable, and let A be cyclic. There exist vectors b,g with $b = Bg \in \{B\}$ such that (A,b) is controllable.

Proof

Let $\alpha(\lambda)$ be the minimal polynomial of A. Since A is cyclic, α is also the characteristic polynomial of A. Let $\beta(\lambda)$ be the minimal polynomial of $\{B\}$ (relative to A). A simple induction shows that β is then the minimal polynomial of the subspace

$$S = \{B,AB,\ldots,A^{n-1}B\}.$$

By controllability, S is the whole space; that is, $\beta = \alpha$, and hence some vector $b\epsilon\{B\}$ has α for its minimal polynomial ([2], p. 180, Theorem 2). But then b is a generator of S; that is, (A,b) is controllable. Clearly $b = g_1b_1 + \cdots + g_mb_m$ for suitable constants g_i; that is, $b = Bg$, for a suitable m-vector g. Q.E.D.

From the viewpoint of control theory it is of interest to note that the control vectors $b_{r+1},\ldots,b_m$, and associated scalar controls

$u_{r+1},\ldots,u_m$, are redundant for stabilization of the system (1). Thus we have arbitrarily set these controls to zero in the foregoing construction. Nevertheless the 'redundant' controls can play an important role in certain problems of stochastic stability, for example in the case of continuously acting additive perturbations of the matrix A.

Another point worth noting is that the canonical form $\tilde{A},\tilde{B}$ adopted here is not unique, and does not exhibit the intrinsic structure of A. Alternative well known canonical representations ([2], Ch. 7) reveal the intrinsic structure by exhibiting S as the direct sum of (absolutely) A-cyclic subspaces, in each of which A is representable by a companion matrix (7). Unfortunately it is not possible in general to form generators of these subspaces as linear combinations of the control vectors b_i alone, even when (A,B) is controllable. In spite of its ad hoc nature, Langenhop's canonical form [1], used here, yields a simple representation of the B matrix -- at the expense, of course, of introducing off-diagonal blocks in the representation of A. It is interesting to note that the triangular representation permits independent modification of the spectrum of each diagonal companion matrix by means of a control which depends on the variables associated with that block alone.

Finally we remark that the condition of controllability is also necessary for the assignment of an arbitrary spectrum to A + BC. The simple proof is omitted.

REFERENCES

[1] Langenhop, C.E., On the stabilization of linear systems, Proc. Amer. Math. Soc. 15 (5), 1964, pp. 735-742.

[2] Gantmacher, F. R., The Theory of Matrices, New York, Chelsea, 1959; Vol. 1.

CHAPTER III

OPTIMAL LINEAR FILTERING AND PREDICTION

1. INTRODUCTION.

In this chapter we study optimal linear filtering and prediction along the lines proposed by Kalman (1960) and Kalman and Bucy (1961). This approach, which amounts to sequential linear regression, represents a simplification of the fundamental studies by Kolmogorov (1941) and Wiener (c.1941-42; published 1948). Briefly stated, the essence of the present method is that an explicit dynamic model of the stochastic processes of interest is assumed at the start, and thus one avoids the difficult problem of spectral factorization. The present method yields a dynamic specification of the optimal filter or predictor which is readily simulated by computer.

We start by recalling some facts from elementary statistics. Let $x = x(\omega)$, $y = y(\omega)$ be random vectors with real-valued components. On the assumption that the indicated moments exist, we write

$$\mathcal{E}\{x\} = m_x\,, \quad \mathcal{E}\{y\} = m_y$$

$$\mathcal{E}\{(x-m_x)(x-m_x)'\} = P,\ \mathcal{E}\{(x-m_x)(y-m_y)'\} = Q,\ \mathcal{E}\{(y-m_y)(y-m_y)'\} = R. \tag{1}$$

Consider the problem: given the value $y(\omega)$, what is the optimal estimate of $x(\omega)$? Let $\hat{x}(\omega)$ be the corresponding estimate. In the special but important case where:

(i) the estimate desired is unbiased: i.e.,

$$\mathcal{E}\{\hat{x}\} = \mathcal{E}\{x\} \ ;$$

(ii) the estimate is linear:

$$x(\omega) = Ay(\omega) + b \quad ; \quad (A,b \text{ are constants})$$

(iii) the criterion of optimality is minimum weighted expected squared error:

$$\mathcal{E}\{(x-\hat{x})'M(x-\hat{x})\} = \min. \ , \quad (M > 0) \ ;$$

then the solution is well known:

$$\hat{x}(\omega) = m_x + QR^{-1}[y(\omega)-m_y] \ . \tag{2}$$

Here we have also assumed $R > 0$. The covariance of error is then

$$\mathcal{E}\{(x-\hat{x})(x-\hat{x})'\} = P - QR^{-1}Q' \tag{3}$$

If x,y are Gaussian random variables, then the estimate (2) is also the conditional expectation of x given y:

$$\hat{x} = \mathcal{E}\{x|\,y\} \tag{4}$$

In that case, the estimate (2) is optimal in the class of all estimators subject to (i) and (iii). Then (3) also gives the a posteriori (i.e. conditional) covariance of x:

$$\begin{aligned} &\mathcal{E}\{(x-\mathcal{E}\{x|\,y\})(x-\mathcal{E}\{x|\,y\})'|\,y\} \\ &\quad = \mathcal{E}\{(x-\mathcal{E}\{x|\,y\})(x-\mathcal{E}\{x|\,y\})'\} \qquad (5) \\ &\quad = P - QR^{-1}Q'. \end{aligned}$$

Observe that in the Gaussian case the a posteriori covariance does not depend on y.

2. DERIVATION OF THE FILTER EQUATIONS.

For $t \geq 0$ let X, Y be Gaussian processes determined by:

$$dx = A(t)x\,dt + B(t)\,dw(t) \qquad (1)$$

$$dy = C(t)x\,dt + D(t)\,dw(t) \qquad (2)$$

together with appropriate initial data.* As usual we assume that the matrices A,B,C,D are Lebesgue measurable and locally bounded for $0 \leq t < \infty$. Assume also that $(DD')^{-1}$ is locally bounded on $[0,\infty)$. Consider the problem of estimating the current state of the process X from observations of past and present values of a sample function of Y.

---

*It is sometimes reasonable to assume that the 'noise' functions acting directly on x and y are independent. For this write $w' = (w_1', w_2')$, $B = [B_1, B_2]$, $D = [D_1, D_2]$ and set $B_2 = 0$, $D_1 = 0$.

Let $\mathcal{Y}_t$ be the smallest σ-algebra of ω-sets relative to which the random variables $y(s)$, $0 \leq s \leq t$, are measurable. For definiteness, we may assume $y(0) = 0$ and $x(0) = x_0$, where x_0 is a Gaussian random variable independent of the dw increments. Our aim is to find the minimal variance unbiased estimate

$$\hat{x}(t) = \mathcal{E}\{x(t) | \mathcal{Y}_t\} \tag{3}$$

Let $P(t)$ be the conditional covariance

$$P(t) = \operatorname{cov}\{x(t) | \mathcal{Y}_t\} = \mathcal{E}\{[x(t)-\hat{x}(t)][x(t)-\hat{x}(t)]' | \mathcal{Y}_t\} \tag{4}$$

It will be convenient to write, for $t > s$,

$$\hat{x}(t|s) = \mathcal{E}\{x(t) | \mathcal{Y}_s\} \tag{5}$$

and

$$P(t|s) = \operatorname{cov}\{x(t) | \mathcal{Y}_s\} \, . \tag{6}$$

Finally, let $\Phi(t,s)$ be the fundamental matrix associated with the matrix A.

We begin by observing that the expectations $\mathcal{E}\{x(t)x(t)'\}$, $\mathcal{E}\{y(t)y(t)'\}$ exist, and are continuous and locally bounded for $t \geq 0$.* Since for any random variable z and sub-σ-field $\mathcal{Y}$,

---

*Equations (1) and (2) together have the form of (2), §2.4, and the theory of Chapt. 2 applies.

$$\mathcal{E}\{|\mathcal{E}\{z|\mathcal{Y}\}|^2\} \leq \mathcal{E}\{\mathcal{E}\{|z|^2|\mathcal{Y}\} \\ = \mathcal{E}\{|z|^2\}, \tag{7}$$

it follows that the existence of $\mathcal{E}\{x(t)x(t)'\}$ implies the existence (finiteness with probability 1) of $\hat{x}(t)$ and $P(t)$.

Our procedure will be to compute a discrete-time approximation to $P(t)$ and from this show that P satisfies a certain differential equation. To justify the method we need some preliminary results.

LEMMA 1 *Let* $\{y(t): t \geq 0\}$ *be a stochastically continuous process in a Euclidean space; let* $S = \{s_n\}$ *be a denumerable set everywhere dense in* $[0,\infty)$, *and let* $\hat{\mathcal{Y}}_t$ *be the smallest* σ*-algebra relative to which the random variables*

$$\{y(s): \quad 0 \leq s \leq t, \quad s \in S\}$$

are measurable. If f *is a random variable such that* $\mathcal{E}\{|f|\} < \infty$, *then*

$$\mathcal{E}\{f|\mathcal{Y}_t\} = \mathcal{E}\{f|\hat{\mathcal{Y}}_t\} \tag{8}$$

for every $t > 0$, *with probability* 1.

PROOF. Since $\mathcal{E}\{f|\hat{\mathcal{Y}}_t\}$ is obviously $\mathcal{Y}_t$-measurable, it is enough

to check that

$$\int_{\Lambda} \mathcal{E}\{f|\hat{y}_t\}dP(\omega) = \int_{\Lambda} f(\omega)dP(\omega) \tag{9}$$

for all sets Λ of the form

$$\Lambda = \{\omega: y(t_1) \in B_1,\ldots,y(t_r) \in B_r\}$$

where $0 \leqq t_i \leqq t$ and B_i is a Borel set of E $(i = 1,\ldots,r)$.* Clearly (9) is true if $\Lambda \in \hat{y}_t$. But if $t' \in (0,t)$ is arbitrary, there is a sequence $\{s_n\}$ in S such that $s_n \to t'$ and hence, by stochastic continuity,

$$P \lim y(s_n) = y(t')$$

As the limit in probability of a sequence of $\hat{y}_t$-measurable functions, $y(t')$ is also $\hat{y}_t$-measurable (more precisely, $y(t')$ is equal with probability one to a random variable which is $\hat{y}_t$-measurable). If $t' = 0$ (resp. $t' = t$) the same argument applies with $s_n \downarrow 0$ (resp. $s_n \uparrow t$). Hence Λ is $\hat{y}_t$-measurable for arbitrary choices of the t_i, and (9) follows.

LEMMA 2 *For $T > 0$ and fixed, let $\{t_j^{(n)} : j = 0,1,\ldots,j(n)\}$, $n = 1,2,\ldots$, be an increasing sequence of partitions of* $[0,T]$

---

* See e.g., the discussion of Doob (1953), Ch. 1, §7, especially pp. 19-20. In (9) P denotes, of course, probability measure on the underlying σ-algebra of ω-sets.

such that

$$0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_{j(n)}^{(n)} = T$$

and

$$\lim_{n \to \infty} \max_{1 \leq j \leq j(n)} (t_j^{(n)} - t_{j-1}^{(n)}) = 0 \qquad (10)$$

Let $\{y(t): t \geq 0\}$, y_t and f be defined as in Lemma 1; and, for $0 \leq t \leq T$, let $y_t^{(n)}$ be the smallest σ-algebra relative to which the random variables

$$\{y(t_j^{(n)}) : 0 \leq j \leq j(n) \quad , \; t_j^{(n)} \leq t \}$$

are measurable. Then

$$\mathcal{E}\{f | y_t^{(n)}\} \to \mathcal{E}\{f | y_t\} \qquad (n \to \infty)$$

with probability 1, and in the mean.

PROOF. Since $\mathcal{E}\{|f|\} < \infty$ the sequence $\mathcal{E}\{f | y_t^{(n)}\}$, $n = 1,2,\ldots,$ is a martingale and, by standard convergence theorems (Doob (1953): Thm.4.1, p.319; Thm.4.3, p.331) the limit

$$\lim_{n \to \infty} \mathcal{E}\{f | y_t^{(n)}\} = \mathcal{E}\{f | \hat{y}_t\}$$

exists with probability 1 and in the mean. Here $\hat{y}_t$ is the σ-algebra of ω-sets generated by the collection of random variables.

$$\{y(t_j^{(n)}) : 0 \leqq j \leqq j(n), \quad t_j^{(n)} \leqq t \; ; \; n = 1,2,\ldots\}$$

Since the set $\overset{\infty}{\underset{1}{U}}\{t_j^{(n)} : 0 \leqq j \leqq j(n) \; , \; t_j^{(n)} \leqq t\}$ is dense in the interval $[0,t]$, it follows by Lemma 1 that $\mathcal{E}\{f|\hat{y}_t\} = \mathcal{E}\{f|y_t\}$.

Q.E.D.

Returning to our filtering problem, we see that if

$$P_n(t) = \operatorname{cov}\{x(t)|y_t^{(n)}\}$$

then $P_n(t) \to P(t)$, as $n \to \infty$, with probability 1. In fact, in the present case, where X and Y are Gaussian processes, we see (from eq. (5), §3.1) that $P_n(t)$ is independent of the values of the random variables involved in the definition of $y_t^{(n)}$. That is, $P_n(t)$ is an ordinary (non-random) function of t, and the limit just written exists as an ordinary limit.

The remainder of the discussion amounts to a somewhat tedious computation. For n fixed and $0 \leqq j \leqq j(n)-1$ we write P_j for $P_n(t_j^{(n)})$, etc. and compute the difference $P_{j+1}-P_j$. To do this, we form the a priori estimate of $x_{j+1} = x(t_{j+1}^{(n)})$ given

the y_i for $i \leq j$ and then use (5), §3.1 to calculate the a posteriori covariance of x_{j+1}.

Let $t > s \geq 0$. Since

$$x(t) = \Phi(t,s)x(s) + \int_s^t \Phi(t,\tau)B(\tau)dw(\tau) \tag{11}$$

we have

$$\hat{x}(t|s) = \mathcal{E}\{x(t)|\, y_s\} = \Phi(t,s)\hat{x}(s) \tag{12}$$

and

$$\begin{aligned} P(t|s) &= \operatorname{cov}\{x(t)|\, y_s\} \\ &= \Phi(t,s)P(s)\Phi(t,s)' + \int_s^t \Phi(t,\tau)B(\tau)B(\tau)'\Phi(t,\tau)'d\tau \end{aligned} \tag{13}$$

For economy of notation we shall write

$$\hat{x}_{j+1}(j) = \mathcal{E}\{x_{j+1}|\, y_{t_j}^{(n)}\} \tag{14}$$

and

$$P_{j+1}(j) = \mathcal{E}\{[x_{j+1}-\hat{x}_{j+1}(j)][x_{j+1}-\hat{x}_{j+1}(j)]'|\, y_{t_j}^{(n)}\} \tag{15}$$

Thus

$$\hat{x}_{j+1}(j) = \Phi(t_{j+1},t_j)\hat{x}_j \tag{16a}$$

and

$$P_{j+1}(j) = \Phi(t_{j+1},t_j)P_j\Phi(t_{j+1},t_j)' + \int_{t_j}^{t_{j+1}} \Phi(t_{j+1},\tau)B(\tau)B(\tau)'\Phi(t_{j+1},\tau)'d\tau. \quad (16b)$$

Write $\delta_j = t_{j+1}^{(n)} - t_j^{(n)}$ and note that

$$|\Phi(t_{j+1},t_j)-I| \leq \text{const. } \delta_j \qquad (\delta_j \downarrow 0) ;$$

so that, expressing the first term on the right of (16b) as the integral of its derivative,

$$P_{j+1}(j) = P_j + \int_{t_j}^{t_{j+1}} [A(\tau)P_j+P_jA(\tau)'+B(\tau)B(\tau)']d\tau + O(\delta_j^2), \qquad \delta_j \downarrow 0. \quad (17)$$

The matrix $P_{j+1}(j)$ will play the role of the matrix P of (5) §3.1. To compute the corresponding matrices Q and R we note that

$$R = \text{cov}\{y_{j+1} | y_{t_j}^{(n)}\} ;$$

$$y_{j+1} = y_j + \int_{t_j}^{t_{j+1}} C(\tau)x(\tau)d\tau + \int_{t_j}^{t_{j+1}} D(\tau)dw(\tau) ; \quad (18)$$

$$\mathcal{E}\{y_{j+1}|y^{(n)}_{t_j}\} = y_j + \int_{t_j}^{t_{j+1}} C(\tau)\hat{x}(\tau|t_j)d\tau$$

$$= y_j + \int_{t_j}^{t_{j+1}} C(\tau)\Phi(\tau,t_j)\hat{x}_j d\tau$$

$$= y_j + \int_{t_j}^{t_{j+1}} C(\tau)\hat{x}_j d\tau + \Omega(\delta_j^2), \qquad (19)$$

where $\Omega(\delta_j^2)$ denotes a random variable z such that $(\mathcal{E}\{|z|^2\})^{1/2} = O(\delta_j^2)$, $\delta_j \downarrow 0$. Using (18) and (19) we find

$$R = \operatorname{cov}(y_{j+1}|y^{(n)}_{t_j}) = \int_{t_j}^{t_{j+1}} D(\tau)D(\tau)'d\tau + \Omega(\delta_j^2) \qquad (20)$$

Finally we compute

$$Q = \operatorname{cov}\{x_{j+1}, y_{j+1}|y^{(n)}_{t_j}\}$$

$$= \mathcal{E}\{[x_{j+1}-\Phi(t_{j+1},t_j)\hat{x}_j][\int_{t_j}^{t_{j+1}} C(\tau)[x(\tau)-\hat{x}(\tau|t_j)]d\tau + \int_{t_j}^{t_{j+1}} D(\tau)dw(\tau)]'|y^{(n)}_{t_j}\}$$

$$= \int_{t_j}^{t_{j+1}} \mathcal{E}\{[x_{j+1}-\Phi(t_{j+1},t_j)\hat{x}_j][x(\tau)'-\hat{x}_j'\Phi(\tau,t_j)']C(\tau)'|y^{(n)}_{t_j}\}d\tau$$

$$+ \mathcal{E}\{[\Phi(t_{j+1},t_j)(x_j-\hat{x}_j)+\int_{t_j}^{t_{j+1}} \Phi(t_{j+1},\tau)B(\tau)dw(\tau)][\int_{t_j}^{t_{j+1}} D(\tau)dw(\tau)]'|y^{(n)}_{t_j}\} \qquad (21)$$

The second expectation on the right side reduces to

$$\mathcal{E}\{[\int_{t_j}^{t_{j+1}} \Phi(t_{j+1},\tau)B(\tau)dw(\tau)][\int_{t_j}^{t_{j+1}} D(\tau)dw(\tau)]'\}$$

$$= \int_{t_j}^{t_{j+1}} \Phi(t_{j+1},\tau)B(\tau)D(\tau)'d\tau$$

$$= \int_{t_j}^{t_{j+1}} B(\tau)D(\tau)'d\tau + O(\delta_j^2) \qquad (\delta_j \downarrow 0) \qquad (22)$$

To evaluate the first expectation we replace Φ in the integrand by $I + O(\delta_j)$, add $-x_j+x_j$ to each of the expressions in $[\cdot]$, and multiply out. By (7) and Schwarz's inequality,

$$\mathcal{E}\{\|\mathcal{E}\{(x_{j+1}-x_j)(x(\tau)-x_j)'|\,y_{t_j}^{(n)}\}\|^2\}$$

$$\leq \mathcal{E}\{\|x_{j+1}-x_j\|^2\}\mathcal{E}\{\|x(\tau)-x_j\|^2\}$$

$$\leq \text{const. } (t-s)^2 = O(\delta_j^2) \ ;$$

$$\mathcal{E}\{\|\mathcal{E}\{(x_{j+1}-x_j)(x_j-\hat{x}_j)'|\,y_{t_j}^{(n)}\}\|^2\}$$

$$\leq \mathcal{E}\{\|x_{j+1}-x_j\|^2\}\mathcal{E}\{\|x_j-\hat{x}_j\|^2\}$$

$$\leq \text{const. } (t-s) = O(\delta_j) \ ;$$

$$\mathcal{E}\{\|\mathcal{E}\{(x_j-\hat{x}_j)(x(\tau)-x_j)'|\,y_{t_j}^{(n)}\}\|^2\}$$

$$\leq \mathcal{E}\{\|x_j-\hat{x}_j\|^2\}\mathcal{E}\{\|x(\tau)-x_j\|^2\}$$

$$\leq \text{const.}\ (\tau-t_j) = O(\delta_j)\ .$$

Thus the first expectation in (21) has the evaluation

$$\int_{t_j}^{t_{j+1}} \mathcal{E}\{(x_j-\hat{x}_j)(x_j-\hat{x}_j)'|\,y_{t_j}^{(n)}\}C(\tau)'d\tau + \Omega(\delta_j^{3/2})$$

and combining this result with (22) we obtain

$$Q = \int_{t_j}^{t_{j+1}} [P_jC(\tau)'+B(\tau)D(\tau)']d\tau + \Omega(\delta_j^{3/2})\ . \tag{23}$$

It is now possible to compute P_{j+1} as

$$P_{j+1} = P_{j+1}(j)-QR^{-1}Q\ .$$

Combining (17),(20),(23), we get

$$P_{j+1}-P_j = \int_{t_j}^{t_{j+1}} [A(\tau)P_j+P_jA(\tau)'+B(\tau)B(\tau)']d\tau -$$

$$-\left(\int_{t_j}^{t_{j+1}} [P_jC(\tau)'+B(\tau)D(\tau)']d\tau\right)\left(\int_{t_j}^{t_{j+1}} D(\tau)D(\tau)'d\tau\right)^{-1}\left(\int_{t_j}^{t_{j+1}} [P_jC(\tau)'+B(\tau)D(\tau)']d\tau\right)' +$$

$$+ \Omega(\delta_j^{3/2})\ .$$

To proceed further we replace the parameter matrices A, B, C, D by continuous approximants. More precisely, with $T < \infty$ fixed there are, for every $\epsilon > 0$, continuous matrices $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$ defined on $[0,T]$, and a set $\theta \subset [0,T]$ of Lebesgue measure $< \epsilon$, such that $\tilde{A}(t) = A(t)$ etc. if $t \in [0,T]-\theta$.* We note also that if $t_j^{(n)} \leqq t < t_{j+1}^{(n)}$ then (cf. (13))

$$P_n(t) = \operatorname{cov}\{x(t) | y_{t_j}^{(n)}\}$$

$$\to P_n(t_j^{(n)}) , \quad t \downarrow t_j^{(n)} ;$$

that is, $P_n(\cdot)$ is continuous from the right at the $t_j^{(n)}$. Therefore (reverting to the original notation)

$$P_n(T) = P_n(0) + \sum_{j=1}^{j(n)} [P_n(t_j^{(n)}) - P_n(t_{j-1}^{(n)})]$$

$$= P_n(0) + \int_0^T \{\tilde{A}(t)P_n(t) + P_n(t)\tilde{A}(t)' + \tilde{B}(t)\tilde{B}(t)' -$$

$$-[P_n(t)\tilde{C}(t)' + \tilde{B}(t)\tilde{D}(t)'][\tilde{D}(t)\tilde{D}(t)']^{-1}[P_n(t)\tilde{C}(t)' + \tilde{B}(t)\tilde{D}(t)']'\}dt$$

$$+ \text{const. } \epsilon + \omega(1) \quad (\max_j \delta_j \downarrow 0) \qquad (24)$$

where $\omega(1)$ denotes a random variable z such that

*See e.g. Munroe (1953) Ch. 4, §21.

$(\mathcal{E}|z|^2)^{1/2} = o(1)$.

Now the $P_n(t)$ are uniformly bounded (by cov{x(t)}, say) and we have seen that $\lim P_n(t) = P(t)$ exists. Hence by the dominated convergence theorem we can take the limit $n \to \infty$, and then the limit $\epsilon \downarrow 0$, to obtain

$$P(T) = P(0) + \int_0^T [AP+PA'+BB'-(PC'+BD')(DD')^{-1}(PC'+BD')']dt \ .$$

Since $T < \infty$ was arbitrary it follows that, for almost all t, $P(t)$ satisfies the <u>matrix Riccati equation</u>

$$\begin{aligned} \frac{dP}{dt} &= AP+PA'+BB'-(PC'+BD')(DD')^{-1}(PC'+BD')' \quad t \geq 0 \\ P(0) &= \text{cov}\{x_o\} \end{aligned} \tag{25}$$

Equation (25) is our first principal result. The second describes the evolution of the estimate $\hat{x}(t)$. Referring to (2) §3.1, we have that

$$\hat{x}_{j+1} = \hat{x}_{j+1}(j) + QR^{-1}(y_{j+1}-\mathcal{E}\{y_{j+1}|\mathcal{Y}_{t_j}^{(n)}\}) \tag{26}$$

Evaluating the quantities on the right from (16a),(19),(20) and (23) we obtain

$$\hat{x}(t_{j+1}) = \Phi(t_{j+1},t_j)\hat{x}(t_j) +$$

$$+\Big(\int_{t_j}^{t_{j+1}} [P_j C(\tau)' + B(\tau)D(\tau)']d\tau\Big)\Big(\int_{t_j}^{t_{j+1}} D(\tau)D(\tau)'\,d\tau\Big)^{-1} \cdot$$

$$\cdot[y(t_{j+1})-y(t_j)-\int_{t_j}^{t_{j+1}} C(\tau)\hat{x}(t_j)d\tau] + \omega(t_{j+1}-t_j) \quad .$$

By the same approximation argument used above, the last relation gives

$$\hat{x}(T) = \hat{x}(0)+\int_0^T A(t)\hat{x}(t)dt+\int_0^T [P(t)C(t)'+B(t)D(t)'][D(t)D(t)']^{-1}\cdot$$

$$[dy(t)-C(t)\hat{x}(t)dt]$$

with probability 1, for each $T < \infty$. That is,

$$d\hat{x} = A\hat{x}dt + (PC'+BD')(DD')^{-1}(dy-C\hat{x}dt), \quad t \geq 0$$

with

$$\hat{x}(0) = \mathcal{E}\{x_0\} \quad . \tag{27}$$

Equations (25) and (27) are the results of this section.

These equations have an important interpretation in technology. Namely, they prescribe the dynamic structure of a filter (analog device) for generating an estimate $\hat{x}(t)$, in real time, from noisy data given by the formal equation

$$\dot{y} = Cx + D\dot{w} \quad . \tag{28}$$

To simulate the filter, one first solves (25) and stores $P(t)$ for $0 \leqq t \leqq T$. The filter itself is specified by (27), or actually, for simulation purposes, by the formal version

$$\frac{d\hat{x}}{dt} = A\hat{x} + (PC'+BD')(DD')^{-1}(\dot{y}-C\hat{x}) \ . \tag{29}$$

Thus (29) is solved in real time as the noisy data $\dot{y}(t)$ becomes available. We note that in practice (25) and (29) are invariably solved in discrete approximation, and that, had we started with a discrete-time stochastic model, the derivation of the discrete-time counterparts of (25) and (27) would have required only a trivial application of the basic relations (2) and (3), §3.1.

EXERCISE: Consider the discrete-time stochastic processes specified by

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k w_k \\ & \qquad\qquad (k = 0,1,\dots) \\ y_{k+1} &= C_k x_k + D_k w_k \end{aligned}$$

where the w_k are independent normalized Gaussian random vectors in a Euclidean space. Derive the difference equations for the (optimal) filter which generates

$$\hat{x}_k = \mathcal{E}\{x_k | y_o,\dots,y_k\}$$

3. QUALITATIVE PROPERTIES OF THE FILTER.

3.1 Markov property of $\hat{x}$.

Consider the process $\{z(t),\ t \geq 0\}$ with differential

$$\begin{aligned} dz(t) &= dy(t) - C(t)\hat{x}(t)dt \\ &= C(t)[x(t)-\hat{x}(t)]dt + D(t)dw(t) \end{aligned} \tag{1}$$

and

$$z(0) = 0.$$

We observe that $z(t)$ is $\mathcal{Y}_t$-measurable; thus if $t > \tau$, the relation

$$z(t) = z(\tau) + \int_\tau^t C(s)[x(s)-\hat{x}(s)]ds + \int_\tau^t D(s)dw(s)$$

implies

$$\mathcal{E}\{z(t)|\mathcal{Y}_\tau\} = z(\tau) + \int_\tau^t C(s)\mathcal{E}\{\mathcal{E}\{x(s)-\hat{x}(s)|\mathcal{Y}_s\}|\mathcal{Y}_\tau\}ds = z(\tau); \tag{2}$$

i.e., the $z(t)$ process is a martingale. Furthermore,

$$\begin{aligned} &\mathcal{E}\{[z(t)-z(\tau)][z(t)-z(\tau)]'|\mathcal{Y}_\tau\} \\ &\qquad \leq \int_\tau^t D(s)D(s)'ds + \text{const.}\ (t-\tau)^2 \ . \end{aligned} \tag{3}$$

In fact, if $e(s) = x(s)-\hat{x}(s)$ then from 3.2(1) and 3.2(27),

$$de = (A-F)e\,dt + (C-G)dw$$

where $F = (PC'+BD')(DD')^{-1}C$, $G = (PC'+BD')(DD')^{-1}D$. If Ψ is the fundamental matrix corresponding to A-F, then

$$z(t)-z(\tau) = \int_t^\tau Ceds + \int_t^\tau D\,dw(s)$$

$$= \int_t^\tau C(s)\{\Psi(s,t)e(t)+ \int_t^s \Psi(s,u)[C(u)-G(u)]dw(u)\}ds + \int_t^\tau D\,dw(s)$$

from which (3) follows by a computation. But then, by a theorem on stochastic differentials [Doob (1953)p.287, Theorem 3.3] the $z(t)$ process has stochastic differential

$$dz(t) = D(t)d\tilde{w}(t) \tag{4}$$

where $\tilde{w}(t)$ is a suitable normalized Brownian motion. Combining (4) with 3.2(27) we get

$$\begin{aligned} d\hat{x} &= A\hat{x}dt + (PC'+BD')(DD')^{-1}Dd\hat{w} \\ \hat{x}(0) &= \mathcal{E}\{x_o\} \ . \end{aligned} \tag{5}$$

Equations (5) show that the filtered process $\hat{x}$ is a Gaussian diffusion in its own right, with differential generator $\hat{\mathcal{L}}$ given by

$$\hat{\mathcal{L}}V(\xi) = \frac{1}{2}\,\mathrm{tr}\,(\hat{B}V_{xx}(\xi)\hat{B}) + (A\xi)'V_x(\xi) \tag{6}$$

where $\hat{B} = (PC'+BD')(DD')^{-1}$.

3.2 Solution of the matrix Riccati equation.

We proceed to show that (25) has a unique solution P, so that (25) can be used in practice to obtain P and thus to simulate the filter explicitly. Now (25) can be written

$$\frac{dP}{dt} = [A-BD'(DD')^{-1}C]P+P[A-(BD')(DD')^{-1}C]'-PC'(DD')^{-1}CP + \\ + BB'-(BD')(DD')^{-1}(BD)'. \tag{1}$$

Write $M = BB'-BD'(DD')^{-1}(BD)'$; then (Exercise) $M \geq 0$. Write $N = DD'$, $\tilde{A} = [A-(BD')(DD')^{-1}C]'$, $\tilde{C} = C'$. Finally, let $\tau = -t$. Then (1) becomes

$$\frac{dP}{d\tau} + \tilde{A}'P + P\tilde{A} - P\tilde{C}N^{-1}\tilde{C}'P + M = 0, \quad \tau < 0 \\ P(0) = \mathrm{cov}\{x_0\} . \tag{2}$$

Now (2) has exactly the form of 2.6(12), with the identifications $\tilde{A} = A$, $\tilde{C} = C$, $0 = T$; the only distinction lies in the terminal (at $\tau = 0$) condition that in general $P(0) \neq 0$. However, the discussion given in §2.6 is easily generalized to the case of a terminal condition of form $V(T,x) = x'Qx$, say; so that $P(T) = Q$, and we now identify $\mathrm{cov}\{x_0\} = Q$. The method of approximation in policy space used in §2.6 then shows that (2) has a unique (and

locally bounded) solution on the halfline $\tau \leqq 0$.

3.3 Asymptotic behavior of the optimal filter.

The discussion just given points up a formal "duality" between the linear regulator problem of 2.6 and the linear filtering problem of the present chapter. This duality can be exploited further by introducing the concept of "observability".

Definition: Let $C(t)$, $A(t)$ be, respectively, $m \times n$ and $n \times n$ (locally bounded, measurable) matrix functions of t defined for $t_o \leqq t \leqq T \leqq \infty$. The pair C, A is observable at $t > 0$ if, for every $t_o \leqq \tau < t$,

$$\int_{\tau}^{t} \Phi(s,t)'C(s)'C(s)\Phi(s,t)ds > 0 \qquad (1)$$

where Φ is the standard fundamental matrix associated with A. Writing

$$\tilde{\Phi}(\tau,s) = \Phi(s,\tau)', \quad \tilde{C}(s) = C(s)'$$

we see that (1) is equivalent to

$$\int_{\tau}^{t} \tilde{\Phi}(t,s)\tilde{C}(s)\tilde{C}(s)'\tilde{\Phi}(t,s)'ds > 0 \qquad (\tau < t) .$$

Comparing with the definition of controllability (§2.4), and recalling that $\tilde{\Phi}(t,s)$ is the fundamental matrix associated with $-A'$, one sees that observability of C, A at t is closely

analogous to the property of controllability of the 'dual' pair $-A', C'$ at t -- our definitions are not quite equivalent, however: for controllability focusses on the immediate future $(t+0)$, observability on the immediate past $(t-0)$.

EXERCISES:

1. If A, C are constants, show that (C,A) is observable if and only if (A', C') is controllable.

2. Let A, C be constants and let B be an arbitrary $n \times m$ constant matrix. Show that (C,A) is observable if and only if $(C, A-BC)$ is observable.

Observability is best understood by considering the simple deterministic system

$$\begin{aligned} \dot{x} &= A(t)x \qquad (A : n \times n) \qquad (2) \\ y(t) &= C(t)x(t) \qquad (C : m \times n) \, . \end{aligned}$$

Here we interpret $y(t)$ as the observed output of the dynamic system (2). The problem is: knowing $A(\cdot)$, $C(\cdot)$, can we determine the state $x(t)$ from a record of $y(s)$ over an (arbitrarily small) interval $(t-\delta, t]$? It is easily seen that the answer is affirmative provided (C,A) is observable at t. Indeed the equation

$$y(s) = C(s)\Phi(s,t)x(t), \quad s < t \, ,$$

implies

$$\left[\int_{t-\delta}^{t} \Phi(s,t)'C(s)'C(s)\Phi(s,t)ds\right]x(t) =$$

$$= \int_{t-\delta}^{t} \Phi(s,t)'C(s)'y(s)ds$$

for $\delta > 0$; and this yields $x(t)$. Let us give a more detailed discussion: from the equation for y above we see that $x(t)$ can be determined if and only if, for every $\delta > 0$, there are points $t-\delta < s_1,\ldots,s_n \leqq t$ such that the $mn \times n$ matrix

$$\begin{bmatrix} C(s_1)\Phi(s_1,t) \\ \ldots\ldots\ldots\ldots \\ C(s_n)\Phi(s_n,t) \end{bmatrix} \qquad (*)$$

has full rank n. If this condition fails, there exists $\delta > 0$ such that, for all n-tuples of points $(s_1,\ldots,s_n)$, $s_i \in (t-\delta,t]$, the matrix (*) has rank less than n. Let the maximal rank of (*) be $k < n$, and choose $(s_1,\ldots,s_k)$ such that the $kn \times n$ matrix $\{C(s_i)\Phi(s_i,t)\}_{i=1,\ldots,k}$ has rank k. Then for all $s \in (t-\delta,t]$, the matrix

$$\begin{bmatrix} C(s_1)\Phi(s_1,t) \\ \ldots\ldots\ldots\ldots \\ C(s_k)\Phi(s_k,t) \\ C(s)\Phi(s,t) \end{bmatrix}$$

has rank no greater than k and hence, choosing $x \neq 0$ such that $C(s_i)\Phi(s_i,t)x = 0$ $(i = 1,\ldots,k)$, we must have

$$C(s)\Phi(s,t)x = 0, \quad t-\delta < s \leqq t \ .$$

Then (1) fails, and so observability entails the rank condition.

On the other hand, the rank condition might hold (with, say, rational s_j's) and yet the integral in (1) vanish for every τ ; this rather artificial situation would be ruled out in "practical" cases where, say, $C(\cdot)$ could be defined so as to be continuous from the left at t: then if (1) fails, the rank condition must fail as well.

We now turn to a discussion of the asymptotic behavior of the Riccati equation (1) as $t \to \infty$. We assume for simplicity that the parameter matrices A,B,C,D are constants. Our result is the

THEOREM <u>If</u> (C,A) <u>is observable, the covariance matrix</u> $P(t)$ <u>tends to a (finite) limit as</u> $t \to \infty$.

PROOF. With the identifications

$$\begin{aligned} \tilde{A} &= [A-BD'(DD')^{-1}C]' \\ \tilde{C} &= C' \\ M &= BB'-BD'(DD')^{-1}(BD')' \\ N &= DD' \\ K(t) &= (DD')^{-1}CP(t) \end{aligned} \tag{3}$$

we can re-write 3.2(25) as

$$-\frac{dP}{dt} + (\tilde{A}-\tilde{C}K)'P + P(\tilde{A}-\tilde{C}K) + M + K'NK = 0, \quad t \geq 0 \tag{4}$$

with

$$P(0) = \text{cov}\{x_o\} = P_o \text{ , say,} \tag{5}$$

and

$$K = K(t) = N^{-1}\tilde{C}'P(t) \tag{6}$$

These equations have the same form as (3), (4), (5) of §2.7. Also, since (C,A) is observable, we have $(C,A-BD'(DD')^{-1}C)$ observable,* and therefore $(\tilde{A},\tilde{C})$ is controllable. Thus our problem is identical with the control problem of §2.6, §2.7, apart from the single fact that P now has, in general, a nonzero initial value P_o. As before, by choosing $\hat{K}$ = constant so that $\tilde{A}-\tilde{C}\hat{K}$ is stable, and then noting that

$$\begin{aligned} 0 \leq P(t) \leq \hat{P}(t) \equiv e^{t(\tilde{A}-\tilde{C}\hat{K})'}P_o e^{t(\tilde{A}-\tilde{C}\hat{K})'} + \\ + \int_o^t e^{s(\tilde{A}-\tilde{C}\hat{K})'}(M+\hat{K}'N\hat{K})e^{s(\tilde{A}-\tilde{C}\hat{K})}ds \end{aligned} \tag{7}$$

we see that $P(t)$ is bounded on $[0,\infty)$. In the following, we use the fact that the set of all symmetric $n \times n$ matrices with real elements forms a conditionally complete lattice, with respect to the usual

---

* See Exercise 2, p. 168.

partial ordering of quadratic forms [see Exercise 3, p. 175]. Thus, from (7) we have

$$\overline{\lim_{t \to \infty}} P(t) \leq \int_0^\infty e^{s(\tilde{A}-\tilde{C}\hat{K})'}(M+\hat{K}'N\hat{K})e^{s(\tilde{A}-\tilde{C}\hat{K})}ds$$

for all constant matrices $\hat{K}$ (of proper dimension); hence

$$\overline{\lim_{t \to \infty}} P(t) \leq \inf_{\hat{K}} \int_0^\infty e^{s(\tilde{A}-\tilde{C}\hat{K})'}(M+\hat{K}'N\hat{K})e^{s(\tilde{A}-\tilde{C}\hat{K})}ds \qquad (8)$$

On the other hand, if Φ is the fundamental matrix associated with $\tilde{A}-\tilde{C}K(t)$, then by (4) and (5),

$$\begin{aligned} P(t) &= \Phi(t,0)'P_0\Phi(t,0) + \\ &\quad + \int_0^t \Phi(s,0)'[M+K(s)'NK(s)]\Phi(s,0)ds \\ &\geq \int_0^t \Phi(s,0)'[M+K(s)'NK(s)]\Phi(s,0)ds \\ &\geq P^*(t), \end{aligned}$$

where $P^*(t)$ is the solution of (4) and (6) corresponding to the special initial condition $P^*(0) = 0$. Here the last inequality holds because P^* is the minimal value of the last written integral with respect to all possible choices of $K(\cdot)$ (and corresponding Φ). Using the fact (§2.7) that P^* is monotone,

$$\underline{\lim}_{t \to \infty} P(t) \geq P^*(\infty) \quad . \tag{9}$$

In the notation of §2.7

$$\begin{aligned} P^*(\infty) = P_\infty^* &= \int_0^\infty e^{t(\tilde{A}-\tilde{C}K_\infty)'}(M+K_\infty' N K_\infty) e^{t(\tilde{A}-\tilde{C}K_\infty)} dt \\ &\geq \inf_K \int_0^\infty e^{t(\tilde{A}-\tilde{C}K)'}(M+K'NK) e^{t(\tilde{A}-\tilde{C}K)} dt \ . \end{aligned} \tag{10}$$

The inequalities (8),(9) and (10) prove the existence of $\lim P(t)$ $(t \to \infty)$.

Q.E.D.

From a practical point of view the significance of this result is that $P(\infty)$ and the associated formal stochastic equation 3.2(29) define a time-invariant linear filter which is optimal by the criterion of stationary mean square error: that is, observability of the pair (C,A) implies not only that the covariance of error is bounded, but even that the error process $\{e(t): t \geq 0\}$ is asymptotically stationary; more precisely, if we suppose $e(0)$ is distributed with mean 0 and covariance matrix $P(\infty)$, then $e(t)$ remains so distributed for all $t > 0$; alternatively, for any $P(0)$, $P(t) \to P(\infty)$ independently of $P(0)$. Observe the interesting and important fact that "the filter tracks the plant" with <u>bounded</u> covariance of error even in the case where the plant (system $\dot{x} = Ax - C\dot{w}$) is inherently unstable; i.e., where $\mathcal{E}\{x(t)x(t)'\} \to \infty$ $(t \to \infty)$. If observability fails, stable

tracking may not be possible.

EXERCISES:

1. Construct a (nontrivial!) example to show that if observability fails then tracking error may be unbounded in mean square as $t \to \infty$.

2. Show that, in the constant parameter case, if $B = 0$ and $DD' = I$, then

$$P(t) = e^{tA}[I+P_o \int_o^t e^{sA'} C'Ce^{sA} ds]^{-1} P_o e^{tA'}$$

Check that the normalization $DD' = I$ is always possible (when $DD' > 0$). Show that if (A,C) is observable, and all eigenvalues of A lie in the closed left half plane, then $P(t) \to 0$ as $t \to \infty$. [Suggestion: to show $P(t) \to 0$, suppose $P_o > 0$ and note

$$P(t) = [e^{-tA'} P_o^{-1} e^{-tA} + H(t)]^{-1}$$

where

$$H(t) = \int_o^t e^{-sA'} C'Ce^{-sA} ds$$

Now if $Q > R > 0$ then $Q^{-1} < R^{-1}$ [Appendix, 1]; thus

$$P(t) < [H(t)]^{-1}$$

$$\Longrightarrow |P(t)| \leq |H(t)^{-1}|$$

$$\leq \{ \inf_{|x|=1} x'H(t)x \}^{-1} \quad \text{[Appendix, (2d)]} .$$

Now $$x'H(t)x = \int_0^t |Ce^{-sA}x|^2 ds$$

$$\geqq kt \qquad (|x| = 1,\ t \geqq 1)$$

for some $k > 0$ (Appendix, 3). Hence

$$|P(t)| \leqq (kt)^{-1} \to 0 \qquad (t \to \infty).$$

For the general case, note $P(t)$ is monotone in P_o, so that, replacing P_o by a majorant pI, $p > 0$, the result follows.]

3. (ON THE LATTICE OF SYMMETRIC MATRICES).

Let $P,Q,\ldots$ be $n \times n$ symmetric matrices. Define $P \geqq Q$ iff $P-Q$ is positive semidefinite (and $P \leqq Q$ if $Q-P$ is positive semidefinite). With $\geqq$ (or $\leqq$), our set is partially ordered. Define

$$P^+ = P \vee 0 = \sup (P,0)$$

as follows: Under some rotation (orthogonal matrix) T we have

$$T'PT = \text{diag}\,(p_1,\ldots,p_n) \ .$$

Define

$$(T'PT)^+ = \text{diag}\,(p_1 \vee 0,\ldots,p_n \vee 0)$$

and then

$$P^+ = T(T'PT)^+T' \quad .$$

Check that $P^+ = \sup(P,0)$ in the sense: $P^+ \geq 0$ and $P^+ \geq P$; and if $\hat{P} \geq 0$ and $\hat{P} \geq P$, then $\hat{P} \geq P^+$. Now define

$$P \vee Q = \sup(P,Q) = (P-Q)^+ + Q \quad .$$

Check that $P \vee Q = (Q-P)^+ + P$ [For this note that if

$$R^- = \inf(R,0) = -\sup(-R,0)$$
$$= -(-R)^+$$

then ("Jordan decomposition")

$$R = R^+ + R^- \quad .]$$

Similarly define

$$P \wedge Q = \inf(P,Q) = -(-P \vee -Q).$$

With these definitions our set is a lattice. A lattice is <u>complete</u>* if every set bounded above has a supremum (and every set bounded below, an infimum). To show completeness here, recall that if $P_n \leq \hat{P}$ and $P_n \uparrow$ then $P = \lim P_n$ exists in the Euclidean (norm) topology of matrices [See Kantorovich and Akilov (1964)p. 189. Theorem 1, where convergence at each x of $P_n x$ is proved, with P_n operators

*or <u>conditionally complete</u>, in Birkhoff's terminology (1940).

in Hilbert space]. Now let $P = \{P_\alpha : \alpha \in A\}$ be an arbitrary set of P's with $P_\alpha < \hat{P}$, $\alpha \in A$. If A is countable let $P = \{P_n\}$; define

$$\overline{P}_n = \sup \{P_1, P_2, \ldots, P_n\} ;$$

and, noting that $\hat{P} \geq \overline{P}_n$, define $\overline{P} = \lim \overline{P}_n$; now check that

$$\overline{P} = \sup \{P_\alpha : \alpha \in A\} .$$

If $P_\alpha > \hat{P}$ and A is countable, $\inf \{P_\alpha : \alpha \in A\}$ is defined similarly. Finally if A is not countable, we use separability of Euclidean space to conclude that for each positive integer k there exists a sequence $\{P_{kn} : n = 1,2,\ldots\}$ such that

$$\sup_{\alpha \in A} \inf_n \|P_\alpha - P_{kn}\| < \frac{1}{k} ,$$

and

$$\sup_n \inf_{\alpha \in A} \|P_{kn} - P_\alpha\| < \frac{1}{k} .$$

Define $\overline{P}_k = \sup_n P_{kn}$ and then $\overline{P} = \lim_k \overline{P}_k$, and verify that $\overline{P} = \sup \{P_\alpha : \alpha \in A\}$.

Reference for the general theory: Garrett Birkhoff (1940), Lattice Theory, Amer. Math. Soc. Colloq. Pub. vol. 25.

[Discussion of lattices of symmetric matrices and application to the proof of the Theorem, p.25, originate with these notes].

REFERENCES

[1] Beckenbach, E., and Bellman, R., Inequalities, Springer-Verlag, Berlin, 1961.

[2] Birkhoff, G., Lattice Theory, Amer. Math. Soc. Colloq. Publ., vol. 25, 1940.

[3] Doob, J.L., Stochastic Processes, New York, Wiley, 1953.

[4] Kalman, R.E., A new approach to linear filtering and prediction problems, J. Basic. Engrg., March, 1960, pp. 35-45.

[5] Kalman, R.E., and Bucy, R.S., New results in linear filtering and prediction theory, J. Basic Engrg., March, 1961, pp. 95-108.

[6] Kolmogorov, A.N., Interpolation und Extrapolation von stationären zufälligen Folgen, Bull. de l'Acad. des Sci. de l'URSS, Ser. Math., 5, 1941.

[7] Munroe, M.E., Introduction to Measure and Integration, Addison-Wesley, Cambridge, Mass., 1953.

[8] Wiener, N., Extrapolation, Interpolation and Smoothing of Stationary Time Series, New York, John Wiley and Sons, 1949.

APPENDIX. Miscellaneous results for positive symmetric matrices.

1. If $P > Q > 0$ then $P^{-1} < Q^{-1}$.

PROOF. There is a nonsingular matrix T such that $T'QT = I$; so $P > Q > 0 \Rightarrow T'PT > I$. But $R > I \Rightarrow x'Rx > |x|^2 \Rightarrow |\sqrt{R}x|^2 > |x|^2 \Rightarrow |y|^2 > |R^{-1/2}y|^2 = y'R^{-1}y \Rightarrow R^{-1} < I$. Thus $(T'PT)^{-1} < I$, or $T^{-1}P^{-1}(T^{-1})' < I \Rightarrow x'T^{-1}P^{-1}(T^{-1})'x < |x|^2$ $\Rightarrow y'P^{-1}y < |T'y|^2 = y'TT'y$, all $y \neq 0$. (*) Now T can be chosen in the form $T = SD^{-1/2}$ where S is orthogonal and D is diagonal; thus $Q = SDS'$ and (*) $\Rightarrow P^{-1} < TT' = SD^{-1}S' = Q^{-1}$. For a neater but less direct proof see Beckenbach and Bellman (1961) p.86.

2. If $P \geq 0$ and $Q \geq 0$ then $(I+PQ)^{-1}$ exists, and $(I+PQ)^{-1}P \geq 0$ and symmetric. [We do not assume P,Q commute and so cannot infer $PQ \geq 0$].

PROOF. (a) $P > 0 \Rightarrow (I+PQ)^{-1} = (P^{-1}+Q)^{-1}P^{-1} \Rightarrow (I+PQ)^{-1}P = (P^{-1}+Q)^{-1} > 0$.

(b) LEMMA: If $P \geq 0$ and $Q \geq 0$, there exist $\tilde{P}$, $\tilde{Q}$ such that $\tilde{P} > 0$, $Q \geq 0$, and

$$PQ = \tilde{P}\tilde{Q}.$$

For the proof let R,N denote range and null space and recall that for a symmetric matrix P, $R(P) \perp N(P)$ and $R(P) \oplus N(P) = E$, where E is the whole space. Now define $\tilde{P},\tilde{Q}$ by setting $\tilde{P}x = Px$,

$x \in R(P)$; $\tilde{P}x = x$, $x \in N(P)$; and $\tilde{Q}x = Qx$, $x \in R(QP)$; $\tilde{Q}x = 0$, $x \in [R(QP)]^{\perp} = N(PQ)$. Then $x \in N(PQ) \Rightarrow \tilde{P}\tilde{Q}x = 0$, and $0 \neq x \in R(QP) \Rightarrow PQx \neq 0$. Thus $x \in R(QP) \Rightarrow Qx \in R(P) \Rightarrow \tilde{P}\tilde{Q}x = \tilde{P}Qx = PQx$. Clearly $x \in N(PQ) \Rightarrow \tilde{P}\tilde{Q}x = 0$.

(c) With $\tilde{P},\tilde{Q}$ as in (b) we have

$$(I+PQ)^{-1} = (I+\tilde{P}\tilde{Q})^{-1}$$
$$= (\tilde{P}^{-1}+Q)^{-1}\tilde{P}^{-1},$$

i.e., the inverse exists, by (a).

(d) LEMMA: If A is an arbitrary nonsingular matrix and if

$$\inf_{|x|=1} |Ax| = \alpha,$$

then

$$|A^{-1}| \leq \alpha^{-1}.$$

For the proof, note that $|x| > \alpha^{-1} \Rightarrow x = \beta y$ with $\beta > \alpha^{-1}$, $|y| = 1$. Then $|Ax| = \beta|Ay| \geq \beta\alpha > 1$; that is, $|Ax| \leq 1 \Rightarrow |x| \leq \alpha^{-1}$, and so

$$|A^{-1}| = \sup_{|y|=1} |A^{-1}y| = \sup_{\{x:|Ax|=1\}} |A^{-1}Ax|$$

$$\leq \sup_{\{x:|Ax|\leq 1\}} |x|$$

$$\leq \sup_{\{x:|x|\leq\alpha^{-1}\}} |x| = \alpha^{-1}.$$

(e) LEMMA. If A^{-1}, A_n^{-1} $(n = 1,2,\ldots)$ exist and $A_n \to A$, then $A_n^{-1} \to A^{-1}$.

For the proof let $\inf_{|x|=1} |Ax| = 2\alpha > 0$; then if $|x| = 1$, $|A_n x| \geq |Ax| - |(A-A_n)x| \geq \alpha$ for all n sufficiently large. By (d), $|A_n^{-1}| < \alpha^{-1}$ for all such n, and then

$$|A_n^{-1} - A^{-1}| \leq |A_n^{-1}||A^{-1}||A-A_n| \to 0 \qquad (n \to \infty)$$

(f) Let $P_n = P + n^{-1} I$, $n = 1,2,\ldots$, and put

$$R_n = (I + P_n Q)^{-1} P_n$$

$$= (P_n^{-1} + Q)^{-1} > 0 .$$

Now $n > m \Rightarrow P_n < P_m \Rightarrow P_n^{-1} > P_m^{-1}$

$$\Rightarrow R_n < R_m .$$

So $\lim_{n \to \infty} R_n = R_\infty$ (say) exists and $R_\infty \geq 0$. We show that $R_\infty = (I + PQ)^{-1} P$. By (b), $(I + P_n Q)^{-1}$ and $(I + PQ)^{-1}$ exist; also

$$I + P_n Q \to I + PQ \qquad (n \to \infty) .$$

Hence by (e),

$$(I + P_n Q)^{-1} \to (I + PQ)^{-1}$$

and clearly

$$P_n \to P \ .$$

Therefore $R_\infty = (I + PQ)^{-1} P$; so

$$(I + PQ)^{-1} P \geq 0 \ .$$

Q.E.D.

CHAPTER IV

ITÔ'S STOCHASTIC INTEGRAL AND APPLICATIONS

1. STOCHASTIC INTEGRAL.

In this section we generalize the stochastic integral introduced in Chapter 2,§3, in assuming now that the integrand is itself a random function. This generalization is indispensable for our later discussion of nonlinear stochastic differential equations. The material of this section is based on Gikhman and Skorokhod (1965) Ch.8,§2; alternative sources are Skorokhod (1965),Ch.2, and Doob (1953),Ch.9,§5.

Let $\{w(t,\omega): t \geq 0\}$ be a Wiener process in R_1 and let $\{f(t,\omega): t \geq 0\}$ be a stochastic process in R_1. In applications, the random variable $f(t)$ will depend, at most, on 'present' and 'past' random variables $w(s)$, $s \leq t$, but not on future values of $w(s)$, $s > t$. To formalize this kind of nonanticipative dependence of f on w, let $\{\mathfrak{F}_t : t \geq 0\}$ be a family of σ-algebras of ω-sets, with the properties:

(i) $\mathfrak{F}_{t_1} \subset \mathfrak{F}_{t_2}$ if $t_1 < t_2$.

(ii) For each t, $w(t)$ is measurable relative to $\mathfrak{F}_t$.

(iii) For $t_1 > t_2 \geq t$, the increments $w(t_1)-w(t_2)$ are independent of the σ-algebra $\mathfrak{F}_t$.

(iv) For each t the random variable $f(t)$ is measurable relative to $\mathfrak{F}_t$.

A typical example of $\{\mathfrak{F}_t\}$ might be the family of σ-algebras $\mathfrak{F}_t$ generated by the random variables $w(s)$ for $0 \leqq s \leqq t$.

We also assume that f is measurable in (t,ω) and that $f(\cdot,\omega)$ is locally $\mathcal{L}_2$ with probability one, that is

$$\int_0^T |f(t)|^2 dt < \infty \qquad (T < \infty) \qquad (1)$$

with probability 1. In the following, M_2 denotes the class of functions f just described; our aim is to define the integral

$$\int_a^b f(t,\omega)dw(t,\omega) \qquad (2)$$

for $0 \leqq a < b < \infty$ and $f \in M_2$. In framing a definition we proceed as before: namely, define (2) for f a step function, then use a completeness argument to obtain the general integral as the limit of an approximating sequence. The first requirement is thus to show that every $f \in M_2$ can be suitably approximated.

Definition. $f = f(t,\omega)$ is a step function if there exists a partition $a = t_0 < t_1 < \ldots < t_r = b$, such that $f(t,\omega) = f(t_i,\omega)$, $t_i \leqq t < t_{i+1}$, $i = 0,1,\ldots,r-1$. [Note: here the t_i are independent of ω].

LEMMA 1:

For each $f \in M_2$ there exists a sequence of step functions $f_n \in M_2$ such that

$$\lim_{n\to\infty} \int_a^b |f(t)-f_n(t)|^2 dt = 0$$

with probability 1.

PROOF. The assertion is obvious if f is continuous: put $f_n(t) = f(k/n)$ for $(k/n) \leqq t \leqq (k+1)/n$. If f is measurable and bounded (by a constant depending on ω), there is a sequence g_n of continuous functions, correspondingly bounded, such that, with probability 1, $g_n(t) \to f(t)$ for almost all t, and thus (by dominated convergence)

$$\lim_{n\to\infty} \int_a^b |f(t)-g_n(t)|^2 dt = 0 \ . \qquad \text{a.s.}$$

[For instance, set $g_n(t) = n \int_a^t e^{n(s-t)} f(s)ds$]. Finally, any $f \in M_2$ can be approximated in mean square by bounded (measurable) functions in M_2. Thus the step functions are dense in M_2, in the sense of $\mathcal{L}_2[a,b]$ convergence. Q.E.D.

To define the integral, if f is a step function, write

$$\int_a^b f(t)dw(t) = \sum_{k=0}^{r-1} f(t_k)[w(t_{k+1})-w(t_k)] \ . \qquad (3)$$

The integral (3) has the properties:

1. *If* $\mathcal{E}\{|f(t)| \mid \mathfrak{F}_a\} < \infty$ *for* $t \in [a,b]$ *then*

$$\mathcal{E}\{\int_a^b f(t)dw(t) \mid \mathfrak{F}_a\} = 0 \qquad \text{(a.s.)} \qquad (4)$$

(Proof : computation)

2. <u>If</u> $\mathcal{E}\{|f(t)|^2 \mid \mathfrak{F}_a\} < \infty$ a.s. then for $t \in [a,b]$

$$\mathcal{E}\{[\int_a^b f(t)dw(t)]^2 \mid \mathfrak{F}_a\} = \int_a^b \mathcal{E}\{f(t)^2 \mid \mathfrak{F}_a\}dt \quad \text{a.s.} \qquad (5)$$

(Proof : computation)

3. For every $N > 0$ and $c > 0$,

$$P\{|\int_a^b f(t)dw(t)| > c\} \leq \frac{N}{c^2} + P\{\int_a^b |f(t)|^2 dt > N\} \qquad (6)$$

PROOF. Noting that $f(t)$ is $\mathfrak{F}_{t_i}$-measurable for $t_i \leq t < t_{i+1}$, we have that

$$\int_a^{t_{i+1}} |f(t)|^2 dt$$

is $\mathfrak{F}_{t_i}$-measurable. Put $f_N(t) = f(t)$ if

$$\int_a^{t_{i+1}} |f(t)|^2 dt \leq N ,$$

where t_{i+1} is such that $t_i \leq t < t_{i+1}$, and $f_N(t) = 0$ for $t_i \leq t \leq t_{i+1}$ if

$$\int_a^{t_{i+1}} |f(t)|^2 dt > N \quad .$$

Then f_N is a step function in $M_2[a,b]$. Since

$$\int_a^b |f_N(t)|^2 dt = \sum_{i=0}^{r-1} |f_N(t_i)|^2 (t_{i+1}-t_i) \leqq N$$

then $|f_N(t_i)|^2 \leqq \frac{N}{t_{i+1}-t_i}$, and so $\mathcal{E}\{|f_N(t)|^2\} < \infty$.

Therefore

$$\mathcal{E}\{|\int_a^b f_N(t)dw(t)|^2\} = \int_a^b \mathcal{E}\{|f_N(t)|^2\}dt \leqq N .$$

Finally

$$P\{\sup_{t\in[a,b]} |f_N(t)-f(t)| > 0\} = P\{\int_a^b |f(t)|^2 dt > N\} .$$

From these relations we have

$$\begin{aligned}
&P\{\int_a^b f(t)dw(t)| > c\} \\
&\quad = P\{|\int_a^b f_N(t)dw(t) + \int_a^b [f(t)-f_N(t)]dw(t)| > c\} \\
&\quad = P\{|\int_a^b f_N(t)dw(t)| > c\} + \\
&\qquad + P\{|\int_a^b [f(t)-f_N(t)]dw(t)| > 0\} \\
&\quad \leqq \frac{\mathcal{E}|\int_a^b f_N(t)dw(t)|^2}{c^2} + P\{\int_a^b |f(t)|^2 dt > N\} ,
\end{aligned}$$

and the proof is complete.

Now suppose f_n are step functions, $f \in M_2$, and

$$P \lim_{n \to \infty} \int_a^b |f(t)-f_n(t)|^2 dt = 0 .$$

The existence of such a sequence is guaranteed by Lemma 1. Then

$$P \lim_{\substack{n \to \infty \\ m \to \infty}} \int_a^b |f_n(t)-f_m(t)|^2 dt = 0 ;$$

that is, for every $\epsilon > 0$,

$$\lim_{n,m \to \infty} P\{\int_a^b |f_n(t)-f_m(t)|^2 dt > \epsilon\} = 0 .$$

Using (6), we have for arbitrary $\epsilon > 0$, $\delta > 0$,

$$\overline{\lim_{n,m \to \infty}} P\{\left|\int_a^b f_n(t)dw(t)-\int_a^b f_m(t)dw(t)\right| > \delta\}$$

$$\leq \frac{\epsilon}{\delta^2} + \overline{\lim_{n,m \to \infty}} P\{\int_a^b |f_n(t) - f_m(t)|^2 dt > \epsilon\}$$

$$= \epsilon/\delta^2 .$$

Since $\epsilon > 0$ was arbitrary, we have for every $\delta > 0$,

$$\overline{\lim_{n,m \to \infty}} P\{\left|\int_a^b f_n(t)dw(t)-\int_a^b f_m(t)dw(t)\right| > \delta\} = 0 .$$

This means that the sequence $\int_a^b f_n(t)dw(t)$ converges in probability to a limit, which we denote by

$$\int_a^b f(t)dw(t) \quad .$$

As usual, this limit is uniquely defined except for its values on an ω-set of P-measure zero. Our definition of the Itô integral is complete.

The integral is, obviously, an additive and homogeneous functional of the argument $f \in M_2$; also, if $a < c < b$,

$$\int_a^b f\,dw = \int_a^c f\,dw + \int_c^b f\,dw$$

with probability 1. We can now regard $\int_a^b f\,dw$ as consistently defined for all $0 \leqq a < b < \infty$.

Property 3 (eq.(6)) can be verified by passing to the limit, through an approximating sequence of step functions, in (6). We can then obtain the property:

4. <u>If</u> $f \in M_2[a,b]$, $f_n \in M_2[a,b]$ <u>and</u>

$$P \lim_{n \to \infty} \int_a^b |f_n(t)-f(t)|^2 dt = 0$$

<u>then</u>

$$P \lim_{n \to \infty} \int_a^b f_n(t)dw(t) = \int_a^b f(t)dw(t)$$

[Proof: exercise]

Next we generalize properties 1 and 2:

2*. <u>If</u> f <u>is such that</u>

$$\int_a^b \mathcal{E}\{|f(t)|^2 \mid \mathfrak{F}_a\}dt < \infty \qquad \text{a.s.}$$

<u>then</u>

$$\mathcal{E}\{\int_a^b f(t)dw(t) \mid \mathfrak{F}_a\} = 0 \qquad \text{a.s.} \tag{6}$$

<u>and</u>

$$\mathcal{E}\{[\int_a^b f(t)dw(t)]^2 \mid \mathfrak{F}_a\} = \int_a^b \mathcal{E}\{|f(t)|^2 \mathfrak{F}_a\}dt \quad \text{a.s.} \tag{7}$$

For the proof choose a sequence of step functions f_N such that

$$P \lim_{n\to\infty} \int_a^b \mathcal{E}\{|f(t)-f_N(t)|^2 \mid \mathfrak{F}_a\} = 0$$

[Let $\bar{f}_n \in M_2[a,b]$ be a sequence of step functions such that

$$P \lim_{n\to\infty} \int_a^b |f(t)-\bar{f}_n(t)|^2 dt = 0 .$$

The existence of such a sequence is guaranteed by Lemma 1. Put $g_N(x) = x$ for $|x| \leq N$, $g_N(x) = Nx/|x|$ for $|x| > N$. Since $|g_N(x)-g_N(y)| \leq |x-y|$,

$$\int_a^b |g_N(f(t))-g_N(\bar{f}_n(t))|^2 dt \leq \int_a^b |f(t)-\bar{f}_n(t)|^2 dt$$

$$\to 0$$

in probability. Since the quantities

$$\int_a^b |g_N(f(t))-g_N(\bar{f}_n(t))|^2 dt$$

are bounded by $4N^2(b-a)$ and tend to zero $(n \to \infty)$ in probability we have by dominated convergence [Lebesgue's Theorem holds under the hypothesis of convergence in measure]

$$\mathcal{E}\{\int_a^b |g_N(f(t))-g_N(\bar{f}_n(t))|^2 dt \mid \mathfrak{F}_a\}$$

$$= \int_a^b \mathcal{E}\{|g_N(f(t))-g_N(\bar{f}_n(t))|^2 \mid \mathfrak{F}_a\} dt \qquad \text{(a.s.)}$$

$$\to 0 \quad (n \to \infty)$$

in probability. On the other hand, since

$$|f(t)-g_N(f(t))|^2 \leq |f(t)|^2 \qquad \text{a.s.}$$

and

$$\int_a^b \mathcal{E}\{|f(t)|^2 \mid \mathfrak{F}_a\} < \infty \qquad \text{a.s.}$$

there follows, again by dominated convergence,

$$P \lim_{N \to \infty} \int_a^b \mathcal{E}\{|g_N(f(t))-f(t)|^2 | \mathfrak{F}_a\}dt = 0 .$$

Thus there exists a sequence n_N such that

$$\int_a^b \mathcal{E}\{|f(t)-g_N(\bar{f}_{n_N}(t))|^2 | \mathfrak{F}_a\}dt$$

$$\leq 2\int_a^b \mathcal{E}\{|f(t)-g_N(f(t))|^2 | \mathfrak{F}_a\}dt$$

$$+ 2\int_a^b \mathcal{E}\{|g_N(f(t))-g_N(\bar{f}_{n_N}(t))|^2 | \mathfrak{F}_a\}dt$$

$$\to 0 \quad \text{in probability.}$$

Finally we write $f_N(t) = g_N(\bar{f}_{n_N}(t))$].

To complete the proof, substitute the f_N in (4) and going to the limit, obtain (6).

To obtain (7), observe that if $P \lim \xi_N = \xi$ and if ξ_N is fundamental in mean square then l.i.m. $\xi_N = \xi$, and of course $\mathcal{E}\{\xi_N^2\} \to \mathcal{E}\{\xi^2\}$. Thus

$$\mathcal{E}\{[\int_a^b f(t)dw(t)]^2 | \mathfrak{F}_a\}$$

$$= \lim_{N \to \infty} \mathcal{E}\{[\int_a^b f_N(t)dw(t)]^2 | \mathfrak{F}_a\} \qquad \text{a.s.}$$

$$= \lim_{N \to \infty} \int_a^b \mathcal{E}\{|f_N(t)|^2 | \mathfrak{F}_a\}dt \qquad \text{a.s.}$$

$$= \int_a^b \mathcal{E}\{|f(t)|^2 | \mathfrak{F}_a\}dt \qquad \text{a.s.}$$

Q.E.D.

Next we consider the integral as a function of the upper limit. The fact that, for each t, the integral

$$x(t,\omega) = \int_a^t f(s,\omega)dw(s,\omega)$$

is uniquely determined except for its values on an ω-set of P-measrue zero, allows us to assume (Theorem 1,§1.1) that $x(t,\cdot)$ is defined in such a way that the $x(t)$ process is separable. From now on we take separability for granted.

Next we note, without proof, the properties:

5. Submartingale inequality:

$$\text{If } \int_a^b \mathcal{E}\{|f(s)|^2 | \mathfrak{F}_a\}ds < \infty, \text{ then}$$

$$P\{\sup_{a\leq t\leq b} |\int_a^t f(s)dw(s)| > c | \mathfrak{F}_a\}$$

$$\leq \frac{1}{c^2}\int_a^b \mathcal{E}\{|f(s)|^2 | \mathfrak{F}_a\}ds \qquad \text{a.s.} \qquad (8)$$

6. The separable process

$$x(t) = \int_a^t f(s,\omega)dw(s)$$

is continuous. [Proof: Use (8), together with previous approximation techniques, and the Borel-Cantelli lemma].

Next we introduce the important concept of stochastic differential. Let $\{x(t,\omega)\}$ be a process (defined for $t \geq 0$) such that for each t the random variable $x(t,\cdot)$ is measurable relative to the σ-algebra $\mathfrak{F}_t$, where the $\mathfrak{F}_t$ are as before. Suppose there exist a function $b(\cdot) \in M_2$ and an $\mathfrak{F}_t$-measurable function $a(\cdot)$, with

$$\int_0^t |a(s)|\,ds < \infty \qquad (t < \infty) \qquad \text{a.s.},$$

such that

$$x(t_2)-x(t_1) = \int_{t_1}^{t_2} a(t)dt + \int_{t_1}^{t_2} b(t)dw(t) \tag{9}$$

for all $0 \leq t_1 < t_2 < \infty$. The symbol $adt + bdw$ is the (Itô) <u>stochastic differential</u> of x, and we write

$$dx(t) = a(t)dt + b(t)dw(t)\ . \tag{10}$$

It is important to keep in mind that (10) is merely a convenient and suggestive notation for (9). However, it is natural to explore the possibility of a 'stochastic differential calculus', and in this direction we derive:

7. <u>Itô's chain rule.</u>

<u>Let $u(t,x)$ be a (nonrandom) continuous function of</u>

(t,x) *for* $t \geq 0$ *and* $x \in R_1$, *with continuous derivatives* $u_t(t,x)$, $u_x(t,x)$, $u_{xx}(t,x)$. *Let* $X = \{x(t): t \geq 0\}$ *be a process with stochastic differential*

$$dx = a(t)dt + b(t)dw \ . \tag{11}$$

Then the process

$$y(t) = u[t,x(t)] \tag{12}$$

has the stochastic differential

$$\begin{aligned} dy(t) = [u_t[t,x(t)] &+ u_x[t,x(t)]a(t) \\ &+ \frac{1}{2} u_{xx}[t,x(t)]b(t)^2]dt + \\ &+ u_x[t,x(t)]b(t)dw(t) \ . \end{aligned} \tag{13}$$

Before proving (13), we note that it is easily obtained heuristically by formal Taylor expansion of $u(t+dt,x+dx)$ and by use of the differential formulas

$$\begin{aligned} dt\ dw &= dx\ dw = 0 \\ (dw)^2 &= dt \ . \end{aligned} \tag{14}$$

Thus our proof will be a rigorous justification of a 'calculus' based on (14).

For the proof, suppose first that a,b are independent of t. Let t_1,t_2 be fixed, and $t_1 = t^{(0)} < \ldots < t^{(n)} = t_2$; then

$$y(t_2)-y(t_1) = u(t_2,x(t_2))-u(t_1,x(t_1))$$

$$= \sum_{k=0}^{n-1} [u(t^{(k+1)},x(t^{(k+1)}))-u(t^{(k)},x(t^{(k)}))\ ].$$

By Taylor's formula,

$$u[t^{(k+1)},x(t^{(k+1)})]-u[t^{(k)},x(t^{(k)})]$$

$$= u_t[t^{(k)}+\theta_k(t^{(k+1)}-t^{(k)}),x(t^{(k+1)})][t^{(k+1)}-t^{(k)}]$$

$$+ u_x[t^{(k)},x(t^{(k)})][x(t^{(k+1)})-x(t^{(k)})] +$$

$$+ \frac{1}{2} u_{xx}[t^{(k)},x(t^{(k)})+\theta'_k\{x(t^{(k+1)})-x(t^{(k)})\}]\cdot$$

$$\cdot[x(t^{(k+1)})-x(t^{(k)})]^2$$

where $0 < \theta_k,\theta'_k < 1$. Since $x(\cdot)$ is continuous and bounded with probability 1, there exist random variables α and β such that,

with probability 1,

$$\left.\begin{array}{l}\alpha \to 0\\ \beta \to 0\end{array}\right\} \quad \text{as} \quad \max_k \,(t^{(k+1)}-t^{(k)}) \to 0$$

and

$$|u_t[t^{(k)}+\theta_k(t^{(k+1)}-t^{(k)}),x(t^{(k)})] -$$

$$-u_t[t^{(k)},x(t^{(k)})]| \leqq \alpha$$

$$|u_{xx}[t^{(k)},x(t^{(k)}+\theta'_k[x(t^{(k+1)})-x(t^{(k)})]] -$$

$$-u_{xx}(t^{(k)},x(t^{(k)}))| \leqq \beta \quad .$$

Then, from the relations

$$\sum [x(t^{(k+1)})-x(t^{(k)})]^2 \leqq 2a^2 \sum [t^{(k+1)}-t^{(k)}]^2 +$$

$$+2b^2\sum [w(t^{(k+1)})-w(t^{(k)})]^2$$

and

$$\varepsilon\sum [w(t^{(k+1)})-w(t^{(k)})]^2 = t_2-t_1 \,,$$

we conclude that

$$|\Sigma u_t[t^{(k)}+\theta_k(t^{(k+1)}-t^{(k)}),x(t^{(k)})](t^{(k+1)}-t^{(k)}) -$$

$$-\Sigma u_t[t^{(k)},x(t^{(k)})](t^{(k+1)}-t^{(k)})| = o(\alpha)$$

$$|\Sigma u_{xx}[t^{(k)},x(t^{(k)})+\theta'_k[x(t^{(k+1)})-x(t^{(k)})][x(t^{(k+1)})-x(t^{(k)})]^2$$

$$-\Sigma u_{xx}[t^{(k)},x(t^{(k)})][x(t^{(k+1)})-x(t^{(k)})]^2|$$

$$\leq \beta \Sigma [x(t^{(k+1)})-x(t^{(k)})]^2 \to 0$$

in probability as $\max (t^{(k+1)}-t^{(k)}) \to 0$.

Now substitute

$$x(t^{(k+1)})-x(t^{(k)}) = a(t^{(k+1)}-t^{(k)}) +$$

$$+ b(w(t^{(k+1)})-w(t^{(k)}))$$

and use:

$$a^2 \sum_{k=0}^{n-1} (t^{(k+1)}-t^{(k)})^2 \to 0$$

$$b \sum_{k=0}^{n-1} (w(t^{(k+1)})-w(t^{(k)})(t^{(k+1)}-t^{(k)}) \to 0$$

a.s. as $\max_k (t^{(k+1)}-t^{(k)}) \to 0$; to obtain

$$y(t_2)-y(t_1) = \operatorname*{P\,lim}_{\max(t^{(k+1)}-t^{(k)})\to 0}\Bigg[\sum_{k=0}^{n-1} u_t(t^{(k)},x(t^{(k)}))\cdot$$

$$\cdot(t^{(k+1)}-t^{(k)}) + \sum_{k=0}^{n-1} u_x(t^{(k)},x(t^{(k)}))a(t^{(k+1)}-t^{(k)})$$

$$+ \sum_{k=0}^{n-1} u_x(t^{(k)},x(t^{(k)}))b[w(t^{(k+1)})-w(t^{(k)})] +$$

$$+ \frac{b^2}{2}\sum_{k=0}^{n-1} u_{xx}(t^{(k)},x(t^{(k)}))(t^{(k+1)}-t^{(k)})$$

$$+ \frac{b^2}{2}\sum_{k=0}^{n-1} u_{xx}(t^{(k)},x(t^{(k)}))[\{w(t^{(k+1)})-w(t^{(k)})\}^2 -$$

$$-(t^{(k+1)}-t^{(k)})]\Bigg].$$

It is clear that all sums except the last tend in the limit to corresponding integrals. The result (13) will follow (for constant a,b) as soon as we show that the limit of the last term is zero. For this put

$$\epsilon_k = [w(t^{(k+1)})-w(t^{(k)})]^2-(t^{(k+1)}-t^{(k)})$$

and denote by $\chi_k^{(N)}$ the indicator of the event $\{|x(t^{(i)})| \leqq N$ for $i \leqq k\}$. Then

$$\mathcal{E}\{[\sum_{k=0}^{n-1} u_{xx}(t^{(k)},x(t^{(k)}))\chi_k^{(N)}\epsilon_k]^2\}$$

$$\leq \sum_{k=0}^{n-1} \mathcal{E}\{u_{xx}(t^{(k)}, x(t^{(k)}))^2 \chi_k^{(N)} \epsilon_k^2\}$$

$$\leq \sup_{\substack{|x|\leq N \\ t}} |u_{xx}(t,x)|^2 \sum \mathcal{E}\{\epsilon_k^2\}$$

$$= 3 \sup_{\substack{t \\ |x|\leq N}} |u_{xx}(t,x)|^2 \sum (t^{(k+1)} - t^{(k)})^2$$

$$\to 0 \qquad \text{as} \quad \max_k (t^{(k+1)} - t^{(k)}) \to 0 \ ;$$

and also

$$P\{\sum_{k=0}^{n-1} u_{xx}(t^{(k)}, x(t^{(k)}))^2 (1-\chi_k^{(N)}) \epsilon_k^2 \neq 0\} \leq$$

$$P\{\sup_{t_1 \leq t \leq t_2} |x(t)| > N\} \to 0 \qquad (N \to \infty) \ .$$

Thus (13) is proved for constant a,b. It is clear, then, that (13) is true when a,b are step functions (see p.3) -- we simply repeat the foregoing discussion for each interval where a,b are both constant. In the general case one chooses* a sequence of step functions a_n, b_n such that, with probability 1,

$$\int_{t_1}^{t_2} |a(t) - a_n(t)| \, dt \to 0$$

$$\int_{t_1}^{t_2} |b_n(t) - b(t)|^2 dt \to 0 \ ,$$

*by Lemma 1, p.3, and its $\mathcal{L}_1$ counterpart.

and the sequence of processes

$$x_n(t) = x(t_1) + \int_{t_1}^{t} a_n(s)ds + \int_{t_1}^{t} b_n(s)dw(s)$$

tends with probability 1 uniformly to $x(t)$. Then $y_n(t)=u(t,x_n(t))$ tends uniformly a.s. to $y(t)$. The general result (13) then follows by passing to the limit in

$$y_n(t_2)-y_n(t_1) = \int_{t_1}^{t_2} [u_t(t,x_n(t))+u_x(t,x_n(t))a_n(t)$$

$$+ \tfrac{1}{2} u_{xx}(t,x_n(t))b_n(t)^2]dt+ \int_{t_1}^{t_2} u_x(t,x_n(t))b_n(t)dw(t).$$

Extension to E_n. All our previous discussion is easily generalized for processes in E_n. In fact if $w(\cdot)$ is a Wiener process in E_m and f is a vector (E_m-valued) function whose components f_i belong to M_2 then

$$\int_{t_1}^{t_2} f(t)'dw(t) = \sum_{i=1}^{m} \int_{t_1}^{t_2} f_i(t)dw_i(t)$$

is defined in the obvious way; and similarly if f is an $n \times m$ matrix-valued function with components $f_{ij} \in M_2$. The properties 1 to 6 remain valid, where $|f|$ now denotes Euclidean norm of f. The stochastic differential generalizes straightforwardly, as does Itô's chain rule. For the latter, let

$$a(t) = [a_i(t)]_{i=1,\ldots,n}$$

$$B(t) = [b_{ij}(t)]_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}}$$

where the functions a_i, b_{ij} satisfy the same conditions as the corresponding scalars a,b of the previous discussion. Let $u = u(t,x)$ be defined and continuous for $t \geq 0$, $x \in E_n$, with continuous derivatives u_t, u_x, u_{xx}. [Here u_x is the vector of first partial x-derivatives, u_{xx} the matrix of second partials]. Let $\{x(t)\}$ have stochastic differential

$$dx(t) = a(t)dt + B(t)dw(t) .$$

Then the process $y(t) = u[t,x(t)]$ has stochastic differential

$$dy(t) = [u_t[t,x(t)] + u_x[t,x(t)]'a(t) +$$

$$+ \frac{1}{2} \operatorname{tr}\{B(t)'u_{xx}[t,x(t)]B(t)\}]dt +$$

$$+ u_x[t,x(t)]'B(t)dw(t).$$

PROBLEMS.

1. Let $x(t)$, $y(t)$ be processes defined for $t_o \leq t \leq T$; and suppose

$$dx(t) = dy(t) + \varphi(t)dt , \qquad t_o \leqq t \leqq T ,$$

$$x(t_o) = y(t_o) = 0 .$$

Denote by $\mathcal{X}_t$ (resp. $\mathcal{Y}_t$) the σ-algebra of ω-sets generated by the $x(s)$ (resp. $y(s)$) for $t_o \leqq s \leqq t$. If $\varphi = \varphi(t,\omega)$ is a measurable process such that $\varphi(t)$ is $\mathcal{Y}_t$-measurable for each t, and

$$\int_{t_o}^{T} |\varphi(t)|\, dt < \infty \qquad \text{a.s.}$$

Show that $y(t)$ is $\mathcal{X}_t$-measurable for each t. [Suggestion: use the counterpart for $L_1[t_o,T]$ of Lemma 1, p.3, to approximate φ by $\mathcal{Y}_t$-measurable step functions].

2. If w is a Wiener process in R_1, evaluate

$$\int_a^b [w(t)]^n dw(t) \qquad , \qquad n = 1,2 .$$

[Suggestion: for n=1, look for a function whose stochastic differential coincides with the given differential; for n=2 integrate formally by parts and verify the result.]

2. INTEGRATION RULE

To extend the class of functions (control laws) within which the linear control law found in §2.6 is optimal, we apply a generalized version of the integration rule obtained in §2.5.

This result is an immediate consequence of Itô's chain rule, but for future reference we state it here as a theorem.

In the following, all processes etc. are defined for $t \in I$ where $I = [t_o,T]$ or $[t_o,\infty)$. Notation and assumptions:

(i) $\{w(t): t \in I\}$ is a Wiener process in E_m

(ii) $\{\mathfrak{F}_t: t \in I\}$ is a family of σ-algebras of ω-sets such such that $\mathfrak{F}_{t_1} \subset \mathfrak{F}_{t_2}$ if $t_1 < t_2$; $w(t)$ is $\mathfrak{F}_t$-measurable; and the random variables $w(t_2)-w(t_1)$ for $t_2 > t_1 \geq t$ are independent of $\mathfrak{F}_t$.

(iii) $f = f(t,\omega)$ is a (t,ω)-measurable, n-vector valued function such that the random variable $f(t) = f(t,\cdot)$ is $\mathfrak{F}_t$-measurable* $(t \in I)$ and, with probability 1,

$$\int_{t_o}^{T} |f(t)|\,dt < \infty \quad ** \qquad (T < \infty\ ;\ T \in I)$$

(iv) $G = G(t,\omega)$ is a (t,ω)-measurable, $(n \times m)$-matrix valued function such that the random variable $G(t) = G(t,\cdot)$ is $\mathfrak{F}_t$-measurable** for each $t \in I$ and, with probability one,

---

* A vector or matrix is measurable if each component is measurable.

** $|\cdot|$ denotes Euclidean norm.

$$\int_{t_o}^{T} |G(t,\omega)|^2 dt < \infty \qquad (T < \infty\ ;\ T \in I)$$

(v) $X = \{x(t):\ t \in I\}$ is the process determined by the initial condition $x(t_o) = x_o$ (where x_o is an E_n-valued, $\mathfrak{S}_{t_o}$-measurable random variable) and the stochastic differential

$$dx(t) = f(t)dt + G(t)dw(t).$$

(vi) $u(t,x)$ is a real-valued function defined for $x \in E_n$, $t \in I$, such that u, u_t, u_x, and u_{xx} are continuous.

(vii) We write:

$$\mathcal{L}u(t,x) \equiv \frac{1}{2} \operatorname{tr}[G(t)' u_{xx}(t,x) G(t)] + f(t)' u_x(t,x)$$

and refer to $\mathcal{L}$ as the differential generator of the X-process.*

THEOREM 1 (Integration Rule). Subject to the above assumptions, if $0 \leq t_1 < t_2 \in I$, then

---

* Properly, this term applies only when the $x(t)$-process is Markov and $\mathcal{L}$ 'generates' a semi-group; but we use it when convenient.

$$u[t_2,x(t_2)]-u[t_1,x(t_1)] =$$

$$= \int_{t_1}^{t_2} \{u_t[t,x(t)] + \mathcal{L}u[t,x(t)]\}dt +$$

$$+ \int_{t_1}^{t_2} u_x[t,x(t)]'G(t)dw(t) \qquad (1)$$

<u>with probability 1</u>. <u>If</u>, <u>furthermore</u>,

$$\mathcal{E}\{|u[t,x(t)]|\} < \infty \qquad (t \in I)$$

<u>and</u>

$$\int_{t_o}^{T} \mathcal{E}\{|u_x[t,x(t)]'G(t)|\}dt < \infty \qquad (t \in I)$$

<u>then</u>

$$\mathcal{E}\{u[t_2,x(t_2)]\}-\mathcal{E}\{u[t_1,x(t_1)]\}$$

$$= \mathcal{E}\{\int_{t_1}^{t_2} (u_t[t,x(t)] + \mathcal{L}u[t,x(t)])dt\} . \qquad (2)$$

Equation (2) is the form of the integration rule we shall use later. To derive it we used (1) and property 1, §4.1, of the stochastic integral.

3. APPLICATION TO THE LINEAR REGULATOR.

We shall now extend the result of §2.6, by showing that the linear control law which is optimal (for the regulator problem) in the class of <u>linear</u> controls is actually optimal in a somewhat

wider class of controls. This wider class is just that for which the integration rule of §4.2 makes sense. Thus, let

$$dx = A(t)xdt - C(t)u(t)dt + B(t)dw(t), \quad t_o \leqq t \leqq T < \infty \qquad (1)$$

where A,B,C are bounded and Lebesgue measurable on $[t_o,T]$. In (1), $u(t)$ will denote a random k-vector $\psi(t,\omega)$. To allow for the fact that the control must be nonanticipative (i.e., the controller is not clairvoyant), we assume that $\psi(t,\cdot)$ is measurable relative to the σ-algebra $\mathfrak{F}_t$ of ω-sets generated by the family of random variables $\{w(t')-w(t'') : t_o \leqq t' < t'' \leqq t\}$. This σ-algebra will play the role of the $\mathfrak{F}_t$ of §4.2. We assume also that

$$\int_{t_o}^{T} \mathcal{E}\{|\psi(s,\omega)|^2\}ds < \infty \quad . \qquad (2)$$

Thus in our present setup, an admissible control law is any random process (in E_k) which satisfies a certain measurability condition and has a t-integrable second moment. We write $\{x^{\psi}(t):t_o \leqq t \leqq T\}$ for the corresponding process determined, in the usual way, by (1). Now let

$$V^o(t,x) = p(t) + x'P(t)x \qquad (3)$$

be the quadratic value function corresponding to the optimal linear feedback control of §2.6. We wish to express $V^o(t,x)$ as the

integral of its stochastic differential along a process of the type determined by (1). To this end write, for any k-vector u,

$$\mathcal{L}_u V^o(t,x) \equiv \frac{1}{2}\,\mathrm{tr}\{B(t)'V^o_{xx}(t,x)B(t)\} + \\ + [A(t)x - C(t)u]'V^o_x(t,x) \qquad (4)$$

and denote by $\mathcal{L}_\psi V^o$ the expression (4) with u replaced by $\psi(t,\cdot)$. Recalling 4.2(2) and the fact that $V^o(T,x) = 0$ we have, formally,

$$V^o(t,x) = -\mathcal{E}_{t,x}\{\int_t^T [V^o_t(s,x^\psi(s)) + \\ + \mathcal{L}_\psi V^o(s,x^\psi(s))]ds\} \ . \qquad (5)$$

In (5), $\mathcal{E}_{t,x}$ denotes expectation on the paths of the process determined by setting $x(t) = x \in E$ a.s. in (1). To justify (5) rigorously we must check the conditions of Theorem 1, §4.2. First, from (1) and the initial condition $x(t) = x$ a.s., we have

$$x^\psi(s) = \Phi(s,t)x - \int_t^s \Phi(s,\tau)C(\tau)\psi(\tau)d\tau \\ + \int_t^s \Phi(s,\tau)B(\tau)dw(\tau) \ ;$$

the second integral on the right side exists, because of (2). Then

$$\mathcal{E}\{|x^{\psi}(s)|^2\} \leqq 3|\Phi(s,t)x|^2 +$$

$$+ 3(s-t)\int_t^s |\Phi(s,\tau)|^2 |C(s)|^2 \mathcal{E}\{|\psi(\tau)|^2\}d\tau$$

$$+ 3\int_t^s |\Phi(s,\tau)B(\tau)|^2 d\tau$$

The second integral on the right side is less than

$$\text{const.} \int_t^s \mathcal{E}\{|\psi(\tau,\omega)|^2\}d\tau$$

and thus

$$\int_t^T \mathcal{E}\{|x^{\psi}(s)|^2\}ds \leqq \text{const.} + \text{const.} \int_t^T \mathcal{E}\{|\psi(\tau)|^2\}d\tau$$

$$< \infty ,$$

in view of (2). This shows that

$$\mathcal{E}_{t,x}\{V^o[s,x^{\psi}(s)]\} < \infty \quad t_o \leqq t \leqq s \leqq T,$$

and also

$$\int_t^T \mathcal{E}_{t,x}\{|V_x^o[s,x^{\psi}(s)]'B(s)|\}ds$$

$$\leqq \text{const.} + \text{const.} \int_t^T \mathcal{E}_{t,x}\{|x^{\psi}(s)|\}ds$$

$$< \infty .$$

Equation (5) is now established.

It remains to verify that the optimal linear control law $\varphi(t,x) = K^{o}(t)x$, of §2.6, is optimal in the class of all admissible controls $\psi(t)$. Obviously φ is admissible in the present sense; but then, using the optimality condition (6)§2.6, we have (in the notation of §2.6)

$$\begin{aligned} V^{o}(t,x) &= -\mathcal{E}_{t,x}\{\int_{t}^{T}(V_{t}^{o}[s,x^{\psi}(s)] + \\ &\quad + \mathcal{L}_{\psi}V^{o}[s,x^{\psi}(s)]ds\} \\ &\leqq -\mathcal{E}_{t,x}\{\int_{t}^{T}(V_{t}^{o}[s,x^{\psi}(s)] + \\ &\quad + \mathcal{L}_{\varphi}V^{o}[s,x^{\psi}(s)] + \\ &\quad + L[s,x^{\psi}(s),\ \varphi(s,x^{\psi}(s))] - \\ &\quad - L[s,x^{\psi}(s),\ \psi(s)])ds\} \\ &= \mathcal{E}_{t,x}\{\int_{t}^{T}L[s,x^{\psi}(s),\ \psi(s)ds\}\ . \end{aligned}$$

The last written quantity is precisely the cost incurred by using the control law ψ, and the argument is complete.

4. APPLICATION TO THE LINEAR REGULATOR WITH NOISY DATA.

4.1 In this section we combine the results of Chapter 3 on optimal linear filtering with those of Chapter 2 on optimal linear

control. The general problem is the following. When observations of the state of a dynamical system to be controlled are perturbed by noise, the Markovian framework of dynamic programming is in general no longer available, and resort must be had to more-or-less ad hoc methods. In a few special cases it is possible to introduce a (finite-dimensional) statistic, or functional on the noisy data, with the properties that (i) the value of this stochastic at any time t represents all information, carried by the data up to time t, that is relevant to the purpose of contol, and (ii) when the control law is made to depend on this statistic alone, then the latter becomes a Markov process. When these features are present it is possible to treat the problem by dynamic programming, where the state vector is now the vector of components of the sufficient statistic.

To date the only problems that have been completely solved by this technique are those involving control of linear systems perturbed by Gaussian white noise, when the observation (feedback) channel is also linear and perturbed by Gaussian white-noise. In this case a sufficient statistic turns out to be the expected value of the system state, and this is conveniently determined by a linear filter, as described in Chapter 3. We shall now discuss the simplest problem of the type just mentioned.

4.2 As in §4.3, consider the linear system described by

$$dx(t) = A(t)x(t)dt - F(t)u(t)dt + B(t)dw(t),$$
$$x(t_o) = x_o \qquad t_o \leqq t \leqq T. \tag{1}$$

We assume t_o, T fixed and that all parameter matrices here and below are measurable and bounded on $[t_o, T]$.*

Now suppose that observations $y(t)$ are made on the system (1), such that

$$dy(t) = C(t)x(t)dt + D(t)dw(t),$$
$$y(t_o) = 0 \qquad t_o \leqq t \leqq T. \tag{2}$$

In physical terms the problem is to control (1) on the basis of data (2), in such a way that a quadratic performance index is minimized:

$$\mathcal{E}\{\int_t^T L[s, x(s), u(s)]ds \,|\, y(\tau),\ t_o \leqq \tau \leqq t\} = \min. \tag{3}$$

More precisely, let $\mathcal{Y}_t$ be the σ-algebra of ω-sets generated by the random variables $y(s)$, $t_o \leqq s \leqq t$; and let $\psi(t,\omega)$ be a measurable E_k-valued random process such that, for each $t \in [t_o, T]$, $\psi(t,\cdot)$ is $\mathcal{Y}_t$-measurable, and

$$\int_{t_o}^T \mathcal{E}\{|\psi(t,\omega)|^2\}dt < \infty. \tag{4}$$

---

* As before, the parameter matrices are nonstochastic.

Now consider the system

$$dx = Axdt - F\psi dt + Bdw$$
$$dy = Cxdt + Ddw \qquad t_o \leq t \leq T. \quad (5)$$

We assume as usual that x_o is a Gaussian random variable independent of the $w(t)$ process, and define $\mathfrak{F}_t$ to be the smallest σ-algebra relative to which x_o and the random variables $w(s)$, $t_o \leq s \leq t$, are measurable. To clarify the meaning of (5), let Φ be the fundamental matrix associated with A: then (5) is interpreted to mean

$$x(t) = \Phi(t,t_o)x_o + \int_{t_o}^{t} \Phi(t,s)\{-F(s)\psi(s)ds + B(s)dw(s)\} \quad (6a)$$

$$y(t) = \int_{t_o}^{t} \{C(s)x(s)ds + D(s)dw(s)\} \quad (6b)$$

The integrals in (6) are well-defined. But to apply the integration rule to functions of the $x(t)$-process, we must show that $x(t)$ is $\mathfrak{F}_t$-measurable. This fact is obvious if $\psi(\cdot)$ is a $\mathcal{Y}_t$-measurable step function: one need only evaluate the integrals in (6a), (6b) alternately in successive intervals $t_i \leq t < t_{i+1}$, $i = 0,1,\ldots$. Using the standard approximation techniques of §4.1 we can express $x(t), y(t)$ as limits in probability of $\mathfrak{F}_t$-measurable random variables, and conclude that

$x(t)$, $y(t)$ are also $\mathfrak{Y}_t$-measurable. In this discussion we have tacitly assumed that $x(t)$, $y(t)$ processes which satisfy (6) exist. It will be seen shortly that this assumption is nonvacuous; but to solve (6) for unique $\{x(t),y(t)\}$ requires further restrictions on ψ and appeal to the theory of functional stochastic differential equations [Itô and Nisio, 1964]. This matter will be taken up in a later chapter.

4.3 Next we turn to a generalization of the linear filtering equations of §3.2. Given that $x(t)$ and $y(t)$ have the stochastic differentials (5), we seek the stochastic differential of the estimate

$$\hat{x}(t) = \mathcal{E}\{x(t)|\mathcal{Y}_t\} \qquad t \in [t_o,T]. \tag{7}$$

This problem is easily reduced to that of §3.2. Let

$$\zeta(t) = \int_{t_o}^{t} \Phi(t,s)F(s)\psi(s)ds \tag{8}$$

$$\xi(t) = x(t) + \zeta(t) \tag{9}$$

$$\begin{aligned} d\eta(t) &= dy(t) + C(t)\zeta(t)dt \\ \eta(t_o) &= 0 \ . \end{aligned} \tag{10}$$

Then

$$\begin{aligned} d\xi &= A\xi dt + Bdw \\ \xi(t_o) &= x_o \end{aligned} \tag{11}$$

and

$$d\eta = C\xi dt + Ddw \ . \tag{12}$$

Since $\zeta(t)$ is $\mathcal{Y}_t$-measurable,

$$\hat{\xi}(t) = \mathcal{E}\{\xi(t)|\mathcal{Y}_t\} = \hat{x}(t) + \zeta(t) \ ; \tag{13}$$

and if $\mathcal{N}_t$ is the σ-algebra generated by the $\eta(s)$, $t_o \leq s \leq t$, then [Problem 1, p.] $y(t)$ is $\mathcal{N}_t$-measurable, and therefore

$$\mathcal{E}\{\xi(t)|\mathcal{Y}_t\} = \mathcal{E}\{\xi(t)|\mathcal{N}_t\} \ . \tag{14}$$

Now (11), (12) have the same form as (1), (2) of §3.2, and so we conclude that

$$d\hat{\xi} = A\hat{\xi}dt + (PC'+BD')(DD')^{-1}(d\eta - C\hat{\xi}dt) \tag{15}$$

where

$$P(t) = \mathrm{cov}\{\xi(t)|\mathcal{N}_t\} = \mathrm{cov}\{x(t)|\mathcal{Y}_t\}$$

and P is given by

$$\begin{aligned} \frac{dP}{dt} &= AP + PA' + BB' - (PC'+BD')(DD')^{-1}(PC'+BD') \quad \text{a.e.} \\ P(t_o) &= \mathrm{cov}\{\xi_o\} = \mathrm{cov}\{x_o\} \ . \end{aligned} \tag{16}$$

Thus on substituting (13) in (15) and using (8), we get

$$d\hat{x} = A\hat{x}dt - F\psi dt + \\ + (PC'+BD')(DD')^{-1}(dy-C\hat{x}dt) \tag{17}$$

Equation (17) shows that the optimal filter for estimating $\hat{x}$ differs from that of §3.2 only by the addition of the $\mathcal{Y}_t$-measurable forcing term $-F\psi dt$ at the filter input. In this way the behavior of the filter is made to imitate that of the plant.

4.4 In §3.3.1 we observed that the differential $dy-C\hat{x}dt$ could be identified with the differential of a certain Wiener process. The argument used there applies equally well here; that is, there exists a Wiener process $\{\hat{w}(t): t_o \leqq t \leqq T\}$ such that $\hat{w}(t)$ is $\mathfrak{V}_t$-measurable and

$$\begin{aligned} d\hat{x} &= A\hat{x}dt-F\psi dt + (PC'+BD')(DD')^{-1}Dd\hat{w} \\ \hat{x}(0) &= \{x_o\} \end{aligned} \tag{18}$$

Equation (18) implies immediately that $\hat{x}(t)$ is a Markov process, provided the random variable $\psi(t)$ is of the special form

$$\psi(t) = K(t)\hat{x}(t) . \tag{19}$$

Let us consider in more detail the relationship of the

$\hat{x}(t)$ process to the $x(t)$ process, when ψ is of the form (19). From (18),(19) it follows that $\{\hat{x}(t)\}$ is a Gaussian diffusion process. If $\tau > t$, $\hat{x}(\tau)$ can be written

$$\hat{x}(\tau) = \hat{\Phi}(\tau,t)\hat{x}(t) + \int_t^\tau \hat{\Phi}(\tau,s)\hat{B}(s)d\hat{w}(s)$$

for suitable $\hat{\Phi},\hat{B}$. Since for each t, $\hat{w}(t)$ is $\mathfrak{F}_t$-measurable, the $\hat{w}(s)$ increments are independent of $\mathcal{Y}_t$ for $s > t$; hence the conditional distribution of $\hat{x}(\tau)$ given $\mathcal{Y}_t$ depends only on the $\mathcal{Y}_t$-measurable statistic $\hat{x}(t)$. Next, let $f(x,\hat{x})$ be a measurable nonnegative function on $E \times E$. Then, if $\tau > t$,

$$\begin{aligned}\mathcal{E}\{f(x_\tau,\hat{x}_\tau)|\mathcal{Y}_t\} &= \mathcal{E}\{\mathcal{E}\{f(x_\tau,\hat{x}_\tau)|\mathcal{Y}_\tau\}|\mathcal{Y}_t\} \\ &= \mathcal{E}\{\mathcal{E}\{f(x_\tau,\hat{x}_\tau)|\hat{x}_\tau\}|\mathcal{Y}_t\} \\ &= \mathcal{E}\{\hat{f}(\hat{x}_\tau)|\hat{x}_t\} \qquad (20)\end{aligned}$$

where

$$\hat{f}(\hat{x}_\tau) = \mathcal{E}\{f(x_\tau,\hat{x}_\tau)|\hat{x}_\tau\} .$$

Equation (20) will be useful later.

4.5 We can now discuss the optimal regulator problem. Let Ψ denote the class of random processes $\{\psi(t,\omega)\}$ and let $\hat{\Psi}$ denote the class of linear controls

$$u(t) = \hat{\psi}[t,\hat{x}(t)] = K(t)\hat{x}(t) . \qquad (21)$$

We will first show that the control optimal in the class $\hat{\Psi}$ is precisely (21) with $K = K^o$, where K^o is the matrix determined in §2.6 for the regulator problem in the absence of observational noise. The proof is easy: with

$$L(t,x,u) = x'M(t)x + u'N(t)u \tag{22}$$

as before, write $x(s) = x_s$ etc. and put

$$V(t,y_t) = \mathcal{E}\{\int_t^T L[s,x_s,\hat{\psi}(s,\hat{x}_s)]ds \,|\, y_t\} \tag{23}$$

In (23), $\{x_s\}$ is the process determined by putting $\psi = \hat{\psi}$ in (5). Next we observe that V can be expressed as a function of $t,\hat{x}_t$. In fact

$$V(t,y_t) = \int_t^T \{L[s,x_s,\hat{\psi}(s,\hat{x}_s)] \,|\, y_t\}ds \quad ;$$

and, recalling (20), we see that $V(t,y_t)$ is a random variable measurable on the sample space of $\hat{x}_t$. From this it follows [see Loève (1963)p.343] that V can be written as $V(t,\hat{x}_t)$; in the present case we apply (20) to find explicitly

$$V(t,\xi) \equiv \mathcal{E}\{\int_t^T L[s,x_s,\hat{\psi}(s,\hat{x}_s)]ds \,|\, \hat{x}_t = \xi\} \tag{24}$$

$$= \mathcal{E}\{\int_t^T \mathcal{E}\{L[s,x_s,\hat{\psi}(s,\hat{x}_s)]\,|\,\hat{x}_s,\hat{x}_t\}ds\,|\,\hat{x}_t = \xi\}$$

$$= \mathcal{E}\{\int_t^T \mathcal{E}\{L[s,x_s,\hat{\psi}_s(s,\hat{x}_s)]\,|\,\hat{x}_s\}ds\,|\,\hat{x}_t = \xi\}\ ;$$

or

$$V(t,\xi) = \mathcal{E}\{\int_t^T \hat{L}[s,\hat{x}_s,\hat{\psi}(s,\hat{x}_s)]ds\,|\,\hat{x}_t = \xi\} \tag{25}$$

where

$$\hat{L}[s,\xi,\hat{\psi}(s,\xi)] = \mathcal{E}\{L[s,x_s,\hat{\psi}(s,\hat{x}_s)]\,|\,\hat{x}_s = \xi\}\ . \tag{26}$$

In particular where $\hat{\psi}$ is given by (19) and L by (22)

$$\begin{aligned}\hat{L}[s,\xi,\hat{\psi}(s,\xi)] &= \mathcal{E}\{x_s'M(s)x_s + \\ &\quad + [K(s)\hat{x}_s]'N(s)K(s)\hat{x}_s\,|\,\hat{x}_s = \xi\} \\ &= \mathrm{tr}\{M(s)P(s)\} + \\ &\quad + \xi'[M(s)+K(s)'N(s)K(s)]\xi\ .\end{aligned} \tag{27}$$

It is now easy to see that the problem of choosing K such that $V(t,\xi)$ is minimal (uniformly for $t_o \leqq t \leqq T$ and $\xi \in E$) is identical with the problem of §2.6, apart from a trivial modification of L. In the present case the basic process is $\{\hat{x}(t) : t_o \leqq t \leqq T\}$ with stochastic differential (18); the control is of type (19); the cost rate function is given by (27) and the value function by (25). Thus the optimal control matrix

K exists and is independent of the coefficient of $d\hat{w}$ in (18). Comparing (18) with (1), we see that only the noise intensity coefficients differ: hence the optimal control 'gains' (elements of K) are the same whether noise is present in the feedback channel or not. When noise is present, the control law is the same function of $\hat{x}$ as it is of x when noise is absent.

The fact that the control law is a function only of $\hat{x}_t$ (and of no additional $\mathcal{Y}_t$-measurable statistics) is a consequence of the linearity of (1) and (2) in x, of the assumption that disturbances are Gaussian, and the assumption that T, the terminal time, is fixed in advance. The fact that the control law is independent of noise intensity results from the restriction that L be quadratic in x and u, and the fact that u is not <u>a priori</u> bounded.

Let us note that

$$V(t,\xi) = \pi(t) + \xi'\Pi(t)\xi \tag{28}$$

where the matrix Π is independent of the noise coefficients B,D ; and π is the sum of two terms, arising respectively from the estimation error $\mathrm{tr}(MP)$ in (27) and from the dynamic disturbance term in the right side of (18). Thus [cf.(11)§2.6]

$$\pi(t) = \int_t^T [\mathrm{tr}\{M(s)P(s)\} + \mathrm{tr}\{Q(s)'\Pi(s)Q(s)\}]ds \tag{29}$$

where

$$Q = [PC' + BD')(DD')^{-1}D \ . \tag{30}$$

PROBLEM 1 For a first order system with constant coefficients, compute $\pi(t)$ explicitly, and find the degradation in performance due to channel noise. [Let x,y be scalars and take B,D to be 1×2 matrices].

PROBLEM 2 In the case where all parameter matrices are constant discuss the asymptotic behavior of the optimal control as $t \to \infty$.

4.6 Finally, we observe that the control $K(t)\hat{x}(t)$ just found is optimal in the class Ψ of controls $\psi(t,\omega)$ defined in §4.4.2. In fact, the quadratic function $V(t,\xi)$ can be expressed as the integral of its stochastic differential along a process $\{\hat{x}^{\psi}(t) : t_o \leqq t \leqq T\}$ of the type described by the system (18). The appropriate integration rule is that of §4.2, with the process $\{x(t)\}$ of that section replaced by $\{\hat{x}(t)\}$. The proof of optimality then closely follows that of §4.3.

PROBLEM 3 Give details of the proof that $K\hat{x}$ is optimal in the class Ψ.

5. ITÔ'S STOCHASTIC DIFFERENTIAL EQUATION.*

So far, we have considered only linear stochastic differential equations. In technology one encounters nonlinear systems perturbed by random disturbances and it is of interest to develop a corresponding theory of nonlinear stochastic differential equations. In this section we present Itô's (1950, 1951) theory for the equation

$$dx = f(t,x)dt + G(t,x)dw$$

(x, f are n-vectors and F an $n \times m$ matrix). The discussion is restricted to existence and uniqueness of the solution and to a derivation of the Markov property. Many further interesting qualitative properties (for instance, of stability type) have been studied, however, and some of these will be reviewed in later chapters.

The exposition below closely follows Gikhman and Skorokhod (1965); at time of writing (January, 1967) the latter work is not available in English.

---

* This section closely follows Gikhman and Skorokhod (1965), Chapter 8, §3.

5.1 Existence and uniqueness of solutions.

Consider the stochastic differential equation

$$dx(t) = a(t,x(t))dt + b(t,x(t))dw(t) \qquad (1)$$

Assume a,b are Borel-measurable in (t,x) for $t_o \leqq t \leqq T$ and $x \in E = R_1$. Equation (1) is equivalent to the integral equation

$$x(t) = x(t_o) + \int_{t_o}^{t} a(s,x(s))ds + \int_{t_o}^{t} b(s,x(s))dw(s) \qquad (2)$$

where we assume $x(t_o)$ is a random variable independent of the $w(t)$-process. In the following, let $\mathfrak{F}_t$ denote the σ-algebra of ω-sets generated by $x(t_o)$ and the increments $w(s)-w(t_o)$ for $t_o \leqq s \leqq t$. A process $\{x(t): t_o \leqq t \leqq T\}$ is a <u>solution</u> of (2) if $x(t)$ is $\mathfrak{F}_t$-measurable and if the integrals in (2) exist and (2) holds for each $t \in [t_o,T]$ with probability 1.

From the inequality (6)§4.1 it follows that if processes $f_1(s)$, $f_2(s)$ are stochastically equivalent then with probability 1,

$$\int_{t_o}^{t} f_1(s)dw(s) = \int_{t_o}^{t} f_2(s)dw(s) \ .$$

Thus if $\{x(t)\}$ satisfies (2) and $\{\tilde{x}(t)\}$ is equivalent to

$\{x(t)\}$, then $\{\tilde{x}(t)\}$ satisfies (2). Also, since the right side of (2) is equivalent to the left, and is a.s. continuous, to each solution of (2) there is associated an equivalent continuous solution. In the following we consider only continuous solutions of (2).

THEOREM 1 *Let* $a(t,x)$, $b(t,x)$ *be Borel-measurable for* $t_o \leqq t \leqq T$, $x \in R_1$, *and suppose there is a constant* K *such that:*

(a) [*uniform Lipschitz condition*]

$$|a(t,x)-a(t,y)| + |b(t,x)-b(t,y)| < K|x-y|$$

(b) [*growth restriction*]

$$|a(t,x)|^2 + |b(t,x)|^2 \leqq K^2(1 + |x|^2) \quad .$$

Then a solution of (2) *exists and, if* $x_1(t)$ *and* $x_2(t)$ *are two continuous solutions (for fixed* $x_o(t)$*) then*

$$P\{\sup_{t_o \leqq t \leqq T} |x_1(t)-x_2(t)| = 0\} = 1 \ .$$

PROOF. First we prove uniqueness. Let $x_1(t)$, $x_2(t)$ be two continuous solutions of (2). Define

$$\chi_N(t) = 1 \quad \text{if} \quad |x_1(s)| < N,\ |x_2(s)| < N \quad \text{for} \quad t_o \leqq s \leqq t$$

$$= 0, \quad \text{otherwise.}$$

Since $\chi_N(t)\chi_N(s) = \chi_N(t)$ for $s < t$, we have

$$\chi_N(t)[x_1(t)-x_2(t)] = \chi_N(t)[\int_{t_o}^{t} \chi_N(s)[a(s,x_1(s))-$$

$$-a(s,x_2(s))]ds + \int_{t_o}^{t} \chi_N(s)[b(s,x_1(s))-b(s,x_2(s))]dw(s)].$$

Since

$$\chi_N(s)[|a(s,x_1(s))-a(s,x_2(s))|+|b(s,x_1(s))-b(s,x_2(s))|]$$

$$\leqq K\,\chi_N(s)|x_1(s)-x_2(s)| \leqq 2KN$$

we derive

$$\mathcal{E}\{\chi_N(t)[x_1(t)-x_2(t)]^2\}$$

$$\leqq 2\mathcal{E}\{\chi_N(t)(\int_{t_o}^{t} \chi_N(s)[a(s,x_1(s))-a(s,x_2(s))]ds)^2\}$$

$$+ 2\mathcal{E}\{\chi_N(t)(\int_{t_o}^{t} \chi_N(s)[b(s,x_1(s))-b(s,x_2(s))]dw(s))^2\}$$

$$\leqq 2(T-t_o)\int_{t_o}^{t} \mathcal{E}\{\chi_N(s)[a(s,x_1(s))-a(s,x_2(s))]^2\}ds$$

$$+ 2\int_{t_o}^{t} \mathcal{E}\{\chi_N(s)[b(s,x_1(s))-b(s,x_2(s)]^2\}ds$$

$$\leq L \int_{t_o}^{t} \mathcal{E}\{\chi_N(s)|x_1(s)-x_2(s)|^2 ds \qquad (3)$$

for some constant L. From Gronwall's inequality we conclude

$$\mathcal{E}\{\chi_N(t)[x_1(t)-x_2(t)]^2\} = 0 \quad ,$$

that is,

$$P\{x_1(t) \neq x_2(t)\} \leq P\{\sup_t |x_1(t)| > N\} + P\{\sup_t |x_2(t)| > N\}$$

[If $\{\chi_N(t) = 0\} = A$ and $\{x_1(t) = x_2(t)\} = B$ then $P(A \cup B) = 1$ $\Longrightarrow P(A' \cup B') = 0 \Longrightarrow P(A') \leq 1-P(B') = P(B)$]

$$\to 0 \quad (N \to \infty)$$

in view of the continuity of $x_1(t)$, $x_2(t)$. Thus $P\{\sup_t |x_1(t)-x_2(t)| > 0\} = 0$.

Next we show existence of a solution of (2). Assume first $\mathcal{E}\{|x(t_o)|^2\} < \infty$. Let B denote the Banach space of functions $z(t,\omega)$ such that z is measurable in (t,ω), $z(t)$ is $\mathfrak{F}_t$-measurable, and $\sup\{\mathcal{E}|z(t)|^2 : t_o \leq t \leq T\} < \infty$; write $\|z\| = [\sup \mathcal{E}\{|z(t)|^2 : t_o \leq t \leq T\}]^{1/2}$. [To verify that B is complete, note that if $\{z_n\}$ is fundamental then $\{z_n(t)\}$ converges in mean square for each t. This defines a limit function z such that, as a limit in measure, z is measurable in (t,ω) and $z(t)$ is $\mathfrak{F}_t$-measurable for each t. Clearly

$\|z\| < \infty$, and so $z \in B$]. In B define the operator S by

$$Sz(t) = x(t_o) + \int_{t_o}^{t} a(s,z(s))ds + \int_{t_o}^{t} b(s,z(s))dw(s) .$$

Here the integrals exist by condition (b), and evidently $z(t)$ is $\mathfrak{F}_t$-measurable. Finally

$$\begin{aligned}
\mathcal{E}\{|Sz(t)|^2\} &< 3\mathcal{E}\{|x(t_o)|^2\} + \\
&+ 3(T-t_o)\mathcal{E}\{\int_{t_o}^{t} K^2(1+|z(s)|^2)ds\} + \\
&+ 3\int_{t_o}^{t} \mathcal{E}\{K^2(1+|z(s)|^2)ds\} \\
&\leq 3\mathcal{E}\{|x(t_o)|^2\} + L_1 + L_2\|z\|^2
\end{aligned}$$

for suitable constants L_1, L_2 and hence $Sz \in B$. Furthermore, we easily find

$$\begin{aligned}
\mathcal{E}\{|Sz_1(t)-Sz_2(t)|^2\} &\leq \\
&\leq L\int_{t_o}^{t}\mathcal{E}\{|z_1(s)-z_2(s)|^2\}ds \\
&\leq L(t-t_o)\|z_1-z_2\|^2 \qquad (L = \text{const.})
\end{aligned}$$

and iterating this inequality,

$$\mathcal{E}\{|S^n z_1(t)-S^n z_2(t)|^2 \leq L\int_{t_o}^{t}\mathcal{E}\{S^{n-1}z_1(u)-S^{n-1}z_2(u)|^2\}du$$

$$\leqq L^n \int_{t_o}^{t} \cdots \int_{t_o}^{t} \mathcal{E}\{\|z_1-z_2\|^2\} du dt_n \cdots dt_1$$

$$\leqq \frac{L^n (t-t_o)^n}{n!} \|z_1-z_2\|^2 \quad .$$

Hence if $z \in B$,

$$\|S^{n+1}z-S^n z\| \leqq \frac{L^n (T-t_o)^n}{n!} \|Sz-z\|^2 \quad .$$

From convergence of

$$\sum_{n=1}^{\infty} \|S^{n+1}z-S^n z\|$$

and the fact that B is complete, it follows that $\lim_{n \to \infty} S^n z(t)$ exists in B. Denote the limit by $x(t)$. Since S is a continuous operator,

$$S(S^n z) \to Sx$$

Since $S(S^n z) = S^{n+1}z \to x$ we conclude that $\|Sx-x\| = 0$: i.e., $x(t) = Sx(t)$ with probability 1 for each t, so that $\{x(t)\}$ is a solution of (2).

In the general case, let $x^N(t_o) = x(t_o)$ for $|x(t_o)| \leqq N$ and $x^N(t_o) = 0$ otherwise. Denote by $\{x^N(t)\}$ the solution of (2) for initial value $x^N(t_o)$. We will show that P lim $x^N(t) = x(t)$ $(N \to \infty)$ when $x(t)$ is a solution of (2). Let $N' > N$ and

let $\eta = 1$ if $|x(t_o)| \leq N$, $\eta = 0$ otherwise. Then $|x^N(t_o) - x^{N'}(t_o)|\eta = 0$. Clearly η is $\mathfrak{F}_{t_o}$-measurable. If follows now from

$$|x^N(t)-x^{N'}(t)|^2\eta \leq 2[\int_o^t [a(s,x^N(s))-a(s,x^{N'}(s))]\eta ds]^2$$

$$+ 2[\int_{t_o}^t [b(s,x^N(s))-b(s,x^{N'}(s))]\eta dw(s)]^2,$$

and from condition (a) and previous estimates of the integrals, that for some constant L,

$$\mathcal{E}\{|x^N(t)-x^{N'}(t)|^2\eta\} \leq L \int_{t_o}^t \mathcal{E}\{|x^N(s)-x^{N'}(s)|^2\eta ds\} .$$

Consequently

$$\mathcal{E}\{|x^N(t)-x^{N'}(t)|^2\eta\} = 0$$

and so

$$P\{|x^N(t)-x^{N'}(t)| > 0\} \leq P\{|x(t_o)| > N\}$$

which implies the existence of P lin $x^N(t) = x(t)$. Since the limit is uniform in t,

$$P \lim_{n \to \infty} \int_{t_o}^T (x^N(t)-x(t))^2 dt = 0 .$$

Now by property (a),

$$\left|\int_{t_o}^{t} a(s,x^N(s))ds - \int_{t_o}^{t} a(s,x(s))\right|$$

$$\leq K[(t-t_o)\int_{t_o}^{t} (x^N(s)-x(s))^2 ds]^{1/2}$$

and

$$P\{\left|\int_{t_o}^{t} b(s,x^N(s))dw(s) - \int_{t_o}^{t} b(s,x(s))dw(s)\right| > \epsilon\}$$

$$\leq \frac{\delta^2}{\epsilon^2} + P\{K^2 \int_{t_o}^{t} (x^N(s)-x(s))^2 ds > \delta^2\} .$$

Hence, we may pass to the limit-in-probability in the right side of the equation

$$x^N(t) = x^N(t_o) + \int_{t_o}^{t} a(s,x^N(s))ds + \int_{t_o}^{t} b(s,x^N(s))dw(s)$$

and the result follows.

5.2 Markov diffusion processes.

We now show that the process determined by (1) is Markov. Before doing so, we recall the definition, that a process $X = \{x(t) : t_o \leq t \leq T\}$ with state space E is a Markov process if

(a) for all $t_1 < t_2 < \ldots < t_n < t$ $(t, t_k \in [t_o, T],\ k = 1,\ldots,n)$ and Borel sets $A \subset E$ we have

$$P\{x(t) \in A \,|\, x(t_1),\ldots,x(t_n)\} = P\{x(t) \in A \,|\, x(t_n)\} \quad \text{a.s.}$$

(b) the transition function

$$P(s,x,t,A) = P\{x(t) \in A \mid x(s) = x\}$$

defined for $t_o \leqq s \leqq t \leqq T$, $x \in E$, $A \in \mathcal{B}(E)$, is a $\mathcal{B}$-measurable function of x for each fixed s,t,A and is a probability measure on $\mathcal{B}$ for fixed s,t,x .

From the basic Markov property (a) one can derive the apparently stronger property that the 'future' and 'past' of $x(\cdot)$ are independent when conditioned on the 'present'. More precisely, if X_t is the σ-algebra of ω-sets generated by $x(t)$, and $\mathfrak{X}_t^+ = \bigvee_{s>t} X_s$, $\mathfrak{X}_t^- = \bigvee_{s<t} X_s$ then for any $A \in \mathfrak{X}_t^+$, $B \in \mathfrak{X}_t^-$,

$$P(A \mid \mathfrak{X}_t^- \vee X_t) = P(A \mid X_t)$$

$$P(AB \mid X_t) = P(A \mid X_t) P(B \mid X_t) \ .$$

A complete discussion of this "Markov dependence" is given by Loève (1963), Ch.12.

Returning to our stochastic differential equation, we have the following result.

THEOREM 2 Let $a(t,x)$, $b(t,x)$ satisfy the conditions of Theorem 1; let $\{x_{t,\xi}(s) : t \leqq s \leqq T\}$, $t \geqq t_o$, $\xi \in E$, be the solution of the equation

$$x_{t,\xi}(s) = \xi + \int_t^s a(u, x_{t,\xi}(u))du + \int_t^s b(u, x_{t,\xi}(u))dw(u) \ . \qquad (4)$$

Then the process $\{x(t): t_o \leq t \leq T\}$ determined by (2) is a Markov process, the transition function of which is defined by the relation

$$P(t,\xi,s,A) = P\{x(s) \in A \mid x(t) = \xi\}$$
$$= P\{x_{t,\xi}(s) \in A\} \ .$$

PROOF. Since $x(t)$ is $\mathfrak{F}_t$-measurable, and $x_{t,\xi}(s)$ is completely determined by the $w(s)-w(t)$ increments for $s \in [t,T]$ (independent of $\mathfrak{F}_t$), we have that $x_{t,\xi}(s)$ is independent of $x(t)$ and of events in $\mathfrak{F}_t$. For $s \in [t,T]$, $x(s)$ is the unique solution (in the sense of Theorem 1) of the equation

$$x(s) = x(t) + \int_t^s a(u,x(u))du + \int_t^s b(u,x(u))dw(u)$$

[This follows from (2) using additivity of the integrals]. Since the process $\{x_{t,x(t)}(s): t \leq s \leq T\}$ is also a solution it follows that

$$x(s) = x_{t,x(t)}(s) \quad , \quad t \leq s \leq T. \qquad (5)$$

We now show that

$$P\{x(s) \in A \mid x(t)\} = P\{x(s) \in A \mid \mathfrak{F}_t\} \ .$$

For this it is sufficient to show that for any $\mathfrak{F}_t$-measurable bounded variable ζ and any bounded continuous function $\lambda(x)$,

$$\mathcal{E}\{\zeta\lambda(x(s))\} = \mathcal{E}\{\zeta\mathcal{E}\{\lambda(x(s))|\, x(t)\}\}. \tag{7}$$

[In particular, let ζ be the indicator of an arbitrary event of $\mathfrak{F}_t$; if α is the indicator of a closed ball A, there is a bounded sequence of continuous functions λ_n such that $\lambda_n(x) \to \alpha(x)$ for all x. Thus (7) implies that (6) is true for every ball, hence every Borel set in E]. Write $\varphi(\xi,\omega) = \lambda[x_{t,\xi}(s,\omega)]$. Clearly φ is measurable in (ξ,ω) and is bounded and continuous in ξ for P-almost all ω. Hence φ can be written as the pointwise a.s. limit of bounded functions φ of the special form

$$\varphi(\xi,\omega) = \sum_{k=1}^{n} \varphi_k(\xi)\psi_k(\omega) \tag{8}$$

where, for instance, the φ_k 'are step functions. Then, since $\psi_k(\omega)$ is independent of $\mathfrak{F}_t$, we have

$$\mathcal{E}\{\zeta \sum_{k=1}^{n} \varphi_k(x(t))\psi_k(\omega)\} = \sum_{k=1}^{n} \mathcal{E}\{\zeta\varphi_k(x(t))\}\mathcal{E}\{\psi_k(\omega)\}$$

$$= \mathcal{E}\{\sum_{k=1}^{n} \zeta\varphi_k(x(t))\mathcal{E}\{\psi_k(\omega)\}$$

and

$$\mathcal{E}\{\sum_{k=1}^{n} \varphi_k(x(t))\psi_k(\omega)|\, x(t)\} = \sum_{k=1}^{n} \varphi_k(x(t))\mathcal{E}\{\psi_k(\omega)\}\ .$$

This establishes (7) when φ has the form (8). In that case, we have also shown, recalling (5), that

$$\mathcal{E}\{\lambda(x(s))|\ \mathfrak{F}_t\} = g\{x(t)\} \tag{9}$$

where $g(\xi) = \{\lambda(x_{t,\xi}(s)\}$. Finally, by passage to the limit and appeal to dominated convergence, (7) and (9) are established for all functions φ of the type considered. Hence

$$P\{x(s) \in A|\ \mathfrak{F}_t\} = P_{t,x(t)}(s,A)$$

where

$$P_{t,\xi}(s,A) = P(x_{t,\xi}(s) \in A)\ .$$

Q.E.D.

5.3 Diffusion in $E = R_n$.

The extension of the foregoing results to the n-dimensional case

$$dx = f(t,x)dt + G(t,x)dw$$

[x,f: n-vectors; G : n m matrix; w: Wiener process in E_m] is straightforward and can be left as an exercise. The previous work is applicable here, with the symbol $|\cdot|$ now interpreted as Euclidean norm.

PROBLEM.

Show that, for the linear Itô equation, the foregoing construction yields the process which was adopted, by definition, for the solution in Chapter II.

REFERENCES

[1] Doob, J.L., Stochastic Processes, New York, Wiley, 1953.

[2] Gikhman, I.I., and Skorokhod, A.V., Introduction to the Theory of Random Processes, (Russ.), Izd.-vo "Nauka", Moscow, 1965.

[3] Itô, K., Stochastic differential equations in a differentiable manifold, Nagoya Math. J., 1, 1950, pp. 35-47.

[4] Itô, K., On Stochastic Differential Equations, Mem. Amer. Math. Soc. 4, 1951.

[5] Itô, K., and Nisio, M., On stationary solutions of a stochastic differential equation, J. Math., Kyoto Univ. 4-1 (1964), pp. 1-75.

[6] Loève, M., Probability Theory, (Third edition), Van Nostrand, Princeton, N.J., 1963.

[7] Skorokhod, A.V., Studies in the Theory of Random Processes, New York, Addison-Wesley, 1965.