Stochastic Differential Equations and Applications

Volume 1

Avner Friedman

Department of Mathematics
Northwestern University
Evanston, Illinois

ACADEMIC PRESS New York San Francisco London 1975

A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1

Library of Congress Cataloging in Publication Data

Friedman, Avner.
Stochastic differential equations and applications.

(Probability and mathematical statistics series)
Bibliography: p.
Includes index.
1. Stochastic differential equations. I. Title.
QA274.23.F74 519.2 74-30808
ISBN 0-12-268201-7

AMS(MOS) 1970 Subject Classifications: 60H05, 60H10, 35J25, 35K15, 93E05, 93E15, 93E20.

PRINTED IN THE UNITED STATES OF AMERICA

Contents

4 The Stochastic Integral

5 Stochastic Differential Equations

6 Elliptic and Parabolic Partial Differential Equations and Their Relations to Stochastic Differential Equations

7 The Cameron–Martin–Girsanov Theorem

8 Asymptotic Estimates for Solutions

9 Recurrent and Transient Solutions

Preface

The object of this book is to develop the theory of systems of stochastic differential equations and then give applications in probability, partial differential equations and stochastic control problems.

In Volume 1 we develop the basic theory of stochastic differential equations and give a few selected topics. Volume 2 will be devoted entirely to applications.

Chapters 1–5 form the basic theory of stochastic differential equations. The material can be found in one form or another also in other texts. Chapter 6 gives connections between solutions of partial differential equations and stochastic differential equations. The material in partial differential equations is essentially self-contained; for some of the proofs the reader is referred to an appropriate text.

In Chapter 7 Girsanov's formula is established. This formula is becoming increasingly useful in the theory of stochastic control.

In Chapters 8 and 9 we study the behavior of sample paths of the solution of a stochastic differential system, as time increases to infinity.

The book is written in a textlike style, namely, the material is essentially self-contained and problems are given at the end of each chapter. The only prerequisite is elementary probability; more specifically, the reader is assumed to be familiar with concepts such as conditional expectation, independence, and with elementary facts such as the Borel–Cantelli lemma. This prerequisite material can be found in any probability text, such as Breiman [1; Chapters 3, 4].

I would like to thank my colleague, Mark Pinsky, for some helpful conversations.

General Notation

All functions are real valued, unless otherwise explicity stated.

In Chapter n, Section m the formulas and theorems are indexed by $(m.k)$ and $m.k$ respectively. When in Chapter l, we refer to such a formula $(m.k)$ (or Theorem $m.k$), we designate it by $(n.m.k)$ (or Theorem $n.m.k$) if $l \neq n$, and by $(m.k)$ (or Theorem $m.k$) if $l = n$.

Similarly, when referring to Section m in the same chapter, we designate the section by m; when referring to Section m of another chapter, say n, we designate the section by $n.m$.

Finally, when we refer to conditions (A), (A_1), (B) etc., these conditions are usually stated earlier in the *same* chapter.

Contents of Volume 2

1

Stochastic Processes

1. The Kolmogorov construction of a stochastic process

We write a.s. for almost surely, a.e. for almost everywhere, and a.a. for almost all.

We shall use the following notation:

R^n is the real euclidean n-dimensional space;

$\mathfrak{B}_n$ is the Borel σ-field of R^n, i.e., the smallest σ-field generated by the open sets of R^n;

R^∞ is the space consisting of all infinite sequences $(x_1, x_2, \ldots, x_n, \ldots)$ of real numbers;

$\mathfrak{B}_\infty$ is the smallest σ-field of subsets of R^∞ containing all k-dimensional rectangles, i.e., all sets of the form

$$\{(x_1, x_2, \ldots); x_1 \in I_1, \ldots, x_k \in I_k\}, \qquad k > 0,$$

where $I_1, \ldots, I_k$ are intervals.

Clearly $\mathfrak{B}_\infty$ is also the smallest σ-field containing all sets

$$\{(x_1, x_2, \ldots); (x_1, \ldots, x_k) \in B\}, \qquad B \in \mathfrak{B}_k.$$

An *n-dimensional distribution function* $F_n(x_1, \ldots, x_n)$ is a real-valued function defined on R^n and satisfying:

(i) for any intervals $I_k = [a_k, b_k)$, $1 \leqslant k \leqslant n$,

$$\Delta_{I_1} \cdots \Delta_{I_n} F_n(x) \geqslant 0 \tag{1.1}$$

where

$$\Delta_{I_k} f(x) = f(x_1, \ldots, x_{k-1}, b_k, x_{k+1}, \ldots, x_n) - f(x_1, \ldots, x_{k-1}, a_k, x_{k+1}, \ldots, x_n);$$

(ii) if $x_j^{(k)} \uparrow x_j$ as $k \uparrow \infty$ $(1 \leqslant j \leqslant n)$, then

$$F_n(x_1^{(k)}, \ldots, x_n^{(k)}) \uparrow F_n(x_1, \ldots, x_n) \quad \text{as} \quad k \uparrow \infty;$$

(iii) if $x_j \downarrow -\infty$ for some j, then $F_n(x_1, \ldots, x_n) \downarrow 0$, and if $x_j \uparrow \infty$ for all j, $1 \leqslant j \leqslant n$, then $F_n(x_1, \ldots, x_n) \uparrow 1$.

Unless the contrary is explicitly stated, random variables are always assumed to be real valued, i.e., they do not assume the values $+\infty$ or $-\infty$.

If $X_1, \ldots, X_n$ are random variables, then $X = (X_1, \ldots, X_n)$ is called an n-dimensional random variable, or a random variable with values in R^n. The function

$$F_n(x_1, \ldots, x_n) = P(X_1 < x_1, \ldots, X_n < x_n) \tag{1.2}$$

is an n-dimensional distribution. Conversely, for any n-dimensional distribution F_n there is a probability space and a random variable X with values in R^n such that its distribution function is F_n, i.e., (1.2) holds; see, for instance, Breiman [1].

A sequence $\{F_n(x_1, \ldots, x_n)\}$ of distribution functions is said to be *consistent* if

$$\lim_{x_n \uparrow \infty} F_n(x_1, \ldots, x_n) = F_{n-1}(x_1, \ldots, x_{n-1}), \qquad n > 1. \tag{1.3}$$

Let $(\Omega, \mathfrak{F}, P)$ be a probability space, and let $\{X_n\}$ be a sequence of random variables. We call such a sequence a *discrete time stochastic process*, or a *countable stochastic process*, or a *stochastic sequence*. The distribution functions of $(X_1, \ldots, X_n)$, $n \geqslant 1$, form a consistent sequence. The converse is also true, namely:

Theorem 1.1. *Let $\{F_n\}$ be a sequence of n-dimensional distribution functions satisfying the consistency condition* (1.3). *Then there exists a probability space* $(\Omega, \mathfrak{F}, P)$ *and a countable stochastic process* $\{X_n\}$ *such that*

$$P(X_1 < x_1, \ldots, X_n < x_n) = F(x_1, \ldots, x_n), \qquad n \geqslant 1.$$

In fact, one can take $\Omega = R^\infty$, $\mathfrak{F} = \mathfrak{B}_\infty$,

$$P\{(x_1, x_2, \ldots); x_1 < y_1, \ldots, x_n < y_n\} = F_n(y_1, \ldots, y_n)$$

and more generally

$$P\{(x_1, x_2, \ldots); a_1 \leqslant x_1 < b_1, \ldots, a_n \leqslant x_n < b_n\}$$

is given by the left-hand side of (1.1), and

$$X_n((x_1, x_2, \ldots)) = x_n.$$

The consistency condition guarantees that P is well defined on all rectangles. The main nontrivial point here is the fact that P can be extended into a probability in $(\Omega, \mathfrak{F})$, and this follows from the ***Kolmogorov extension theorem***:

If P is defined on all finite-dimensional rectangles, and:

(i) $P \geqslant 0$, $P(R^\infty) = 1$;
(ii) if $S_1, \ldots, S_m$ are disjoint n-rectangles and $S = S_1 \cup \cdots \cup S_m$ then $P(S) = \sum_{j=1}^m P(S_j)$;
(iii) if $\{S_j\}$ is a nondecreasing sequence of n-dimensional rectangles and $S_j \uparrow S$ as $j \uparrow \infty$, then $\lim_{j\to\infty} P(S_j) = P(S)$;

then there is a unique extension of P into a probability on $(R^\infty, \mathfrak{B}_\infty)$.

For proof, see Breiman [1].

A *stochastic process* is a family of random variables $\{X(t)\}$ defined on a probability space $(\Omega, \mathfrak{F}, P)$, where t varies in a real interval I (I is open, closed, or half-closed). We denote it also by $\{X(t), t \in I\}$, $\{X(t)\}$ $(t \in I)$ or $X(t)$, $t \in I$.

A stochastic process defines a set of distribution functions

$$F_{t_1 \cdots t_n}(x_1, \ldots, x_n) = P[X(t_1) < x_1, \ldots, X(t_n) < x_n] \tag{1.4}$$

where $t_1 < \cdots < t_n$, $t_j \in I$, $n = 1, 2, \ldots$; they are called the finite-dimensional joint distribution functions of the process.

A set of distribution functions

$$\{F_{t_1 \cdots t_n}(x_1, \ldots, x_n)\} \tag{1.5}$$

defined for all $t_1 < \cdots < t_n$, $t_j \in I$, $1 \leqslant j \leqslant n$, $n = 1, 2, \ldots$, is said to be *consistent* if

$$\lim_{x_k \uparrow \infty} F_{t_1 \cdots t_n}(x_1, \ldots, x_n) = F_{t_1 \cdots t_{k-1} t_{k+1} \cdots t_n}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n).$$

The set given by (1.4) is consistent. We shall now prove the converse.

Theorem 1.2. *Given a consistent set of distribution functions as in* (1.5), *there exists a stochastic process* $\{X(t), t \in I\}$ *defined on a probability space* $(\Omega, \mathfrak{F}, P)$ *such that the given distribution functions are the joint distribution functions of the process, i.e., such that* (1.4) *holds.*

Proof. Denote by R^I the set of all real valued functions $x(\cdot)$ on I. A set of the form

$$B = \{x(\cdot); (x(t_1), x(t_2), \ldots) \in D\}, \qquad D \in \mathfrak{B}_\infty \tag{1.6}$$

is called a cylinder set with countable base $\{t_j\}$. Denote by $\mathfrak{B}_I$ the class of all cylinder sets with countable base. This class is clearly a σ-field, and it coincides with the σ-field generated by all the finite-dimensional rectangles in R^I,

$$\{x(\cdot); x(t_1) \in I_1, \ldots, x(t_n) \in I_n\},$$

where $I_1, \ldots, I_n$ are intervals.

We take $\Omega = R^I$, $\mathfrak{F} = \mathfrak{B}_I$. Given a sequence $T = \{t_j\}$ in I, denote by $\mathfrak{B}_T$ the class of all sets

$$\{x(\cdot); (x(t_1), x(t_2), \dots) \in D\}, \qquad D \in \mathfrak{B}_\infty.$$

By Theorem 1.1 and the remark following it, there exists a unique probability measure P_T on $(\Omega, \mathfrak{B}_T)$ such that

$$P_T\{x(\cdot); x(t_1) < x_1, \dots, x(t_n) < x_n\} = F_{t_1, \dots, t_n}(x_1, \dots, x_n), \qquad n \geq 1. \tag{1.7}$$

Set $P(B) = P_T(B)$ if $B \in \mathfrak{B}_T$. For this definition to make sense we must show that if $B \in \mathfrak{B}_{T'}$ where $T' = \{t_j'\}$ then

$$P_T(B) = P_{T'}(B). \tag{1.8}$$

Setting $S = T \cup T'$ we have $T \subset S$, $\mathfrak{B}_T \subset \mathfrak{B}_S$. Since $P_T = P_S$ on all finite-dimensional rectangles with base in T, $P_T = P_S$ on any set of $\mathfrak{B}_T$. Thus $P_T(B) = P_S(B)$. Similarly, $P_{T'}(B) = P_S(B)$, and (1.8) follows.

Having defined the probability space $(\Omega, \mathfrak{F}, P)$, we now take

$$X(t, x(\cdot)) = x(t);$$

(1.4) then immediately follows.

We refer to the stochastic process constructed in this proof as the process obtained by the *Kolmogorov construction*.

Definition. A *ν-dimensional countable stochastic process* (or a *ν-dimensional stochastic sequence*) is a sequence $\{X_n\}$ of ν-dimensional random variables. A family $\{X(t), t \in I\}$ of ν-dimensional random variables is called a *ν-dimensional stochastic process*.

A ν-dimensional stochastic sequence $\{X_n\}$ gives rise to functions

$$F_n(\bar{x}_1, \dots, \bar{x}_n) = P(X_1 < \bar{x}_1, \dots, X_n < \bar{x}_n) \tag{1.9}$$

where $\bar{x}_i \in R^\nu$ and $a < b$ means that $a_j < b_j$ for $1 \leq j \leq \nu$, where $a = (a_1, \dots, a_\nu)$, $b = (b_1, \dots, b_\nu)$. Setting

$$G_{(n-1)\nu+k}(x_1, \dots, x_{(n-1)\nu+k}) = F_n(\bar{x}_1, \dots, \bar{x}_{n-1}, \bar{x}_n^*) \qquad (1 \leq k \leq \nu) \tag{1.10}$$

where

$$\bar{x}_i = (x_{(i-1)\nu+1}, x_{(i-1)\nu+2}, \dots, x_{i\nu}) \qquad \text{if} \quad 1 \leq i \leq n-1,$$

$$\bar{x}_n^* = (x_{(n-1)\nu+1}, \dots, x_{(n-1)\nu+k}, \infty, \dots, \infty),$$

it is clear that the sequence $\{G_i(x_1, \dots, x_i)\}$ is a consistent sequence of distribution functions.

Definition. A function $F_n(\bar{x}_1, \ldots, \bar{x}_n)$ defined for all $\bar{x}_i \in R^\nu$ is called an *n-dimensional distribution function with respect to* R^ν if:

(i) condition (1.1) holds with a_k, b_k replaced by $\bar{a}_k = (a_{k1}, \ldots, a_{k\nu})$, $\bar{b}_k = (b_{k1}, \ldots, b_{k\nu})$, where $\bar{a}_k < \bar{b}_k$, and x_j replaced by $\bar{x}_j$ in the definition of $\Delta_{I_k} f$;

(ii) $F_n(\bar{x}_1^{(k)}, \ldots, \bar{x}_n^{(k)}) \uparrow F_n(\bar{x}_1, \ldots, \bar{x}_n)$ if $\bar{x}_i^{(k)} \uparrow \bar{x}_i$ for $1 \leqslant k \leqslant n$ (i.e., each component of $\bar{x}_i^{(k)}$ increases to the corresponding component of $\bar{x}_i$);

(iii) if some kth component of some $\bar{x}_i$ decreases to $-\infty$, then $F_n(\bar{x}_1, \ldots, \bar{x}_n) \downarrow 0$; and if all components of all the $\bar{x}_j$ increase to ∞, then $F_n(\bar{x}_1, \ldots, \bar{x}_n) \uparrow 1$.

Set

$$F_n(\bar{x}_1, \ldots, \bar{x}_{n-1}, \bar{x}_n^*) = \lim_{k\to\infty} F_n(\bar{x}_1, \ldots, \bar{x}_{n-1}, \bar{x}_n^k) \tag{1.11}$$

where the first k components of $\bar{x}_n^*$, $\bar{x}_n^k$ agree, whereas the last $\nu - k$ components of $\bar{x}_n^*$ and $\bar{x}_n^k$ are equal, respectively, to ∞ and k. Then, obviously, the set of distribution functions $G_{(n-1)\nu+1}, \ldots, G_{n\nu}$ defined by (1.10) is consistent.

A sequence of distributions $\{F_n\}$ with respect to R^ν is said to be *consistent* if

$$\lim_{\bar{x}_n \uparrow \infty} F_n(\bar{x}_1, \ldots, \bar{x}_n) = F_{n-1}(\bar{x}_1, \ldots, \bar{x}_{n-1}),$$

where $\bar{x}_n \uparrow \infty$ means that each component of $\bar{x}_n$ increases to ∞. This is clearly the case if and only if the corresponding sequence of distributions $G_i(x_1, \ldots, x_i)$, defined by (1.10), (1.11), is consistent.

If we apply Theorem 1.1 to the latter sequence, we conclude that for any consistent sequence of distribution functions, with respect to R^ν, there exists a countable ν-dimensional stochastic process $\{X_n\}$ whose distribution functions are the F_n, i.e., (1.9) holds. The probability space $(\Omega, \mathcal{F}, P)$ can be taken such that $\Omega = R^{\nu,\infty}$ = the space of all infinite sequences $(\bar{x}_1, \bar{x}_2, \ldots)$, $\bar{x}_n \in R^\nu$, and $\mathcal{F} = \mathcal{B}_{\nu,\infty}$ the smallest σ-field containing all finite-dimensional rectangles

$$\{(\bar{x}_1, \bar{x}_2, \ldots); \bar{x}_1 \in \bar{I}_1, \bar{x}_2 \in \bar{I}_2, \ldots\}$$

where $\bar{I}_j$ is an interval in R^ν.

A set of distribution functions $\{F_{t_1 \cdots t_n}(\bar{x}_1, \ldots, \bar{x}_n)\}$ where $t_1 < \cdots < t_n$, $t_i \in I$, $n = 1, 2, \ldots,$ is said to be *consistent* if

$$\lim_{\bar{x}_k \uparrow \infty} F_{t_1 \cdots t_n}(\bar{x}_1, \ldots, \bar{x}_n) = F_{t_1 \cdots t_{k-1} t_{k+1} \cdots t_n}(\bar{x}_1, \ldots, \bar{x}_{k-1}, \bar{x}_{k+1}, \ldots, \bar{x}_n).$$

Theorem 1.3. *Given a consistent set of distribution functions with respect to* R^ν, $\{F_{t_1 \cdots t_n}(\bar{x}_1, \ldots, \bar{x}_n);\ t_1 < \cdots < t_n,\ t_i \in I\}$, *there exists* *probability space* $(\Omega, \mathcal{F}, P)$ *and a* ν*-dimensional stochastic process* $X(t)$ $(t \in I)$ *defined on it such that*

$$P[X(t_1) < \bar{x}_1, \ldots, X(t_n) < \bar{x}_n] = F_{t_1 \cdots t_n}(\bar{x}_1, \ldots, \bar{x}_n). \qquad (1.12)$$

The functions $F_{t_1 \cdots t_n}$ are called the finite-dimensional joint distribution functions (with respect to R^ν) of the process $X(t)$.

The proof is similar to the proof of Theorem 1.2; it is based on the extension of Theorem 1.1 to distribution functions with respect to R^ν, mentioned above. Here Ω is the space $R^{\nu, I}$ of all functions $x(\cdot)$ from I into R^ν and $\mathcal{F} = \mathcal{B}_{\nu, I}$ consists of all sets (1.6) with $D \in \mathcal{B}_{\nu, \infty}$. The construction of the process $X(t)$ in this proof is called the *Kolmogorov construction*.

2. Separable and continuous processes

Definition. A ν-dimensional stochastic process $\{X(t), t \in I\}$ is called *separable* if there exists a countable sequence $T = \{t_j\}$ that is a dense subset of I and a subset N of Ω with $P(N) = 0$ such that, if $\omega \not\in N$,

$$\{X(t, \omega) \in F \text{ for all } t \in J\} = \{X(t_j, \omega) \in F \text{ for all } t_j \in J\}$$

for any open subset J of I and for any closed subset F of R^ν. T is called a *set of separability*, or briefly, a *separant*.

Two ν-dimensional stochastic processes $\{X(t)\}$ and $\{X'(t)\}$ $(t \in I)$ defined on the same probability space are said to be *stochastically equivalent* if

$$P[X(t) \neq X'(t)] = 0 \qquad \text{for all} \quad t \in I.$$

We then say that $\{X'(t)\}$ is a *version* of $\{X(t)\}$.

Theorem 2.1. *Any* ν*-dimensional process is stochastically equivalent to a separable stochastic process.*

The random variables of the separable process may actually take the values $\pm\infty$, even though the random variables of the original process are real valued.

For proof the reader is referred to Doob [1]. Theorem 2.1 will not be used in the text of this book.

From Theorem 2.1 it follows that when one constructs a stochastic process from a given set of distribution functions (as in the Kolmogorov construction), one may always take this process to be separable.

Let $X(t)$ be a ν-dimensional stochastic process for $t \in I$. Thus, for each t, $X(t)$ is a random variable $X(t, \omega)$. The functions $t \to X(t, \omega)$ (ω fixed) are called the *sample paths* of the process. If for a.a. ω the sample paths are continuous functions for all $t \in I$, then we say that the stochastic process is *continuous*. If for a.a. ω the sample paths are right (left) continuous, then we say that the stochastic process is *right (left) continuous*.

It is easy to see that a right (or left) continuous stochastic process is separable and any dense sequence in I is a set of separability.

A stochastic process $\{X(t), t \in I\}$ is said to be *continuous in probability* if for any $s \in I$ and $\epsilon > 0$,

$$P[|X(t) - X(s)| > \epsilon] \to 0 \qquad \text{if} \quad t \in I, \quad t \to s.$$

The following theorem is due to Kolmogorov.

Theorem 2.2. *Let $X(t)$ $(t \in I)$ be a ν-dimensional stochastic process such that*

$$E|X(t) - X(s)|^\beta \leqslant C|t - s|^{1+\alpha} \tag{2.1}$$

for some positive constants C, α, β. Then there is a version of $X(t)$ that is a continuous process.

Proof. Take for simplicity $I = [0, 1]$. Let $0 < \gamma < \alpha/\beta$ and let δ be a positive number such that

$$(1 - \delta)(\alpha + 1 - \beta\gamma) > 1 + \delta. \tag{2.2}$$

Then, by (2.1),

$$\begin{aligned} &P\{|X(j2^{-n}) - X(i2^{-n})| > [(j - i)2^{-n}]^\gamma \\ &\qquad\qquad\qquad \text{for some } 0 \leqslant i \leqslant j \leqslant 2^n, j - i < 2^{n\delta}\} \\ &\leqslant \sum_{\substack{0 \leqslant i \leqslant j < 2^n \\ j - i < 2^{n\delta}}} [(j - i)2^{-n}]^{-\beta\gamma} E|X(j2^{-n}) - X(i2^{-n})|^\beta \\ &\leqslant C_1 \sum_{\substack{0 \leqslant i \leqslant j < 2^n \\ j - i < 2^{n\delta}}} [(j - i)2^{-n}]^{1+\alpha-\beta\gamma} \leqslant C_2 2^{n[1+\delta-(1-\delta)(1+\alpha-\beta\gamma)]} \\ &= C_2 2^{-n\mu} \end{aligned}$$

where $\mu > 0$ by (2.2), and C_1, C_2 are positive constants. Since $\sum 2^{-n\mu} < \infty$, the Borel–Cantelli lemma implies that for a.a. ω

$$|X(j2^{-n}) - X(i2^{-n})| \leqslant h((j - i)2^{-n}) \qquad \text{if} \quad 0 \leqslant i \leqslant j \leqslant 2^n, \quad j - i \leqslant 2^{n\delta}, \quad n \geqslant m(\omega) \tag{2.3}$$

where $h(t) = t^\gamma$ and $m(\omega)$ is some sufficiently large positive integer.

Let t_1, t_2 be any rational numbers in the interval (0, 1) such that $t_1 < t_2$,

$t = t_2 - t_1 < 2^{-m(1-\delta)}$ where $m = m(\omega)$ is as in (2.3). Choose $n \geq m$ such that

$$2^{-(n+1)(1-\delta)} \leq t < 2^{-n(1-\delta)}.$$

We can expand t_1, t_2 in the form

$$t_1 = i2^{-n} - 2^{-p_1} - \cdots - 2^{-p_k},$$
$$t_2 = j2^{-n} + 2^{-q_1} + \cdots + 2^{-q_l}$$

where $n < p_1 < \cdots < p_k$, $n < q_1 < \cdots < q_l$. Then $t_1 \leq i2^{-n} \leq j2^{-n} \leq t_2$ and $j - i \leq t2^n < 2^{n\delta}$.

By (2.3),

$$|X(i2^{-n} - 2^{-p_1} - \cdots - 2^{-p_r}) - X(i2^{-n} - 2^{-p_1} - \cdots - 2^{-p_{r-1}})| \leq h(2^{-p_r}).$$

Hence

$$|X(t_1) - X(i2^{-n})| \leq \sum_{r=1}^{k} h(2^{-p_r}) \leq C_3 h(2^{-n}) \qquad (C_3 \text{ constant}).$$

Similarly,

$$|X(t_2) - X(j2^{-n})| \leq C_3 h(2^{-n}).$$

Finally, by (2.3),

$$|X(j2^{-n}) - X(i2^{-n})| \leq h((j-i)2^{-n}) \leq h(t).$$

Hence

$$|X(t_2) - X(t_1)| \leq Ch(t_2 - t_1) \qquad (C \text{ constant}). \tag{2.4}$$

Let $T = \{t_j\}$ be the sequence of all rational numbers in the interval (0, 1). From (2.4) it follows that

$$|X(t_j) - X(t_i)| \leq Ch(t_j - t_i) \qquad \text{if} \quad 0 \leq t_j - t_i \leq 2^{-m(1-\delta)}, \quad m = m(\omega). \tag{2.5}$$

Thus $X(t, \omega)$, when restricted to $t \in T$ is uniformly continuous. Let $\tilde{X}(t, \omega)$ be its unique continuous extension to $0 \leq t \leq 1$. Then the process $\tilde{X}(t)$ is continuous. We shall complete the proof of Theorem 2.2 by showing that $\tilde{X}(t)$ is a version of $X(t)$.

Let $t^* \in [0, 1]$ and let $t_{j'} \in T$, $t_{j'} \to t^*$ if $j' \to \infty$. Since $\tilde{X}(t)$ is a continuous process,

$$\tilde{X}(t_{j'}, \omega) \to \tilde{X}(t^*, \omega) \quad \text{a.s.} \qquad \text{as} \quad j' \to \infty.$$

From (2.1) it also follows that

$$X(t_{j'}, \omega) \to X(t^*, \omega) \quad \text{in probability,} \qquad \text{as} \quad j' \to \infty.$$

Since $\tilde{X}(t_{j'}, \omega) = X(t_{j'}, \omega)$ a.s., we get $\tilde{X}(t^*, \omega) = X(t^*, \omega)$ a.s.

From (2.5) we see that

$$|\tilde{X}(t) - \tilde{X}(s)| \leqslant Ch(t-s) = C(t-s)^{\gamma}$$

if $t-s<2^{-m(1-\delta)}$, $m=m(\omega)$. Thus:

Corollary 2.3. *Let $X(t)$ $(t\in I)$ be a ν-dimensional process satisfying (2.1). Then for any bounded subinterval $J\subset I$ and for any $\epsilon>0$, a.a. sample paths of the continuous version of $X(t)$ satisfy a Hölder condition with exponent $(\alpha/\beta)-\epsilon$ on J.*

3. Martingales and stopping times

Definition. A stochastic sequence $\{X_n\}$ is called a *martingale* if $E|X_n|<\infty$ for all n, and

$$E(X_{n+1}|X_1, \ldots, X_n) = X_n \quad \text{a.s.} \qquad (n = 1, 2, \ldots).$$

It is a *submartingale* if $E|X_n| < \infty$ for all n, and

$$E(X_{n+1}|X_1, \ldots, X_n) \geqslant X_n \quad \text{a.s.} \qquad (n = 1, 2, \ldots).$$

Theorem 3.1. *If $\{X_n\}$ is a submartingale, then*

$$P\Big(\max_{1\leqslant k\leqslant n} X_k > \lambda\Big) \leqslant \frac{1}{\lambda} EX_n^+ \qquad \text{for any} \quad \lambda > 0, \quad n \geqslant 1. \tag{3.1}$$

This inequality is called the *martingale inequality*.

Proof. Let

$$A_k = \bigcap_{j<k} (X_j < \lambda)\cap(X_k > \lambda) \qquad (1 \leqslant k \leqslant n),$$

$$A = \Big(\max_{1\leqslant k\leqslant n} X_k > \lambda\Big).$$

Then A is a disjoint union $\bigcup_{k=1}^{n} A_k$. Therefore

$$\lambda P(A) = \sum_{k=1}^{n} \lambda P(A_k) \leqslant \sum_{k=1}^{n} E(X_k\chi_{A_k}) \tag{3.2}$$

where χ_B is the indicator function of a set B. Since A_k belongs to the σ-field of $X_1, \ldots, X_k$,

$$\begin{aligned} EX_n^+ &\geq \sum_{k=1}^n E(X_n^+ \chi_{A_k}) = \sum_{k=1}^n EE(X_n^+ \chi_{A_k} | X_1, \ldots, X_k) \\ &= \sum_{k=1}^n E\chi_{A_k} E(X_n^+ | X_1, \ldots, X_k) \\ &\geq \sum_{k=1}^n E\chi_{A_k} E(X_n | X_1, \ldots, X_k) \geq \sum_{k=1}^n E(\chi_{A_k} X_k). \end{aligned}$$

Comparing this with (3.2), the assertion (3.1) follows.

Corollary 3.2. *If $\{X_n\}$ is a martingale and $E|X_n|^\alpha < \infty$ for some $\alpha \geq 1$ and all $n \geq 1$, then*

$$P\Big(\max_{1 \leq k \leq n} |X_k| > \lambda\Big) \leq \frac{1}{\lambda^\alpha} E|X_n|^\alpha \qquad \textit{for any} \quad \lambda > 0, \quad n \geq 1.$$

This follows from Theorem 3.1 and the fact that $\{|X_n|^\alpha\}$ is a submartingale (see Problem 8).

Definition. A stochastic process $\{X(t), t \in I\}$ is called a *martingale* if $E|X(t)| < \infty$ for all $t \in I$, and

$$E\{X(t) | X(\tau), \tau \leq s, \tau \in I\} = X(s) \qquad \text{for all} \quad t \in I, \quad s \in I, \quad s \leq t.$$

It is called a *submartingale* if $E|X(t)| < \infty$ and

$$E\{X(t) | X(\tau), \tau \leq s, \tau \in I\} \geq X(s) \qquad \text{for all} \quad t \in I, \quad s \in I, \quad s \leq t.$$

Corollary 3.3. (i) *If $\{X(t)\}$ $(t \in I)$ is a separable submartingale, then*

$$P\Big\{\max_{s \leq t, s \in I} X(s) > \lambda\Big\} \leq \frac{1}{\lambda} EX^+(t) \qquad \textit{for all} \quad \lambda > 0, \quad t \in I.$$

(ii) *If $\{X(t)\}$ $(t \in I)$ is a separable martingale and if $E|X(t)|^\alpha < \infty$ for some $\alpha \geq 1$ and all $t \in I$, then*

$$P\Big\{\max_{s \leq t, s \in I} X(s) > \lambda\Big\} \leq \frac{1}{\lambda^\alpha} E|X(t)|^\alpha \qquad \textit{for all} \quad \lambda > 0, \quad t \in I.$$

This follows by applying Theorem 3.1 and Corollary 3.2 with $X_j = X(s_j)$ if $1 \leq j \leq n-1$, $X_j = X(t)$ if $j \geq n$, where $s_1 < \cdots < s_{n-1} < t$, taking the set $\{s_1, \ldots, s_{n-1}\}$ to increase to the set of all t_j with $t_j < t$, where $\{t_j\}$ is a set of separability (cf. Problem 1(b)).

We denote by $\mathfrak{F}(X(\lambda), \lambda \in I)$ the σ-field generated by the random variables $X(\lambda)$, $\lambda \in I$.

Definition. Let $X(t)$, $\alpha \leqslant t \leqslant \beta$, be a ν-dimensional stochastic process. (If $\beta = \infty$, we take $\alpha \leqslant t < \infty$.) A (finite-valued) random variable τ is called a *stopping time* with respect to the process $X(t)$ if $\alpha \leqslant \tau \leqslant \beta$ and if, for any $\alpha \leqslant t \leqslant \beta$, $\{\tau \leqslant t\}$ belongs to $\mathcal{F}(X(\lambda), \alpha \leqslant \lambda \leqslant t)$. If $\beta = \infty$ and τ may also assume the value $+\infty$, then we call τ an *extended stopping time*.

Notice that the sets $\{\tau > t\}$, $\{s < \tau \leqslant t\}$ (where $s < t$) also belong to $\mathcal{F}(X(\lambda), \alpha \leqslant \lambda \leqslant t)$.

Lemma 3.4. *If τ is a stopping time with respect to a ν-dimensional process $X(t)$, $t \geqslant 0$, then there exists a sequence of stopping times τ_n such that*

(i) *τ_n has a discrete range,*
(ii) *$\tau_n \geqslant \tau$ everywhere,*
(iii) *$\tau_n \downarrow \tau$ everywhere as $n \uparrow \infty$.*

Proof. Define $\tau_n = 0$ if $\tau = 0$ and

$$\tau_n = \frac{i+1}{n} \qquad \text{if} \quad \frac{i}{n} < \tau \leqslant \frac{i+1}{n} \qquad (i = 1, 2, \dots).$$

Then clearly (i)–(iii) hold. Finally, τ_n is a stopping time since

$$\{\tau_n \leqslant t\} = \bigcup_{(i+1)/n \leqslant t} \left\{\tau_n = \frac{i+1}{n}\right\}$$

$$= \bigcup_{(i+1)/n \leqslant t} \left\{\frac{i}{n} < \tau \leqslant \frac{i+1}{n}\right\} \in \mathcal{F}(X(\lambda), \lambda \leqslant t).$$

If $X(t)$ is right continuous, then

$$X(t \wedge \tau_n) \to X(t \wedge \tau) \quad \text{a.s.} \qquad \text{as} \quad n \to \infty.$$

Since $X(t \wedge \tau_n)$ is obviously a random variable, the same is true of $X(t \wedge \tau)$. We shall prove:

Theorem 3.5. *If τ is a stopping time with respect to a right continuous martingale $X(t)$, $t \geqslant 0$, then the process $Y(t) = X(t \wedge \tau)$, $t \geqslant 0$, is also a right continuous martingale.*

Before giving the proof, we need some preliminaries.

Definition. A sequence of random variables X_n is said to be *uniformly integrable* if $E|X_n| < \infty$ for any $n \geqslant 1$, and

$$\overline{\lim_{n\to\infty}} \int_{|X_n|>\lambda} |X_n| \, dP \to 0 \qquad \text{as} \quad \lambda \to \infty. \tag{3.3}$$

Lemma 3.6. *If $\{X_n\}$ is uniformly integrable, then* $\overline{\lim}\ E|X_n| < \infty$. *If further $X_n \to X$ a.s. or in probability, then $E|X - X_n| \to 0$ and $EX_n \to EX$.*

Proof. For any $\lambda > 0$,

$$E|X_n| \leqslant \lambda + \int_{|X_n|>\lambda} |X_n|\, dP.$$

From this and (3.3) it follows that $E|X_n| \leqslant C$, where C is a constant independent of n. To prove the last assertion, notice, by Fatou's lemma, that if $X_n \to X$ a.s., then

$$\int |X|\, dP \leqslant \underline{\lim} \int |X_n|\, dP < \infty.$$

Next, by Egoroff's theorem, $X_n \to X$ almost uniformly, i.e., for any $\epsilon > 0$ there is an event A with $P(A) < \epsilon$ such that $X_n \to X$ uniformly on $\Omega \backslash A$. Hence

$$\overline{\lim_{n\to\infty}} \int |X_n - X|\, dP = \overline{\lim_{n\to\infty}} \int_A |X_n - X|\, dP \leqslant \overline{\lim_{n\to\infty}} \int_A |X_n|\, dP + \int_A |X|\, dP.$$

Now, by (3.3),

$$\int_A |X_n|\, dP \leqslant \int_{|X_n|>\lambda} |X_n|\, dP + \int_{(|X_n|<\lambda)\cap A} |X_n|\, dP \leqslant \eta(\lambda) + \lambda\epsilon,$$

where $\eta(\lambda) \to 0$ if $\lambda \to \infty$, and $\eta(\lambda)$ does not depend on n. By Fatou's lemma, the same inequality holds for X. Taking $\epsilon \downarrow 0$, we get

$$\overline{\lim_{n\to\infty}} \int |X_n - X|\, dP \leqslant 2\eta(\lambda) \to 0 \qquad \text{if} \quad \lambda \to \infty.$$

If $X_n \to X$ in probability, then any subsequence $X_{n'}$ of X_n has a subsubsequence $X_{n''}$ which converges a.s. to X. By what we have proved, $E|X_{n''} - X| \to 0$ if $n'' \to \infty$. But this implies that $E|X_n - X| \to 0$ as $n \to \infty$.

Proof of Theorem 3.5. Let $0 \leqslant t_1 < t_2 < \cdots < t_{m-1} < t_m = s < t$,

$$A = \bigcap_{i=1}^{m} [X(t_i \wedge \tau) \in B_i], \qquad B_i \text{ Borel sets in } R^1, \tag{3.4}$$

and let τ_n be the stopping times constructed in the proof of Lemma 3.4. Assume first that there is an integer j such that $s = (j+1)/n$. On the set $\tau \leqslant (j+1)/n$, $\tau_n \leqslant (j+1)/n$ so that

$$X(s \wedge \tau_n) = X(\tau_n) = X(t \wedge \tau_n).$$

Consequently

$$\int_{A\cap[\tau\leqslant(j+1)/n]} X(s \wedge \tau_n)\, dP = \int_{A\cap[\tau\leqslant(j+1)/n]} X(t \wedge \tau_n)\, dP. \tag{3.5}$$

Next,

$$\int_{A\cap[\tau>(j+1)/n]} X(t\wedge\tau_n)\,dP=\sum_{l=j+1}^{M}\int_{A\cap[l/n<\tau\leq(l+1)/n]} X\left(\frac{l+1}{n}\right)dP + \int_{A\cap[\tau>(M+1)/n]} X(t)\,dP \tag{3.6}$$

where $(M+1)/n<t\leq(M+2)/n$.

Notice that

$$A\cap\left[\tau>\frac{M+1}{n}\right]=\left[\bigcap_{i=1}^{m}(X(t_i)\in B_i)\right]\cap\left[\tau>\frac{M+1}{n}\right].$$

Since each set on the right belongs to $\mathcal{F}(X(\lambda), \lambda\leq(M+1)/n)$, we can use the martingale property of $X(t)$ to deduce that

$$\int_{A\cap[\tau>(M+1)/n]} X(t)\,dP=\int_{A\cap[\tau>(M+1)/n]} X\left(\frac{M+1}{n}\right)dP.$$

Hence, from (3.6),

$$\int_{A\cap[\tau>(j+1)/n]} X(t\wedge\tau_n)\,dP=\sum_{l=j+1}^{M-1}\int_{A\cap[l/n<\tau\leq(l+1)/n]} X\left(\frac{l+1}{n}\right)dP + \int_{A\cap[\tau>M/n]} X\left(\frac{M+1}{n}\right)dP.$$

Proceeding in this manner step by step, we arrive at

$$\int_{A\cap[\tau>(j+1)/n]} X(t\wedge\tau_n)\,dP=\int_{A\cap[\tau>(j+1)/n]} X\left(\frac{j+2}{n}\right)dP$$
$$=\int_{A\cap[\tau>(j+1)/n]} X(s)\,dP=\int_{A\cap[\tau>(j+1)/n]} X(s\wedge\tau_n)\,dP,$$

since $s=(j+1)/n$,

$$A\cap\left[\tau>\frac{j+1}{n}\right]=\left[\bigcap_{i=1}^{m}(X(t_i)\in B_i)\right]\cap\left[\tau>\frac{j+1}{n}\right]$$

is in $\mathcal{F}(X(\lambda), \lambda\leq(j+1)/n)$, and $s\wedge\tau_n=s$ on this set.

Recalling (3.5) we conclude that

$$\int_A X(s\wedge\tau_n)\,dP=\int_A X(t\wedge\tau_n)\,dP. \tag{3.7}$$

So far we have assumed that the number s is in the range of τ_n. If this is not the case, i.e., if $j/n<s<(j+1)/n$ for some j, then we modify the definition of τ_n, by taking

$$\tau_n = \begin{cases} s & \text{if } \dfrac{j}{n} < \tau \leqslant s, \\ \dfrac{j+1}{n} & \text{if } s < \tau \leqslant \dfrac{j+1}{n}. \end{cases}$$

The assertion (3.7) then follows as before, and the τ_n satisfy the properties asserted in Lemma 3.4.

We shall now complete the proof of the theorem in case τ is bounded, i.e., $\tau \leqslant N < \infty$ (N constant).

The range $\{s_{in}\}$ of τ_n lies in $[0, N+1)$.

Notice now that $|X(t)|$ is a submartingale. If we apply the proof of (3.7) to $|X(t)|$, instead of $X(t)$, we obtain

$$\int_A |X(s \wedge \tau_n)|\, dP \leqslant \int_A |X(t \wedge \tau_n)|\, dP$$

whenever $t > s$. Taking, in particular,

$$\tau = \tau_n, \qquad A = \{|X(s \wedge \tau_n)| > \lambda\}, \qquad t = N+1,$$

we get

$$\int_{|X(s\wedge\tau_n)|>\lambda} |X(s \wedge \tau_n)|\, dP \leqslant \int_{|X(s\wedge\tau_n)|>\lambda} |X(\tau_n)|\, dP.$$

But since $|X(t)|$ is a submartingale,

$$\int_{\substack{|X(s\wedge\tau_n)|>\lambda \\ \tau_n = s_{in}}} |X(\tau_n)|\, dP \leqslant \int_{\substack{|X(s\wedge\tau_n)|>\lambda \\ \tau_n = s_{in}}} |X(N+1)|\, dP.$$

Summing over i,

$$\int_{|X(s\wedge\tau_n)|>\lambda} |X(s \wedge \tau_n)|\, dP \leqslant \int_{|X(s\wedge\tau_n)|>\lambda} |X(N+1)|\, dP. \tag{3.8}$$

Now, since $X(t)$ is right continuous and τ is bounded, the random variable

$$\sup_n \sup_{s \geqslant 0} |X(s \wedge \tau_n)|$$

is bounded in probability. This implies that the right-hand side of (3.8) converges to 0 if $\lambda \to \infty$, uniformly with respect to n. Thus the sequence $\{X(s \wedge \tau_n)\}$ is uniformly integrable. Similarly, one shows that the sequence $\{X(t \wedge \tau_n)\}$ is uniformly integrable. By Lemma 3.6 it follows that $E|X(t \wedge \tau)| < \infty$.

If we now take $n \uparrow \infty$ in (3.7) and apply Lemma 3.6, we get

$$\int_A X(s \wedge \tau)\, dP = \int_A X(t \wedge \tau)\, dP. \tag{3.9}$$

Since this is true for any set A of the form (3.4), $\{X(t \wedge \tau)\}$ is a martingale.

We have thus completed the proof in case τ is bounded. If τ is unbounded, then $\tau \wedge N$ is a bounded stopping time. Hence $E|X(t \wedge \tau)| < \infty$ if $N > t$. Further, (3.9) holds with τ replaced by $\tau \wedge N$, and taking $N > t$ the relation (3.9) follows. This completes the proof.

Remark. Let $\{X(t), 0 \leqslant t \leqslant T\}$ be a right continuous process and let $\mathcal{F}_t$ $(0 < t \leqslant T)$ be an increasing family of σ-fields such that $\mathcal{F}(X(\lambda), 0 < \lambda < t)$ is a subset of $\mathcal{F}_t$. If $E|X(t)| < \infty$ and

$$E(X(t)|\mathcal{F}_s) = X(s) \qquad \text{for all} \quad 0 \leqslant s \leqslant t \leqslant T,$$

then we say that $X(t)$ is a *martingale with respect to* $\mathcal{F}_t$ $(0 \leqslant t \leqslant T)$. This clearly implies that $\{X(t), 0 \leqslant t \leqslant T\}$ is a martingale. The proof of Theorem 3.5 shows that if $X(t)$ is a martingale with respect to $\mathcal{F}_t$ $(0 \leqslant t \leqslant T)$, then the process $X(t \wedge \tau)$ is also a martingale with respect to $\mathcal{F}_t$; here τ is any random variable such that $\tau \geqslant 0$ and $\{\tau \leqslant t\} \in \mathcal{F}_t$ for all $0 \leqslant t \leqslant T$.

PROBLEMS

1. Let $X(t)$, $t \geqslant 0$ be a stochastic process and let $\{t_j\}$ be a set of separability. Prove that for a.a. ω:

(a) $$\lim_{n\to\infty} \inf_{|t_j - t| < 1/n} X(t_j, \omega) \leqslant X(t, \omega) \leqslant \lim_{n\to\infty} \sup_{|t_j - t| < 1/n} X(t_j, \omega) \qquad (t > 0).$$

(b) $$\sup_{a<t<b} X(t, \omega) = \sup_{a<t_j<b} X(t_j, \omega), \quad \inf_{a<t<b} X(t, \omega) = \inf_{a<t_j<b} X(t_j, \omega).$$

(Notice that, as a consequence, $\sup_{a<t<b} X(t)$ and $\inf_{a<t<b} X(t)$ are finite- or infinite-valued random variables.)

(c) $\sup_{a\leqslant t\leqslant b}|X(t)|$, $\limsup_{t\to s} X(t)$, and $\liminf_{t\to s} X(t)$ are finite- or infinite-valued random variables.

(d) For any positive numbers ϵ, δ,

$$\sup_{|t_i - t_j| < \delta} |X(t_i, \omega) - X(t_j, \omega)| \leqslant \epsilon$$

if and only if

$$\sup_{|t - s| < \delta} |X(t, \omega) - X(s, \omega)| \leqslant \epsilon.$$

2. Let $X(t)$, $X'(t)$ $(t \geqslant 0)$ be ν-dimensional stochastic processes in $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', P')$ respectively, having the same finite-dimensional distribution functions. Then for any sequence $\{t_n\}$, $t_n \geqslant 0$, and for any set $D \in \mathcal{B}_{\nu,\infty}$,

$$P\{(X(t_1), X(t_2), \ldots) \in D\} = P'\{(X'(t_1), X'(t_2), \ldots) \in D\}.$$

3. Suppose two ν-dimensional stochastic processes $X(t)$ and $X'(t)$ ($t \geqslant 0$) defined in probability spaces $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', P')$, respectively, are separable and have the same finite-dimensional distribution functions. If $X(t)$ is continuous, then $X'(t)$ is continuous.

[*Hint*: Let $\nu = 1$. It suffices to prove continuity for $0 \leqslant t \leqslant t^*$, $t^* < \infty$. Let $T = \{t_j\}$ be a set of separability for both processes for $0 \leqslant t \leqslant t^*$. The continuity of $X(t)$ implies that

$$P\left\{ \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \sup_{|t_i - t_j| < 1/m} |X(t_i) - X(t_j)| < \frac{1}{n} \right\} = 1.$$

By Problem 2, with

$$D = \bigcap_{n} \bigcup_{m} \bigcap_{|t_i - t_j| < 1/m} D_{mn}^{ij}, \qquad D_{mn}^{ij} = \left\{ (x_1, x_2, \ldots), |x_i - x_j| < \frac{1}{n} \right\},$$

$$P'\left\{ \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \sup_{|t_i - t_j| < 1/m} |X'(t_i) - X'(t_j)| < \frac{1}{n} \right\} = 1.$$

Now use Problem 1(d).]

4. Let $X(t)$ and $X'(t)$ ($t \geqslant 0$) be ν-dimensional separable stochastic processes defined in probability spaces $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', P')$ respectively, and having the same finite-dimensional distribution functions. Then:

(a) $P\{|X(t)| \to \infty \text{ as } t \to \infty\} = P'\{|X'(t)| \to \infty \text{ as } t \to \infty\}$;

(b) $P\{|X(t)|/t^{\alpha} \to M \text{ as } t \to \infty\} = P'\{|X'(t)|/t^{\alpha} \to M \text{ as } t \to \infty\}$,

for any $\alpha > 0$, $M \geqslant 0$.

[*Hint*: $\{|X(t)| \to \infty \text{ as } t \to \infty\} = \{\bigcap_n \bigcup_m \inf_{t_j > m} |X(t_j)| > n\}$.]

5. If two stochastic processes are stochastically equivalent, then they have the same finite-dimensional distribution functions.

6. Let X_n, $n \geqslant 1$ be independent random variables with $E|X_n| < \infty$, $EX_n = 0$. Then the sequence of partial sums $X_1 + \cdots + X_n$ is a martingale.

7. Let X be a random variable in $(\Omega, \mathcal{F}, P)$ with $E|X| < \infty$, and let $\phi(x)$ be a convex function defined on the real line such that $E|\phi(X)| < \infty$. Prove:

(a) $\phi(EX) \leqslant E\phi(X)$ (Jensen's inequality);

(b) $\phi(E(X|\mathcal{F}_0)) \leqslant E(\phi(X)|\mathcal{F}_0)$ a.e., where $\mathcal{F}_0$ is any σ-subfield of $\mathcal{F}$.

8. If $\{X_n\}$ is a martingale and $\phi(x)$ is a convex function on the real line, and if $E|\phi(X_n)| < \infty$ for all n, then $\{\phi(X_n)\}$ is a submartingale.

9. If $\{X_n\}$ is a submartingale and $\phi(x)$ is a convex function and monotone nondecreasing on the range of the X_n, and if $E|\phi(X_n)| < \infty$ for all n, then $\{\phi(X_n)\}$ is a submartingale.

10. If τ is a bounded stopping time with respect to right continuous martingale $X(t)$, then $EX(\tau) = EX(0)$.

11. Prove that a continuous process is continuous in probability. Give an example showing that the converse is not true in general.

12. Let $X(t)$, $\tilde{X}(t)$ $(t \in I)$ be two stochastically equivalent and separable processes. Show that if one of them is continuous, then the other is also continuous, and

$$P\{X(t) = \tilde{X}(t) \text{ for all } t \in I\} = 1.$$

13. A stochastic process $\{X(t), t \in I\}$ is said to be *measurable* if the function $(t, \omega) \to X(t, \omega)$ (from the product measure space into R^1) is measurable. Show that a continuous stochastic process is measurable.

2

Markov Processes

1. Construction of Markov processes

If $x(t)$ is a stochastic process, we denote by $\mathcal{F}(x(\lambda), \lambda \in J)$ the smallest σ-field generated by the random variables $x(t)$, $t \in J$.

Definition. Let $p(s, x, t, A)$ be a nonnegative function defined for $0 \leqslant s < t < \infty$, $x \in R^\nu$, $A \in \mathcal{B}_\nu$ and satisfying:

(i) $p(s, x, t, A)$ is Borel measurable in x, for fixed s, t, A;
(ii) $p(s, x, t, A)$ is a probability measure in A, for fixed s, x, t;
(iii) p satisfies the *Chapman-Kolmogorov equation*

$$p(s, x, t, A) = \int_{R^\nu} p(s, x, \lambda, dy) p(\lambda, y, t, A) \qquad \text{for any} \quad s < \lambda < t. \tag{1.1}$$

Then we call p a *Markov transition function*, a *transition probability function*, or a *transition probability*.

Theorem 1.1. *Let p be a transition probability function. Then, for any $s \geqslant 0$ and for any probability distribution $\pi(dx)$ on $(R^\nu, \mathcal{B}_\nu)$, there exists a ν-dimensional stochastic process $\{x(t), s \leqslant t < \infty\}$ such that*:

$$P[x(s) \in A] = \pi(A), \tag{1.2}$$

$$P\{x(t) \in A \mid \mathcal{F}(x(\lambda), s \leqslant \lambda \leqslant \bar{s})\} = p(\bar{s}, x(\bar{s}), t, A) \quad a.s. \qquad (s \leqslant \bar{s} < t). \tag{1.3}$$

Proof. Let $M[[0, \infty); R^\nu]$ be the space of all functions ω from $[0, \infty)$ into R^ν. Let $\mathcal{F}_t^s$ be the smallest σ-field with respect to which all functions $\omega(u)$ are measurable, $s \leqslant u \leqslant t$, i.e., $\mathcal{F}_t^s$ is the σ-field generated by all the sets $\{\omega;\ \omega(u) \in A\}$ where $s \leqslant u \leqslant t$, $A \in \mathcal{B}_\nu$.

Let $\mathcal{F}^s_\infty$ be the smallest σ-field with respect to which all functions $\omega(u)$ are measurable, $s \leqslant u < \infty$. We take $\Omega = M[[0, \infty); R^\nu]$, $\mathcal{F} = \mathcal{F}^s_\infty$.

We next define a family of finite-dimensional distribution functions with respect to R^ν: if $s = t_0 < t_1 < \cdots < t_n$ and $\bar{x}_i \in R^\nu$ $(0 \leqslant i \leqslant n)$, then

$$F_{t_0t_1\cdots t_n}(\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_n) = \int_{-\infty}^{\bar{x}_n} \cdots \int_{-\infty}^{\bar{x}_1} \int_{-\infty}^{\bar{x}_0} p(t_{n-1}, y_{n-1}, t_n, dy_n) \cdots p(t_1, y_1, t_2, dy_2)p(t_0, y_0, t_1, dy_1)\pi(dy_0). \quad (1.4)$$

It is easy to verify that the F_n are in fact n-dimensional distribution functions with respect to R^ν. Using the Chapman–Kolmogorov equation it follows that

$$F_{t_0\cdots t_n}(\bar{x}_0, \ldots, \bar{x}_n) \to F_{t_0\cdots t_{k-1}t_{k+1}\cdots t_n}(\bar{x}_0, \ldots, \bar{x}_{k-1}, \bar{x}_{k+1}, \ldots, \bar{x}_n) \quad \text{if} \quad \bar{x}_k \uparrow \infty;$$

thus the family of distribution functions in (1.4) is consistent.

By the Kolmogorov construction, there exists a probability P_s in $(\Omega, \mathcal{F}^s_\infty)$ and a ν-dimensional stochastic process $x(t)$, $s \leqslant t < \infty$, such that the finite-dimensional distribution functions of this process coincide with the given distributions functions $F_{t_0\cdots t_n}$.

Notice that $x(t, \omega) = \omega(t)$.

Now, if $s = t_0 < t_1 < \cdots < t_n$,

$$P_s\{x(t_0) \in B_1, \ldots, x(t_n) \in B_n\} = \int_{B_n} \cdots \int_{B_1} \int_{B_0} p(t_{n-1}, y_{n-1}, t_n, dy_n) \cdots p(t_0, y_0, t_1, dy_1)\pi(dy_0) \quad (1.5)$$

provided each B_i is a ν-dimensional interval $(-\infty, \bar{x}_i)$. It is easily seen that both sides of (1.5) coincide also when the B_i are any ν-dimensional intervals $(\tilde{x}_i, \bar{x}_i)$, or disjoint unions of such intervals. Since two probabilities that agree on a field $\mathcal{Q}$ agree also on the σ-field generated by $\mathcal{Q}$, it follows that (1.5) holds for all B_i in $\mathcal{B}_\nu$.

Taking $n = 0$ in (1.5) we get the assertion (1.2). Thus it remains to prove (1.3). Since the random variable on the right-hand side of (1.3) is measurable with respect to $\mathcal{F}(x(\lambda), s \leqslant \lambda \leqslant \bar{s})$, it remains to prove that

$$\int_D \chi_{[x(t)\in A]}\, dP_s = \int_D p(\bar{s}, x(\bar{s}), t, A)\, dP_s \qquad \text{if} \quad D \in \mathcal{F}(x(\lambda), s \leqslant \lambda \leqslant \bar{s}). \quad (1.6)$$

If we prove (1.6) for any set D of the form

$$D = \{x(t_0) \in B_0, x(t_1) \in B_1, \ldots, x(t_n) \in B_n\} \quad (1.7)$$

where $s = t_0 < t_1 < \cdots < t_n = \bar{s}$, then, by the countable additivity of both sides of (1.6), the assertion (1.6) will follow for any $D \in \mathcal{F}(x(\lambda), s \leqslant \lambda \leqslant \bar{s})$.

Set $t_{n+1} = t$, $B_{n+1} = A$. When D is given by (1.7), the left-hand side of (1.6) becomes

$$P_s[D \cap (x(t) \in A)] = P_s\{x(t_0) \in B_0, \ldots, x(t_n) \in B_n, x(t_{n+1}) \in B_{n+1}\}$$
$$= \int_{B_{n+1}} \int_{B_n} \cdots \int_{B_1} \int_{B_0} p(t_n, y_n, t_{n+1}, dy_{n+1}) p(t_{n-1}, y_{n-1}, t_n, dy_n)$$
$$\cdots p(t_0, y_0, t_1, dy_1) \pi(dy_0)$$
$$= \int_{B_n} \cdots \int_{B_1} \int_{B_0} p(t_n, y_n, t_{n+1}, A) p(t_{n-1}, y_{n-1}, t_n, dy_n)$$
$$\cdots p(t_0, y_0, t_1, dy_1) \pi(dy_0).$$

Therefore it suffices to show that

$$\int_{B_n} \cdots \int_{B_1} \int_{B_0} f(y_n) p(t_{n-1}, y_{n-1}, t_n, dy_n) \cdots p(t_0, y_0, t_1, dy_1) \pi(dy_0)$$
$$= \int_{x(t_n) \in B_n} \cdots \int_{x(t_0) \in B_0} f(x(t_n)) \, dP_s \tag{1.8}$$

where f is a Borel measurable function. It clearly suffices to prove (1.8) when f is any indicator function χ_F. Setting $B_n' = B_n \cap F$, (1.8) becomes

$$\int_{B_n'} \int_{B_{n-1}} \cdots \int_{B_1} \int_{B_0} p(t_{n-1}, y_{n-1}, t_n, dy_n) \cdots p(t_0, y_0, t_1, dy_1) \pi(dy_0)$$
$$= \int_{x(t_n) \in B_n'} \int_{x(t_{n-1}) \in B_{n-1}} \cdots \int_{x(t_0) \in B_0} dP_s. \tag{1.9}$$

Since the right-hand side is equal to

$$P_s\{x(t_0) \in B_0, \ldots, x(t_{n-1}) \in B_{n-1}, x(t_n) \in B_n'\},$$

(1.9) is a consequence of (1.5). This completes the proof of (1.3).

For any $x \in R^n$ let π_x be defined by

$$\pi_x(A) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Denote by $P_{x,s}$ the probability P_s constructed on the measure space $(\Omega, \mathcal{F}_\infty)$ when $\pi = \pi_x$, and denote the expectation by $E_{x,s}$. Set $\mathcal{F} = \mathcal{F}_\infty^0$. Then we have the following situation:

(a) $(\Omega, \mathcal{F})$ is a measure space, $\mathcal{F}_t^s$ are σ-subfields such that $\mathcal{F}_t^s \subset \mathcal{F}_{t'}^{s'}$ if $s' \leqslant s$, $t < t'$, and $\mathcal{F}$ is the smallest σ-field containing all the $\mathcal{F}_t^s$.

(b) There is a function $x(t, \omega)$ from $[0, \infty) \times \Omega$ into R^ν such that $x(t, \omega)$ is $\mathcal{F}_t^s$ measurable for each $t > s$.

(c) For each $x \in R^\nu$, $s > 0$ there is a probability $P_{x,s}$ defined on $(\Omega, \mathcal{F}_\infty)$

such that

$$P_{x,s}\{x(s,\omega)=x\}=1, \tag{1.10}$$

$$P_{x,s}\{x(t+h,\omega)\in A|\mathcal{F}_t^s\}=p(t,x(t,\omega),t+h,A) \quad \text{a.s.} \qquad (t\geqslant s, h>0). \tag{1.11}$$

Definition. A collection $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ satisfying (a), (b), (c), where p is a transition probability function, is called a (ν-dimensional) *Markov process* corresponding to p.

The property (1.11) is called the *Markov property*.

Notice that in the model constructed in Theorem 1.1 $\mathcal{F}_t^s$ is the smallest σ-field generated by $x(u)$, $s \leqslant u \leqslant t$. Notice also that $x(t,\omega)=\omega(t)$.

The model constructed in Theorem 1.1 will be called the *Kolmogorov model*.

Taking the expectation $E_{x,s}$ of both sides of (1.11), we get (when $t=s$ and $t+h$ is denoted by t),

$$E_{x,s}f(x(t))=\int_{R^\nu} p(s,x,t,dy)f(y) \tag{1.12}$$

if $f=\chi_A$. By approximation, this relation holds also for any bounded Borel measurable function f.

We can therefore write (1.11) in the form

$$\begin{aligned} E_{x,s}\{f(x(t+h))|\mathcal{F}_t^s\} &= \int_{R^\nu} p(t,x(t),t+h,dy)f(y) \\ &= E_{x(t),t}f(x(t+h)) \quad \text{a.s.} \end{aligned} \tag{1.13}$$

if $f=\chi_A$. By approximation, this relation holds also for any bounded Borel measurable function f.

We have the following generalization of (1.5) (in case $\pi=\pi_x$):

$$\begin{aligned} &P_{x,s}\{(x(t_1),\ldots,x(t_n))\in D\} \\ &\quad =\int\cdot\underset{D}{\cdot}\cdot\int p(t_{n-1},y_{n-1},t_n,dy_n)\cdots p(s,x,t_1,dy_1) \end{aligned} \tag{1.14}$$

where D is any set in $\mathcal{B}_{\nu,n}$ ($\mathcal{B}_{\nu,n}$ is the σ-field of Borel sets in $R^\nu \times R^\nu \times \cdots \times R^\nu$ (n times)). Indeed, if we denote the left-hand and right-hand sides of (1.14) by $P(D)$ and $Q(D)$ respectively, then $P(D)$ and $Q(D)$ define probability measures on $\mathcal{B}_{\nu,n}$. Since, by (1.5), these measures agree on all rectangular sets, they must agree on $\mathcal{B}_{\nu,n}$.

Theorem 1.2. *If* $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{s,x}\}$ *is a ν-dimensional Markov process, then*

$$P_{x,s}(B|\mathcal{F}_t^s)=P_{x,s}(B|x(t)) \quad \textit{a.s.} \qquad \textit{for any} \quad B\in\mathcal{F}(x(\lambda),\lambda\geqslant t). \tag{1.15}$$

Proof. Let $0 < h_1 < h_2$ and let f_1, f_2 be bounded Borel measurable functions in R^ν. Using (1.13), we have

$$\begin{aligned}
&E_{x,s}\{f_1(x(t+h_1))f_2(x(t+h_2))|\mathcal{F}_t^s\}\\
&\quad = E_{x,s}\{f_1(x(t+h_1))E_{x,s}[f_2(x(t+h_2))|\mathcal{F}_{t+h_1}^s]|\mathcal{F}_t^s\}\\
&\quad = E_{x,s}\left\{\left[f_1(x(t+h_1))\int p(t+h_1, x(t+h_1), t+h_2, dy_2)f_2(y_2)\right]\Big|\mathcal{F}_t^s\right\}\\
&\quad = \int p(t, x(t), t+h_1, dy_1)f_1(y_1)\int p(t+h_1, y_1, t+h_2, dy_2)f_2(y_2).
\end{aligned}$$

Similarly one can prove by induction on n that, if $0 < h_1 < \cdots < h_n$,

$$\begin{aligned}
&E_{x,s}[f_1(x(t+h_1))\cdots f_n(x(t+h_n))|\mathcal{F}_t^s]\\
&\quad = \int\cdots\int p(t, x(t), t+h_1, dy_1)f_1(y_1)\\
&\qquad \cdots p(t+h_{n-1}, y_{n-1}, t+h_n, dy_n)f_n(y_n).
\end{aligned} \tag{1.16}$$

For any set $D \in \mathcal{B}_{\nu,n}$ its indicator function $f(x_1, \ldots, x_n) = \chi_D(x_1, \ldots, x_n)$ can be uniformly approximated by finite linear combinations of bounded measurable functions of the form $f_1(x_1)\cdots f_n(x_n)$ (each x_i varies in R^ν). Employing (1.16) we deduce (by the Lebesgue bounded convergence theorem) that

$$\begin{aligned}
&E_{x,s}[f(x(t+h_1), \ldots, x(t+h_n))|\mathcal{F}_t^s]\\
&\quad = \int\cdots\int f(y_1, \ldots, y_n)p(t, x(t), t+h_1, dy_1)\\
&\qquad \cdots p(t+h_{n-1}, y_{n-1}, t+h_n, dy_n) \quad \text{a.s.}
\end{aligned} \tag{1.17}$$

Setting

$$B = \{(x(t_1), \ldots, x(t_n)) \in D\}, \tag{1.18}$$

we conclude that $P_{x,s}(B|\mathcal{F}_t^s)$ is $\mathcal{F}(x(t))$ measurable and (1.15) holds for this B; here $\mathcal{F}(x(t))$ denotes the smallest σ-subfield with respect to which $x(t)$ is measurable.

Since the class $\mathcal{Q}$ of sets B for which (1.15) holds is a monotone class containing all the sets of the form (1.18), it must contain the σ-field generated by these sets, i.e., $\mathcal{Q} = \mathcal{F}(x(\lambda), \lambda \geq t)$.

Theorem 1.3. *Let $x(t)$ and $y(t)$ be ν-dimensional processes defined for $t \geq 0$ and having the same set of joint distribution functions. Suppose $x(t)$ satisfies the assertions in Theorem* 1.1. *Then $y(t)$ also satisfies the same assertions.*

An outline of the proof is given in Problem 4.

2. The Feller and the strong Markov properties

Denote by $M(R^\nu)$ the space of all bounded Borel measurable functions f from R^ν into R^1 with the norm $\|f\| = \text{ess sup}_{x\in R^\nu}|f(x)|$. Consider the mappings $T_{s,t}$ from $M(R^\nu)$ into $M(R^\nu)$:

$$(T_{s,t}f)(x) = \int_{R^\nu} p(s, x, t, dy)f(y) \qquad (0 \leqslant s < t < \infty).$$

It is easily seen that $\|T_{s,t}\| = 1$, and

$$T_{s,t}T_{t,u} = T_{s,u} \qquad \text{if} \quad 0 \leqslant s < t < u.$$

We define $T_{t,t}$ = identity. The family $\{T_{s,t}\}$ is called the *semigroup* associated with the transition probability function p (or with the corresponding Markov process).

Definition. If for any bounded continuous function $f(x)$, the function

$$(t, z)\to(T_{t,t+\lambda}f)(z) = \int_{R^\nu} p(t, z, t + \lambda, dy)f(y)$$

is continuous, for any $\lambda > 0$, then we say that the transition probability p (or the corresponding Markov process) satisfies the *Feller property*.

We denote by $\mathcal{F}^s_{t+}$ the intersection of the σ-fields $\mathcal{F}^s_{t'}$, $t' > t$.
Similarly we denote by $\mathcal{F}^s_{t-}$ the smallest σ-field containing $\mathcal{F}^s_{t'}$, $t' < t$.

Theorem 2.1. *Let $\{\Omega, \mathcal{F}, \mathcal{F}^s_t, x(t), P_{x,s}\}$ be a ν-dimensional Markov process, right continuous and satisfying the Feller property. Then $\{\Omega, \mathcal{F}, \mathcal{F}^s_{t+}, x(t), P_{x,s}\}$ is also a Markov process with the same transition probability function.*

Proof. Let f be a bounded continuous function, and let $\epsilon > 0$, $h > 0$. By (1.13),

$$E_{x,s}\{f(x(t+h+\epsilon))|\mathcal{F}^s_{t+\epsilon}\} = \int p(t+\epsilon, x(t+\epsilon), t+h+\epsilon, dy)f(y).$$

Taking the conditional expectation with respect to $\mathcal{F}^s_{t+}$, we get

$$E_{x,s}\{f(x(t+h+\epsilon))|\mathcal{F}^s_{t+}\}$$
$$= E_{x,s}\Big\{\int p(t+\epsilon, x(t+\epsilon), t+h+\epsilon, dy)f(y)|\mathcal{F}^s_{t+}\Big\}. \tag{2.1}$$

Let $\epsilon \downarrow 0$. Since $x(t+\epsilon)\to x(t)$ a.s., the Feller property implies that

$$\int p(t+\epsilon, x(t+\epsilon), t+h+\epsilon, dy)f(y)\to \int p(t, x(t), t+h, dy)f(y) \quad \text{a.s.}$$

By the Lebesgue bounded convergence theorem for conditional expectation, we then get

$$E_{x,s}\left\{ \int p(t+\epsilon, x(t+\epsilon), t+h+\epsilon, dy) f(y) | \mathcal{F}^s_{t+} \right\}$$

$$\to E_{x,s}\left\{ \int p(t, x(t), t+h, dy) f(y) | \mathcal{F}^s_{t+} \right\} = \int p(t, x(t), t+h, dy) f(y),$$

since $x(t)$ is $\mathcal{F}^s_{t+}$ measurable.

On the other hand

$$f(x(t+h+\epsilon)) \to f(x(t)) \qquad \text{as} \quad \epsilon \downarrow 0,$$

so that

$$E_{x,s}\{ f(x(t+h+\epsilon)) | \mathcal{F}^s_{t+} \} \to E_{x,s}\{ f(x(t+h)) | \mathcal{F}^s_{t+} \}.$$

Thus, upon taking $\epsilon \downarrow 0$ in (2.1), we get

$$E_{x,s}\{ f(x(t+h)) | \mathcal{F}^s_{t+} \} = \int p(t, x(t), t+h, dy) f(y).$$

This completes the proof.

From Theorem 2.1 it follows that whenever one has constructed, from a given transition probability function, a Markov process that is right continuous and satisfies the Feller property, one may assume (without loss of generality) that $\mathcal{F}^s_{t+} = \mathcal{F}^s_t$.

Remark. If a Markov process is left continuous, and if $\mathcal{F}^s_t = \mathcal{F}\{x(\lambda); s \leqslant \lambda \leqslant t\}$, then $\mathcal{F}^s_{t-} = \mathcal{F}^s_t$. This follows by noting that, for any $\alpha \in R^\nu$, the set $\{\alpha \cdot x(t) \leqslant \lambda\}$ coincides with the set

$$\bigcap_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left[\alpha \cdot x\left(t - \frac{1}{k}\right) \leqslant \lambda + \frac{1}{m} \right]$$

which belongs to $\mathcal{F}^s_{t-}$.

Definition. An extended real-valued random variable τ defined on Ω is called an *s-Markov time* or an *s-stopping time* with respect to a given ν-dimensional Markov process $\{\Omega, \mathcal{F}, \mathcal{F}^s_t, x(t), P_{x,s}\}$ if $\tau \geqslant s$ and the sets $\{\tau \leqslant t\}$ belong to $\mathcal{F}^s_t$ for every $t \geqslant s$.

Any constant function $\tau = t$ with $t \geqslant s$ is an s-stopping time.

Since $\{\tau \leqslant t\}$ is the complement of $\{\tau > t\}$, τ is an s-stopping time if and only if $\tau \geqslant s$ and the sets $\{\tau > t\}$ belong to $\mathcal{F}^s_t$, $t \geqslant s$. Notice that the sets

$$\{\lambda < \tau \leqslant t\} = \{\tau \leqslant t\} \backslash \{\tau \leqslant \lambda\}$$

also belong to $\mathcal{F}^s_t$.

Consider the events $A \subset \mathcal{F}_\infty^s$ such that

$$A \cap \{\tau \leqslant t\} \in \mathcal{F}_t^s \qquad \text{for all} \quad t \geqslant s.$$

They form a σ-field $\mathcal{F}_\tau^s$.

Definition. A ν-dimensional Markov process $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ with transition probability p is said to have the *strong Markov property* if for every $x \in R^\nu$ and for every real-valued s-Markov time τ,

$$P_{x,s}\{x(t+\tau) \in A \mid \mathcal{F}_\tau^s\} = p(\tau, x(\tau), t+\tau, A) \quad \text{a.e.} \tag{2.2}$$

For $\tau = h$ (h constant), this coincides with the Markov property (1.11).

Using (1.12) it is clear that the strong Markov property is equivalent to

$$E_{x,s}\{f(t+\tau) \mid \mathcal{F}_\tau^s\} = \int p(\tau, x(\tau), t+\tau, dy) f(y) = E_{x(\tau),\tau} f(x(t+\tau)) \tag{2.3}$$

for any $f = \chi_A$; hence for any bounded Borel measurable function f.

We shall give some examples of s-stopping times.

Theorem 2.2. *Let $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ be a ν-dimensional continuous Markov process and let F be a nonempty closed subset of R^ν. Let $\tau(\omega) =$ first time $t \geqslant s$ such that $x(t, \omega) \in F$. Then τ is an s-stopping time, for any $s \geqslant 0$.*

If $s = 0$, τ is called the (*first*) *hitting time* of the set F or the *exit time* from the open set $R^\nu \backslash F$; if $s > 0$, τ is called hitting time, after s, of F (or the exit time, after s, from $R^\nu \backslash F$).

Proof. Let $\{t_j\}$ be a dense sequence in the interval $[s, \infty)$. Denote by $\rho(x, F)$ the distance from x to F. Let $\{\zeta_j\}$ be a dense sequence in F. Since

$$\{|x(t) - \zeta_j| < \alpha\} = \{\omega; x(t, \omega) \in D\} \qquad (\alpha > 0)$$

where $D = \{x \in R^\nu, |x - \zeta_j| < \alpha\}$, it follows that the set $\{|x(t) - \zeta_j| < \alpha\}$ is in $\mathcal{F}_t^s$, if $t \geqslant s$. Consequently, also the set

$$\bigcap_{n=1}^{\infty} \bigcup_{t_j < t} \left\{\rho(x(t_j), F) \leqslant \frac{1}{n}\right\} = \bigcap_{n=1}^{\infty} \bigcup_{t_j < t} \bigcup_{k} \left\{|x(t_j) - \zeta_k| < \frac{1}{n}\right\}$$

is in $\mathcal{F}_t^s$. But as is easily seen, this set coincides with the set $\{\tau \leqslant t\}$.

Let G be a nonempty open set in R^ν. Let

$$\tau(\omega) = \inf\{t; t \geqslant s, x(t, \omega) \in G\}.$$

Then τ is called the (*first*) *hitting time*, after s, of G. Notice that, for a continuous process,

$$\begin{aligned} \tau(\omega) &= \inf\{t; t > s, x(t, \omega) \in G\} \\ &= \inf\{t; t > s, \text{meas}\,[s < t' < t; x(t', \omega) \in G] > 0\} \quad \text{a.s.} \end{aligned}$$

For any set G, the extended random variable defined by the last expression is called the (*first*) *penetration time*, after s, of G.

Theorem 2.3. *Let $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ be a ν-dimensional continuous Markov process with $\mathcal{F}_{t+}^s = \mathcal{F}_t^s$, and let τ be the penetration time, after s, of an open set G. Then τ is an s-stopping time, for any $s \geqslant 0$.*

Proof. Let $\{t_j\}$ be a dense sequence in $[s, \infty)$, and let

$$A_n = \bigcup_{t_j < t+1/n} \bigcup_{k=1}^{\infty} \left[\rho(x(t_j), R^\nu \setminus G) > \frac{1}{k}\right].$$

It is clear that if $\tau(\omega) \leqslant t$, then $\omega \in A_n$; and if $\omega \in A_n$, then $\tau(\omega) < t + 1/n$. Hence $(\tau \leqslant t) = \lim A_n$. Since $A_n \in \mathcal{F}_{t+1/n}^s$, it follows that the set $(\tau \leqslant t)$ belongs to $\mathcal{F}_{t+}^s$.

Remark. Theorems 2.2, 2.3 are clearly valid not only for a continuous Markov process, but also for any continuous process $x(t)$; the definition of (extended) stopping time is given in Chapter 1, Section 3.

Theorem 2.4. *Let $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ be a ν-dimensional Markov process, right continuous and satisfying the Feller property, and let $\mathcal{F}_{t+}^s = \mathcal{F}_t^s$. Then it satisfies the strong Markov property.*

We begin with a special case.

Lemma 2.5. *The strong Markov property* (2.2) *holds for any Markov process provided τ is an s-stopping time with a countable range.*

Proof. Denote by $\{t_j\}$ the range of τ. Since $\{\tau \leqslant t_j\}$ and

$$\{\tau < t_j\} = \bigcup_{t_l < t_j} (\tau \leqslant t_l)$$

belong to $\mathcal{F}_{t_j}^s$, it follows that

$$G_j \equiv \{\tau = t_j\} \in \mathcal{F}_{t_j}^s.$$

Note next that G_j belongs to $\mathcal{F}_\tau^s$, since

$$\begin{aligned} \{\tau = t_j\} \cap \{\tau \leqslant t\} &\in \mathcal{F}_t^s \quad \text{if } t \geqslant t_j, \\ &= \phi \in \mathcal{F}_t^s \quad \text{if } t < t_j. \end{aligned}$$

Hence, for any $A \in \mathcal{B}_\nu$,

$$\{x(\tau) \in A\} = \bigcup_{x(t_j) \in A} \{\tau = t_j\} \qquad \text{is in } \mathcal{F}_\tau^s.$$

This shows that $x(\tau)$ is $\mathcal{F}_\tau^s$ measurable.

Now let $E \in \mathcal{F}_\tau^s$. Then

$$E \cap G_j = E \cap (\tau \leqslant t_j) \cap G_j \in \mathcal{F}_{t_j}^s.$$

The Markov property gives

$$P_{x,s}\{x(t + t_j) \in A | \mathcal{F}_{t_j}^s\} = p(t_j, x(t_j), t_j + t, A) \quad \text{a.s.,}$$

and since $E \cap G_j \in \mathcal{F}_{t_j}^s$,

$$P_{x,s}\{[x(t + t_j) \in A] \cap E \cap G_j\} = \int_{E \cap G_j} P(t_j, x(t_j), t_j + t, A)\, dP_{x,s}.$$

Noting that $\tau = t_j$ on G_j and summing over j, we get, upon recalling that $\bigcup G_j = \Omega$,

$$P_{x,s}\{[x(t + \tau) \in A] \cap E\} = \int_E p(\tau, x(\tau), \tau + t, A)\, dP_{x,s}. \tag{2.4}$$

To complete the proof of (2.2) we have to show that the integrand on the right-hand side of (2.2) is $\mathcal{F}_\tau^s$ measurable. Since τ and $x(\tau)$ are $\mathcal{F}_\tau^s$ measurable, the Feller property implies that

$$\int p(\tau, x(\tau), \tau + t, dy) f(y)$$

is $\mathcal{F}_\tau^s$ measurable if $f(y)$ is a bounded continuous function. Taking a sequence $f = f_m$ of uniformly bounded and continuous functions that converge a.e. to χ_A, we conclude that

$$\lim_{m \to \infty} \int p(\tau, x(\tau), \tau + t, dy) f_m(y) = p(\tau, x(\tau), \tau + t, A)$$

is $\mathcal{F}_\tau^s$ measurable.

Proof of Theorem 2.4. Choose τ_n as in Lemma 1.3.4. We claim that $\mathcal{F}_\tau^s \subset \mathcal{F}_{\tau_n}^s$. Indeed, since $\tau_n \geqslant \tau$,

$$A \cap (\tau_n \leqslant t) = [A \cap (\tau \leqslant t)] \cap (\tau_n \leqslant t)$$

for any set A and for any $t \geqslant s$. If $A \in \mathcal{F}_\tau^s$, then $A \cap (\tau \leqslant t) \in \mathcal{F}_t^s$. Since also $(\tau_n \leqslant t) \in \mathcal{F}_t^s$,

$$A \cap (\tau_n \leqslant t) \in \mathcal{F}_t^s, \qquad \text{i.e.,} \qquad A \in \mathcal{F}_{\tau_n}.$$

If $f = \chi_B$, then the strong Markov property for τ_n gives

$$E_{x,s}[f(x(t + \tau_n)) | \mathcal{F}_{\tau_n}^s] = \int p(\tau_n, x(\tau_n), t + \tau_n, dy) f(y) \quad \text{a.s.} \tag{2.5}$$

But then this relation holds also for any bounded continuous function f.

Let $E \in \mathcal{F}_\tau^s$. Then (as proved above) $E \in \mathcal{F}_{\tau_n}^s$ and therefore, by (2.5),

$$\int_E f(x(t + \tau_n))\, dP_{x,s} = \int_E \left[\int p(\tau_n, x(\tau_n), t + \tau_n, dy) f(y) \right] dP_{x,s}. \tag{2.6}$$

The function $\int p(\lambda, x, t+\lambda, dy)f(y)$ is continuous in λ, x. Also, if $n \uparrow \infty$, $\tau_n \downarrow \tau$ and $x(\tau_n) \to x(\tau)$ (by the right continuity of the $x(t)$). Hence, if we let $n \uparrow \infty$ in (2.6) and employ the Lebesgue bounded convergence theorem, we get

$$\int_E f(x(t+\tau))\, dP_{x,s} = \int_E \left[\int p(\tau, x(\tau), t+\tau, dy) f(y) \right] dP_{x,s}. \tag{2.7}$$

This relation holds also if $f = \chi_A$, $A \in \mathfrak{B}_\nu$.

If we show that $x(\tau)$ is $\mathfrak{F}^s_\tau$ measurable, then the proof of the theorem follows from (2.7) and the argument following (2.4).

Let $\alpha \in R^\nu$. Then, by the right continuity of $x(t)$,

$$[\alpha \cdot x(\tau) \leqslant \lambda] = \bigcap_{m=N}^{\infty} \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \left[\alpha \cdot x(\tau_n) < \lambda + \frac{1}{m} \right]$$

for any positive integer N. We also have

$$[\tau \leqslant t] = \bigcap_{p=N}^{\infty} \bigcap_{q=1}^{\infty} \bigcup_{r=q}^{\infty} \left[\tau_r < t + \frac{1}{p} \right]$$

and, more generally,

$$\begin{aligned} Q &\equiv [\alpha \cdot x(\tau) \leqslant \lambda] \cap [\tau \leqslant t] \\ &= \bigcap_{m=N}^{\infty} \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \left[\alpha \cdot x(\tau_n) \leqslant \lambda + \frac{1}{m} \right] \cap \left[\tau_n < t + \frac{1}{m} \right]. \end{aligned}$$

Since $x(\tau_n) \in \mathfrak{F}^s_{\tau_n}$,

$$\left[\alpha \cdot x(\tau_n) \leqslant \lambda + \frac{1}{m} \right] \cap \left[\tau_m < t + \frac{1}{m} \right] \in \mathfrak{F}^s_{t+1/m} \subset \mathfrak{F}^s_{t+1/N}.$$

It follows $Q \in \mathfrak{F}^s_{t+1/N}$ for any N. Hence $Q \in \mathfrak{F}^s_{t+}$. Since $\mathfrak{F}^s_{t+} = \mathfrak{F}^s_t$, we conclude that $[\alpha \cdot x(\tau) \leqslant \lambda] \in \mathfrak{F}^s_\tau$. But since α and λ are arbitrary, $x(\tau)$ is $\mathfrak{F}^s_\tau$ measurable.

Corollary 2.6. *Let $\{\Omega, \mathfrak{F}, \mathfrak{F}^s_t, x(t), P_{x,s}\}$ be a ν-dimensional continuous Markov process satisfying the Feller property. Then it satisfies the strong Markov property.*

Notice that we do not assume here that $\mathfrak{F}^s_{t+} = \mathfrak{F}^s_t$.

Proof. The proof of Theorem 2.4 applies here except for the last step, namely, the proof that $x(\tau)$ is $\mathfrak{F}^s_\tau$ measurable. To prove this, suppose for simplicity that $s = 0$. Let $\{t_l\}$ be a dense sequence in $[0, \infty)$. For any $\alpha \in R^\nu$ and λ real,

$$[\alpha \cdot x(\tau) \leqslant \lambda] \cap [\tau \leqslant t] = \bigcap_{n=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{i=0}^{n-1}{}^{*} \bigcup_{(i/n)t < t_l < ((i+1)/n)t} \left[\alpha \cdot x(t_l) \leqslant \lambda + \frac{1}{m}\right] \cap \left[\frac{i}{n} t < \tau \leqslant \frac{i+1}{n} t\right]$$

where * indicates that for $i = 0$ one takes $0 \leqslant \tau \leqslant t/n$. Since each set $[\cdots]$ on the right-hand side is in $\mathcal{F}_t^0$, it follows that $[\alpha \cdot x(\tau) \leqslant \lambda] \in \mathcal{F}_\tau^0$. But since α and λ are arbitrary, $x(\tau)$ is $\mathcal{F}_\tau^0$ measurable.

It can be shown (see Stroock and Varadhan [1]) that, for a continuous Markov process, $\mathcal{F}_\tau^s$ is actually generated by $x(t \wedge \tau)$, $t \geqslant s$.

The next result is the *Blumenthal zero–one law.*

Theorem 2.7. *Let $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ be a ν-dimensional Markov process, right continuous and satisfying the Feller property. If, for some $s \geqslant 0$, $x(s) = x$ a.s. (x a point in R^n), then*

$$P_{x,s}(B) = 0 \text{ or } 1 \qquad \text{for any} \quad B \in \mathcal{F}_{s+}^s. \tag{2.8}$$

Proof. By Theorem 2.1, formula (1.17) remains valid when $\mathcal{F}_t^s$ is replaced by $\mathcal{F}_{t+}^s$. Taking $t = s$, we get

$$P_{x,s}\{[(x(s+h_1), \ldots, x(s+h_n)) \in D] \mid \mathcal{F}_{s+}^s\} = \int_D \cdots \int p(s, x, s+h_1, dy_1) \cdots p(s+h_{n-1}, y_{n-1}, s+h_n, dy_n) \quad \text{a.s.}$$

By (1.14), the right-hand side is equal to

$$P_{x,s}\{(x(s+h_1), \ldots, x(s+h_n)) \in D\}.$$

Hence,

$$P_{x,s}(B \mid \mathcal{F}_{s+}^s) = P_{x,s}(B) \quad \text{a.s.} \tag{2.9}$$

for any set B of the form (1.18) with $t_i > s$. But then (2.9) follows also for any set B in $\mathcal{F}_\infty^s$. Taking, in particular, $B \in \mathcal{F}_{s+}^s$ and noting that $P_{x,s}(B \mid \mathcal{F}_{s+}^s) = \chi_B$, we get

$$\chi_B(\omega) = P_{x,s}(B) \quad \text{a.s.}$$

Since the right-hand side is constant, $\chi_B = \text{const}$ a.s. Hence either $\chi_B = 0$ a.s. or $\chi_B = 1$ a.s. This gives (2.8).

3. Time-homogeneous Markov processes

Definition. A Markov process is said to have a *stationary transition probability function* p if

$$p(s, x, t, A) = p(0, x, t - s, A).$$

We then also say that the Markov process is *time-homogeneous*, or *temporally homogeneous*.

It is clear that, when p is stationary, $P_{x,s} = P_{x,0}$, $E_{x,s} = E_{x,0}$ for all $s \geq 0$. Therefore, from the probabilistic point of view it suffices to work just with $P_{x,s}$, $E_{x,s}$, $\mathcal{F}_t^s$ when $s = 0$. We set

$$P_x = P_{x,0}, \qquad E_x = E_{x,0}, \qquad p(t, x, A) = p(0, x, t, A), \qquad \mathcal{F}_t = \mathcal{F}_t^0.$$

A 0-stopping time τ will be called, simply, a *stopping time* or a *Markov time*. (Notice that τ is actually an extended stopping time according to the definition of Chapter 1, Section 3.) The Markov property now takes the form

$$E_x\{f(x(t + h))|\mathcal{F}_t\} = E_{x(t)}f(x(h)) \quad \text{a.s.}; \tag{3.1}$$

(1.12) for $s = 0$ becomes

$$E_x f(x(t)) = \int p(t, x, dy) f(y), \tag{3.2}$$

and the strong Markov property takes the form

$$E_x\{f(x(t + \tau))|\mathcal{F}_\tau\} = E_{x(\tau)}f(x(t)) \tag{3.3}$$

where τ is a stopping time and $\mathcal{F}_\tau = \mathcal{F}_\tau^0$.

We shall denote a time-homogeneous Markov process by $\{\Omega, \mathcal{F}, \mathcal{F}_t, x(t), P_x\}$.

For a time-homogeneous Markov process, the Feller property states that the function

$$z \to \int p(t, z, dy) f(y)$$

is continuous for any bounded continuous function f and for any $t > 0$.

The semigroup associated with a stationary transition probability function is given by

$$(T_t f)(x) = \int p(t, x, dy) f(y). \tag{3.4}$$

Denote by M_0 the space of all functions from $M(R^\nu)$ into R^1 for which

$$\|T_t f - f\| \to 0 \quad \text{if} \quad t \to 0,$$

where $\|f\| = \text{ess sup}_{x \in R^\nu} |f(x)|$.

Definition. The *infinitesimal generator* $\mathcal{Q}$ of the time-homogeneous Markov process is defined by

$$\mathcal{Q}f = \lim_{t \to 0} \frac{T_t f - f}{t}. \tag{3.5}$$

Its domain $D_{\mathcal{Q}}$ consists of all the functions in M_0 for which the limit in (3.5) exists for a.a. x.

Theorem 3.1. *Let $\mathcal{Q}$ be the infinitesimal generator of a continuous time-homogeneous Markov process satisfying the Feller property. If $f \in D_{\mathcal{Q}}$ and $\mathcal{Q}f$ is continuous at a point y, then*

$$(\mathcal{Q}f)(y) = \lim_{U \downarrow \{y\}} \frac{E_y f(x(\tau_U)) - f(y)}{E_y \tau_U} \tag{3.6}$$

where U is an open neighborhood of y and τ_U is the exit time from U.

This is called *Dynkin's formula*. Since we shall not use this formula in the sequel, we relegate the proof to the problems.

For nonstationary transition probability function, one defines the infinitesimal generators by

$$\mathcal{Q}_s f = \lim_{t \downarrow 0} \frac{T_{s,\, t+s} f - f}{t}. \tag{3.7}$$

The Dynkin formula extends to this case (with obvious changes in the proof).

PROBLEMS

1. Let $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ be a continuous Markov process. Let τ be the first time that $t \geqslant s$ and $(t, x(t))$ hits a closed set F in $R^{\nu+1}$. Prove that τ is an s-stopping time, for any $s \geqslant 0$.
2. Let $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ be a ν-dimensional Markov process with transition probability p. Let X denote a variable point (x_0, x) in $R^{\nu+1}$ and let $X(t, \omega) = (t, x(t, \omega))$. Prove that $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, X(t), P_{X,s}\}$ is a Markov process with transition probability $\tilde{p}$,
$$\tilde{p}(s, X, t, I \times A) = \chi_I(x_0) p(s, x, t, A)$$
where I is any Borel set in $[0, \infty)$.
3. If $\{\Omega, \mathcal{F}, \mathcal{F}_t, x(t), P_x\}$ is a time-homogeneous Markov process and if
$$\int |y_1 - y_2|^\alpha p(t, x, dy_1) p(s, y_1, dy_2) \leqslant C|t - s|^{1+\epsilon}$$
for some $\alpha > 0$, $\epsilon > 0$, $C > 0$, then $x(t)$ has a continuous version.
[*Hint*: Use Theorem 1.2.2.]
4. Prove Theorem 1.3.

[*Hint*: Since $p(\bar{s}, y(\bar{s}), t, A)$ is $\mathcal{F}(y(\lambda), s \leqslant \lambda < \bar{s})$ measurable, it suffices to show that if $t_1 < \cdots < t_n = \bar{s} < t_{n+1} = t$,

$$P[y(t_1) \in A_1, \ldots, y(t_n) \in A_n, y(t_{n+1}) \in A]$$
$$= Ep(t_n, y(t_n), t_{n+1}, A)\chi_{A_1}(y(t_1)) \cdots \chi_{A_n}(y(t_n)).$$

Denote by Q_1 the probability distribution of $(x(t_1), \ldots, x(t_n), x(t_{n+1}))$ and by Q_2 the probability distribution of $(y(t_1), \ldots, y(t_n), y(t_{n+1}))$. Make use of the relations

$$\int \chi_{A_1}(x_1) \cdots \chi_{A_n}(x_n)\chi_A(x_{n+1})\, dQ_1(x_1, \ldots, x_n, x_{n+1})$$
$$= \int p(t_n, x_n, t_{n+1}, A)\chi_{A_1}(x_1) \cdots \chi_{A_n}(x_n)\, dQ_1(x_1, \ldots, x_n, x_{n+1}),$$
$$\int \Phi(x_1, \ldots, x_{n+1})\, dQ_1 = \int \Phi(x_1, \ldots, x_{n+1})\, dQ_2,$$

where Φ is bounded and measurable.]

5. A ν-dimensional stochastic process $x(t)$ $(t \geqslant 0)$ is said to satisfy the *Markov property* if

$$P[x(t) \in A | \mathcal{F}(x(\lambda)), \lambda \leqslant s)] = P[x(t) \in A | x(s)] \quad \text{a.s.} \qquad (\star)$$

for any $0 \leqslant s < t$ and for any Borel set A. It is well known (see Breiman [1]) that there exists a regular conditional probability $p(s, x, t, A)$ of $(\star)$, i.e., $p(s, x, t, A)$ is a probability in A for fixed (s, x, t) and Borel measurable in x for (s, t, A) fixed. Prove that p satisfies

$$\int p(s, x, \lambda, dy) p(\lambda, y, t, A) = p(s, x, t, A) \quad \text{a.s.} \qquad (s < \lambda < t).$$

6. Let $x(t)$ be a process satisfying the Markov property. Then, for any $0 = t_0 < t_1 < \cdots < t_n$ and for any Borel sets $B_0, \ldots, B_n$,

$$P\{x(t_n) \in B_n, \ldots, x(t_1) \in B_1, x(t_0) \in B_0\}$$
$$= \int_{B_n} \cdots \int_{B_1} \int_{B_0} p(t_{n-1}, x_{n-1}, t_n, dx_n) \cdots p(t_0, x_0, t_1, dx_1)\pi(dx_0)$$

where $\pi(dx)$ is the probability distribution of $x(0)$.

7. Let $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, x(t), P_{x,s}\}$ be a ν-dimensional continuous Markov process satisfying the Feller property. Let τ be a real-valued s-stopping time. Denote by $\mathcal{F}_\infty^\tau$ the σ-field generated by $x(t + \tau)$, $t > 0$. Prove that

$$P_{x,s}(B | \mathcal{F}_\tau^s) = P_{x,s}(B | x(\tau)) \qquad \text{for any} \quad B \in \mathcal{F}_\infty^\tau. \tag{3.8}$$

[*Hint*: Prove that for any bounded measurable function $f(x_1, \ldots, x_n)$,

$$E_{x,s}[f(x(t_1 + \tau), \ldots, x(t_n + \tau)) | \mathcal{F}_\tau^s] = E_{x(\tau),\tau} f(x(t_1 + \tau), \ldots, x(t_n + \tau)) \tag{3.9}$$

where $s \leqslant t_1 < \cdots < t_n$; cf. the proof of Theorem 1.2.]

In Problems 8–16, $\{\Omega, \mathcal{F}, \mathcal{F}_t, x(t), P_x\}$ is a continuous time-homogeneous ν-dimesional Markov process satisfying the Feller property. By Corollary 2.6, the process also satisfies the strong Markov property.

8. Prove that for every open (closed) set G, $p(t, x, G)$ is lower (upper) semicontinuous.

9. Let A be an open set if R^ν with boundary ∂A, and let Γ be a closed subset of ∂A. Assume that there is an increasing sequence of closed subsets Γ'_m of ∂A such that $\Gamma'_m \uparrow \Gamma' = \partial A \setminus \Gamma$ as $m \uparrow \infty$. Assume that $P_x\{x(t)$ hits $\partial A\} = 1$, and let $E = \{x(t)$ hits Γ before it hits $\Gamma'\}$. Then there exists a sequence $f_n(x_1, \ldots, x_n)$ of bounded measurable functions such that

$$f_n(x(t_1), \ldots, x(t_n)) \to \chi_E \quad \text{a.s.},$$

and the functions $f_n(x_1, \ldots, x_n)$ do not depend on the particular process $x(t)$.

[*Hint*: Let $\{\lambda_i\}$ be a dense sequence in $[0, \infty)$. Check that

$$E = \bigcup_j \bigcap_m \left\{ \bigcap_p \bigcup_{t_i < \lambda_j} \left[\rho(x(t_i), \Gamma) < \frac{1}{p} \right] \right\} \cap \left\{ \bigcup_r \bigcap_{t_l < \lambda_j} \left[\rho(x(t_l), \Gamma'_m) \geqslant \frac{1}{r} \right] \right\}$$

where j, m, p, i, r, l vary over $1, 2, \ldots$. Show that $P(E) = \lim P(E_\alpha)$ where E_α is defined as in E but with j, m, p, i, r, l varying only over a finite set. Deduce that $\chi_E = \lim \chi_{E_\alpha}$, $\chi_{E_\alpha} = f_n(x(t_1), \ldots, x(t_n))$ for some $n = n(\alpha)$, and $f_n(x_1, \ldots, x_n)$ does not depend on the particular process $x(t)$.]

10. Let A and B be open sets with boundaries ∂A and ∂B respectively, and assume that $A \cup \partial A$ is contained in B, and that ∂A consists of a finite number of manifolds. Denote by t_A and t_B the exit times from A and B respectively. Suppose that $P_x(t_B < \infty) = 1$ for all $x \in A \cup \partial A$. Prove that for any bounded measurable function f,

$$E_x f(x(t_B)) = E_x E_{x(t_A)} f(x(t_B)) \qquad (x \in A). \tag{3.10}$$

[*Hint*: For time-homogeneous processes, (3.9) becomes

$$E_x[f(x(t_1 + \tau), \ldots, x(t_n + \tau)) | \mathcal{F}_\tau] = E_{x(\tau)} f(x(t_1), \ldots, x(t_n)). \tag{3.11}$$

Suppose $f = \chi_\Gamma$, Γ closed, and introduce the events

$$E = \{x(t) \text{ hits } \Gamma \cap \partial B \text{ before it hits } \partial B \setminus \Gamma\},$$

$$E' = \{x(t + t_A) \text{ hits } \Gamma \cap \partial B \text{ before it hits } \partial B \setminus \Gamma\}.$$

By Problem 9,

$$\chi_E = \lim f_n(x(t_1), \ldots, x(t_n)), \qquad \chi_{E'} = \lim f_n(x(t_1 + t_A), \ldots, x_n(t_n + t_A)),$$

Apply (3.11) to f_n to obtain

$$E_x f(x(t_B)) = E_x \chi_{E'} = E_x E_{x(t_A)} \chi_E = E_x E_{x(t_A)} f(x(t_B)).]$$

11. Let τ be the exit time from an open set G. Suppose $P_x(\tau > \alpha) \leqslant \beta$ for all $x \in G$. Prove that

(i) $P_x(\tau > n\alpha) \leqslant \beta^n$;
(ii) $E_x \tau \leqslant \alpha/(1 - \beta)$;
(iii) $E_x e^{\lambda\tau} \leqslant e^{\lambda\alpha}/(1 - e^{\lambda\alpha}\beta)$ if $\beta < 1, \lambda < (-\alpha^{-1} \log \beta)$.

[*Hint*: Write

$$\chi_{[x(t) \in G \text{ if } t < \alpha]} = \lim f_m(x(t_1), \ldots, x(t_m)) \qquad (t_i < \alpha),$$

$$\chi_{[x(t) \in G \text{ if } t < \alpha + \beta]} = \lim f_m(x(t_1 + \beta), \ldots, x(t_m + \beta))\chi_{[x(t) \in G \text{ if } t < \alpha]}$$

for any $\beta > 0$. Use (3.11) and induction to verify that

$$P_x[\tau > (n+1)\alpha] = E_x \chi_{\tau > \alpha} \chi_{\tau > (n+1)\alpha} = E_x \chi_{\tau > \alpha}(E_x \chi_{\tau > (n+1)\alpha} | \mathcal{F}_\alpha)$$
$$= E_x \chi_{\tau > \alpha} E_{x(\alpha)} \chi_{\tau > n\alpha} \leqslant \beta^{n+1}].$$

12. Suppose $\nu = 1$ and $\tau = \inf\{s, x(s) = y\}$ is finite valued, where y is a real number. Prove, for any closed set A,

$$P[x(t) \in A, \tau \leqslant t] = E_x \chi_{x(t) \in A} \chi_{\tau \leqslant t} = E_x E_x[\chi_{x(t) \in A} \chi_{\tau \leqslant t} | \mathcal{F}_\tau]$$
$$= E_x \chi_{\tau \leqslant t} E_x[\chi_{x(t) \in A} | \mathcal{F}_\tau] = E_x \chi_{\tau \leqslant t} E_{x(\tau)} \chi_{x(t-\tau) \in A}$$
$$= E_x \chi_{\tau \leqslant t} P_{x(\tau)}[x(t - \tau) \in A] = \int_0^t P_x(\tau \in ds) P_y[x(t - s) \in A].$$

[*Hint*: To prove the fourth equality, choose a sequence of functions as in Problem 10 with $E = \{x(t - \tau) \in A\}$.]

13. A point y is called *absorbing* if $p(t, y, \{y\}) = 1$ for all $t \geqslant 0$. It follows that $(T_t f)(y) = f(y)$. Hence, if $f \in D_{\mathcal{A}}$, then $(\mathcal{A} f)(y) = 0$. Show that if y is nonabsorbing and U is a sufficiently small neighborhood of y, τ_U = exit time from U, then $\sup_{x \in U} E_x \tau_U < \infty$.

[*Hint*: $p(t_0, y, N) < 1$ for some $t_0 > 0$ and a closed neighborhood N of y. Hence $p(t_0, x, R^\nu \backslash N) \geqslant \delta > 0$ if $|x - y|$ is sufficiently small, and then $P_x(\tau_U > t_0) \leqslant 1 - \delta$.]

14. Prove that M_0 is a closed subspace of $M(R^\nu)$, that $T_t M_0 \subset M_0$ if $t \geqslant 0$, and that $T_t f$ is a continuous function in $t \in [0, \infty)$ for any $f \in M_0$.

15. From the theory of semigroups (see, for instance, Dynkin [2]) one knows that:

(i) $D_{\mathcal{A}}$ is dense in M_0, $\mathcal{A} f \in M_0$ if $f \in D_{\mathcal{A}}$;

(ii) $T_t D_{\mathfrak{A}} \subset D_{\mathfrak{A}}$;
(iii) if $f \in D_{\mathfrak{A}}$, $dT_t f/dt = T_t \mathfrak{A} f = \mathfrak{A} T_t f$ for $t \geqslant 0$;
(iv) the operator $R_\lambda f = \int_0^\infty e^{-\lambda t}(T_t f)\,dt$ $(f \in M_0)$ is defined for all $\lambda > 0$ and satisfies:
(a) $R_\lambda f \in D_{\mathfrak{A}}$ if $f \in M_0$;
(b) $\|R_\lambda f\| \leqslant \lambda^{-1}\|f\|$;
(c) $(\lambda I - \mathfrak{A})R_\lambda f = f$ if $f \in M_0$, $R_\lambda(\lambda I - \mathfrak{A})f = f$ if $f \in D_{\mathfrak{A}}$.

Now let y be nonabsorbing and take a neighborhood U of y with $\sup_{x \in U} E_x \tau_U < \infty$. Let $h_\lambda = \lambda f - \mathfrak{A} f$. Check that

$$f(x) = R_\lambda h_\lambda(x)$$

$$= E_x \int_0^{\tau_U} e^{-\lambda t} h_\lambda(x(t))\,dt + E_x\left[e^{-\lambda \tau_U} \int_0^\infty e^{-\lambda t} h_\lambda(x(t + \tau_U))\,dt\right].$$

Use the strong Markov property to deduce

$$E_x f(x(\tau_U)) - f(x) = E_x\left[(e^{-\lambda \tau_U} - 1) f(x(\tau_U))\right] - E_x\left[\int_0^{\tau_U} e^{-\lambda t} h_\lambda(x(t))\,dt\right].$$

16. Use Problems 13 and 15 to complete the proof of Dynkin's formula.

3

Brownian Motion

1. Existence of continuous Brownian motion

Definition. A *Brownian motion* or a *Wiener process* is a stochastic process $x(t)$, $t \geqslant 0$ satisfying;

(i) $x(0) = 0$;
(ii) for any $0 \leqslant t_0 < t_1 < \cdots < t_n$, the random variables $x(t_k) - x(t_{k-1})$ $(1 \leqslant k \leqslant n)$ are independent;
(iii) if $0 \leqslant s < t$, $x(t) - x(s)$ is normally distributed with

$$E(x(t) - x(s)) = (t - s)\mu, \qquad E(x(t) - x(s))^2 = (t - s)\sigma^2$$

where μ, σ are real constants, $\sigma \neq 0$.
μ is called the *drift* and σ^2 is called the *variance*.

Property (ii) implies that $x(t) - x(s)$ is independent of $\mathcal{F}(x(\lambda), \lambda \leqslant s)$ or, more generally, that $\mathcal{F}(x(\mu) - x(s), \mu \geqslant s)$ is independent of $\mathcal{F}(x(\lambda), \lambda \leqslant s)$.

If $x(t)$ is a Brownian motion with drift μ and variance σ^2, and if $0 \leqslant t_0 < t_1 < \cdots < t_n$, then

$$\Gamma(t_j, t_k) \equiv E\big(x(t_j) - \mu t_j\big)\big(x(t_k) - \mu t_k\big) = \sigma^2 \min(t_j, t_k). \tag{1.1}$$

Indeed, if $t_j > t_k$,

$$\Gamma(t_j, t_k) = E\big[x(t_j) - x(t_k) - \mu(t_j - t_k) + x(t_k) - \mu t_k\big]\big[x(t_k) - \mu t_k\big]$$
$$= E\big(x(t_k) - \mu t_k\big)^2 = t_k\sigma^2.$$

If $\mu = 0$, $\sigma^2 = 1$, then we speak of *normalized* Brownian motion. Notice that for any Brownian motion with drift μ and variance σ^2, $(x(t) - \mu t)/\sigma$ is a normalized Brownian motion.

From now on we consider only normalized Brownian motion and refer to it briefly as Brownian motion.

Let $X_1, \ldots, X_n$ be random variables with joint normal distribution $N(0, \Gamma)$ where $\Gamma = (\Gamma_{ij})$, $\Gamma_{ij} = \min(t_i, t_j)$, $0 < t_1 < \cdots < t_n$. It is easily checked that $\det \Gamma \neq 0$. Let (Γ_{ij}^{-1}) be the inverse matrix to Γ. Let X_0 be the random variable $X_0 = 0$; its distribution function is

$$X_0(x_0) = \begin{cases} 0 & \text{if } x_0 \leqslant 0, \\ 1 & \text{if } x_0 > 0. \end{cases}$$

Then (see Problem 1) the joint distribution function $F_{t_0 t_1 \cdots t_n}$ of X_0, $X_1, \ldots, X_n$ (where $t_0 = 0$) is given by

$$F_{t_0 t_1 \cdots t_n}(x_0, x_1, \ldots, x_n) = \frac{\chi(x_0)}{(2\pi)^{n/2}(\det \Gamma)^{1/2}} \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} \exp\left[-\tfrac{1}{2}\sum_{i,j=1}^{n} \Gamma_{ij}^{-1} y_i y_j\right] dy_1 \cdots dy_n. \tag{1.2}$$

It follows that

$$F_{t_0 \cdots t_n}(x_0, \ldots, x_n) \uparrow F_{t_0 \cdots t_{k-1} t_{k+1} \cdots t_n}(x_0, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \quad \text{if} \quad x_k \uparrow \infty.$$

Thus, the functions $F_{t_0 t_1 \cdots t_n}$ given by (1.2) form a consistent family of distribution functions. By the Kolmogorov construction there exists a probability space $(\Omega, \mathcal{F}, P)$ and a process $x(t)$ such that the joint distribution functions of $x(t)$ are given by (1.2). But then, by Problem 2, $x(t)$ is a Brownian motion.

We have thus proved that the Kolmogorov construction produces a Brownian motion.

It is easily seen that a Brownian motion is a martingale. Thus the martingale inequality can be applied to Brownian motion. Another important inequality is given in the following theorem.

Theorem 1.1. *Let $x(t)$ be a Brownian motion and let $0 = t_0 < t_1 < \cdots < t_n$. Then, for any $\lambda > 0$,*

$$P\Big[\max_{0 \leqslant j \leqslant n} x(t_j) > \lambda\Big] \leqslant 2P[x(t_n) > \lambda], \tag{1.3}$$

$$P\Big[\max_{0 \leqslant j \leqslant n} |x(t_j)| > \lambda\Big] \leqslant 2P[|x(t_n)| > \lambda]. \tag{1.4}$$

Proof. Let $j^* =$ first j such that $x(t_j) > \lambda$. Since $x(t_n) - x(t_j)$ $(0 \leqslant j < n)$ has normal distribution, its probability distribution is symmetric about the origin, i.e., $P[x(t_n) - x(t_j) > 0] = P[x(t_n) - x(t_j) < 0]$. Using also the facts that $x(t_n) - x(t_j)$ is independent of $\mathcal{F}(x(t_0), \ldots, x(t_j))$ and that $\{j^* = j\}$

belongs to this σ-field, we can write

$$P\Big[\max_{0\leqslant j\leqslant n} x(t_j) > \lambda,\ x(t_n) < \lambda\Big] = \sum_{k=1}^{n-1} P[j^* = k,\ x(t_n) < \lambda]$$
$$\leqslant \sum_{k=1}^{n} P[j^* = k,\ x(t_n) - x(t_k) < 0] = \sum_{k=1}^{n} P(j^* = k)P[x(t_n) - x(t_k) < 0]$$
$$= \sum_{k=1}^{n} P(j^* = k)P[x(t_n) - x(t_k) > 0] = \sum_{k=1}^{n} P[j^* = k,\ x(t_n) - x(t_k) > 0]$$
$$\leqslant \sum_{k=1}^{n} P[j^* = k,\ x(t_n) > \lambda] \leqslant P[x(t_n) > \lambda].$$

On the other hand,

$$P\Big[\max_{0\leqslant j\leqslant n} x(t_j) > \lambda,\ x(t_n) > \lambda\Big] = P[x(t_n) > \lambda].$$

Adding these inequalities, (1.3) follows. Noting that $-x(t)$ is also a Brownian motion and employing (1.3), we get

$$P\Big[\max_{0\leqslant j\leqslant n} |x(t_j)| > \lambda\Big] \leqslant P\Big[\max_{0\leqslant j\leqslant n} x(t_j) > \lambda\Big] + P\Big[\max_{0\leqslant j\leqslant n} (-x(t_j)) > \lambda\Big]$$
$$\leqslant 2P[x(t_n) > \lambda] + 2P[(-x(t_n)) > \lambda] = 2P[|x(t_n)| > \lambda].$$

If X is a random variable with normal distribution $N(0, \sigma^2)$, then

$$EX^{2n} = \frac{(2n)!}{2^n n!}\sigma^{2n}, \qquad EX^{2n+1} = 0 \qquad (n = 0, 1, 2, \ldots). \tag{1.5}$$

In particular, for any Brownian motion $x(t)$,

$$E|x(t) - x(s)|^{2n} = C_n|t - s|^n \qquad (C_n \text{ constant}) \tag{1.6}$$

for $n = 1, 2, \ldots$. Taking $n = 2$ and applying Theorem 1.2.2, we get:

Theorem 1.2. *There is a continuous version of a Brownian motion.*

Corollary 1.3. *Every separable Brownian motion is continuous.*

This follows from Theorem 1.2 and Problem 3, Chapter 1.

From now on, when we speak of a Brownian motion, it is always tacitly assumed that we speak of a continuous version.

From (1.6) we see that Corollary 1.2.3 can be applied to a Brownian motion with $\beta = 2n$, $\alpha = n - 1$, for any $n \geqslant 1$. Consequently:

Corollary 1.4. *For any $\alpha < \frac{1}{2}$, almost all the sample paths of a Brownian motion are Hölder continuous, with exponent α, on every bounded t-interval.*

2. Nondifferentiability of Brownian motion

We have proved the existence of a Hölder continuous Brownian motion $x(t)$, with any exponent $\alpha < \frac{1}{2}$.

Lévy [1] has proved that the sample paths of $x(t)$ satisfy:

$$\varlimsup_{\substack{0 \leq s < t \leq T \\ t-s \downarrow 0}} \frac{|x(t) - x(s)|}{\{2(t-s)\log 1/(t-s)\}^{1/2}} = 1 \quad \text{a.s.,} \qquad \text{for any} \quad T > 0;$$

for proof see also McKean [1].

We shall prove here only a weaker result:

Theorem 2.1. *For any $\alpha > \frac{1}{2}$, almost all simple paths of a Brownian motion are nowhere Hölder continuous with exponent α.*

Proof. Let T be any positive integer and let N be a positive integer such that $N(\alpha - \frac{1}{2}) > 1$. If $x(t, \omega)$ is Hölder continuous at s ($s \in [0, T]$) with exponent α, then

$$|x(t, \omega) - x(s, \omega)| < \beta_0 |t - s|^\alpha \qquad \text{if} \quad |t - s| \leq N/n$$

for some $\beta_0 > 0$ and some positive integer n.

Let β be a positive number and consider the event

$$A_n = \{\text{there exists an } s \in [0, T] \text{ such that } |x(t) - x(s)| < \beta |t - s|^\alpha \\ \text{if } |t - s| \leq N/n\}.$$

Then $A_n \uparrow A$ as $n \uparrow \infty$. If we prove that, for any β, T, $P(A) = 0$, then the assertion of the theorem follows.

Let

$$Z_k = \max_{1 \leq i \leq N} \left| x\left(\frac{k+i}{n}\right) - x\left(\frac{k+i-1}{n}\right) \right| \qquad (k = 0, 1, \ldots, nT),$$

$$B_n = \{\omega;\ Z_k(\omega) \leq 2\beta N^\alpha / n^\alpha \text{ for some } k\}.$$

Note that if $\omega \in A_n$ and $|x(t, \omega) - x(s, \omega)| < \beta |t - s|^\alpha$ when $|t - s| \leq N/n$, and if we take k to be the largest integer such that $k/n \leq s$, then

$$Z_k(\omega) \leq 2\beta (N/n)^\alpha.$$

Consequently $A_n \subset B_n$, so that

$$\begin{aligned} P(A) &= \lim_{n \to \infty} P(A_n) \\ &\leq \varliminf_{n \to \infty} P(B_n). \end{aligned} \tag{2.1}$$

Next,

$$P(B_n) = P\left\{ \bigcup_{k=0}^{nT} \left[Z_k \leqslant \frac{2\beta N^\alpha}{n^\alpha} \right] \right\}$$

$$\leqslant \sum_{k=0}^{nT} P\left[Z_k \leqslant \frac{2\beta N^\alpha}{n^\alpha} \right] = nT\, P\left[Z_0 \leqslant \frac{2\beta N^\alpha}{n^\alpha} \right].$$

Since the random variables

$$x\left(\frac{i}{n}\right) - x\left(\frac{i-1}{n}\right) \qquad (1 \leqslant i \leqslant N)$$

are independent and identically distributed,

$$P(B_n) \leqslant nT\left\{ P\left[\left| x\left(\frac{1}{n}\right) \right| \leqslant \frac{2\beta N^\alpha}{n^\alpha} \right] \right\}^N$$

$$= nT\left\{ \sqrt{\frac{n}{2\pi}} \int_{-\gamma/n^\alpha}^{\gamma/n^\alpha} e^{-nx^2/2}\, dx \right\}^N \qquad (\gamma = 2\beta N^\alpha).$$

Substituting $n^\alpha x = y$, we get

$$P(B_n) = nT\left\{ \frac{1}{\sqrt{2\pi}\, n^{\alpha-1/2}} \int_{-\gamma}^{\gamma} e^{-y^2/(2n^{2\alpha-1})}\, dy \right\}^N \to 0 \qquad \text{if} \quad n \to \infty,$$

since $N(\alpha - 1/2) > 1$. Using (2.1) we then conclude that $P(A) = 0$. This completes the proof.

Corollary 2.2. (i) *Almost all sample paths of a Brownian motion are nowhere differentiable.*

(ii) *Almost all sample paths of a Brownian motion have infinite variation on any finite interval.*

The assertion (i) follows from the fact that if a function is differentiable at a point, then it is Lipschitz continuous at that point. Since a function $f(t)$ with finite variation is almost everywhere differentiable, (ii) is a consequence of (i).

3. Limit theorems

Theorem 3.1. *For a Brownian motion $x(t)$,*

$$\overline{\lim_{t \downarrow 0}} \frac{x(t)}{\sqrt{2t \log\log (1/t)}} = 1 \quad a.s., \tag{3.1}$$

$$\overline{\lim_{t \uparrow \infty}} \frac{x(t)}{\sqrt{2t \log\log t}} = 1 \quad a.s. \tag{3.2}$$

These formulas are called the *laws of the iterated logarithm.*

Proof. We first prove that

$$\overline{\lim_{t \downarrow 0}} \frac{x(t)}{\sqrt{2t \log \log(1/t)}} \leqslant 1. \tag{3.3}$$

Let $\delta > 0$, $\phi(t) = \sqrt{2t \log \log(1/t)}$, and take a sequence $t_n \downarrow 0$. Consider the event

$$A_n = \{x(t) > (1 + \delta)\phi(t) \text{ for at least one } t \in [t_{n+1}, t_n]\}.$$

Since $\phi(t) \uparrow$ if $t \uparrow$,

$$A_n \subset \Big\{ \sup_{0 \leqslant t \leqslant t_n} x(t) > (1 + \delta)\phi(t_{n+1}) \Big\}.$$

By Theorem 1.1, if $x > 0$

$$P\Big[\sup_{0 \leqslant t \leqslant t_n} x(t) > x\sqrt{t_n}\Big] \leqslant 2P\big[x(t_n) > x\sqrt{t_n}\big] = 2P\left[\frac{x(t_n)}{\sqrt{t_n}} > x\right]$$

$$= \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-z^2/2}\, dz \leqslant \sqrt{\frac{2}{\pi}}\, \frac{1}{x} \int_x^\infty z e^{-z^2/2}\, dz = \sqrt{\frac{2}{\pi}}\, \frac{e^{-x^2/2}}{x}.$$

Taking $x = x_n = (1 + \delta)\phi(t_{n+1})/\sqrt{t_n}$, we get

$$P(A_n) \leqslant \sqrt{\frac{2}{\pi}}\, \frac{\exp(-x_n^2/2)}{x_n}.$$

Now take $t_n = q^n$ where $0 < q < 1$ and $\lambda = q(1 + \delta)^2 > 1$. Then

$$x_n = \{2 \log[\alpha(n + 1)^\lambda]\}^{1/2}, \qquad \alpha = \log(1/q).$$

It follows that

$$P(A_n) \leqslant \frac{C}{(n + 1)^\lambda \log(n + 1)} \qquad (C \text{ constant}).$$

Hence $\Sigma\, P(A_n) < \infty$, so that, by the Borel–Cantelli lemma, $P(A_n \text{ i.o.}) = 0$ (where "i.o." stands for infinitely often). But this means that

$$\overline{\lim_{t \downarrow 0}} \frac{x(t)}{\sqrt{2t \log \log(1/t)}} \leqslant 1 + \delta.$$

Since δ is arbitrary, (3.3) follows.

We shall next prove that

$$\overline{\lim_{t \downarrow 0}} \frac{x(t)}{\sqrt{2t \log \log(1/t)}} \geqslant 1. \tag{3.4}$$

We again start with a sequence $t_n \downarrow 0$. Let $Z_n = x(t_n) - x(t_{n+1})$. For any $x > 0$, $\epsilon > 0$,

$$P\left(Z_n > x\sqrt{t_n - t_{n+1}}\right) = P\left(\frac{x(t_n) - x(t_{n+1})}{\sqrt{t_n - t_{n+1}}} > x\right) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-z^2/2}\, dz.$$

By integration by parts,

$$\frac{1}{x} e^{-x^2/2} = \int_x^\infty \left(1 + \frac{1}{z^2}\right) e^{-z^2/2}\, dz \leqslant \left(1 + \frac{1}{x^2}\right) \int_x^\infty e^{-z^2/2}\, dz,$$

so that

$$\int_x^\infty e^{-z^2/2}\, dz \geqslant \left(x + \frac{1}{x}\right)^{-1} e^{-x^2/2}.$$

Thus

$$P\left(Z_n > x\sqrt{t_n - t_{n+1}}\right) \geqslant \frac{1}{2\sqrt{2\pi}\, x} e^{-x^2/2} \qquad \text{if} \quad x > 1.$$

Taking $t_n = q^n$, $0 < q < 1$,

$$x = x_n = (1 - \epsilon) \frac{\phi(t_n)}{\sqrt{t_n - t_{n+1}}} = \frac{1 - \epsilon}{\sqrt{1 - q}} \sqrt{2 \log[n \log(1/q)]} = \sqrt{\beta \log(\alpha n)}$$

where $\alpha = \log(1/q)$, $\beta = 2(1 - \epsilon)^2/(1 - q)$, and choosing q sufficiently small so that $\beta < 2$, we get

$$P[Z_n > (1 - \epsilon)\phi(t_n)] \geqslant \frac{c}{n^{\beta/2}\sqrt{\log n}} \qquad (c \text{ positive constant})$$

and, consequently,

$$\sum_{n=1}^\infty P[Z_n > (1 - \epsilon)\phi(t_n)] = \infty.$$

Since the events $[Z_n > (1 - \epsilon)\phi(t_n)]$ are independent, the Borel–Cantelli lemma implies that

$$P[Z_n > (1 - \epsilon)\phi(t_n) \text{ i.o.}] = 1. \tag{3.5}$$

By the proof of (3.1) applied to the Brownian motion $-x(t)$,

$$P[x(t_{n+1}) < -(1 + \epsilon)\phi(t_{n+1}) \text{ i.o.}] = 0.$$

Putting it together with (3.5), we deduce that a.s.

$$\begin{aligned} x(t_n) &= Z_n + x(t_{n+1}) > (1 - \epsilon)\phi(t_n) - (1 + \epsilon)\phi(t_{n+1}) \\ &= \phi(t_n)\left[1 - \epsilon - (1 + \epsilon)\frac{\phi(t_{n+1})}{\phi(t_n)}\right] \quad \text{i.o.} \end{aligned}$$

For any given $\delta > 0$, if we choose ϵ and q so small that

$$1 - \epsilon - (1 + \epsilon)\sqrt{q} > 1 - \delta,$$

and note that $\phi(t_{n+1})/\phi(t_n) \to \sqrt{q}$ if $n \to \infty$, we get that

$$P[x(t_n) > (1 - \delta)\phi(t_n) \text{ i.o.}] = 1.$$

Since δ is arbitrary, this completes the proof of (3.4).

In order to prove (3.2), consider the process

$$\tilde{x}(t) = \begin{cases} tx\left(\dfrac{1}{t}\right) & \text{if } t > 0, \\ 0 & \text{if } t = 0. \end{cases}$$

It is easy to check that $\tilde{x}(t)$ satisfies the conditions (i)–(iii) in the definition of a Brownian motion. Since it is clearly also a separable process, it is a (continuous) Brownian motion. If we now apply (3.1) to $\tilde{x}(t)$, then we obtain (3.2).

Corollary 3.2. *For a Brownian motion $x(t)$,*

$$\varliminf_{t \downarrow 0} \frac{x(t)}{\sqrt{2t \log \log(1/t)}} = -1 \quad \text{a.s.,} \tag{3.6}$$

$$\varliminf_{t \uparrow \infty} \frac{x(t)}{\sqrt{2t \log \log t}} = -1 \quad \text{a.s.} \tag{3.7}$$

This follows by applying Theorem 3.1 to the Brownian motion $-x(t)$.

If $f(t)$ is a continuously differentiable function in an interval $[a, b]$, and if

$$\Pi_n = \{t_{n,1}, \ldots, t_{n,m_n}\} \tag{3.8}$$

is a sequence of partitions of $[a, b]$ with mesh $|\Pi_n| = \max(t_{n,j} - t_{n,j-1})$ converging to 0, then

$$\sum_{j=1}^{m_n} [f(t_{n,j}) - f(t_{n,j-1})]^2 \to 0 \quad \text{if} \quad n \to \infty.$$

Such a conclusion does not hold for a Brownian motion. Instead we have:

Theorem 3.3. *Let $x(t)$ be a Brownian motion, and Π_n a sequence of partitions (3.8) of a finite closed interval $[a, b]$ with mesh $|\Pi_n| \to 0$ if $n \to \infty$. Let*

$$S_n = \sum_{j=1}^{m_n} [x(t_{n,j}) - x(t_{n,j-1})]^2.$$

Then $S_n \to b - a$ in the mean.

Proof. Write $t_j = t_{n,j}$, $m = m_n$. Then

$$S_n - (b-a) = \sum_{j=1}^{m} \left[\left(x(t_j) - x(t_{j-1})\right)^2 - (t_j - t_{j-1}) \right].$$

Since the summands are independent and of mean zero,

$$E[S_n - (b-a)]^2 = E \sum_{j=1}^{n} \left[\left(x(t_j) - x(t_{j-1})\right)^2 - (t_j - t_{j-1}) \right]^2$$

$$= \sum_{j=1}^{m} E\left[\left(Y_j^2 - 1\right)(t_j - t_{j-1}) \right]^2.$$

where

$$Y_j = \frac{x(t_j) - x(t_{j-1})}{(t_j - t_{j-1})^{1/2}} .$$

Since the Y_j are equally distributed with normal distribution,

$$E[S_n - (b-a)]^2 = E\left(Y_1^2 - 1\right)^2 \sum_{j=1}^{m} (t_j - t_{j-1})^2$$

$$\leqslant E\left(Y_1^2 - 1\right)^2 \cdot (b-a)|\Pi_n| \to 0 \qquad \text{if} \quad n \to \infty.$$

4. Brownian motion after a stopping time

Let τ be a stopping time for a Brownian motion $x(t)$. Denote by $\mathcal{F}_\tau$ the σ-field of all events A such that

$$A \cap (\tau \leqslant t) \qquad \text{is in} \qquad \mathcal{F}(x(\lambda), 0 \leqslant \lambda \leqslant t).$$

From the considerations of Section 1.3 we know that $x(\tau)$ is a random variable.

Theorem 4.1. *If τ is a stopping time for a Brownian motion $x(t)$, then the process*

$$y(t) = x(t + \tau) - x(\tau), \qquad t \geqslant 0$$

is a Brownian motion, and $\mathcal{F}(y(t), t \geqslant 0)$ is independent of $\mathcal{F}_\tau$.

Thus the assertion is that a Brownian motion starts afresh at any stopping time.

Proof. Notice that if $\tau = s$ (a constant), then the assertion is obvious. Suppose now that the range of τ is a countable set $\{s_j\}$. Let $B \in \mathcal{F}_\tau$ and

$0 \leqslant t_1 < t_2 < \cdots < t_n$. Then, for any Borel sets $A_1, \ldots, A_n$,

$$\begin{aligned} &P[y(t_1) \in A_1, \ldots, y(t_n) \in A_n, B] \\ &\quad = \sum_k P[y(t_1) \in A_1, \ldots, y(t_n) \in A_n, \tau = s_k, B] \\ &\quad = \sum_k P[(x(t_1 + s_k) - x(s_k)) \in A_1, \ldots, (x(t_n + s_k) - x(s_k)) \\ &\qquad\qquad \in A_n, \tau = s_k, B]. \end{aligned} \tag{4.1}$$

Since

$$(\tau = s_k) \cap B = [B \cap (\tau \leqslant s_k)] \cap (\tau = s_k) \quad \text{is in} \quad \mathcal{F}(x(\lambda), 0 \leqslant \lambda \leqslant s_k)$$

and $\mathcal{F}(x(\lambda + s_k) - x(s_k)), \lambda \geqslant 0)$ is independent of $\mathcal{F}(x(\lambda), 0 \leqslant \lambda \leqslant s_k)$, and since the assertion of the theorem is true if $\tau \equiv s_k$, the kth term on the right-hand side of (4.1) is equal to

$$\begin{aligned} &P[(x(t_1 + s_k) - x(s_k)) \in A_1, \ldots, (x(t_n + s_k) - x(s_k)) \in A_k]P[\tau = s_k, B] \\ &\quad = P[x(t_1) \in A_1, \ldots, x(t_n) \in A_n]P[\tau = s_k, B]. \end{aligned}$$

Summing over k, we get

$$P[y(t_1) \in A_1, \ldots, y(t_n) \in A_n, B] = P[x(t_1) \in A_1, \ldots, x(t_n) \in A_n]P(B). \tag{4.2}$$

Taking $B = \Omega$, it follows that the joint distribution functions of the process $y(t)$ are the same as those for $x(t)$. Since $y(t)$ is clearly a continuous process, it is a (continuous) Brownian motion.

From (4.2) we further deduce that the σ-field $\mathcal{F}(y(t), t \geqslant 0)$ is independent of $\mathcal{F}_\tau$.

In order to prove Theorem 4.1 for a general stopping time, we approximate τ by a sequence of stopping times τ_n, defined in Lemma 1.3.4. We have (see the first paragraph of the proof of Theorem 2.2.4)

$$\mathcal{F}_\tau \subset \mathcal{F}_{\tau_n}.$$

Set $y_n(t) = x(t + \tau_n) - x(\tau_n)$. By what we have already proved, if $B \in \mathcal{F}_\tau$, then

$$P[y_n(t_1) < x_1, \ldots, y_n(t_k) < x_k, B] = P[x(t_1) < x_1, \ldots, x(t_k) < x_k]P(B)$$

for any $0 \leqslant t_1 < \cdots < t_k$. Notice that $y_n(t) \to y(t)$ a.s. for all $t \geqslant 0$, as $n \to \infty$. Hence, if $(x_1, \ldots, x_k)$ is a point of continuity of the k-dimensional distribution function $F_k(x_1, \ldots, x_k)$ of $(x(t_1), \ldots, x(t_k))$, then

$$P[y(t_1) < x_1, \ldots, y(t_k) < x_k, B] = P[x(t_1) < x_1, \ldots, x(t_k) < x_k]P(B). \tag{4.3}$$

Since (see Problem 1) $F_k(x_1, \ldots, x_k)$ is actually an integral

$$\int_{-\infty}^{x_k} \cdots \int_{-\infty}^{x_1} \rho(z_1, \ldots, z_k)\, dz_1 \cdots dz_k,$$

it is continuous everywhere. Thus (4.3) holds for all $x_1, \ldots, x_k$. This implies that

$$P[y(t_1) \in A_1, \ldots, y(t_k) \in A_k, B] = P[x(t_1) \in A_1, \ldots, x(t_k) \in A_k] P(B)$$

for any Borel sets $A_1, \ldots, A_k$, and the proof of the theorem readily follows.

5. Martingales and Brownian motion

If $x(t)$ is a Brownian motion and $\mathscr{F}_t = \mathscr{F}(x(\lambda), 0 \leq \lambda \leq t)$, then

$$E[(x(t) - x(s)) | \mathscr{F}_s] = 0, \tag{5.1}$$

$$E\big[(x(t) - x(s))^2 | \mathscr{F}_s\big] = t - s \tag{5.2}$$

a.s. for any $0 \leq s < t$. Note that (5.1), (5.2) hold if and only if $x(t)$ and $x^2(t) - t$ are martingales.

We shall now prove the converse.

Theorem 5.1. *Let $x(t)$, $t \geq 0$ be a continuous process and let $\mathscr{F}_t$ $(t \geq 0)$ be an increasing family of σ-fields such that $x(t)$ is $\mathscr{F}_t$ measurable and* (5.1), (5.2) *hold a.s. for all* $0 \leq s < t$. *Then $x(t)$ is a Brownian motion.*

Proof. For any $\epsilon > 0$ and a positive integer n, let τ_0 be the first value of t such that

$$\max_{\substack{|s'-s''| \leq 1/n \\ 0 \leq s', s'' \leq t}} |x(s') - x(s'')| = \epsilon;$$

if no such t exists, set $\tau_0 = \infty$. Let $\tau = \tau_0 \wedge 1$. Since, for $0 < s < 1$,

$$\{\tau > s\} = \bigcup_{m=1}^{\infty} \bigcap_{\substack{|s'-s''| \leq 1/n \\ 0 < s', s'' \leq s}} \left\{ |x(s') - x(s'')| \leq \epsilon - \frac{1}{m} \right\} \quad \text{is in } \mathscr{F}_s,$$

τ is a stopping time.

Let $y(t) = x(t \wedge \tau)$. Since $x(t)$ and $x^2(t) - t$ are continuous martingales, Theorem 1.3.5 implies that $y(t)$ and $y^2(t) - t$ are also continuous martingales. Hence,

$$\int_A \big[y^2(s) - s \wedge \tau \big]\, dP = \int_A \big[y^2(t) - t \wedge \tau \big]\, dP \qquad \text{if } A \in \tilde{\mathscr{F}}_s, \quad s < t$$

where $\tilde{\mathfrak{F}}_s = \mathfrak{F}(y(\lambda), 0 \leq \lambda \leq s)$. It follows that

$$\int_A \{E[y^2(t) - y^2(s)] | \tilde{\mathfrak{F}}_s\}\, dP = \int_A [y^2(t) - y^2(s)]\, dP$$
$$= \int_A [t \wedge \tau - s \wedge \tau]\, dP \leq \int_A (t - s)\, dP.$$

Since the integrand on the left is $\tilde{\mathfrak{F}}_s$ measurable, we conclude that

$$E\{[y^2(t) - y^2(s)] | \tilde{\mathfrak{F}}_s\} \leq t - s \quad \text{a.s.}, \qquad s < t. \tag{5.3}$$

Our aim is to prove that if $0 \leq t_0 < t_1 < \cdots < t_k$, then

$$E \exp\left\{ i \sum_{j=1}^{k} \lambda_j [x(t_j) - x(t_{j-1})] \right\} = \exp\left[-\sum_{j=1}^{k} \tfrac{1}{2} \lambda_j^2 (t_j - t_{j-1}) \right] \tag{5.4}$$

for any real numbers λ_j. In view of the result stated in Problem 1(a), this will show that $x(t)$ is a Brownian motion.

We begin with the special case $k = 1$, $t_0 = 0$, $t_1 = 1$. Thus, we want to prove that

$$Ee^{i\lambda x(1)} = e^{-\lambda^2/2}. \tag{5.5}$$

Our approach will be to write

$$x(1) = \sum_{j=1}^{n} \left[x\left(\frac{j}{n}\right) - x\left(\frac{j-1}{n}\right) \right]$$

and make use of (5.1), (5.2) while at the same time, letting n increase to infinity. However, a technical difficulty arises due to the fact that the random variables $x(j/n) - x((j-1)/n)$ are not uniformly bounded. It is for this reason that we shall work, instead, with the process $y(t)$ and with the random variables

$$\zeta_j = y\left(\frac{j}{n}\right) - y\left(\frac{j-1}{n}\right) \qquad (1 \leq j \leq n).$$

In view of the definitions of τ and $y(t)$,

$$|\zeta_j| \leq \epsilon. \tag{5.6}$$

Since $y(t)$ is a martingale and since $y^2(t)$ satisfies (5.3),

$$E\zeta_1 = 0, \qquad E\{\zeta_j | \zeta_1, \ldots, \zeta_{j-1}\} = 0 \qquad (2 \leq j \leq n),$$
$$\sigma_1^2 = E\zeta_1^2 \leq \frac{1}{n}, \qquad \sigma_j^2 = E\{\zeta_j^2 | \zeta_1, \ldots, \zeta_{j-1}\} \leq \frac{1}{n} \qquad (2 \leq j \leq n) \tag{5.7}$$

a.s. Hence, if $j \geqslant 1$,

$$\begin{aligned} Ee^{i\lambda y(j/n)} &= E\left\{e^{i\lambda y((j-1)/n)}E\left[e^{i\lambda\zeta_j}|\zeta_1, \ldots, \zeta_{j-1}\right]\right\} \\ &= Ee^{i\lambda y((j-1)/n)}\left[1 - \sigma_j^2 \frac{\lambda^2}{2}(1 + o(1))\right] \\ &= E\exp\left[i\lambda y\left(\frac{j-1}{n}\right) - \sigma_j^2 \frac{\lambda^2}{2}(1 + o(1))\right] \end{aligned}$$

where, in view of (5.6), $o(1) \to 0$ if $\epsilon \to 0$, uniformly with respect to λ, n provided λ varies in a bounded interval.

It follows that

$$\begin{aligned} &\left|E\exp\left[i\lambda y\left(\frac{j}{n}\right)\right] - E\exp\left[i\lambda y\left(\frac{j-1}{n}\right) - \frac{\lambda^2}{2n}\right]\right| \\ &= \left|E\left\{\left\{\exp\left[i\lambda y\left(\frac{j-1}{n}\right) - \frac{\lambda^2}{2n}\right]\right\}\right.\right. \\ &\qquad \left.\left.\times\left\{\exp\left[-\frac{\lambda^2}{2}\left(\sigma_j^2(1 + o(1)) - \frac{1}{n}\right)\right] - 1\right\}\right\}\right| \\ &\leqslant E\left\{\left|\exp\left[\frac{\lambda^2}{2}\left(\frac{1}{n} - \sigma_j^2\right) + \sigma_j^2 o(1)\right] - 1\right|\right\} \\ &\leqslant C\,E\left[\left(\frac{1}{n} - \sigma_j^2\right) + \frac{o(1)}{n}\right] \\ &\leqslant C\left[\frac{1}{n} - E(\zeta_j^2)\right] + \frac{o(1)}{n}; \end{aligned}$$

we have used here (5.7). The symbol C denotes any one of various different positive constants which do not depend on n, ϵ, λ, provided λ is restricted to bounded intervals, $\epsilon < 1$, $n \geqslant 1$.

The last estimate implies that

$$\begin{aligned} &\left|E\exp\left[i\lambda y\left(\frac{j}{n}\right) + \frac{j}{2n}\lambda^2\right] - E\exp\left[i\lambda y\left(\frac{j-1}{n}\right) + \frac{j-1}{2n}\lambda^2\right]\right| \\ &\leqslant C\left[\frac{1}{n} - E(\zeta_j^2)\right] + \frac{o(1)}{n}. \end{aligned}$$

Summing over j and noting that

$$1 > \sum_{j=1}^{n} E(\zeta_j^2) = E(y^2(1)), \tag{5.8}$$

we get

$$|E[e^{i\lambda y(1)+\lambda^2/2} - 1]| \leq C\big[1 - E\big(y^2(1)\big)\big] + o(1).$$

Consequently,

$$|E[e^{i\lambda y(1)} - e^{-\lambda^2/2}]| \leq C\big[1 - E\big(y^2(1)\big)\big] + o(1). \tag{5.9}$$

Since $x(t)$ is a continuous process,

$$P(\tau = 1) \to 1 \quad \text{and} \quad P[y(1) = x(1)] \to 1 \qquad \text{if} \quad n \to \infty.$$

Hence, by the Lebesgue bounded convergence theorem,

$$Ee^{i\lambda y(1)} \to Ee^{i\lambda x(1)} \qquad \text{if} \quad n \to \infty.$$

Therefore from (5.9) it follows that, for any $\gamma > 0$, if $n \geq n_0(\epsilon, \gamma)$, then

$$|E[e^{i\lambda x(1)} - e^{-\lambda^2/2}]| \leq C\big[1 - E\big(y^2(1)\big)\big] + o(1) + \gamma. \tag{5.10}$$

Note next that

$$1 - E(y^2(1)) = E[x^2(1) - y^2(1)] = \int_{\tau<1} [x^2(1) - y^2(\tau)]\, dP$$
$$\leq \int_{\tau<1} x^2(1)\, dP \to 0$$

if $P(\tau = 1) \to 1$. Thus, if $n \geq n_1(\epsilon, \gamma)$,

$$C\big[1 - E\big(y^2(1)\big)\big] \leq \gamma.$$

Using this in (5.10) we get

$$|Ee^{i\lambda x(1)} - e^{-\lambda^2/2}| \leq o(1) + 2\gamma.$$

Taking $\gamma \to 0$ (and $n \geq n_0(\epsilon, \gamma)$, $n \geq n_1(\epsilon, \gamma)$) and then $\epsilon \to 0$, the assertion (5.5) follows.

Similarly one can prove that

$$E \exp[i\lambda_j(x(t_j) - x(t_{j-1}))] = \exp[-\tfrac{1}{2}\lambda_j^2] \tag{5.11}$$

for any $0 \leq t_{j-1} < t_j$, λ_j real.

By reviewing the proof of (5.11) one finds that the same proof with minor changes actually gives a better result, namely,

$$E\left\{\exp[i\lambda_j(x(t_j) - x(t_{j-1}))] \,|\, \mathcal{F}(x(\lambda), 0 \leq \lambda \leq t_{j-1}\right\} = \exp\big[-\tfrac{1}{2}\lambda_j^2\big] \quad \text{a.s.} \tag{5.12}$$

(Notice that if τ_{j-1} is defined analogously to τ, with $t = 0$ replaced by $t = t_{j-1}$, and if $y(t) = x(t \wedge \tau_{j-1})$, then

$$\mathcal{F}(y(\lambda), 0 \leq \lambda \leq t_{j-1}) = \mathcal{F}(x(\lambda), 0 \leq \lambda \leq t_{j-1})$$

since $\tau_{j-1} > t_{j-1}$.)

Now let $0 \leq t_0 < \cdots < t_{k-1} < t_k$ and let $\lambda_1, \ldots, \lambda_k$ be real numbers.

Applying (5.12) successively, we find that

$$
\begin{aligned}
E \exp&\left[i \sum_{j=1}^{k} \lambda_j (x(t_j) - x(t_{j-1}))\right] \\
&= E\left\{ \exp\left[i \sum_{j=1}^{k-1} \lambda_j (x(t_j) - x(t_{j-1}))\right] \right. \\
&\quad \left. \times E\left\{ \exp[i\lambda_k(x(t_k) - x(t_{k-1}))] \,|\, \mathcal{F}(x(\lambda), 0 \leqslant \lambda \leqslant t_{k-1})\right\}\right\} \\
&= E\left\{ \exp\left[i \sum_{j=1}^{k-1} \lambda_j (x(t_j) - x(t_{j-1}))\right]\right\} \exp\left[-\tfrac{1}{2}\lambda_k^2 (t_k - t_{k-1})\right] \\
&= \cdots = \exp\left[-\sum_{j=1}^{k} \tfrac{1}{2}\lambda_j^2 (t_j - t_{j-1})\right].
\end{aligned}
$$

This completes the proof of (5.4).

Remark. By the remark at the end of Section 1.3 we have that $y(t)$ and $y^2(t) - t$ (introduced in the last proof) are martingales with respect to $\mathcal{F}_t$. Hence instead of (5.3) we actually have

$$E\{ y^2(t) - y^2(s) \,|\, \mathcal{F}_s\} \leqslant t - s \quad \text{a.s.}, \qquad s < t.$$

This implies

$$E\{\exp[i\lambda(x(t) - x(s))] \,|\, \mathcal{F}_s\} = \exp\left[-\tfrac{1}{2}\lambda^2 (t - s)\right]. \tag{5.13}$$

6. Brownian motion in n dimensions

Definition. An n-dimensional process $w(t) = (w_1(t), \ldots, w_n(t))$ is called an n-dimensional *Brownian motion* (or *Wiener process*) if each process $w_i(t)$ is a Brownian motion and if the σ-fields $\mathcal{F}(w_i(t), t \geqslant 0)$, $1 \leqslant i \leqslant n$, are independent.

Theorem 6.1. *Let $w(t)$ be an n-dimensional Brownian motion. Then*

$$P \varlimsup_{t \downarrow 0} \frac{|w(t)|}{\sqrt{2t \log \log(1/t)}} = 1 \quad \text{a.s.}, \tag{6.1}$$

$$P \varlimsup_{t \uparrow \infty} \frac{|w(t)|}{\sqrt{2t \log \log t}} = 1 \quad \text{a.s.} \tag{6.2}$$

Proof. It is enough to prove (6.1). If (6.1) is false, then there is a $\delta > 0$ such that

$$\overline{\lim_{t \downarrow 0}} \frac{|w(t)|}{\sqrt{2t \log \log(1/t)}} \geqslant 1 + \delta \quad \text{with positive probability.}$$

Given any $\epsilon > 0$, cover the unit sphere by a finite number of spherical regions K_i with opening ϵ with respect to the origin. Then, for some i,

$$P\left\{ \overline{\lim_{t \downarrow 0}} \frac{|w(t)|}{\sqrt{2t \log \log(1/t)}} \geqslant 1 + \delta, \frac{w(t)}{|w(t)|} \in K_i \right\} > 0.$$

But if z is the center of K_i and $(1 + \delta) \cos \epsilon > 1$, then

$$P\left\{ \overline{\lim_{t \to 0}} \frac{z \cdot w(t)}{\sqrt{2t \log \log(1/t)}} \geqslant (1 + \delta) \cos \epsilon > 1 \right\} > 0.$$

This is impossible since $z \cdot w(t)$ is a Brownian motion (cf. Problem 7).

Theorem 4.1 and its proof extend immediately to n-dimensional Brownian motion.

We shall now extend Theorem 5.1 to n-dimensions.

Theorem 6.2. *Let $x(t) = (x_1(t), \ldots, x_n(t))$, $t \geqslant 0$ be a continuous, n-dimensional process and let $\mathcal{F}_t$ $(t \geqslant 0)$ be an increasing family of σ-fields such that $x(t)$ is $\mathcal{F}_t$ measurable and, for all $0 \leqslant s < t < \infty$,*

$$E[x(t) - x(s) | \mathcal{F}_s] = 0 \quad \text{a.s.},$$

$$E[(x_i(t) - x_i(s))(x_j(t) - x_j(s))] = \delta_{ij}(t - s) \quad \text{a.s.}$$

where $\delta_{ij} = 0$ if $i \neq j$, $\delta_{ii} = 1$. Then $x(t)$ is a Brownian motion.

Proof. For any $\gamma \in R^n$, $|\gamma| = 1$, the conditions of Theorem 5.1 are satisfied for $\gamma \cdot x(t)$. Hence $\gamma \cdot x(t)$ is a Brownian motion. Thus, in particular, each $x_i(t)$ is a Brownian motion. It remains to show that these motions are mutually independent. For this it suffices to show that $\mathcal{F}(x_i(t), t \geqslant 0)$ is independent of $\mathcal{F}(x_j(t), t \geqslant 0)$ if $i \neq j$. Take, for simplicity, $i = 1$, $j = 2$.

Since $\gamma_1 x_1 + \gamma_i x_2$ is a Brownian motion if $\gamma_1^2 + \gamma_2^2 = 1$,

$$E[\gamma_1 x_1(t) + \gamma_2 x_2(t)]^2 = t.$$

Since also $Ex_i^2(t) = t$,

$$Ex_1(t)x_2(t) = 0.$$

But then

$$Ex_1(t + s)x_2(t) = E\{x_2(t)E[x_1(t + s)|\mathcal{F}_t]\} = Ex_2(t)x_1(t) = 0,$$

i.e., $x_1(t+s)$ is independent of $x_2(t)$. Similarly $x_2(t+s)$ is independent of $x_1(t)$. Since t and s are arbitrary nonnegative numbers, it follows that $\mathcal{F}(x_1(\lambda), \lambda \geqslant 0)$ is independent of $\mathcal{F}(x_2(\lambda), \lambda \geqslant 0)$. This completes the proof.

One can easily check that an n-dimensional Brownian motion satisfies the Markov property with the stationary transition probability function

$$p(t, x, A) = \int_A \frac{1}{(2\pi t)^{n/2}} \exp\left[-\frac{|x-y|^2}{2t}\right] dy. \tag{6.3}$$

By the method of Section 2.1 we can construct a time-homogeneous Markov process corresponding to the transition probability (6.3). The sample space consists of all R^n-valued functions on $[0, \infty)$. But we can also construct another model $\{\Omega_0, \mathfrak{M}, \mathfrak{M}_t, \xi(t), P_x\}$ where Ω_0 is the space of all continuous functions $x(\cdot)$ from $[0, \infty)$ into R^n, $\mathfrak{M}_t$ is the smallest σ-field such that the process

$$\xi(s, x(\cdot)) = x(s) \tag{6.4}$$

is measurable for all $0 \leqslant s \leqslant t$, and

$$\begin{aligned} P_x\{x(\cdot); \xi(t_1) \in A_1, \ldots, \xi(t_k) \in A_k\} \\ = P\{\omega; x + w(t_1) \in A_1, \ldots, x + w(t_k) \in A_k\}. \end{aligned} \tag{6.5}$$

Theorem 6.3. *$\{\Omega_0, \mathfrak{M}, \mathfrak{M}_t, \xi(t), P_x\}$ is a time-homogeneous Markov process with the transition probability function given by* (6.3).

Proof. For any set B in $\mathfrak{M}$, define

$$P_x(B) = P\{\omega; x + w(\cdot, \omega) \text{ is in } B\}.$$

It is easily seen that P_x is a measure on $(\Omega_0, \mathfrak{M})$, satisfying (6.5). The rest follows from Theorem 2.1.3.

To any continuous ν-dimensional Markov process $\{\Omega^*, \mathcal{F}^*, \mathcal{F}^{*s}_t, x^*(t), P^*_{x,s}\}$ we can correspond a Markov process $\{\Omega, \mathfrak{M}, \mathfrak{M}^s_t, \xi(t), P_{x,s}\}$ having the same transition probability function p. Ω is the space of all continuous functions $x(\cdot)$ from $0 \leqslant t < \infty$ into R^ν, $\mathfrak{M}^s_t$ is the smallest σ-field such that $x(\lambda)$ is measurable for any $s \leqslant \lambda \leqslant t$, $\mathfrak{M} = \mathfrak{M}^0_\infty$, $\xi(t, x(\cdot)) = x(t)$ and, finally, if $s \leqslant t_1 < \cdots < t_k$,

$$P_{x,s}\{x(\cdot); \xi(t_1) \in A_1, \ldots, \xi(t_k) \in A_k\} = P^*_{x,s}\{\omega; x^*(t_1) \in A_1, \ldots, x^*(t_k) \in A_k\}. \tag{6.6}$$

In fact, for any set B in $\mathfrak{M}$, define

$$P_{x,s}(B) = P^*_{x,s}\{\omega; x^*(\cdot, \omega) \in B\}.$$

It is easily seen that $P_{x,s}$ is a measure on $(\Omega, \mathfrak{M})$, and it satisfies (6.6). In view of Theorem 2.1.3, $\{\Omega, \mathfrak{M}, \mathfrak{M}_t^s, \xi(t), P_{x,s}\}$ is then a Markov process with the same transition probability function as for $\{\Omega^*, \mathcal{F}^*, \mathcal{F}_t^{*s}, x^*(t), P_{x,s}^*\}$.

PROBLEMS

1. Recall that n random variables $X_1, \ldots, X_n$ are said to have *joint normal distribution* (μ, Γ), where $\mu = (\mu_1, \ldots, \mu_n)$, $\Gamma = (\Gamma_{ij})_{i,j=1}^n$ symmetric, if the characteristic function

$$f(u) = E \exp\left[i \sum_{j=1}^{n} u_j X_j\right] \qquad (u = (u_1, \ldots, u_n))$$

has the form

$$f(u) = \exp\left[\sum_{j=1}^{n} \mu_j u_j - \tfrac{1}{2} \sum_{j,k=1}^{n} \Gamma_{jk} u_j u_k\right].$$

Prove that if this is the case then:

(a) The random variables $X_1, \ldots, X_n$ are independent if and only if $\Gamma_{ij} = 0$ whenever $i \neq j$.

(b) If X_i, X_j are independent for any pair (i, j), $i \neq j$, then $X_1, \ldots, X_n$ are mutually independent.

(c) If $Y_i = \sum_{j=1}^{n} a_{ij} X_j$ $(1 \leqslant i \leqslant m)$ where a_{ij} are constants, then $Y_1, \ldots, Y_n$ have joint normal distribution.

(d) $\mu_i = EX_i$, $\Gamma_{ij} = E(X_i - \mu_i)(X_j - \mu_j)$.

(e) If $\det \Gamma \neq 0$, then $(X_1 - \mu_1, \ldots, X_n - \mu_n)$ has a distribution function $\rho(y)$ given by

$$\rho(y) = \frac{1}{(2\pi)^{n/2} (\det \Gamma)^{1/2}} \exp\left[-\tfrac{1}{2} \sum_{i,j=1}^{n} \Gamma_{ij}^{-1} y_i y_j\right]$$

where (Γ_{ij}^{-1}) is the inverse matrix to Γ.

2. If $x(t)$, $t \geqslant 0$ is a process with $x(0) = 0$ such that for any $0 \leqslant t_0 < t_1 < \cdots < t_n$ the random variables $x(t_0), \ldots, x(t_n)$ have joint normal distribution $N(\hat{\mu}, \Gamma)$ where $\hat{\mu} = (\mu, \ldots, \mu)$, $\Gamma = (\Gamma_{jk})$, $\Gamma_{jk} = \sigma^2 \min(t_j, t_k)$, then $x(t)$ is a Brownian motion with drift μ and variance σ^2.

3. Verify (1.5) for a random variable X with normal distribution $N(0, \sigma^2)$.

4. Give a direct proof of (3.2). [*Hint*: Apply the method of proof of (3.1) with $q > 1$.]

5. If $x(t)$ is a Brownian motion, $[x(t)/t] \to 0$ a.s. as $t \to \infty$.

6. Let $x(t)$ be a Brownian motion and let I be an interval in $[0, \infty)$. Then for any $\delta > 0$, $P\{x(n\delta) \in I \text{ i.o.}\} = 1$.

7. If $\tilde{w}_i(t) = \sum_{j=1}^{n} a_{ij} w_j(t)$ where (a_{ij}) is an orthogonal matrix and

$(w_1(t), \ldots, w_n(t))$ is an n-dimensional Brownian motion, then $(\tilde{w}_1, \ldots, \tilde{w}_n)$ is also an n-dimensional Brownian motion.

8. If $w(t)$ is a Brownian motion then $\alpha w(t/\alpha^2)$ is also a Brownian motion, for any positive constant α.

9. If $w(t)$ is a Brownian motion and $z(t) = \exp[\alpha w(t) - \alpha^2 t/2]$, α positive constant, then

(i) $Ez(t) = 1$.

(ii) $z(t)$ is a martingale; in fact, if $\mathcal{F}_s = \mathcal{F}(w(\lambda), 0 \leqslant \lambda \leqslant s)$, then $E(z(t)|\mathcal{F}_s) = z(s)$ if $t > s$.

(iii) $P\{\max_{0 \leqslant s \leqslant t}[w(s) - \alpha s/2] > \beta \leqslant e^{-\alpha\beta}$ if $\alpha \geqslant 0$, $\beta > 0$.

10. Let $u_n(t)$ be Brownian motion defined on the same probability space, and suppose that $u_n(t) \xrightarrow{P} u(t)$ for each $t \geqslant 0$, as $n \to \infty$. Prove that $u(t)$ is a Brownian motion.

11. If $w(t)$ is a Brownian motion and A is a positive constant, then, for any $T > 0$,

$$Ee^{A|w(t)|} \leqslant C \qquad \text{if} \quad 0 \leqslant t \leqslant T,$$

where C is a positive constant depending on A, T.

12. Let $\mathcal{F}_t$, $t \geqslant 0$ be an increasing family of σ-fields and let $x(t)$ be an n-dimensional process such that $x(t)$ is $\mathcal{F}_t$ measurable and $E[x(t+s)|\mathcal{F}_t] = x(t)$ for all $t \geqslant 0$, $s \geqslant 0$. Prove that if $\gamma \cdot x(t)$ is a Brownian motion for any $\gamma \in R^n$, $|\gamma| = 1$, then $x(t)$ is an n-dimensional Brownian motion. [*Hint*: Cf. the proof of Theorem 6.2.]

4

The Stochastic Integral

In this chapter we shall define the integral

$$I(T) = \int_0^T f(t)\, dw(t)$$

where $w(t)$ is a Brownian motion and $f(t)$ is a stochastic function, and study its basic properties. One may define

$$I(T) = f(T)w(T) - \int_0^T f'(t)w(t)\, dt$$

if f is absolutely continuous for each ω. However, if f is only continuous, or just integrable, this definition does not make sense.

Since $w(t)$ is nowhere differentiable with probability 1, the integral $\int_0^T f(t)\, dw(t)$ cannot be defined in the usual Lebesgue–Stieltjes sense.

1. Approximation of functions by step functions

We shall call a stochastic process also a stochastic function or, briefly, a function.

Let $w(t)$, $t \geqslant 0$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{F}_t$ $(t \geqslant 0)$ be an increasing family of σ-fields, i.e., $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ if $t_1 < t_2$, such that $\mathcal{F}_t \subset \mathcal{F}$, $\mathcal{F}(w(s), 0 \leqslant s \leqslant t)$ is in $\mathcal{F}_t$, and

$$\mathcal{F}(w(\lambda + t) - w(t), \lambda \geqslant 0) \qquad \text{is independent of } \mathcal{F}_t$$

for all $t > 0$. One can take, for instance, $\mathcal{F}_t = \mathcal{F}(w(s), 0 \leqslant s \leqslant t)$.

Let $0 \leqslant \alpha < \beta < \infty$. A stochastic process $f(t)$ defined for $\alpha \leqslant t \leqslant \beta$ is called a *nonanticipative* function with respect to $\mathcal{F}_t$ if:

(i) $f(t)$ is a separable process;

(ii) $f(t)$ is a measurable process, i.e., the function $(t, \omega) \to f(t, \omega)$ from $[\alpha, \beta] \times \Omega$ into R^1 is measurable;
(iii) for each $t \in [\alpha, \beta]$, $f(t)$ is $\mathcal{F}_t$ measurable.

When (iii) holds we say that $f(t)$ is *adapted* to $\mathcal{F}_t$. We denote by $L_w^p[\alpha, \beta]$ $(1 \leqslant p \leqslant \infty)$ the class of all nonanticipative functions $f(t)$ satisfying:

$$P\left\{ \int_\alpha^\beta |f(t)|^p \, dt < \infty \right\} = 1 \qquad \left(P\left\{ \operatorname*{ess\,sup}_{\alpha \leqslant t \leqslant \beta} |f(t)| < \infty \right\} = 1 \text{ if } p = \infty \right).$$

We denote by $M_w^p[\alpha, \beta]$ the subset of $L_w^p[\alpha, \beta]$ consisting of all functions f with

$$E \int_\alpha^\beta |f(t)|^p \, dt < \infty \qquad \left(E\left[\operatorname*{ess\,sup}_{\alpha \leqslant t \leqslant \beta} |f(t)| \right] < \infty \text{ if } p = \infty \right).$$

Definition. A stochastic process $f(t)$ defined on $[\alpha, \beta]$ is called a *step function* if there exists a partition $\alpha = t_0 < t_1 < \cdots < t_r = \beta$ of $[\alpha, \beta]$ such that

$$f(t) = f(t_i) \qquad \text{if} \quad t_i \leqslant t < t_{i+1}, \quad 0 \leqslant i \leqslant r-1.$$

Lemma 1.1. *Let $f \in L_w^2[\alpha, \beta]$. Then:*

(i) *there exists a sequence of continuous functions g_n in $L_w^2[\alpha, \beta]$ such that*

$$\lim_{n \to \infty} \int_\alpha^\beta |f(t) - g_n(t)|^2 \, dt = 0 \quad a.s.; \tag{1.1}$$

(ii) *there exists a sequence of step functions f_n in $L_w^2[\alpha, \beta]$ such that*

$$\lim_{n \to \infty} \int_\alpha^\beta |f(t) - f_n(t)|^2 \, dt = 0 \quad a.s. \tag{1.2}$$

Proof. Let

$$\rho(t) = \begin{cases} c \exp\left[-1/(1 - t^2) \right] & \text{if} \quad |t| \leqslant 1, \\ 0 & \text{if} \quad |t| > 1 \end{cases}$$

where c is a positive constant determined by the condition $\int_{-\infty}^{\infty} \rho(t) \, dt = 1$.
Define $f(t) = 0$ if $t < \alpha$ and let

$$(J_\epsilon f)(t) = \frac{1}{\epsilon} \int_{\alpha-1}^\beta \rho\left(\frac{t - s - \epsilon}{\epsilon} \right) f(s) \, ds \qquad (2\epsilon < 1). \tag{1.3}$$

Then $J_\epsilon f$ is clearly continuous, and

$$(J_\epsilon f)(t) = \frac{1}{\epsilon}\int_{t-2\epsilon}^{t} \rho\left(\frac{t-s-\epsilon}{\epsilon}\right) f(s)\,ds = \int_{-\epsilon}^{\epsilon} \rho(z) f(t-z-\epsilon)\,dz. \tag{1.4}$$

By Schwarz's inequality,

$$\int_\alpha^\beta (J_\epsilon f)^2\,dt \leqslant \int_\alpha^\beta \left\{ \int_{-\epsilon}^{\epsilon} \rho(z)\,dz \int_{-\epsilon}^{\epsilon} \rho(z) f^2(t-z-\epsilon)\,dz \right\} dt$$
$$\leqslant \int_{-\epsilon}^{\epsilon} \rho(z) \left\{ \int_{\alpha-1}^{\beta} f^2(t)\,dt \right\} dz.$$

Hence

$$\int_\alpha^\beta (J_\epsilon f)^2\,dt \leqslant \int_\alpha^\beta f^2(t)\,dt. \tag{1.5}$$

For fixed ω for which $\int_\alpha^\beta f^2(t,\omega)\,dt < \infty$, let u_n be nonrandom continuous functions such that $u_n(t) = 0$ if $t < \alpha$ and

$$\int_\alpha^\beta |u_n(t) - f(t,\omega)|^2\,dt \to 0 \quad \text{if} \quad n \to \infty. \tag{1.6}$$

Since u_n is continuous, it is clear that

$$(J_\epsilon u_n)(t) \to u_n(t) \quad \text{uniformly in} \quad t \in [\alpha, \beta], \qquad \text{as} \quad \epsilon \to 0.$$

Writing

$$\int_\alpha^\beta |(J_\epsilon f)(t,\omega) - f(t,\omega)|^2\,dt \leqslant \int_\alpha^\beta |J_\epsilon(f(\cdot,\omega) - u_n(\cdot))(t)|^2\,dt$$
$$+ \int_\alpha^\beta |(J_\epsilon u_n)(t) - u_n(t)|^2\,dt + \int_\alpha^\beta |u_n(t) - f(t,\omega)|^2\,dt$$

and using (1.5) with f replaced by $f - u_n$, we get, after taking $\epsilon \to 0$,

$$\overline{\lim_{\epsilon \downarrow 0}} \int_\alpha^\beta |(J_\epsilon f)(t,\omega) - f(t,\omega)|^2\,dt \leqslant 2\int_\alpha^\beta |u_n(t) - f(t,\omega)|^2\,dt.$$

Taking $n \to \infty$ and using (1.6), we obtain

$$\overline{\lim_{\epsilon \downarrow 0}} \int_\alpha^\beta |(J_\epsilon f)(t) - f(t)|^2\,dt = 0 \quad \text{a.s.}$$

Since the integrand on the right-side of (1.4) is a separable process that is $\mathfrak{F}_t$ measurable, the integral is also $\mathfrak{F}_t$ measurable. Consequently the assertion (i) holds with $g_n = J_{1/n}f$.

To prove (ii), let

$$h_{n,m}(t) = g_n\left(\frac{k}{m}\right) \quad \text{if } \alpha + \frac{k}{m} \leqslant t < \alpha + \frac{k+1}{m} \qquad (0 \leqslant k < m(\beta - \alpha)).$$

Then

$$\lim_{m\to\infty} \int_\alpha^\beta |h_{n,m}(t) - g_n(t)|^2\, dt = 0 \quad \text{a.s.} \tag{1.7}$$

Now, for any $\delta > 0$,

$$P\left\{ \int_\alpha^\beta |f(t) - g_n(t)|^2\, dt > \frac{\delta}{2} \right\} < \frac{\delta}{2} \qquad \text{for some} \quad n = n_0.$$

From (1.7) with $n = n_0$ we get

$$P\left\{ \int_\alpha^\beta |g_{n_0}(t) - h_{n_0,m}(t)|^2\, dt > \frac{\delta}{2} \right\} < \frac{\delta}{2} \qquad \text{for some} \quad m = m_0.$$

Hence

$$P\left\{ \int_\alpha^\beta |f(t) - h_{n_0,m_0}(t)|^2\, dt > \delta \right\} < \delta.$$

Taking $\delta = 1/k$ and denoting the corresponding h_{n_0,m_0} by ψ_k, it follows that $\psi_k \in L_w^2[\alpha, \beta]$ and

$$\int_\alpha^\beta |f(t) - \psi_k(t)|^2\, dt \xrightarrow{P} 0.$$

But then there is a subsequence $\{f_n\}$ of $\{\psi_k\}$ satisfying the assertion (ii).

Lemma 1.2. *Let* $f \in M_w^2[\alpha, \beta]$. *Then:*

(i) *there exists a sequence of continuous functions* k_n *in* $M_w^2[\alpha, \beta]$ *such that*

$$E \int_\alpha^\beta |f(t) - k_n(t)|^2\, dt \to 0 \qquad \text{if} \quad n \to \infty; \tag{1.8}$$

(ii) *there exists a sequence of bounded step functions* l_n *in* $M_w^2[\alpha, \beta]$ *such that*

$$E \int_\alpha^\beta |f(t) - l_n(t)|^2\, dt \to 0 \qquad \text{if} \quad n \to \infty. \tag{1.9}$$

Proof. Let g_n be as in Lemma 1.1. For any $N > 0$, let

$$\phi_N(t) = \begin{cases} t & \text{if } |t| \leqslant N, \\ Nt/|t| & \text{if } |t| > N. \end{cases}$$

Notice that $|\phi_N(t)-\phi_N(s)| \leqslant |t-s|$. Therefore

$$\int_\alpha^\beta |\phi_N(f(t)) - \phi_N(g_n(t))|^2\,dt \leqslant \int_\alpha^\beta |f(t)-g_n(t)|^2\,dt \to 0 \quad \text{a.s.}$$

Since

$$\int_\alpha^\beta |\phi_N(f(t)) - \phi_N(g_n(t))|^2\,dt \leqslant 4N^2(\beta-\alpha),$$

the Lebesgue bounded convergence theorem gives

$$E\int_\alpha^\beta |\phi_N(f(t)) - \phi_N(g_n(t))|^2\,dt \to 0 \qquad \text{if} \quad n\to\infty. \tag{1.10}$$

Next,

$$E\int_\alpha^\beta |\phi_N(f(t)) - f(t)|^2\,dt \leqslant 4 \iint_{\{|f|>N\}} |f(t)|^2\,dt\,dP \to 0$$

as $N\to\infty$, since $f\in M_w^2[\alpha,\beta]$. From this and (1.10) it follows that for every positive integer k there are $N=N(k)$ and $n=n(k,N)$ such that

$$E\int_\alpha^\beta |\phi_N(g_n(t)) - f(t)|^2\,dt < \frac{1}{k}\,.$$

Taking $h_k=\phi_N(g_n)$ where $N=N(k)$, $n=n(N,k)$ and noting that $h_k(t)$ is nonanticipative, the assertion (i) follows.

The proof of (ii) is similar. The l_n are of the form $\phi_N(f_n)$ where the f_n are as in Lemma 1.1. Notice that these are in fact step functions.

2. Definition of the stochastic integral

Definition. Let $f(t)$ be a step function in $L_w^2[\alpha,\beta]$, say $f(t)=f_i$ if $t_i\leqslant t<t_{i+1}$, $0\leqslant i\leqslant r-1$ where $\alpha=t_0<t_1<\cdots<t_r=\beta$. The random variable

$$\sum_{k=0}^{r-1} f(t_k)[w(t_{k+1})-w(t_k)]$$

is denoted by

$$\int_\alpha^\beta f(t)\,dw(t)$$

and is called the *stochastic integral* of f with respect to the Brownian motion w; it is also called the *Itô integral*.

Lemma 2.1. *Let f_1, f_2 be two step functions in $L_w^2[\alpha, \beta]$ and let λ_1, λ_2 be real numbers. Then $\lambda_1 f_1 + \lambda_2 f_2$ is in $L_w^2[\alpha, \beta]$ and*

$$\int_\alpha^\beta [\lambda_1 f_1(t) + \lambda_2 f_2(t)]\, dw(t) = \lambda_1 \int_\alpha^\beta f_1(t)\, dw(t) + \lambda_2 \int_\alpha^\beta f_2(t)\, dw(t). \tag{2.1}$$

The proof is left to the reader.

Lemma 2.2. *If f is a step function in $M_w^2[\alpha, \beta]$, then*

$$E \int_\alpha^\beta f(t)\, dw(t) = 0, \tag{2.2}$$

$$E\left| \int_\alpha^\beta f(t)\, dw(t) \right|^2 = E \int_\alpha^\beta f^2(t)\, dt. \tag{2.3}$$

Proof. Since

$$E \int_\alpha^\beta f^2(t)\, dt = \sum_{i=0}^{r-1} Ef^2(t_i)(t_{i+1} - t_i) \tag{2.4}$$

is finite, by assumption, we deduce that $Ef^2(t_i) < \infty$. In particular, $E|f(t_i)| < \infty$. Also, $E|w(t_{i+1}) - w(t_i)| < \infty$. But since $f(t_i)$ is $\mathcal{F}_{t_i}$ measurable whereas $w(t_{i+1}) - w(t_i)$ is independent of $\mathcal{F}_{t_i}$,

$$Ef(t_i)(w(t_{i+1}) - w(t_i)) = Ef(t_i)E(w(t_{i+1}) - w(t_i)) = 0.$$

Summing over i, (2.2) follows.

Next, since $f^2(t_i)$ and $(w(t_{i+1}) - w(t_i))^2$ are independent and have finite expectation, also $f^2(t_i)(w(t_{i+1}) - w(t_i))^2$ has finite expectation. By Schwarz's inequality it follows that

$$E|f(t_k)f(t_i)(w(t_{i+1}) - w(t_i))| < \infty.$$

If $k > i$, then $w(t_{k+1}) - w(t_k)$ is independent of $f(t_k)f(t_i)(w(t_{i+1}) - w(t_i))$. In view of the last inequality and the finiteness of $E|w(t_{k+1}) - w(t_k)|$, we deduce that

$$Ef(t_k)f(t_i)(w(t_{i+1}) - w(t_i))(w(t_{k+1}) - w(t_k)) = 0.$$

Hence

$$E\left|\int_\alpha^\beta f(t)\,dw(t)\right|^2 = \sum_{i=0}^{r-1} Ef^2(t_i)(w(t_{i+1}) - w(t_i))^2$$
$$= \sum_{i=0}^{r-1} Ef^2(t_i)E(w(t_{i+1}) - w(t_i))^2$$
$$= \sum_{i=0}^{r-1} Ef^2(t_i)(t_{i+1} - t_i) = E\int_\alpha^\beta f^2(t)\,dt$$

by (2.4), and (2.3) is proved.

Lemma 2.3. *For any step function f in $L_w^2[\alpha, \beta]$ and for any $\epsilon > 0$, $N > 0$,*

$$P\left\{\left|\int_\alpha^\beta f(t)\,dw(t)\right| > \epsilon\right\} \leq P\left\{\int_\alpha^\beta f^2(t)\,dt > N\right\} + \frac{N}{\epsilon^2}. \tag{2.5}$$

Proof. Let

$$\phi_N(t) = \begin{cases} f(t) & \text{if } t_k \leq t < t_{k+1} \text{ and } \sum_{j=0}^{k} f^2(t_j)(t_{j+1} - t_j) \leq N, \\ 0 & \text{if } t_k \leq t < t_{k+1} \text{ and } \sum_{j=0}^{k} f^2(t_j)(t_{j+1} - t_j) > N, \end{cases}$$

where $f(t) = f(t_j)$ if $t_j \leq t < t_{j+1}$; $t_0 = \alpha < t_1 < \cdots < t_r = \beta$. Then $\phi_N \in L_w^2[\alpha, \beta]$ and

$$\int_\alpha^\beta \phi_N^2(t)\,dt = \sum_{j=0}^{\nu} f^2(t_j)(t_{j+1} - t_j)$$

where ν is the largest integer such that

$$\sum_{j=0}^{\nu} f^2(t_j)(t_{j+1} - t_j) \leq N, \qquad \nu \leq r - 1.$$

Hence

$$E\int_\alpha^\beta \phi_N^2(t)\,dt \leq N.$$

Further, $f(t) - \phi_N(t) = 0$ for all $t \in [\alpha, \beta)$ if $\int_\alpha^\beta f^2(t)\,dt < N$. Therefore

$$P\left\{\left|\int_\alpha^\beta f(t)\,dw(t)\right| > \epsilon\right\} \leq P\left\{\left|\int_\alpha^\beta \phi_N(t)\,dw(t)\right| > \epsilon\right\} + P\left\{\int_\alpha^\beta f^2(t)\,dt > N\right\}.$$

Since, by Chebyshev's inequality, the first integral on the right is bounded by

$$\frac{1}{\epsilon^2} E\left| \int_\alpha^\beta \phi_N(t)\, dw(t) \right|^2 \leq \frac{N}{\epsilon^2},$$

the assertion (2.5) follows.

We shall now proceed to define the stochastic integral for any function f in $L_w^2[\alpha, \beta]$.

By Lemma 1.1 there is a sequence of step functions f_n in $L_w^2[\alpha, \beta]$ such that

$$\int_\alpha^\beta |f_n(t) - f(t)|^2\, dt \xrightarrow{P} 0 \quad \text{if} \quad n \to \infty. \tag{2.6}$$

Hence

$$\lim_{n,m\to\infty} P \int_\alpha^\beta |f_n(t) - f_m(t)|^2\, dt \xrightarrow{P} 0.$$

By Lemma 2.3, for any $\epsilon > 0$, $\rho > 0$,

$$P\left\{ \left| \int_\alpha^\beta f_n(t)\, dw(t) - \int_\alpha^\beta f_m(t)\, dw(t) \right| > \epsilon \right\}$$
$$\leq \rho + P\left\{ \int_\alpha^\beta |f_n(t) - f_m(t)|^2\, dt > \epsilon^2 \rho \right\}.$$

It follows that the sequence

$$\left\{ \int_\alpha^\beta f_n(t)\, dw(t) \right\}$$

is convergent in probability. We denote the limit by

$$\int_\alpha^\beta f(t)\, dw(t)$$

and call it the *stochastic integral* (or the *Itô integral*) of $f(t)$ with respect to the Brownian motion $w(t)$.

The above definition is independent of the particular sequence $\{f_n\}$. For if $\{g_n\}$ is another sequence of step functions in $L_w^2[\alpha, \beta]$ converging to f in the sense that

$$\int_\alpha^\beta |g_n(t) - f(t)|^2\, dt \xrightarrow{P} 0,$$

then the sequence $\{h_n\}$ where $h_{2n} = f_n$, $h_{2n+1} = g_n$ is also convergent to f in the same sense. But then, by what we have proved, the sequence

$$\left\{ \int_\alpha^\beta h_n(t)\, dw(t) \right\}$$

is convergent in probability. It follows that the limits (in probability) of $\int_\alpha^\beta f_n\,dw$ and of $\int_\alpha^\beta g_n\,dw$ are equal a.s.

Lemmas 2.1–2.3 extend to any functions from $L_w^2[\alpha, \beta]$:

Theorem 2.4. *Let f_1, f_2 be functions from $L_w^2[\alpha, \beta]$ and let λ_1, λ_2 be real numbers. Then $\lambda_1 f_1 + \lambda_2 f_2$ is in $L_w^2[\alpha, \beta]$ and*

$$\int_\alpha^\beta [\lambda_1 f_1(t) + \lambda_2 f_2(t)]\,dw(t) = \lambda_1 \int_\alpha^\beta f_1(t)\,dw(t) + \lambda_2 \int_\alpha^\beta f_2(t)\,dw(t). \tag{2.7}$$

Theorem 2.5. *If f is a function in $M_w^2[\alpha, \beta]$, then*

$$E\int_\alpha^\beta f(t)\,dw(t) = 0, \tag{2.8}$$

$$E\left|\int_\alpha^\beta f(t)\,dw(t)\right|^2 = E\int_\alpha^\beta f^2(t)\,dt. \tag{2.9}$$

Theorem 2.6. *If f is a function from $L_w^2[\alpha, \beta]$, then, for any $\epsilon > 0$, $N > 0$,*

$$P\left\{\left|\int_\alpha^\beta f(t)\,dw(t)\right| > \epsilon\right\} \leqslant P\left\{\int_\alpha^\beta f^2(t)\,dt > N\right\} + \frac{N}{\epsilon^2}. \tag{2.10}$$

The proof of Theorem 2.4 is left to the reader.

Proof of Theorem 2.5. By Lemma 1.2 there exists a sequence of step functions f_n in $M_w^2[\alpha, \beta]$ such that

$$E\int_\alpha^\beta |f_n(t) - f(t)|^2\,dt \to 0 \quad \text{if} \quad n \to \infty.$$

This implies that

$$E\int_\alpha^\beta f_n^2(t)\,dt \to E\int_\alpha^\beta f^2(t)\,dt. \tag{2.11}$$

By Lemma 2.2,

$$E\int_\alpha^\beta f_n(t)\,dw(t) = 0, \tag{2.12}$$

$$E\left|\int_\alpha^\beta f_n(t)\,dw(t)\right|^2 = E\int_\alpha^\beta f_n^2(t)\,dt, \tag{2.13}$$

$$E\left|\int_\alpha^\beta f_n(t)\,dw(t) - \int_\alpha^\beta f_m(t)\,dw(t)\right|^2 = E\int_\alpha^\beta |f_n(t) - f_m(t)|^2\,dt \to 0$$
$$\text{if} \quad n, m \to \infty. \tag{2.14}$$

From the definition of the stochastic integral,

$$\int_\alpha^\beta f_n(t)\,dw(t) \xrightarrow{P} \int_\alpha^\beta f(t)\,dw(t).$$

Using (2.14) we conclude that actually

$$\int_\alpha^\beta f_n(t)\,dw(t) \to \int_\alpha^\beta f(t)\,dw(t) \qquad \text{in} \quad L^2(\Omega).$$

Hence, in particular,

$$E \int_\alpha^\beta f(t)\,dw(t) = \lim_{n\to\infty} E \int_\alpha^\beta f_n(t)\,dw(t),$$

$$E\left|\int_\alpha^\beta f(t)\,dw(t)\right|^2 = \lim_{n\to\infty} E\left|\int_\alpha^\beta f_n(t)\,dw(t)\right|^2,$$

and using (2.12) and (2.13), (2.11), the assertions (2.8), (2.9) follow.

Proof of Theorem 2.6. By Lemma 1.1 there exists a sequence of step functions f_n in $L_w^2[\alpha, \beta]$ such that

$$\int_\alpha^\beta |f_n(t) - f(t)|^2\,dt \xrightarrow{P} 0. \tag{2.15}$$

By definition of the stochastic integral,

$$\int_\alpha^\beta f_n(t)\,dw(t) \xrightarrow{P} \int_\alpha^\beta f(t)\,dw(t). \tag{2.16}$$

Applying Lemma 2.3 to f_n we have

$$P\left\{\left|\int_\alpha^\beta f_n(t)\,dw(t)\right| > \epsilon'\right\} \leqslant P\left\{\int_\alpha^\beta f_n^2(t)\,dt > N'\right\} + \frac{N'}{(\epsilon')^2}.$$

Taking $n \to \infty$ and using (2.15), (2.16), we get

$$P\left\{\left|\int_\alpha^\beta f(t)\,dw(t)\right| > \epsilon\right\} \leqslant P\left\{\int_\alpha^\beta f^2(t)\,dt > N\right\} + \frac{N'}{(\epsilon')^2}$$

for any $\epsilon > \epsilon'$, $N < N'$. Taking $\epsilon' \uparrow \epsilon$, $N' \downarrow N$, (2.10) follows.

Theorem 2.7. *Let f, f_n be in $L_w^2[\alpha, \beta]$ and suppose that*

$$\int_\alpha^\beta |f_n(t) - f(t)|^2\,dt \xrightarrow{P} 0 \qquad \textit{as} \quad n \to \infty. \tag{2.17}$$

Then

$$\int_\alpha^\beta f_n(t)\,dw(t) \xrightarrow{P} \int_\alpha^\beta f(t)\,dw(t) \qquad \textit{as} \quad n \to \infty. \tag{2.18}$$

Proof. By Theorem 2.6, for any $\epsilon > 0$, $\rho > 0$,

$$P\left\{\left|\int_\alpha^\beta (f_n(t) - f(t))\,dw(t)\right| > \epsilon\right\} \leq P\left\{\int_\alpha^\beta |f_n(t) - f(t)|^2\,dt > \epsilon^2\rho\right\} + \rho.$$

Taking $n \to \infty$ and using (2.17), the assertion (2.18) follows.

The next theorem improves Theorem 2.5.

Theorem 2.8. *Let $f \in M_w^2[\alpha, \beta]$. Then*

$$E\left\{\int_\alpha^\beta f(t)\,dw(t)\Big|\mathcal{F}_\alpha\right\} = 0, \tag{2.19}$$

$$E\left\{\left|\int_\alpha^\beta f(t)\,dw(t)\right|^2\Big|\mathcal{F}_\alpha\right\} = E\left\{\int_\alpha^\beta f^2(t)\,dt\Big|\mathcal{F}_\alpha\right\} = \int_\alpha^\beta E[f^2(t)|\mathcal{F}_\alpha]\,dt. \tag{2.20}$$

We first need a simple lemma.

Lemma 2.9. *If $f \in L_w^2[\alpha, \beta]$ and ζ is a bounded and $\mathcal{F}_\alpha$ measurable function, then ζf is in $L_w^2[\alpha, \beta]$ and*

$$\int_\alpha^\beta \zeta f(t)\,dw(t) = \zeta \int_\alpha^\beta f(t)\,dw(t). \tag{2.21}$$

Proof. It is clear that ζf is in $L_w^2[\alpha, \beta]$. If f is a step function, then (2.21) follows from the definition of the stochastic integral. For general f in $L_w^2[\alpha, \beta]$, let f_n be step functions in $L_w^2[\alpha, \beta]$ satisfying (2.17). Applying (2.21) to each f_n and taking $n \to \infty$, the assertion (2.21) follows.

Proof of Theorem 2.8. Let ζ be a bounded and $\mathcal{F}_\alpha$ measurable function. Then ζf belongs to $M_w^2[\alpha, \beta]$ and, by Theorem 2.6,

$$E\int_\alpha^\beta \zeta f(t)\,dw(t) = 0.$$

Hence, by (2.21),

$$E\left\{\zeta \int_\alpha^\beta f(t)\,dw(t)\right\} = 0,$$

i.e.,

$$E\left\{\zeta E\left[\int_\alpha^\beta f(t)\,dw(t)|\mathcal{F}_\alpha\right]\right\} = 0.$$

This implies (2.19). The proof of (2.20) is similar.

Theorem 2.10. *If $f \in L^2_w[\alpha, \beta]$ and f is continuous, then, for any sequence Π_n of partitions $\alpha = t_{n,0} < t_{n,1} < \cdots < t_{n,m_n} = \beta$ of $[\alpha, \beta]$ with mesh $|\Pi_n| \to 0$,*

$$\sum_{k=0}^{m_n - 1} f(t_{n,k})[w(t_{n,k+1}) - w(t_{n,k})] \xrightarrow{P} \int_\alpha^\beta f(t)\, dw(t) \qquad \text{as} \quad n \to \infty. \tag{2.22}$$

Proof. Introduce the step functions g_n:

$$g_n(t) = f(t_{n,k}) \qquad \text{if} \quad t_{n,k} \leqslant t < t_{n,k+1}, \quad 0 \leqslant k \leqslant m_n - 1.$$

For a.a. ω, $g_n(t) \to f(t)$ uniformly in $t \in [\alpha, \beta)$ as $n \to \infty$. Hence

$$\int_\alpha^\beta |g_n(t) - f(t)|^2\, dt \to 0 \quad \text{a.s.}$$

By Theorem 2.7 we then have

$$\int_\alpha^\beta g_n(t)\, dw(t) \xrightarrow{P} \int_\alpha^\beta f(t)\, dw(t).$$

Since

$$\int_\alpha^\beta g_n(t)\, dw(t) = \sum_{k=0}^{m_n - 1} f(t_{n,k})[w(t_{n,k+1}) - w(t_{n,k})],$$

the assertion (2.22) follows.

Lemma 2.11. *Let f, g belong to $L^2_w[\alpha, \beta]$ and assume that $f(t) = g(t)$ for all $\alpha \leqslant t \leqslant \beta$, $\omega \in \Omega_0$. Then*

$$\int_\alpha^\beta f(t)\, dw(t) = \int_\alpha^\beta g(t)\, dw(t) \qquad \text{for a.a. } \omega \in \Omega_0. \tag{2.23}$$

Proof. Let ψ_k be the step function in $L^2_w[\alpha, \beta]$ constructed in the proof of Lemma 1.1, satisfying

$$\int_\alpha^\beta |f(t) - \psi_k(t)|^2\, dt \xrightarrow{P} 0.$$

Similarly let ϕ_k be step functions in $L^2_w[\alpha, \beta]$ satisfying

$$\int_\alpha^\beta |g(t) - \phi_k(t)|^2\, dt \xrightarrow{P} 0.$$

From the construction in Lemma 1.1 we deduce that we can choose the sequences ϕ_k, ψ_k so that, if $\omega \in \Omega_0$, $\phi_k(t, \omega) = \psi_k(t, \omega)$ for $\alpha \leqslant t \leqslant \beta$. Hence, by the definition ff the integral of a step function,

$$\int_\alpha^\beta \psi_k(t)\, dw(t) = \int_\alpha^\beta \phi_k(t)\, dw(t) \qquad \text{if} \quad \omega \in \Omega_0.$$

Taking $k \to \infty$, the assertion (2.23) follows.

3. The Indefinite Integral

Let $f \in L_w^2[0, T]$ and consider the integral

$$I(t) = \int_0^t f(s)\, dw(s), \qquad 0 \leqslant t \leqslant T \tag{3.1}$$

where, by definition, $\int_0^0 f(s)\, dw(s) = 0$. We refer to $I(t)$ as the indefinite integral of f. Notice that $I(t)$ is $\mathcal{F}_t$ measurable.

If f is a step function, then clearly

$$\int_\alpha^\beta f(s)\, dw(s) + \int_\beta^\gamma f(s)\, dw(s) = \int_\alpha^\gamma f(s)\, dw(s) \quad \text{if} \quad 0 \leqslant \alpha < \beta < \gamma \leqslant T. \tag{3.2}$$

By approximation we find that (3.2) holds for any f in $L_w^2[0, T]$.

Theorem 3.1. *If $f \in M_w^2[0, T]$, then the indefinite integral $I(t)$ $(0 \leqslant t \leqslant T)$ is a martingale.*

Proof. Let $0 \leqslant t' < t \leqslant T$. By (3.2) and Theorem 2.8,

$$E(I(t) | \mathcal{F}_{t'}) = E(I(t') | \mathcal{F}_{t'}) + E\Big(\int_{t'}^t f(s)\, ds | \mathcal{F}_{t'}\Big) = I(t').$$

Theorem 3.2. *If $f \in L_w^2[0, T]$, then the indefinite integral $I(t)$ $(0 \leqslant t \leqslant T)$ has a continuous version.*

Proof. Suppose first that $f \in M_w^2[0, T]$. Let $\{f_n\}$ be a sequence of step functions in $M_w^2[0, T]$ such that

$$E \int_0^T (f(t) - f_n(t))^2\, dt \to 0 \qquad \text{if} \quad n \to \infty. \tag{3.3}$$

Notice that the indefinite integrals

$$I_n(t) = \int_0^t f_n(s)\, dw(s) \qquad (0 \leqslant t \leqslant T)$$

are continuous functions. By Theorem 2.8, $I_n(t) - I_m(t)$ is a martingale, for each n, m. Hence, by the martingale inequality (Corollary 1.3.3)

$$P\Big\{\sup_{0\leq t\leq T} |I_n(t) - I_m(t)| > \epsilon\Big\}$$
$$\leq \frac{1}{\epsilon^2} E\left| \int_0^T (f_n(s) - f_m(s))\, dw(s)\right|^2$$
$$\leq \frac{1}{\epsilon^2} E \int_0^T |f_n(s) - f_m(s)|^2\, ds \to 0 \qquad \text{if} \quad n, m\to\infty.$$

Taking $\epsilon = 1/2^k$ it follows that for some n_k sufficiently large,

$$P\Big\{\sup_{0\leq t\leq T} |I_{n_k}(t) - I_m(t)| > \frac{1}{2^k}\Big\} < \frac{1}{k^2} \qquad \text{if} \quad m \geq n_k.$$

We can choose the n_k in such a way that $n_k \uparrow$ if $k \uparrow$. Hence,

$$P\Big\{\sup_{0\leq t\leq T} |I_{n_k}(t) - I_{n_{k+1}}(t)| > \frac{1}{2^k}\Big\} < \frac{1}{k^2} \qquad (k = 1, 2, \ldots).$$

Since $\sum k^{-2} < \infty$, the Borel–Cantelli lemma inplies that

$$P\Big\{\sup_{0\leq t\leq T} |I_{n_k}(t) - I_{n_{k+1}}(t)| > \frac{1}{2^k} \text{ i.o.}\Big\} = 0,$$

i.e., for a.a. ω

$$|I_{n_k}(t) - I_{n_{k+1}}(t)| \leq \frac{1}{2^k} \qquad \text{for all} \quad 0 \leq t \leq T, \quad \text{if} \quad k \geq k_0(\omega).$$

But then, with probability one, $\{I_{n_k}(t)\}$ is uniformly convergent in $t\in[0, T]$. The limit $J(t)$ is therefore a continuous function in $t\in[0, T]$ for a.a. ω. Since (3.3) implies that

$$\int_0^t f_n(s)\, dw(s) \to \int_0^t f(s)\, dw(s) \qquad \text{in } L^2(\Omega),$$

it follows that

$$J(t) = \int_0^t f(s)\, dw(s) \quad \text{a.s.}$$

Thus, the indefinite integral has a continuous version.

Consider now the general case where $f\in L_w^2[0, T]$. For any $N > 0$, let

$$\chi_N(z) = \begin{cases} 1 & \text{if} \quad z \leq N, \\ 0 & \text{if} \quad z > N, \end{cases} \tag{3.4}$$

and introduce the function

$$f_N(t) = f(t)\chi_N\Big(\int_0^t f^2(s)\, ds\Big). \tag{3.5}$$

It is easily checked that f_N belongs to $M_w^2[0, T]$. Hence, by what was already proved, a version of

$$J_N(t) = \int_0^t f_N(s)\, dw(s) \qquad (0 \leqslant t \leqslant T)$$

is a continuous process.

Let

$$\Omega_N = \left\{ \int_0^T f^2(t)\, dt < N \right\}.$$

If $\omega \in \Omega_N$, then $f_N(t) = f_M(t)$ for $0 \leqslant t \leqslant T$, $M > N$. By Lemma 2.11 it follows that for a.a. $\omega \in \Omega_N$

$$J_N(t) = J_M(t) \qquad \text{if} \quad 0 \leqslant t \leqslant T.$$

Therefore

$$\tilde{J}(t) = \lim_{M \to \infty} J_M(t)$$

is continuous in $t \in [0, T]$ for a.a. $\omega \in \Omega_N$.

Since $\Omega_N \uparrow$, $P(\Omega_N) \uparrow 1$ if $N \uparrow \infty$, $\tilde{J}(t)$ $(0 \leqslant t \leqslant T)$ is a continuous process.

But since for each $t \in (0, T]$,

$$P\left\{ \int_0^t |f(s) - f_M(s)|^2\, ds > 0 \right\} = P\left\{ \int_0^t f^2(s)\, ds > M \right\} \to 0$$

as $M \to \infty$, we have, by Theorem 2.7,

$$J_M(t) \xrightarrow{P} \int_0^t f(s)\, dw(s) = I(t).$$

Consequently, $I(t)$ has the continuous version $\tilde{J}(t)$.

Remark. From now on, when we speak of the indefinite integral (3.1) of a function $f \in L_w^2[0, T]$ we always mean a continuous version of it.

Theorem 3.3. *Let $f \in L_w^2[0, T]$. Then, for any $\epsilon > 0$, $N > 0$,*

$$P\left\{ \sup_{0 \leqslant t \leqslant T} \left| \int_0^t f(s)\, dw(s) \right| > \epsilon \right\} \leqslant P\left\{ \int_0^T f^2(t)\, dt > N \right\} + \frac{N}{\epsilon^2}. \tag{3.6}$$

Proof. With the notation of (3.4), (3.5) we have

$$\begin{aligned} &P\left\{ \sup_{0 \leqslant t \leqslant T} \left| \int_0^t f(s)\, dw(s) \right| > \epsilon \right\} \\ &\quad \leqslant P\left\{ \sup_{0 \leqslant t \leqslant T} \left| \int_0^t f(s)\, dw(s) - \int_0^t f_N(s)\, dw(s) \right| > 0 \right\} \\ &\qquad + P\left\{ \sup_{0 \leqslant t \leqslant T} \left| \int_0^t f_N(s)\, dw(s) \right| > \epsilon \right\} \\ &\quad \equiv A + B. \end{aligned}$$

By Theorem 3.1, $\int_0^t f_N(s)\, dw(s)$ is a martingale. Hence, by the martingale inequality,

$$B \leq \frac{1}{\epsilon^2} E\left| \int_0^T f_N(s)\, dw(s) \right|^2 = \frac{1}{\epsilon^2} E \int_0^T (f_N(s))^2\, ds \leq \frac{N}{\epsilon^2}.$$

Next, on the set

$$\Omega_N = \left\{ \int_0^T f^2(s)\, ds \leq N \right\}$$

we have $f(t) = f_N(t)$ for $0 \leq t \leq T$. Hence, by Lemma 2.11, for each fixed t in $[0, T]$

$$\int_0^t f(s)\, dw(s) = \int_0^t f_N(s)\, dw(s) \qquad \text{for a.a.} \quad \omega \in \Omega_N.$$

Since both integrals are continuous processes, the last relation holds for all $t \in [0, T]$ and for all $\omega \in \Omega'_N \subset \Omega_N$ where $P(\Omega_N \backslash \Omega'_N) = 0$, i.e.,

$$A \leq P(\Omega \backslash \Omega_N) = P\left\{ \int_0^T f^2(s)\, ds > N \right\}.$$

Combining the estimates on A, B, (3.6) follows.

Theorem 3.4. *Let f_n, f belong to $L_w^2[0, T]$ and assume that $\int_0^T |f_n - f|^2\, dt \xrightarrow{P} 0$ if $n \to \infty$. Then*

$$\sup_{0 \leq t \leq T} \left| \int_0^t f_n(s)\, dw(s) - \int_0^t f(s)\, dw(s) \right| \xrightarrow{P} 0 \qquad \text{if} \quad n \to \infty.$$

This is a consequence of Theorem 3.3.

Theorem 3.5. *Let $f \in M_w^2[0, T]$. Then for any $\lambda > 0$*

$$P\left\{ \sup_{0 \leq t \leq T} \left| \int_0^t f(s)\, dw(s) \right| > \lambda \right\} \leq \frac{1}{\lambda^2} E \int_0^T f^2(s)\, ds. \tag{3.7}$$

Proof. By Theorems 3.1, 3.2, the indefinite integral of f is a continuous martingale. The inequality (3.7) then follows from the martingale inequality.

Another estimate is given in the following theorem.

Theorem 3.6. *Let $f \in M_w^2[0, T]$. Then*

$$E\left\{ \sup_{0 \leq t \leq T} \left| \int_0^t f(s)\, dw(s) \right|^2 \right\} \leq 4E\left| \int_0^T f(t)\, dw(t) \right|^2 = 4E \int_0^T f^2(t)\, dt. \tag{3.8}$$

Like Theorem 3.5, the present estimate is also a consequence of the fact that $\int_0^t f(s)\,dw(s)$ is a continuous martingale. The general result for martingales is given in the following theorem.

Theorem 3.7. *If $X(t)$ $(0 \leqslant t \leqslant T)$ is a separable martingale, then, for any $\alpha > 1$,*

$$E\Big\{\sup_{0\leqslant t\leqslant T} |X(t)|^\alpha\Big\} \leqslant \Big(\frac{\alpha}{\alpha-1}\Big)^\alpha E|X(T)|^\alpha. \tag{3.9}$$

The proof follows immediately from the following lemma.

Lemma 3.8. *If $\{X_n\}$ is a submartingale and $X_n \geqslant 0$ for all n, then, for any $\alpha > 1$,*

$$E\Big[\max_{1\leqslant i\leqslant n} (X_i)^\alpha\Big] \leqslant \Big(\frac{\alpha}{\alpha-1}\Big)^\alpha E(X_n)^\alpha. \tag{3.10}$$

Proof. Let $Y = \max_{1\leqslant i\leqslant n} X_i$. For any $\lambda > 0$, set $\chi_k(\lambda) = 1$ if $X_i < \lambda$ for $1 \leqslant i \leqslant k-1$, $X_k \geqslant \lambda$, and $\chi_k(\lambda) = 0$ otherwise. Then $\sum_{k=1}^n \chi_k(\lambda) = 1$ if $\lambda \leqslant Y$ and $\sum_{k=1}^n \chi_k(\lambda) = 0$ if $\lambda > Y$. It follows that, for any $\beta > 0$,

$$Y^\beta = \beta \int_0^\infty \lambda^{\beta-1}\Big[\sum_{k=1}^n \chi_k(\lambda)\Big]\,d\lambda. \tag{3.11}$$

Since $\lambda\chi_k(\lambda) \leqslant X_k\chi_k(\lambda)$,

$$\lambda^{\alpha-1}\sum_{k=1}^n \chi_k(\lambda) \leqslant \sum_{k=1}^n \lambda^{\alpha-2} X_k\chi_k(\lambda). \tag{3.12}$$

Observing that $\chi_k(\lambda)$ is measurable with respect to $\mathcal{F}(X_1, \ldots, X_k)$ and using the submartingale assumption, we get

$$E\{X_n\chi_k(\lambda)|\mathcal{F}(X_1, \ldots, X_k)\} \geqslant \chi_k(\lambda)E\{X_n|\mathcal{F}(X_1, \ldots, X_k)\} \geqslant \chi_k(\lambda)X_k.$$

Taking the expectation in (3.12) and using the last inequality, we find that

$$E\lambda^{\alpha-1}\sum_{k=1}^n \chi_k(\lambda) \leqslant E\lambda^{\alpha-2}\sum_{k=1}^n \chi_k(\lambda)X_n.$$

Integrating with respect to λ and using (3.11), we get

$$\frac{1}{\alpha}EY^\alpha \leqslant \frac{1}{\alpha-1}EY^{\alpha-1}X_n.$$

By Hölder's inequality,

$$EY^\alpha \leqslant \frac{\alpha}{\alpha-1}(EY^\alpha)^{(\alpha-1)/\alpha}(EX_n^\alpha)^{1/\alpha},$$

and (3.10) follows.

Remark. It is clear that all the results derived so far in this section regarding the indefinite integral (3.1) extend to indefinite integrals

$$\int_\alpha^t f(s)\, dw(s).$$

Theorem 3.9. *Let $f \in M_w^2[\alpha, \beta]$. Then*

$$E\left\{ \sup_{\alpha < t < \beta} \left| \int_\alpha^t f(s)\, dw(s) \right|^2 \middle| \mathcal{F}_\alpha \right\} \leqslant 4E\left\{ \int_\alpha^\beta f^2(t)\, dt \middle| \mathcal{F}_\alpha \right\}. \quad (3.13)$$

The proof is left to the reader.

Definition. Let τ_0 be a random variable, $0 \leqslant \tau_0 \leqslant T$. If $f \in L_w^2[0, T]$, we define

$$\int_0^{\tau_0} f(s)\, dw(s)$$

to be the random variable $I(\tau_0)$, where $I(t)$ is given by (3.1).

From Theorems 3.1, 3.2, and 1.3.5 we deduce:

Theorem 3.10. *If $f \in M_w^2[0, \tau]$ and τ is a stopping time with respect to $\mathcal{F}_t$, $0 \leqslant \tau \leqslant T$, i.e., $\{\tau \leqslant t\} \in \mathcal{F}_t$ for all $0 \leqslant t \leqslant T$, then the process*

$$\int_0^{\tau \wedge t} f(s)\, dw(s), \qquad 0 \leqslant t \leqslant T$$

is a martingale and

$$E \int_0^{\tau \wedge t} f(s)\, dw(s) = 0. \quad (3.14)$$

4. Stochastic integrals with stopping time

If $f \in L_w^2[\alpha, T]$ for all $T > 0$, then we say that f belongs to $L_w^2[\alpha, \infty)$. Similarly we define $f \in M_w^2[\alpha, \infty)$.

Let $f \in L_w^2[0, T]$ and let ζ_1, ζ_2 be random variables, $0 \leqslant \zeta_1 \leqslant \zeta_2 \leqslant T$. We define

$$\int_{\zeta_1}^{\zeta_2} f(t)\, dw(t) = \int_0^{\zeta_2} f(t)\, dw(t) - \int_0^{\zeta_1} f(t)\, dw(t).$$

Lemma 4.1. *Define $\chi_i(t) = 1$ if $t < \zeta_i$, $\chi_i(t) = 0$ if $t \geqslant \zeta_i$ $(i = 1, 2)$. Then the $\chi_i(t)$ are $\mathcal{F}_t$ measurable and*

$$\int_{\zeta_1}^{\zeta_2} f(t)\, dw(t) = \int_0^T \chi_2(t) f(t)\, dw(t) - \int_0^T \chi_1(t) f(t)\, dw(t). \quad (4.1)$$

Proof. It is enough to prove the lemma in case $\zeta_1 \equiv 0$. It is clear that $\chi_2(t)$ is $\mathfrak{F}_t$ measurable; in fact, it is a nonanticipative function. In case f is a step function and ζ_2 is a simple function, (4.1) follows directly from the definition of the integral. In the general case, let f_n be step functions such that

$$\int_0^T |f_n(t) - f(t)|^2 \, dt \xrightarrow{P} 0 \quad \text{if} \quad n \to \infty,$$

and let ζ_{2n} be simple stopping times such that $\zeta_{2n} \downarrow \zeta_2$ everywhere if $n \uparrow \infty$ (cf. Lemma 1.3.4). We have

$$\int_0^{\zeta_{2n}} f_n(t) \, dw(t) = \int_0^T \chi_{2n}(t) f_n(t) \, dw(t) \tag{4.2}$$

where $\chi_{2n}(t) = 1$ if $t < \zeta_{2n}$, $\chi_{2n}(t) = 0$ if $t \geqslant \zeta_{2n}$. It is clear that $\chi_{2n}(t) \to \chi_2(t)$ for all ω and $t \neq \zeta_2(\omega)$, as $n \to \infty$. Hence

$$\int_0^T |\chi_{2n}(t) - \chi_2(t)|^2 f(t) \, dt \to 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty,$$

by the Lebesgue bounded convergence theorem. We also have

$$\int_0^T |\chi_{2n}(t)|^2 |f(t) - f_n(t)|^2 \, dt \xrightarrow{P} 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty.$$

Putting these together, we find that

$$\int_0^T |\chi_{2n}(t) f_n(t) - \chi_2(t) f(t)|^2 \, dt \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Hence, by Theorem 2.7,

$$\int_0^T \chi_{2n}(s) f_n(s) \, dw(s) \xrightarrow{P} \int_0^T \chi_2(s) f(s) \, dw(s). \tag{4.3}$$

By Theorem 3.4,

$$\sup_{0<t<T} \left| \int_0^t f_n(s) \, dw(s) - \int_0^t f(s) \, dw(s) \right| \xrightarrow{P} 0 \quad \text{if} \quad n \to \infty.$$

It follows that

$$\int_0^{\zeta_{2n}} f_n(s) \, dw(s) - \int_0^{\zeta_{2n}} f(s) \, dw(s) \xrightarrow{P} 0 \quad \text{if} \quad n \to \infty.$$

Since from the continuity of the indefinite integral we also have

$$\int_0^{\zeta_{2n}} f(s) \, dw(s) \to \int_0^{\zeta_2} f(s) \, dw(s) \quad \text{a.s.} \quad \text{as} \quad n \to \infty,$$

we find that

$$\int_0^{\zeta_{2n}} f_n(s) \, dw(s) \xrightarrow{P} \int_0^{\zeta_2} f(s) \, dw(s) \quad \text{if} \quad n \to \infty.$$

Combining this with (4.2), (4.3), the formula (4.1) (in case $\zeta_1 \equiv 0$) follows.

Theorem 4.2. *Let $f \in M_w^2[0, T]$ and let ζ_1, ζ_2 be stopping times, $0 \leq \zeta_1 \leq \zeta_2 \leq T$. Then*

$$E \int_{\zeta_1}^{\zeta_2} f(t)\, dw(t) = 0, \tag{4.4}$$

$$E\left[\int_{\zeta_1}^{\zeta_2} f(t)\, dw(t) \right]^2 = E \int_{\zeta_1}^{\zeta_2} f^2(t)\, dt. \tag{4.5}$$

Proof. Let $\chi_i(t) = 1$ if $t < \zeta_i$, $\chi_i(t) = 1$ if $t \geq \zeta_i$. By Lemma 4.1,

$$\int_{\zeta_1}^{\zeta_2} f(t)\, dw(t) = \int_0^T (\chi_2(t) - \chi_1(t)) f(t)\, dw(t).$$

Applying Theorem 2.5 to the right-hand side, the assertions (4.4), (4.5) readily follow.

Note that (4.4) follows also from Theorem 3.10.

Theorem 4.2 is a generalization of Theorem 2.5. The next theorem is a generalization of Theorem 2.8.

Theorem 4.3. *Let $f \in M_w^2[0, T]$ and let ζ_1, ζ_2 be stopping times (with respect to $\mathcal{F}_t$), $0 \leq \zeta_1 \leq \zeta_2 \leq T$. Then*

$$E\left\{ \int_{\zeta_1}^{\zeta_2} f(t)\, dw(t) | \mathcal{F}_{\zeta_1} \right\} = 0, \tag{4.6}$$

$$E\left\{ \left| \int_{\zeta_1}^{\zeta_2} f(t)\, dw(t) \right|^2 | \mathcal{F}_{\zeta_1} \right\} = E\left\{ \int_{\zeta_1}^{\zeta_2} f^2(t)\, dt | \mathcal{F}_{\zeta_1} \right\}. \tag{4.7}$$

Here $\mathcal{F}_\zeta$ denotes the σ-field of all events A such that

$$A \cap (\zeta \leq s) \quad \text{is in } \mathcal{F}_s, \quad \text{for all} \quad s \geq 0.$$

Denote by $\mathcal{F}_\zeta^*$ the σ-field generated by all sets of the form

$$A \cap (\zeta > t), \qquad A \in \mathcal{F}_t, \quad t \geq 0.$$

We shall need the following lemma.

Lemma 4.4. *For any stopping time ζ, $\mathcal{F}_\zeta = \mathcal{F}_\zeta^*$.*

Proof. Let $B = A \cap (\zeta > t)$, $A \in \mathcal{F}_t$. For any $s \geq 0$, let

$$C = B \cap (\zeta \leq s) = A \cap (\zeta > t) \cap (\zeta \leq s).$$

If $s \leq t$, then $C = \emptyset \in \mathcal{F}_s$. If $s > t$, then

$$C = [A^c \cup (\zeta \leq t)]^c \cap (\zeta \leq s)$$

where D^c is the complement of D. Since $A^c \in \mathcal{F}_t$ and $(\zeta \leq t) \in \mathcal{F}_t$, it follows that $[A^c \cup (\zeta \leq t)]^c \in \mathcal{F}_t$. Hence $C \in \mathcal{F}_s$. We have thus proved that $B \cap (\zeta \leq s)$ is in $\mathcal{F}_s$ for all $s \geq 0$, i.e., $B \in \mathcal{F}_\zeta$. It follows that $\mathcal{F}_\zeta^* \subset \mathcal{F}_\zeta$.

Conversely let $B \in \mathcal{F}_\zeta$. Then

$$B_s = B \cap (\zeta \leqslant s) \qquad \text{is in } \mathcal{F}_s, \quad \text{for all} \quad s \geqslant 0.$$

Hence $B_s^c = B^c \cup (\zeta > s)$ is in $\mathcal{F}_s$, and since

$$B_s^c = B_s^c \cap (\zeta > s),$$

it follows that B_s^c is one of the sets that generate $\mathcal{F}_\zeta^*$. Therefore $B_s^c \in \mathcal{F}_\zeta^*$. Since $\mathcal{F}_\zeta^*$ is a σ-field, B_s belongs to $\mathcal{F}_\zeta^*$, and also $\lim_{s \uparrow \infty} B_s \in \mathcal{F}_\zeta^*$. But this limit is the set B. Hence $B \in \mathcal{F}_\zeta^*$.

***Proof of Theorem* 4.3.** Let $C = A \cap (\zeta_1 > s)$, $A \in \mathcal{F}_s$. Its indicator function $\chi_C = \chi_A \chi_1(s)$ is $\mathcal{F}_s$ measurable. Consider the function

$$\chi_C(\chi_2(t) - \chi_1(t)) = \chi_A \chi_1(s)(\chi_2(t) - \chi_1(t)).$$

If $s \leqslant t$, each factor on the left is in $\mathcal{F}_t$, and so is then the product. If $s > t$, then $\chi_1(t) = 1$, so that $\chi_2(t) - \chi_1(t) = 0$. Thus the product is again in $\mathcal{F}_t$. We have thus proved that

$$\chi_C(\chi_2(t) - \chi_1(t)) \qquad \text{is } \mathcal{F}_t \text{ measurable for any set } C. \tag{4.8}$$

Let $B \in \mathcal{F}_{\zeta_1}$. From the proof of Lemma 4.4 we have that B_s^c has the form of the set C for which (4.8) holds. Hence also

$$\chi_{B_s}(\chi_2(t) - \chi_1(t)) \qquad \text{is } \mathcal{F} \text{ measurable.}$$

Taking $s \uparrow \infty$ we conclude that

$$\chi_B(\chi_2(t) - \chi_1(t)) \qquad \text{is } \mathcal{F}_t \text{ measurable.}$$

We can now proceed as in the proof of Theorem 4.2 to prove that

$$E\chi_B \int_{\zeta_1}^{\zeta_2} f(t)\, dw(t) = 0,$$

$$E\left|\chi_B \int_{\zeta_1}^{\zeta_2} f(t)\, dw(t)\right|^2 = E \int_{\zeta_1}^{\zeta_2} \chi_B f^2(t)\, dt,$$

and the assertions (4.6), (4.7) follow.

As an application of Theorem 4.3 we shall prove:

Theorem 4.5. *Let $f \in L_w^2[0, \infty)$ and assume that $\int_0^\infty f^2(t)\, dt = \infty$ with probability* 1. *Let*

$$\tau(t) = \inf\left\{ s;\ \int_0^s f^2(\lambda)\, d\lambda = t \right\},$$

then the process

$$u(t) = \int_0^{\tau(t)} f(s)\, dw(s)$$

is a Brownian motion.

Let

$$t^*(t) = \int_0^t f^2(s)\, ds.$$

Then τ is the left-continuous inverse of t^*, i.e., $\tau(t) = \min(s, t^*(s) = t)$. t^* is called the *intrinsic time* (or *intrinsic clock*) for $I(t) = \int_0^t f(s)\, dw(s)$. Theorem 4.5 asserts that there is a Brownian motion $u(t)$ such that $u(t^*(t)) = I(t)$.

Proof. It is easily verified that $\tau(t)$ is a stopping time. Notice next that $\mathcal{F}_{\tau(t_1)} \subset \mathcal{F}_{\tau(t_2)}$ if $t_1 < t_2$. Indeed, if $A \in \mathcal{F}_{\tau(t_1)}$, then $A \cap [\tau(t_1) \leqslant s]$ is in $\mathcal{F}_s$ for all $s \geqslant 0$. But since $\tau(t_2) \geqslant \tau(t_1)$,

$$A \cap [\tau(t_2) \leqslant s] = \{A \cap [\tau(t_1) \leqslant s]\} \cap [\tau(t_2) \leqslant s] \qquad \text{is in } \mathcal{F}_s.$$

Thus $A \in \mathcal{F}_{\tau(t_2)}$. We shall now assume that $f \in M_w^2[0, \infty)$ and $f(t) = 1$ if $t > \lambda$, for some $\lambda > 0$. Then $\tau(t_2) < t_2 + \lambda$ and, by Theorem 4.3,

$$E\left\{ \int_{\tau(t_1)}^{\tau(t_2)} f(s)\, dw(s) | \mathcal{F}_{\tau(t_1)} \right\} = 0,$$

$$E\left\{ \left| \int_{\tau(t_1)}^{\tau(t_2)} f(s)\, dw(s) \right|^2 | \mathcal{F}_{\tau(t_1)} \right\} = E\left\{ \int_{\tau(t_1)}^{\tau(t_2)} f^2(s)\, ds | \mathcal{F}_{\tau(t_1)} \right\} = t_2 - t_1.$$

If we prove that $u(t)$ is continuous, then Theorem 3.5.1 implies that $u(t)$ is a Brownian motion. From Theorem 3.6 we deduce that

$$E\left\{ \sup_{\substack{t' < t < t'+\epsilon \\ t' < s < t'+\epsilon}} |u(t) - u(s)|^2 \right\} \leqslant 4E\left\{ \int_{\tau(t')}^{\tau(t'+\epsilon)} f^2(s)\, ds \right\} = 4\epsilon.$$

Since t' is arbitrary,

$$P\left\{ \sup_{\substack{|t-s| < \epsilon \\ t > 0,\, s > 0}} |u(t) - u(s)| > \epsilon^\alpha \right\} \leqslant \frac{4\epsilon}{\epsilon^{2\alpha}}.$$

Taking $\alpha = \frac{1}{4}$, $\epsilon = 1/k^4$ and using the Borel–Cantelli lemma, we find that

$$P\left\{ \sup_{\substack{|t-s| < 1/k^4 \\ t > 0, s > 0}} |u(t) - u(s)| > 1/k \quad \text{i.o.} \right\} = 0, \qquad \text{i.e.} \quad u \text{ is continuous.}$$

For a general $f \in M_w^2[0, \infty)$, apply the preceding result to f_λ where $f_\lambda(t) = f(t)$ if $t < \lambda$, $f_\lambda(t) = 1$ if $t > \lambda$. Since (by Theorem 4.7 below)

$$\int_0^{\tau_\lambda(t)} f_\lambda(s)\, dw(s) = \int_0^{\tau(t)} f(s)\, ds \quad \text{a.e. on the set } \lambda > \tau(t),$$

where τ_λ is the τ function of f_2, Problem 10, Chapter 3 implies that $u(t)$ is a Brownian motion.

So far we have proved the theorem only in case $f \in M_w^2[0, \infty)$. Now let

$f \in L_w^2[0, \infty)$ and define χ_N, f_N as in (3.4), (3.5). Then $f_N \in M_w^2[0, \infty)$, so that

$$\int_0^{\tau_N(t)} f_N(s)\, dw(s)$$

is a Brownian motion. But $\tau_N(t) = \tau(t)$, $f_N(s) = f(s)$ if $0 < s \leqslant \tau(t)$, $t < N$. Hence (by Theorem 4.7 below)

$$\int_0^{\tau_N(t)} f_N(s)\, dw(s) = \int_0^{\tau(t)} f(s)\, dw(s)$$

if $N > t$. It follows that the function on the right is Brownian.

Corollary 4.6. *Let $f \in L_w^2[0, \infty)$, $|f| \leqslant K$ where K is a constant. Then, for any $\alpha > \frac{1}{2}$,*

$$\frac{1}{t^\alpha} \int_0^t f(s)\, dw(s) \to 0 \qquad \text{if} \quad t \to \infty. \tag{4.9}$$

Proof. Let $M > K$ and define

$$t^*(t) = \int_0^t (f(s) + M)^2\, ds,$$

$$a(t) = \int_0^{\tau(t)} (f(s) + M)^2\, dw(s) \qquad (\tau = \text{inverse of } t^*).$$

Since $|f + M| \geqslant M - K > 0$, the function $t^*(t)$ is strictly monotone increasing to ∞ as $t \uparrow \infty$. Hence $a(t)$ is defined for all $t \geqslant 0$ and, by Theorem 4.5, it is a Brownian motion. But then, by the law of the iterated logarithm,

$$\frac{a(t)}{t^\alpha} \to 0 \quad \text{a.s.} \quad \text{if} \quad t \to \infty,$$

or

$$\frac{a(t^*(t))}{(t^*(t))^\alpha} \to 0 \quad \text{a.s.} \qquad \text{if} \quad t \to \infty.$$

Noting that $t^*(t) \leqslant (K + M)^2 t$, we get

$$\frac{1}{t^\alpha} \int_0^t (f(s) + M)\, dw(s) \to 0 \quad \text{a.s.} \qquad \text{if} \quad t \to \infty.$$

Using the fact that $(w(t)/t^\alpha) \to 0$ if $t \to \infty$, the assertion (4.9) follows.

Theorem 4.7. *Let $f \in L_w^2[0, \infty)$, $g \in L_w^2[0, \infty)$, and set*

$$h(t) = \int_0^t f(s)\, dw(s), \qquad k(t) = \int_0^t g(s)\, dw(s).$$

Let τ be a random variable, $\tau > 0$, and assume that $f(s) = g(s)$ if $s < \tau$. Then $h(t) = k(t)$ a.s. if $t < \tau$.

Proof. Construct, by Lemma 1.1, step functions f_n, g_n which belong to $L_w^2[0, T]$ and which satisfy

$$\int_0^T |f_n(t) - f(t)|^2 \, dt \xrightarrow{P} 0, \qquad \int_0^T |g_n(t) - g(t)|^2 \, dt \xrightarrow{P} 0$$

as $n \to \infty$. That construction is such that

$$f_n(s, \omega) = g_n(s, \omega) \qquad \text{if} \quad s < \tau(\omega) \wedge T. \tag{4.10}$$

Let

$$h_n(t) = \int_0^t f_n(s) \, dw(s), \qquad k_n(t) = \int_0^t g_n(s) \, dw(s).$$

From the definition of the integral of a step function and from (4.10) it easily follows that

$$h_n(t, \omega) = k_n(t, \omega) \qquad \text{if} \quad t < \tau(\omega) \wedge T. \tag{4.11}$$

By Theorem 3.4,

$$\sup_{0 \leqslant t \leqslant T} |h_n(t) - h(t)| \xrightarrow{P} 0, \qquad \sup_{0 \leqslant t \leqslant T} |k_n(t) - k(t)| \xrightarrow{P} 0$$

if $n \to \infty$. Hence, for some subsequence $\{n'\}$,

$$\sup_{0 \leqslant t \leqslant T} |h_{n'}(t) - h(t)| \to 0, \qquad \sup_{0 \leqslant t \leqslant T} |k_{n'}(t) - k(t)| \to 0 \quad \text{a.s.}$$

if $n' \to \infty$. Recalling (4.11), we find that for a.a. ω,

$$h(t, \omega) = k(t, \omega) \qquad \text{if} \quad t < \tau(\omega) \wedge T.$$

The proof is now completed by taking a sequence $T = T_n \uparrow \infty$.

5. Itô's formula

Definition. Let $\xi(t)$ $(0 \leqslant t \leqslant T)$ be a process such that for any $0 \leqslant t_1 < t_2 \leqslant T$

$$\xi(t_2) - \xi(t_1) = \int_{t_1}^{t_2} a(t) \, dt + \int_{t_1}^{t_2} b(t) \, dw(t)$$

where $a \in L_w^1[0, T]$, $b \in L_w^2[0, T]$. Then we say that $\xi(t)$ has *stochastic differential* $d\xi$, on $[0, T]$, given by

$$d\xi(t) = a(t) \, dt + b(t) \, dw(t).$$

Observe that $\xi(t)$ is a nonanticipative function. It is also a continuous process. Hence, in particular, it belongs to $L_w^\infty[0, T]$.

Example 1. If $0 \leqslant t_1 < t_2$ and $\Pi_n = \{t_1 = t_{n,1}, t_{n,2}, \ldots, t_{n,n} = t_2\}$ is a sequence of partitions of $[t_1, t_2]$ with mesh $|\Pi_n| \to 0$ then, by Theorem 2.10,

$$\int_{t_1}^{t_2} w(t)\, dw(t) = \lim_{n\to\infty} \sum_{k=1}^{n-1} w(t_{n,k})(w(t_{n,k+1}) - w(t_{n,k}))$$

$$= \tfrac{1}{2} \lim_{n\to\infty} \sum_{k=1}^{n-1} \{[(w(t_{n,k+1}))^2 - (w(t_{n,k}))^2] - (w(t_{n,k+1}) - w(t_{n,k}))^2\}$$

$$= \tfrac{1}{2}(w(t_2))^2 - \tfrac{1}{2}(w(t_1))^2 - \tfrac{1}{2} \lim_{n\to\infty} \sum_{k=1}^{n-1} (w(t_{n,k+1}) - w(t_{n,k}))^2$$

where $\lim_{n\to\infty}$ is taken as the limit in probability.

By Theorem 3.3.3, the last limit in probability is equal to $t_2 - t_1$. Hence

$$\int_{t_1}^{t_2} w(t)\, dw(t) = \tfrac{1}{2}(w(t_2))^2 - \tfrac{1}{2}(w(t_1))^2 - \tfrac{1}{2}(t_2 - t_1), \tag{5.1}$$

or

$$d(w(t))^2 = dt + 2w(t)\, dw(t). \tag{5.2}$$

Example 2. By Theorem 2.10,

$$\int_{t_1}^{t_2} t\, dw(t) = \lim_{n\to\infty} \sum_{k=1}^{n-1} t_{n,k}[w(t_{n,k+1}) - w(t_{n,k})] \quad \text{in probability.}$$

Clearly

$$\int_{t_1}^{t_2} w(t)\, dt = \lim_{n\to\infty} \sum_{k=1}^{n-1} w(t_{n,k+1})(t_{n,k+1} - t_{n,k})$$

for all ω for which $w(t, \omega)$ is continuous. The sum of the right-hand sides is equal to

$$\lim_{n\to\infty} \sum_{k=1}^{n-1} [t_{n,k+1} w(t_{n,k+1}) - t_{n,k} w(t_{n,k})] = t_2 w(t_2) - t_1 w(t_1).$$

Hence

$$d(tw(t)) = w(t)\, dt + t\, dw(t). \tag{5.3}$$

Definition. Let $\xi(t)$ be as in the definition above and let $f(t)$ be a function in $L_w^\infty[0, T]$. We define

$$f(t)\, d\xi(t) = f(t)a(t)\, dt + f(t)b(t)\, dw(t).$$

Example 3. $f(t)\, d\xi(t)$ is a stochastic differential $d\eta$, where

$$\eta(t) = \int_0^t f(s)a(s)\, ds + \int_0^t f(s)b(s)\, dw(s).$$

Theorem 5.1. *If* $d\xi_i(t) = a_i(t)\,dt + b_i(t)\,dw(t)$ $(i = 1, 2)$, *then*

$$d(\xi_1(t)\xi_2(t)) = \xi_1(t)\,d\xi_2(t) + \xi_2(t)\,d\xi_1(t) + b_1(t)b_2(t)\,dt. \tag{5.4}$$

The integrated form of (5.4) asserts that, for any $0 \leq t_1 < t_2 \leq T$,

$$\begin{aligned}\xi_1(t_2)\xi_2(t_2) - \xi_1(t_1)\xi_2(t_1) = &\int_{t_1}^{t_2} \xi_1(t)a_2(t)\,dt + \int_{t_1}^{t_2} \xi_1(t)b_2(t)\,dw(t)\\ &+ \int_{t_1}^{t_2} \xi_2(t)a_1(t)\,dt + \int_{t_1}^{t_2} \xi_2(t)b_1(t)\,dw(t)\\ &+ \int_{t_1}^{t_2} b_1(t)b_2(t)\,dt.\end{aligned} \tag{5.5}$$

Proof. Suppose first that a_i, b_i are constants in the interval $[t_1, t_2)$. Then (5.5) follows from (5.2), (5.3). Next, if a_i, b_i are step functions in $[t_1, t_2)$, constants on successive intervals $I_1, I_2, \ldots, I_l$, then (5.5) holds with t_1, t_2 replaced by the end points of each interval I_j. Taking the sum we obtain (5.5).

Consider now the general case. Approximate a_i, b_i by nonanticipative step functions $a_{i,n}$, $b_{i,n}$ in such a way that

$$\int_0^T |a_{i,n}(t) - a_i(t)|\,dt \to 0 \quad \text{a.s.,}$$

$$\int_0^T |b_{i,n}(t) - b_i(t)|^2\,dt \to 0 \quad \text{a.s.}$$

Let

$$\xi_{i,n}(t) = \xi_i(0) + \int_0^t a_{i,n}(s)\,ds + \int_0^t b_{i,n}(s)\,dw(s).$$

By Theorem 3.4,

$$\sup_{0 \leq t \leq T} |\xi_{i,n}(t) - \xi_i(t)| \xrightarrow{P} 0 \qquad \text{if} \quad n \to \infty;$$

hence there is a subsequence $\xi_{i,n'}$ such that

$$\xi_{i,n'}(t) \to \xi_i(t) \quad \text{uniformly in} \quad t \in [0, T], \quad \text{a.s.} \tag{5.6}$$

Using (5.6) and Theorem 2.7 it is easily seen that

$$\int_{t_1}^{t_2} \xi_{i,n}(t)b_{j,n}(t)\,dw(t) \xrightarrow{P} \int_{t_1}^{t_2} \xi_i(t)b_j(t)\,dw(t)$$

as $n = n' \to \infty$. Clearly also

$$\int_{t_1}^{t_2} \xi_{i,n}(t)a_{j,n}(t)\,dt \to \int_{t_1}^{t_2} \xi_i(t)a_j(t)\,dt,$$

$$\int_{t_1}^{t_2} b_{1,n}(t)b_{2,n}(t)\,dt \to \int_{t_1}^{t_2} b_1(t)b_2(t)\,dt$$

a.s. Writing (5.5) for $a_{i,n}$, $b_{i,n}$, $\xi_{i,n}$ and taking $n \to \infty$, the assertion (5.5) follows. Since t_1, t_2 are arbitrary, the proof of the theorem is complete.

Theorem 5.2. *Let $d\xi(t) = a(t)\,dt + b(t)\,dw(t)$, and let $f(x, t)$ be a continuous function in $(x, t) \in R^1 \times [0, \infty)$ together with its derivatives f_x, f_{xx}, f_t. Then the process $f(\xi(t), t)$ has a stochastic differential, given by*

$$df(\xi(t), t) = [f_t(\xi(t), t) + f_x(\xi(t), t)a(t) + \tfrac{1}{2}f_{xx}(\xi(t), t)b^2(t)]\,dt + f_x(\xi(t), t)b(t)\,dw(t). \tag{5.7}$$

This is called *Itô's formula.* Notice that if $w(t)$ were continuously differentiable in t, then (by the standard calculus formula for total derivatives) the term $\frac{1}{2}f_{xx}b^2\,dt$ would not appear.

Proof. The proof will be divided into several steps.

Step 1. For any integer $m \geqslant 2$,

$$d(w(t))^m = m(w(t))^{m-1} + \tfrac{1}{2}m(m-1)(w(t))^{m-2}\,dt. \tag{5.8}$$

Indeed, this follows by induction, using Theorem 5.1.

By linearity of the stochastic differential we then get

$$dQ(w(t)) = Q'(w(t))\,dw(t) + \tfrac{1}{2}Q''(w(t))\,dt \tag{5.9}$$

for any polynomial Q.

Step 2. Let $G(x, t) = Q(x)g(t)$ where $Q(x)$ is a polynomial and $g(t)$ is continuously differentiable for $t \geqslant 0$. By Theorem 5.1 and (5.9),

$$\begin{aligned} dG(w(t), t) &= f(w(t))dg(t) + g(t)df(w(t)) \\ &= [f(w(t))g'(t) + \tfrac{1}{2}g(t)f''(w(t))]\,dt + g(t)f'(w(t))\,dw(t), \end{aligned}$$

i.e., for any $0 \leqslant t_1 < t_2 \leqslant T$,

$$G(w(t_2), t_2) - G(w(t_1), t_1) = \int_{t_1}^{t_2} [G_t(w(t), t) + \tfrac{1}{2}G_{xx}(w(t), t)]\,dt + \int_{t_1}^{t_2} G_x(w(t), t)\,dw(t). \tag{5.10}$$

Step 3. Formula (5.10) remains valid if

$$G(x, t) = \sum_{i=1}^{m} f_i(x)g_i(t)$$

where $f_i(x)$ are polynomials and $g_i(t)$ are continuously differentiable. Now let $G_n(x, t)$ be polynomials in x and t such that

$$G_n(x, t) \to f(x, t),$$

$$\frac{\partial}{\partial x} G_n(x, t) \to f_x(x, t), \qquad \frac{\partial^2}{\partial x^2} G_n(x, t) \to f_{xx}(x, t), \qquad \frac{\partial}{\partial t} G_n(x, t) \to f_t(x, t)$$

uniformly on compact subsets of $x, t) \in R^1 \times [0, \infty)$; see Problem 11 for the proof of the existence of such a sequence. We have

$$G_n(w(t_2), t_2) - G_n(w(t_1), t_1) = \int_{t_1}^{t_2} \left[\frac{\partial}{\partial t} G_n(w(t), t) + \frac{1}{2} \frac{\partial^2}{\partial x^2} G_n(w(t), t) \right] dt + \int_{t_1}^{t_2} \frac{\partial}{\partial x} G_n(w(t), t)\, dw(t). \tag{5.11}$$

It is clear that

$$\int_{t_1}^{t_2} \left[\frac{\partial}{\partial t} G_n(w(t), t) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (w(t), t) \right] dt \to \int_{t_1}^{t_2} [f_t(w(t), t) - \tfrac{1}{2} f_{xx}(w(t), t)]\, dt \quad \text{a.s.},$$

$$\int_{t_1}^{t_2} \left| \frac{\partial}{\partial x} G_n(w(t), t) - f_x(w(t), t) \right|^2 dt \to 0 \quad \text{a.s.}$$

Hence, taking $n \to \infty$ in (5.11), we get the relation

$$f(w(t_2), t_2) - f(w(t_1), t_1) = \int_{t_1}^{t_2} [f_t(w(t), t) + \tfrac{1}{2} f_{xx}(w(t), t)]\, dt + \int_{t_1}^{t_2} f_x(w(t), t)\, dw(t). \tag{5.12}$$

Step 4. Formula (5.12) extends to the process

$$\Phi(w(t), t) = f(\xi_1 + a_1 t + b_1 w(t), t)$$

where ξ_1, a_1, b_1 are random variables measurable with respect to $\mathcal{F}_{t_1}$, i.e.,

$$\Phi(w(t_2), t_2) - \Phi(w(t_1), t_1) = \int_{t_1}^{t_2} [f_t(\tilde{\xi}(t), t) + f_x(\tilde{\xi}(t), t) a_1 + \tfrac{1}{2} f_{xx}(\tilde{\xi}(t), t) b_1^2]\, dt + \int_{t_1}^{t_2} f_x(\tilde{\xi}(t), t) b_1\, dw(t) \tag{5.13}$$

where $\tilde{\xi}(t) = \xi_1 + a_1 t + b_1 w(t)$.

The proof of (5.13) is a repetition of the proof of (5.12) with obvious changes resulting from the formula

$$d(\tilde{\xi}(t))^m = m(\tilde{\xi}(t))^{m-1}[a_1\, dt + b_1\, dw(t)] + \tfrac{1}{2} m(m-1)(\tilde{\xi}(t))^{m-2} b_1^2\, dt, \tag{5.14}$$

which replaces (5.8). The details are left to the reader.

Step 5. If $a(t)$, $b(t)$ are step functions, then

$$f(\xi(t_2), t_2) - f(\xi(t_1), t_1) = \int_{t_1}^{t_2} \left[f_t(\xi(t), t) + f_x(\xi(t), t)a(t) + \tfrac{1}{2} f_{xx}(\xi(t), t)b^2(t) \right] dt + \int_{t_1}^{t_2} f_x(\xi(t), t)b(t)\, dw(t). \tag{5.15}$$

Indeed, denote by $I_1, \ldots, I_k$ the successive intervals in $[t_1, t_2]$ in which a, b are constants. If we apply (5.13) with t_1, t_2 replaced by the end points of I_l, and sum over l, the formula (5.15) follows.

Step 6. Let a_i, b_i be nonanticipative step functions such that

$$\int_0^T |a_i(t) - a(t)|\, dt \to 0 \quad \text{a.s.,} \tag{5.16}$$

$$\int_0^T |b_i(t) - b(t)|^2\, dt \xrightarrow{P} 0, \tag{5.17}$$

and let

$$\xi_i(t) = \xi(0) + \int_0^t a_i(s)\, ds + \int_0^t b_i(s)\, dw(s).$$

Then

$$\sup_{0 \le t \le T} |\xi_i(t) - \xi(t)| \xrightarrow{P} 0.$$

Hence, for a subsequence $\{i'\}$,

$$\sup_{0 \le t \le T} |\xi_i(t) - \xi(t)| \to 0 \quad \text{a.s.} \qquad \text{if } i = i' \to \infty. \tag{5.18}$$

This and (5.17) imply that

$$\int_0^T |f_x(\xi_i(t), t)b_i(t) - f_x(\xi(t), t)b(t)|^2\, dt \xrightarrow{P} 0 \qquad \text{if } i = i' \to \infty.$$

It follows that

$$\int_{t_1}^{t_2} f_x(\xi_i(t), t)b_i(t)\, dw(t) \xrightarrow{P} \int_{t_1}^{t_2} f_x(\xi(t), t)b(t)\, dw(t) \qquad \text{if } i = i' \to \infty.$$

It is clear from (5.16)–(5.18) that also

$$\int_{t_1}^{t_2} \left[f_t(\xi_i(t), t) + f_x(\xi_i(t), t)a_i(t) + \tfrac{1}{2} f_{xx}(\xi_i(t), t)(b_i(t))^2 \right] dt$$

$$\xrightarrow{P} \int_{t_1}^{t_2} \left[f_t(\xi(t), t) + f_x(\xi(t), t)a(t) + \tfrac{1}{2} f_{xx}(\xi(t), t)b^2(t) \right] dt$$

if $i = i' \to \infty$.

Writing (5.15) for $a = a_i$, $b = b_i$, $\xi = \xi_i$ and taking $i = i' \to \infty$, the formula (5.15) follows for general a, b. This completes the proof of the theorem.

Theorem 5.3. *Let $d\xi_i(t) = a_i(t)\,dt + b_i(t)\,d\xi$ $(1 \leqslant i \leqslant m)$ and let $f(x_1, \ldots, x_m, t)$ be a continuous function in (x, t) where $x = (x_1, \ldots, x_m) \in R^m$, $t \geqslant 0$, together with its first t-derivative and second x-derivatives. Then $f(\xi_1(t), \ldots, \xi_m(t), t)$ has a stochastic differential, given by*

$$df(X(t), t) = \left[f_t(X(t), t) + \sum_{i=1}^{m} f_{x_i}(X(t), t) a_i(t) + \frac{1}{2} \sum_{i,j=1}^{m} f_{x_i x_j}(X(t), t) b_i(t) b_j(t) \right] dt + \sum_{i=1}^{m} f_{x_i}(X(t), t) b_i(t)\, dw(t), \tag{5.19}$$

where $X(t) = (\xi_1(t), \ldots, \xi_m(t))$.

Formula (5.19) is also called *Itô's formula*. It includes both Theorems 5.1 and 5.2.

The proof is left to the reader (see Problem 15).

Remark. Itô's formula (5.7) asserts that the two processes $f(\xi(t), t) - f(\xi(0), 0)$ and

$$\int_0^t [f_s(\xi(s), s) + f_x(\xi(s), s) a(s) + \tfrac{1}{2} f_{xx}(\xi(s), s) b^2(s)]\, ds + \int_0^t f_x(\xi(s), s) b(s)\, dw(s)$$

are stochastically equivalent. Since they are continuous, their sample paths coincide a.s. Consequently

$$f(\xi(\tau), \tau) - f(\xi(0), 0) = \int_0^\tau [f_t(\xi(t), t) + f_x(\xi(t), t) a(t) + \tfrac{1}{2} f_x(\xi(t), t) b^2(t)]\, dt + \int_0^\tau f_x(\xi(t), t) b(t)\, dw(t) \tag{5.20}$$

for any random variable τ, $0 < \tau < T$.

If, in particular, τ is a stopping time, when, taking the expectation and using Theorem 3.10, we find that

$$Ef(\xi(\tau), \tau) - Ef(\xi(0), 0) = E \int_0^\tau (Lf)(\xi(t), t)\, dt \tag{5.21}$$

where

$$Lf \equiv f_t + af_x + \tfrac{1}{2} b^2 f_{xx},$$

provided

$$b(t) f_x(\xi(t, t)) \quad \text{belongs to} \quad M_w^2[0, T],$$
$$(Lf)(\xi(t), t) \quad \text{belongs to} \quad M_w^1[0, T].$$

6. Applications of Itô's formula

Itô's formula will become a standard tool in the sequel. In the present section we give a few straightforward applications. First we need two lemmas.

Lemma 6.1. *If $f \in L_w^p[\alpha, \beta]$ for some $p \geq 1$, then there exists a sequence of step functions f_n in $L_w^p[\alpha, \beta]$ such that*

$$\lim_{n\to\infty} \int_\alpha^\beta |f(t) - f_n(t)|^p\, dt \quad a.s.$$

Proof. The proof is similar to the proof of Lemma 1.1. Instead of (1.5) we now have

$$\int_\alpha^\beta |J_\epsilon f|^p\, dt \leq \int_\alpha^\beta |f|^p\, dt. \tag{6.1}$$

Indeed, from (1.4) and Hölder's inequality,

$$\int_\alpha^\beta |J_\epsilon f|^p\, dt \leq \int_\alpha^\beta \left[\int_{-\epsilon}^\epsilon \rho(z)\, dz\right]^{p/q} \left[\int_{-\epsilon}^\epsilon \rho(z) |f(t - z - \epsilon)|^p\, dz\right] dt$$

$$\leq \int_{-\epsilon}^\epsilon \rho(z) \left[\int_{\alpha-1}^\beta |f(t)|^p\, dt\right] dz \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right)$$

and (6.1) follows. (If $p = 1$, we do not use Hölder's inequality.)

Using (6.1) we can proceed to show that

$$\int_\alpha^\beta |J_\epsilon f - f|^p\, dt \to 0 \qquad \text{if} \quad \epsilon \downarrow 0, \tag{6.2}$$

by the argument given the proof of Lemma 1.1. The rest of the proof is the same as for Lemma 1.1.

Lemma 6.2. *If $f \in M_w^p[\alpha, \beta]$ for some $p \geqslant 1$, then there exists a sequence of bounded step functions f_n in $M_w^p[\alpha, \beta]$ such that*

$$E \int_\alpha^\beta |f(t) - f_n(t)|^p \, dt \to 0 \qquad \text{if} \quad n \to \infty.$$

This follows by obvious changes in the proof of Lemma 1.2, exploiting Lemma 6.1.

Theorem 6.3. *Let $f \in M_w^{2m}[0, T]$ where m is a positive integer. Then*

$$E\left[\int_0^T f(t)\, dw(t)\right]^{2m} \leqslant [m(2m-1)]^m T^{m-1}\left[E \int_0^T f^{2m}(t)\, dt\right]. \tag{6.3}$$

For $m = 1$, there is equality.

Proof. We apply Itô's formula with $f(x, t) = x^{2m}$, $\xi(t) = \int_0^t f(s)\, dw(s)$, and obtain

$$\left[\int_0^t f(s)\, dw(s)\right]^{2m} = 2m \int_0^t \left[\int_0^s f(\lambda)\, dw(\lambda)\right]^{2m-1} f(s)\, dw(s)$$
$$+ m(2m-1)\int_0^t \left[\int_0^s f(\lambda)\, dw(\lambda)\right]^{2m-2} f^2(s)\, ds. \tag{6.4}$$

Suppose $f(t)$ is a bounded step function. From the definition of $\int_0^s f(\lambda)\, dw(\lambda)$ and the fact that

$$E|w(t)|^r < C_{r,t} \qquad (C_{r,t} \text{ constant})$$

for any $t > 0$, $r > 0$, we find that

$$\left[\int_0^s f(\lambda)\, dw(\lambda)\right]^{2m-1} f(s) \qquad \text{belongs to} \qquad M_w^2[0, T].$$

Hence, taking the expectation in (6.4) and using Theorem 2.5, we get

$$E\left[\int_0^T f(s)\, dw(s)\right]^{2m} = m(2m-1)\int_0^T E\left[\int_0^s f(\lambda)\, dw(\lambda)\right]^{2m-2} f^2(s)\, ds. \tag{6.5}$$

By Hölder's inequality, the right-hand side is bounded by

$$m(2m-1)\left\{\int_0^T E\left[\int_0^s f(\lambda)\, dw(\lambda)\right]^{2m} ds\right\}^{(2m-2)/2m} \left\{\int_0^T Ef^{2m}(s)\, ds\right\}^{2/2m}. \tag{6.6}$$

From (6.5) we see that the function

$$t \to E\left[\int_0^t f(s)\, dw(s)\right]^{2m}$$

is monotone increasing in t. Hence

$$\int_0^T E\left[\int_0^s f(\lambda)\, dw(\lambda)\right]^{2m} ds \leqslant \int_0^T E\left[\int_0^T f(\lambda)\, dw(\lambda)\right]^{2m} ds.$$

Using this to estimate the first expression $\{\cdots\}$ in (6.6), we then obtain from (6.5),

$$E\left[\int_0^T f(s)\, dw(s)\right]^{2m}$$
$$\leqslant m(2m-1)\left\{TE\left[\int_0^T f(s)\, dw(s)\right]^{2m}\right\}^{(m-1)/m}\left\{\int_0^T Ef^{2m}(s)\, ds\right\}^{1/m},$$

and (6.3) follows.

To prove (6.3) in general, we can take, by Lemma 6.2, a sequence of bounded step functions f_n such that

$$E\int_0^T |f(t)-f_n(t)|^{2m}\, dt \to 0 \quad \text{if} \quad n\to\infty. \tag{6.7}$$

Then

$$\int_0^T f_n(t)\, dw(t) \xrightarrow{P} \int_0^T f(t)\, dw(t).$$

We may assume that the convergence is a.s., for otherwise we can take a subsequence of the f_n. Writing (6.3) for f_n, and using Fatou's lemma and (6.7), the assertion (6.3) follows.

From Theorems 3.7 and 6.3 we obtain:

Corollary 6.4. *If $f\in M_w^{2m}[0,T]$ where m is a positive integer, then*

$$E\left\{\sup_{0<t<T}\left|\int_0^t f(s)\, dw(s)\right|^{2m}\right\} \leqslant C_m T^{m-1} E\int_0^T |f(t)|^{2m}\, dt$$

where $C_m = [4m^3/(2m-1)]^m$.

Theorem 6.5. *Let $f\in L_w^2[0,T]$, and let α, β be any positive numbers. Then*

$$P\left\{\max_{0<t<T}\left[\int_0^t f(\lambda)\, dw(\lambda) - \frac{\alpha}{2}\int_0^t f^2(\lambda)\, d\lambda\right] > \beta\right\} \leqslant e^{-\alpha\beta}. \tag{6.8}$$

Proof. Consider first the case where f is a bounded step function, and set

$$\xi(t) = \int_0^t f(\lambda)\, dw(\lambda) - \tfrac{1}{2}\int_0^t f^2(\lambda)\, d\lambda,$$
$$\zeta(t) = \exp[\xi(t)].$$

Applying Itô's formula to $f(x) = e^x$ and the process $\xi(t)$, we get

$$\zeta(t) = \zeta(s) + \int_s^t e^{\xi(\lambda)} f(\lambda)\, dw(\lambda).$$

Since f is a bounded step function,

$$e^{\xi(\lambda)} \leqslant A_0 \prod_{i=1}^{l} e^{A_i |w(t_i)|}$$

for some positive numbers A_i, t_i. It follows that $[\exp(\xi(\lambda))] f(\lambda)$ belongs to $M_w^2[0, T]$. Hence, by Theorem 2.8,

$$E[\zeta(t) | \mathcal{F}_s] = \zeta(s), \tag{6.9}$$

i.e., $\zeta(t)$ is a martingale. Note also that $E\zeta(0) = 1$.

Set

$$\xi_\alpha(t) = \int_0^t f(\lambda)\, dw(\lambda) - \frac{\alpha}{2} \int_0^t f^2(\lambda)\, d\lambda,$$

$$\hat{\xi}_\alpha(t) = \int_0^t \alpha f(\lambda)\, dw(\lambda) - \frac{\alpha^2}{2} \int_0^t f^2(\lambda)\, d\lambda.$$

By what we have already proved, $\exp(\hat{\xi}_\alpha(t))$ is a martingale with expectation 1. Hence, by the martingale inequality,

$$P\Big\{ \max_{0 \leqslant t \leqslant T} \xi_\alpha(t) > \beta \Big\} = P\Big\{ \max_{0 \leqslant t \leqslant T} [\exp \hat{\xi}_\alpha(t)] > e^{\alpha\beta} \Big\} \leqslant e^{-\alpha\beta}.$$

This proves (6.8) in case f is a bounded step function.

Now let f be any function in $L_w^2[0, T]$ and take bounded step functions f_n such that

$$\int_0^T |f_n(t) - f(t)|^2\, dt \to 0 \quad \text{a.s.}$$

Theorem 3.4 implies that

$$\max_{0 \leqslant t \leqslant T} \left| \int_0^t f_n(s)\, dw(s) - \int_0^t f(s)\, dw(s) \right| \to 0 \quad \text{in probability.}$$

By taking a subsequence, if necessary, we obtain convergence a.s. It follows that a.s.

$$\exp\left[\int_0^t f_n(\lambda)\, dw(\lambda) - \frac{\alpha}{2} \int_0^t f_n^2(\lambda)\, d\lambda \right]$$

$$\to \exp\left[\int_0^t f(\lambda)\, dw(\lambda) - \frac{\alpha}{2} \int_0^t f^2(\lambda)\, d\lambda \right]$$

uniformly with respect to t, $0 \leqslant t \leqslant T$. Hence, by writing (6.8) for each f_n and taking $n \to \infty$, the assertion (6.8) follows.

The inequality (6.8) will be referred to as the *exponential martingale inequality.*

Corollary 6.6. *If $f \in L_w^2[0, T]$, then the process*

$$\zeta(t) = \exp\left\{ \int_0^t f(\lambda)\, dw(\lambda) - \tfrac{1}{2} \int_0^t f^2(\lambda)\, d\lambda \right\} \tag{6.10}$$

is a supermartingale, i.e., $-\zeta(t)$ is a submartingale.

Proof. Define

$$\xi_n(t) = \int_0^s f(\lambda)\, dw(\lambda) + \int_s^t f_n(\lambda)\, dw(\lambda) - \tfrac{1}{2} \int_0^s f^2(\lambda)\, d\lambda - \tfrac{1}{2} \int_s^t f_n^2(\lambda)\, d\lambda,$$

$$\zeta_n(t) = \exp(\xi_n(t))$$

for $s \leqslant t \leqslant T$, where the f_n are step functions as in the preceding proof. By the proof of (6.9), $\zeta_n(t)$ is a martingale for $s \leqslant t \leqslant T$. Hence, for any event $A \in \mathcal{F}_s$,

$$\int_A \zeta_n(t)\, dP = \int_A \zeta_n(s)\, dP. \tag{6.11}$$

Notice that $\zeta_n(s) = \zeta(s)$ where $\zeta(t)$ is defined by (6.10). If $n = n' \to \infty$ ($\{n'\}$ a suitable subsequence of $\{n\}$), then $\zeta_n(t) \to \zeta(t)$ a.s. Hence, taking $n \to \infty$ in (6.11), and using Fatou's lemma, we get

$$\int_A \zeta(t)\, dP \leqslant \int_A \zeta(s)\, dP,$$

and the assertion follows.

7. Stochastic integrals and differentials in n dimensions

Let $w(t) = (w_1(t), \ldots, w_n(t))$ be an n-dimensional Brownian motion. Let $\mathcal{F}_t$ ($t \geqslant 0$) be an increasing family of σ-fields such that $w(t)$ is $\mathcal{F}_t$ measurable and $\mathcal{F}(w(\lambda + t) - w(t), \lambda \geqslant 0)$ is independent of $\mathcal{F}_t$, for any $t \geqslant 0$.

We shall say that a matrix of functions belongs to $L_w^p[\alpha, \beta]$ (or to $M_w^p[\alpha, \beta]$) if each of its elements belongs to $L_w^p[\alpha, \beta]$ (or to $M_w^p[\alpha, \beta]$).

Let $b = (b_{ij})$ be an $m \times n$ matrix that belongs to $L_w^2[\alpha, \beta]$. The *stochastic integral* $\int_\alpha^\beta b(t)\, dw(t)$ is an m-vector defined by

$$\int_\alpha^\beta b(t)\, dw(t) = \left\{ \sum_{j=1}^n \int_\alpha^\beta b_{ij}(t)\, dw_j(t) \right\}_{i=1,\ldots,m}$$

where each integral on the right is defined as in Section 2.

If we substitute $\alpha = \int_{t_1}^{t_2} f\, dw_i$, $\beta = \int_{t_1}^{t_2} g\, dw_i$ in the identity $4\alpha\beta = (\alpha + \beta)^2 - (\alpha - \beta)^2$, and use Theorem 2.5, we find that

$$E \int_{t_1}^{t_2} f(t)\, dw_i(t) \int_{t_1}^{t_2} g(t)\, dw_i(t) = E \int_{t_1}^{t_2} f(t) g(t)\, dt \tag{7.1}$$

provided f and g belong to $M_w^2[t_1, t_2]$.

We also have

$$E \int_{t_1}^{t_2} f(t)\, dw_i(t) \int_{t_1}^{t_2} g(t)\, dw_j(t) = 0 \qquad \text{if} \quad i \neq j, \tag{7.2}$$

because the integrals are independent and with zero expectation.

Using (7.1), (7.2) we easily see that if $b = (b_{ij})$ is an $m \times n$ matrix in $M_w^2[t_1, t_2]$, then

$$E\left|\int_{t_1}^{t_2} b(t)\, dw(t)\right|^2 = E\int_{t_1}^{t_2} |b(t)|^2\, dt \tag{7.3}$$

where

$$|b|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (b_{ij})^2.$$

Definition. Let $\xi(t)$ be an m-dimensional process for $0 \leqslant t \leqslant T$, and suppose that, for any $0 \leqslant t_1 < t_2 \leqslant T$,

$$\xi(t_2) - \xi(t_1) = \int_{t_1}^{t_2} a(t)\, dt + \int_{t_1}^{t_2} b(t)\, dw(t)$$

where $a = (a_1, \ldots, a_m)$ and the $m \times n$ matrix $b = (b_{ij})$ belong to $L_w^1[0, T]$ and $L_w^2[0, T]$ respectively. Then we say that $\xi(t)$ has a *stochastic differential* $d\xi(t)$ given by

$$d\xi(t) = a(t)\, dt + b(t)\, dw(t). \tag{7.4}$$

We shall now state *Itô's formula.*

Theorem 7.1. *Let $u(x, t)$ be a continuous function in $(x, t) \in R^m \times [0, \infty)$ together with its derivatives u_t, u_{x_i}, $u_{x_i x_j}$. Let $\xi(t)$ be an m-dimensional process having a stochastic differential*

$$d\xi(t) = a(t)\, dt + b(t)\, dw(t)$$

where $a = (a_1, \ldots, a_m)$ and $b = (b_{ij})$ $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$ belong to $L_w^1[0, T]$ and $L_w^2[0, T]$ respectively. Then $u(\xi(t), t)$ has a stochastic differential

$$\begin{aligned} du(\xi(t), t) = \Bigg[& u_t(\xi(t), t) + \sum_{i=1}^{m} u_{x_i}(\xi(t), t) a_i(t) \\ & + \frac{1}{2} \sum_{l=1}^{n} \sum_{i,j=1}^{m} u_{x_i x_j}(\xi(t), t) b_{il}(t) b_{jl}(t) \Bigg] dt \\ & + \sum_{l=1}^{n} \sum_{i=1}^{m} u_{x_i}(\xi(t), t) b_{il}(t)\, dw_l(t). \end{aligned} \tag{7.5}$$

We begin with

Lemma 7.2. *If w_1, w_2 are independent Brownian motions, then*

$$d(w_1 w_2) = w_1\, dw_2 + w_2 dw_1. \tag{7.6}$$

Proof. It is easily seen that $w(t) = (w_1 + w_2)/\sqrt{2}$ is a Brownian motion. Hence, by (5.2)

$$d(w^2) = dt + 2w\, dw.$$

Since $d(w_1^2)$ and $d(w_2^2)$ are also given by the same formula, $d(w_1 w_2)$ exists and is equal to

$$d(w^2 - \tfrac{1}{2} w_1^2 - \tfrac{1}{2} w_2^2) = dw^2 - \tfrac{1}{2}\, dw_1^2 - \tfrac{1}{2}\, dw_2^2 = w_1\, dw_2 + w_2\, dw_1.$$

Lemma 7.3. *If $d\xi_i = a_i(t)\, dt + \sum_{j=1}^n b_{ij}(t)\, dw_j$ $(i = 1, 2)$ where $w_1, \ldots, w_n$ are independent Brownian motions, then*

$$d(\xi_1 \xi_2) = \xi_1\, d\xi_2 + \xi_2\, d\xi_1 + \sum_{j=1}^n b_{1j} b_{2j}.$$

Proof. The proof is similar to the proof of Theorem 5.1. It is based on the special cases (7.6), (5.2), (5.3) and the approximation procedure employed in the proof of Theorem 5.1.

Proof of Theorem 7.1. We begin with the special case of $u(\xi_* + a_* t + b_* w(t), t)$ where $t_1 < t < t_2$ and ξ_*, a_*, b_* are random variables measurable with respect to $\mathcal{F}_{t_1}$. The special case where $u = x_i^k$ follows by induction on m, using Lemma 7.3. The more general case where $u = x_1^{k_1} \cdots x_m^{k_m}$ follows by using the previous special case and Lemma 7.3. The case where $u = g(t)x_1^{k_1} \cdots x_m^{k_m}$ follows by again using Lemma 7.3. Now approximate general u by linear combinations of u's of the last special form.

Once Theorem 7.1 has been proved in the special case where $\xi(t) = \xi_* + a_* t + b_* w(t)$, we can obtain the general case by a proof similar to that of Theorem 5.2.

Let

$$a_{ij} = \sum_{l=1}^n b_{il} b_{jl},$$

$$Lu = \tfrac{1}{2} \sum_{i,j=1}^m a_{ij} \frac{\partial^2 u}{\partial x_i\, \partial x_j} + \sum_{i=1}^m a_i \frac{\partial u}{\partial x_i} + \frac{\partial u}{\partial t}.$$

Then we can write Itô's formula (7.5) in the form

$$du(\xi(t), t) = Lu(\xi(t), t)\, dt + u_x(\xi(t), t) \cdot b(t)\, dw(t). \tag{7.7}$$

Let us define formally a multiplication table:

$$\begin{aligned} dw_i\, dt &= 0, \\ dt\, dt &= 0, \\ dw_i\, dw_j &= 0 \quad \text{if } i \neq j, \\ dw_i\, dw_j &= dt, \end{aligned}$$

so that

$$d\xi_i\, d\xi_j = \sum_{l=1}^{n} b_{il} b_{jl}\, dt. \tag{7.8}$$

Then Itô's formula (7.5) takes the form

$$\begin{aligned} du(\xi(t), t) = u_t(\xi(t), t)\, dt &+ \sum_{i=1}^{m} u_{x_i}(\xi(t), t)\, d\xi_i \\ &+ \tfrac{1}{2} \sum_{i,j=1}^{m} u_{x_i x_j}(\xi(t), t)\, d\xi_i\, d\xi_j. \end{aligned} \tag{7.9}$$

From Itô's formula we obtain (cf. (5.20))

$$\begin{aligned} u(\xi(\tau), \tau) - u(\xi(0), 0) &= \int_0^\tau (Lu)(\xi(s), s)\, ds \\ &\quad + \int_0^\tau u_x(\xi(s), s) \cdot b(s)\, dw(s) \end{aligned} \tag{7.10}$$

for any random variable τ, $0 \leqslant \tau \leqslant T$. If τ is a stopping time and Lu, $u_x \cdot b$ are in $M_w^1[0, T]$ and $M_w^2[0, T]$ respectively, then (cf. (5.21))

$$Eu(\xi(\tau), \tau) - Eu(\xi(0), 0) = E \int_0^\tau Lu(\xi(s), s)\, ds. \tag{7.11}$$

The results of Sections 2–4 extend with obvious changes to stochastic integrals in n dimensions. We shall give here the generalization of Theorem 4.5.

Theorem 7.4. *Let $f = (f_1, \ldots, f_n)$ belong to $L_w^2[0, \infty)$ and suppose that $\int_0^\infty |f|^2\, dt = \infty$ with probability* 1, *and define*

$$\tau(t) = \inf\left\{ s;\ \int_0^s |f(\lambda)|^2\, d\lambda = t \right\}.$$

Then the process

$$u(t) = \int_0^{\tau(t)} f(s)\, dw(s)$$

is a Brownian motion.

Proof. The proof is similar to the proof of Theorem 4.5 once we have established that if $f \in M_w^2[0, \infty)$ and ζ_1, ζ_2 are bounded stopping times, $\zeta_1 \leq \zeta_2$, then

$$E\left\{ \int_{\zeta_1}^{\zeta_2} f(s)\, dw(s) | \mathcal{F}_{\zeta_1} \right\} = 0,$$

$$E\left\{ \left| \int_{\zeta_1}^{\zeta_2} f(s)\, dw(s) \right|^2 | \mathcal{F}_{\zeta_1} \right\} = E\left\{ \int_{\zeta_1}^{\zeta_2} |f|^2\, ds | \mathcal{F}_{\zeta_1} \right\}$$

a.s. The proof of these formulas is similar to the proof of Theorem 4.3. It employs Theorem 2.8 which remains valid for n-dimensional stochastic integrals.

We conclude this section with an extension of Theorem 6.5 to n dimensions.

Theorem 7.5. *Let $f = (f_1, \ldots, f_n)$ belong to $L_w^2[0, T]$, and let α, β be positive numbers. Then*

$$P\left\{ \max_{0 \leq t \leq T} \left[\int_0^t f(\lambda)\, dw(\lambda) - \frac{\alpha}{2} \int_0^t |f(\lambda)|^2\, d\lambda \right] > \beta \right\} \leq e^{-\alpha\beta}. \quad (7.12)$$

The proof is similar to the proof of Theorem 6.5. First we prove (7.12) in case $f(t)$ is a step function, using the martingale inequality, and then proceed to general f by approximation.

The inequality (7.12) is referred to as the *exponential martingale inequality*.

Corollary 6.6 also extends to the present n-dimensional case, i.e.,

$$\exp\left\{ \int_0^t f\, dw - \tfrac{1}{2} \int_0^t |f|^2\, ds \right\}$$

is a supermartingale.

PROBLEMS

1. Prove (2.20).
2. Prove Theorem 3.9 [*Hint*: Apply Theorem 3.6 to $\xi f(t)$, ξ bounded and $\mathcal{F}_\alpha$ measurable.]
3. Suppose $f \in L_w^2[0, \infty)$ and ζ is a stopping time such that $E \int_0^\zeta f^2(t)\, dt < \infty$. Prove that

$$E \int_0^\zeta f(t)\, dw(t) = 0, \qquad E\left| \int_0^\zeta f(t)\, dw(t) \right|^2 = E \int_0^\zeta f^2(t)\, dt.$$

4. Let

$$\rho(x) = \begin{cases} c \exp[1/(|x|^2 - 1)] & \text{if } |x| < 1 \\ 0 & \text{if } |x| > 1 \end{cases}$$

for $x \in R^n$, where c is a positive constant such that $\int_{R^n} \rho(x)\,dx = 1$. If f is a function locally integrable, then

$$(J_\epsilon f)(x) = \frac{1}{\epsilon^n} \int_{R^n} \rho\left(\frac{x-y}{\epsilon}\right) f(y)\,dy$$

is called a *mollifier* of f. Prove:

(i) $J_\epsilon f$ is in $C^\infty(R^n)$;

(ii) If K is a compact set and Ω a bounded open set containing K, then

$$(J_\epsilon f)(x) = \frac{1}{\epsilon^n} \int_{\Omega} \rho\left(\frac{x-y}{\epsilon}\right) f(y)\,dy$$

$$= \int_{|z|<1} \rho(z) f(x - \epsilon z)\,dz \qquad (x \in K),$$

provided $\epsilon < \operatorname{dist}(K, R^n \backslash \Omega)$.

(iii) If $f \in L^p(\Omega)$ for some $p \geqslant 1$, then

$$\left\{ \int_K |J_\epsilon f|^p\,dx \right\}^{1/p} \leqslant \left\{ \int_\Omega |f|^p\,dx \right\}^{1/p}.$$

(iv) If $f \in L^p(\Omega)$ for some $p \geqslant 1$, then

$$\int_K |J_\epsilon f - f|^p\,dx \to 0 \quad \text{if} \quad \epsilon \to 0.$$

5. Let $f(x)$ be a continuous function for $\alpha \leqslant x \leqslant \beta$, and let

$$(P_k f)(x) = \frac{\int_\alpha^\beta \left[1 - (x-y)^2\right]^k f(y)\,dy}{\int_{-1}^1 (1-y^2)^k\,dy} \qquad (k = 1, 2, \ldots).$$

Let δ be any positive number. Prove that $(P_k f)(x) \to f(x)$ uniformly in $x \in [\alpha + \delta, \beta - \delta]$ as $k \to \infty$. [*Hint*: $[\int_\epsilon^1 (1-y^2)^k\,dy / \int_0^1 (1-y^2)^k\,dy] \to 0$ if $k \to \infty$, for any $\epsilon > 0$.]

6. Let $f(x)$ be a continuous function in an n-dimensional interval $\mathrm{I} \equiv \{\mathrm{x};\ \alpha_\mathrm{i} \leqslant \mathrm{x} \leqslant \beta_\mathrm{i},\ 1 \leqslant \mathrm{i} \leqslant \mathrm{n}\}$, and let

$$(P_k f)(x) = \frac{\int_{\alpha_1}^{\beta_1} \cdots \int_{\alpha_n}^{\beta_n} \prod_{i=1}^n \left[1 - (x_i - y_i)^2\right]^k f(y)\,dy_n \cdots dy_1}{\left[\int_{-1}^1 (1-y^2)^k\,dy\right]^n}$$

$$(k = 1, 2, \ldots).$$

Let I_0 be any subset lying in the interior of I. Prove that, as $k \to \infty$,

$$(P_k f)(x) \to f(x) \quad \text{uniformly in} \quad x \in I_0.$$

Notice that $P_k f$ is a polynomial. It is called a ***polynomial mollifier*** of f.

7. If in the preceding problem f belongs to $C^m(I)$ and f vanishes in a neighborhood of the boundary of I, then

$$\frac{\partial^{i_1+\cdots+i_n}}{\partial x_1^{i_1}\cdots\partial x_n^{i_n}}(P_kf)(x)\to\frac{\partial^{i_1+\cdots+i_n}}{\partial x_1^{i_1}\cdots\partial x_n^{i_n}}f(x)\qquad\text{if}\quad k\to\infty,$$

uniformly in $x\in I_0$, for any $(i_1,\ldots,i_n)$ such that $0\leqslant i_1+\cdots+i_n\leqslant m$.

8. If $f\in C^m(R^n)$, then there exists a sequence of polynomials Q_k such that, as $k\to\infty$,

$$\frac{\partial^{i_1+\cdots+i_n}}{\partial x_1^{i_1}\cdots\partial x_n^{i_n}}Q_k(x)\to\frac{\partial^{i_1+\cdots+i_n}}{\partial x_1^{i_1}\cdots\partial x_n^{i_n}}f(x)\qquad\text{for}\quad 0\leqslant i_1+\cdots+i_n\leqslant m,$$

uniformly in x in compact subsets of R^n. [*Hint*: Approximate f by functions with compact support, and apply Problem 7 to these functions.]

9. If in the previous problem it is assumed that f, f_{x_i} $(1\leqslant i\leqslant n)$ and $f_{x_ix_j}$ $(2\leqslant i,j\leqslant n)$ are continuous in R^n (instead of $f\in C^m(R^n)$), then

$$Q_k\to f,\qquad \frac{\partial}{\partial x_i}Q_k\to\frac{\partial f}{\partial x_i}\qquad(1\leqslant i\leqslant n),$$

$$\frac{\partial^2}{\partial x_i\,\partial x_j}Q_k\to\frac{\partial^2 f}{\partial x_i\,\partial x_j}\qquad(2\leqslant i,j\leqslant n)$$

uniformly on compact subsets of R^n.

10. Let $f(x,t)$ be a continuous function in $(x,t)\in \mathrm{R}^n\times[0,\infty)$ together with its derivatives f_t, f_{x_i}, $f_{x_ix_j}$. Prove that there exists a function F continuous in $(x,t)\in R^n\times R^1$ together with its derivatives F_t, F_{x_i}, $F_{x_ix_j}$, such that $F(x,t)=f(x,t)$ if $x\in R^n$, $t\geqslant 0$.

11. Let $f(x,t)=f(x_1,\ldots,x_n,t)$ be a continuous function in $(x,t)\in R^n\times[0,\infty)$ together with its derivatives f_t, f_{x_i}, $f_{x_ix_j}$. Then there exists a sequence of polynomials $Q_m(x,t)$ such that, as $m\to\infty$,

$$Q_m\to f,\qquad\frac{\partial}{\partial t}Q_m\to f_t,\qquad\frac{\partial}{\partial x_i}Q_m\to f_{x_i},\qquad\frac{\partial^2}{\partial x_i\,\partial x_j}Q_m\to f_{x_ix_j}$$

uniformly in compact subsets. [*Hint*: Combine Problems 9, 10.]

12. Prove (5.8).

13. Prove (5.14) and complete the proof of (5.13).

14. Let $f\in L^2_w[0,\infty)$, $|f|\leqslant K$ (K constant) and let $d\xi(t)=f(t)\,dw(t)$, $\xi(0)=0$ where $w(t)$ is a Brownian motion. Prove:

(i) if $f\leqslant\beta$, then $E|\xi(t)|^2\leqslant\beta^2t$;

(ii) if $f\geqslant\alpha>0$, then $E|\xi(t)|^2\geqslant\alpha^2t$.

15. Prove Theorem 5.3. [*Hint*: Proceed as in the proof of Theorem 5.2, but with

$$\Phi(w(t),t)=f(\xi_{10}+a_1t+b_1w(t),\ldots,\xi_{m0}+a_mt+b_mw(t))$$

where ξ_{i0}, a_i are random variables and the b_i are random n-vectors; cf. Step 4.]

16. Let $\xi(t) = \int_0^t b(t)\,dw(t)$ where b is an $n \times n$ matrix belonging to $L_w^2[0, \infty)$. Suppose that $d\xi_i\,d\xi_j = 0$ if $i \neq j$, $d\xi_i\,d\xi_i = dt$ (see (7.8) for the definition of $d\xi_i\,d\xi_j$), for all $1 \leqslant i, j \leqslant n$. Prove that $\xi(t)$ is an n-dimensional Brownian motion. [*Hint*: First proof: Use Theorem 3.6.2. Second proof: Suppose the elements of b are bounded step functions and let $\zeta(t) = \exp[i\gamma \cdot \xi(t) + \gamma^2 t/2]$. By Itô's formula $d\zeta = i\zeta\gamma\,dw$. By Theorem 2.8

$$E[e^{i\gamma \cdot \xi(t)} | \mathcal{F}_s] = e^{i\gamma \cdot \xi(s)} e^{-\gamma^2(t-s)/2}.$$

Use Problem 2, Chapter 3.]

17. Let $\gamma > 0$, $a > 0$, $\tau = \min\{t;\, w(t) = a\}$ where $w(t)$ is one-dimensional Brownian motion. Prove that $P(\tau < \infty) = 1$ and

$$Ee^{-\gamma\tau} = \exp\left(-\sqrt{2\gamma}\,a\right).$$

[*Hint*: For any $c > 0$,

$$P\Big[\max_{0 \leqslant s \leqslant t} w(s) > c\Big] \leqslant P\Big[\max_{0 \leqslant s \leqslant t}\Big(w(s) - \frac{\alpha}{2}s\Big) > \beta\Big] < e^{-c^2/2t}$$

where $\alpha = c/t$, $\beta = c/2$. Hence $P(\tau < \infty) = 1$. Since $y(t) = \exp[\gamma w(t) - \gamma^2 t/2]$ is a martingale, so is $y(t \wedge \tau)$. Hence

$$E\exp\left[\gamma w(t \wedge \tau) - \tfrac{1}{2}\gamma^2(t \wedge \tau)\right] = 1.$$

Take $t \uparrow \infty$.]

18. Under the conditions of the previous problem

$$P(\tau \in dt) = \frac{a}{(2\pi t^3)^{1/2}} \exp\left(-\frac{a^2}{2t}\right) dt.$$

[*Hint*: Use the fact (see, for instance, Feller [1]) that if the Laplace transform of two probability distributions concentrated on $[0, \infty)$ coincide, then the probability distributions coincide.]

19. If $w(t)$ is a Brownian motion and $0 \leqslant y$, $x < y$, then

$$P\Big(w(t) \in dx,\ \max_{0 \leqslant s \leqslant t} w(s) \in dy\Big) = \left(\frac{2}{\pi t^3}\right)^{1/2} (2y - x) \exp\left[-\frac{(2y-x)^2}{2t}\right] dx\,dy.$$

[*Hint*: Use Problem 12, Chapter 2 and Theorem 3.6.3 to deduce that

$$P\Big[w(t) \in dx,\ \max_{0 \leqslant s \leqslant t} w(s) \geqslant y\Big] = \int_0^t P(\tau \in ds) P[w(t-s) + y \in dx]$$

where $\tau = \min\{t;\, w(t) = y\}$, and apply the preceding problem.]

20. Let $f(t)$ be a continuous process in $L_w^2[0, T]$ and let $\Pi_n : t_{n,0} = 0 < t_{n,1} < \cdots < t_{n,n} = T$ be a partition with mesh $|\Pi_n| \to 0$ as $n \to \infty$. Define

$$g_n(t) = \sum_{i=0}^{j-1} f(t_{n,i})(w(t_{n,i+1}) - w(t_{n,i})) + f(t_{n,j})(t - t_{n,j})$$
$$\text{if} \quad t_{n,j} \leqslant t < t_{n,j+1}.$$

Prove that for some subsequence $\{n'\}$ of $\{n\}$,

$$\sup_{0 \leqslant t \leqslant T} \left| \int_0^t f(s)\, dw(s) - g_{n'}(t) \right| \to 0 \quad \text{a.s.} \qquad \text{if} \quad n' \to \infty.$$

21. Let $\sigma(x, t)$ be a measurable function in $(x, t) \in R^n$ such that

$$|\sigma(x, t) - \sigma(\bar{x}, t)| \leqslant \eta(|x - \bar{x}|), \qquad \eta(\delta) \downarrow 0 \quad \text{if } \delta \downarrow 0,$$

and let $f(t)$ be an n-dimensional continuous process in $L_w^2[0, T]$. Let

$$\sigma_\epsilon(x, t) = \frac{1}{\epsilon} \int_{-1}^{T} \rho\left(\frac{t - s - \epsilon}{\epsilon} \right) \sigma(x, s)\, ds \qquad (2\epsilon < 1)$$

where $\rho(t)$ is defined as in Lemma 1.1 and $\sigma(x, s) = \sigma(x, 0)$ if $-1 < s < 0$. Prove:

(i) $\int_0^T |\sigma(x, t) - \sigma_\epsilon(x, t)|^2\, dt \to 0$ uniformly in x in bounded sets, as $\epsilon \to 0$.

(ii) $\int_0^T |\sigma(f(t), t) - \sigma_\epsilon(f(t), t)|^2\, dt \to 0$ a.s. as $\epsilon \to 0$.

(iii) $\sup_{0 \leqslant t \leqslant T} |\int_0^t \sigma(f(s), s)\, dw(s) - \int_0^t \sigma_{\epsilon_n}(f(s), s)\, dw(s)| \to 0$ a.s. for some sequence $\epsilon_n \downarrow 0$.

[*Hint*: for (i), use the uniform continuity in x of $\int \sigma(x, t)\, dt$ and of $\int \sigma_\epsilon(x, t)\, dt$.]

5

Stochastic Differential Equations

1. Existence and uniqueness

If $\sigma = (\sigma_{ij})$ is a matrix, we write $|\sigma|^2 = \sum_{i,j}|\sigma_{ij}|^2$.

Let $b(x, t) = (b_1(x, t), \ldots, b_n(x, t))$, $\sigma(x, t) = (\sigma_{ij}(x, t))_{i,j=1}^n$ and suppose the functions $b_i(x, t)$, $\sigma_{ij}(x, t)$ are measurable in $(x, t) \in R^n \times [0, T]$. If $\xi(t)$ $(0 \leqslant t \leqslant T)$ is a stochastic process such that

$$d\xi(t) = b(\xi(t), t)\, dt + \sigma(\xi(t), t)\, dw(t), \tag{1.1}$$

$$\xi(0) = \xi_0 \quad \text{a.s.}, \tag{1.2}$$

then we say that $\xi(t)$ satisfies the *system of stochastic differential equations* (1.1) and the *initial condition* (1.2). Note that it is implicitly assumed that $b(\xi(t), t)$ belongs to $L_w^1[0, T]$ and $\sigma(\xi(t), t)$ belongs to $L_w^2[0, T]$.

Theorem 1.1. *Suppose $b(x, t)$, $\sigma(x, t)$ are measurable in $(x, t) \in R^n \times [0, T]$ and*

$$\begin{gathered} |b(x, t) - b(\bar{x}, t)| \leqslant K_*|x - \bar{x}|, \quad |\sigma(x, t) - \sigma(\bar{x}, t)| \leqslant K_*|x - \bar{x}|, \\ |b(x, t)| \leqslant K(1 + |x|), \quad |\sigma(x, t)| \leqslant K(1 + |x|) \end{gathered} \tag{1.3}$$

where K_*, K are constants. Let ξ_0 be any n-dimensional random vector independent of $\mathcal{F}(w(t), 0 \leqslant t \leqslant T)$, such that $E|\xi_0|^2 < \infty$. Then there exists a unique solution of (1.1), (1.2) *in $M_w^2[0, T]$.*

The assertion of uniqueness means that if $\xi_1(t)$, $\xi_2(t)$ are two solutions of (1.1), (1.2) and if they belong to $M_w^2[0, T]$, then

$$P\{\xi_1(t) = \xi_2(t) \text{ for all } 0 \leqslant t \leqslant T\} = 1.$$

Proof. To prove uniqueness, suppose $\xi_1(t)$ and $\xi_2(t)$ are two solutions belonging to $M_w^2[0, T]$. Then

$$\begin{aligned} \xi_1(t) - \xi_2(t) = & \int_0^t [b(\xi_1(s), s) - b_2(\xi_2(s), s)]\, ds \\ & + \int_0^t \sigma(\xi_1(s), s)\, dw(s) - \int_0^t \sigma(\xi_2(s), s)\, dw(s). \end{aligned} \tag{1.4}$$

Set $f_i(s) = \sigma(\xi_i(s), s)$ and note that the stochastic integral $\int_0^t f_i(s)\, dw(s)$ is defined with respect to an increasing family of σ-fields $\mathcal{F}_t^i$ which may depend on i. If $f_i(s)$ is a step function, for $i = 1, 2$, then using the definition of the stochastic integral we get (cf. the proof of Lemma 4.2.2 and formula (4.7.3))

$$E\left|\int_0^t f_1(s)\, dw(s) - \int_0^t f_2(s)\, dw(s)\right|^2 = E\int_0^t |f_1(s) - f_2(s)|^2\, ds. \quad (1.5)$$

By approximation we find that (1.5) is true for any pair f_1, f_2 of nonanticipative functions with respect to $\mathcal{F}_t^1$ and $\mathcal{F}_t^2$ respectively, provided $E\int_0^t |f_i(s)|^2\, ds < \infty$ $(i = 1, 2)$.

Taking the expectation of the squares of the absolute values on both sides of (1.4) and using (1.3) and (1.5) with $f_i(s) = \sigma(\xi_i(s), s)$, we find that

$$E|\xi_1(t) - \xi_2(t)|^2$$
$$\leqslant 2K_*^2 t \int_0^t E|\xi_1(s) - \xi_2(s)|^2\, ds + 2K_*^2 \int_0^t E|\xi_1(s) - \xi_2(s)|^2\, ds.$$

Thus the function $\phi(t) = E|\xi_1(t) - \xi_2(t)|^2$ satisfies

$$\phi(t) \leqslant C\int_0^t \phi(s)\, ds, \qquad \phi(0) = 0,$$

where C is a positive constant. Therefore $\phi(t) \equiv 0$, and the assertion of uniqueness is proved.

To prove the existence of a solution we introduce an increasing family of σ-fields $\mathcal{F}_t$ $(0 \leqslant t \leqslant T)$ such that ξ_0 is $\mathcal{F}_0$ measurable and $w(t)$ is $\mathcal{F}_t$ measurable, and such that $\mathcal{F}(w(t+s) - w(t), 0 \leqslant s \leqslant T - t)$ is independent of $\mathcal{F}_t$, for all $t \geqslant 0$. We can take for instance $\mathcal{F}_t$ to be the σ-field generated by ξ_0 and $\mathcal{F}(w(s), s \leqslant t)$; here we use the assumption that ξ_0 is independent of $\mathcal{F}(w(s), 0 \leqslant s \leqslant T)$.

Define $\xi_0(t) = \xi_0$ and

$$\xi_{m+1}(t) = \xi_0 + \int_0^t b(\xi_m(s), s)\, ds + \int_0^t \sigma(\xi_m(s), s)\, dw(s). \quad (1.6)$$

The inductive assumption is that $\xi_m \in M_w^2[0, T]$ and that

$$E|\xi_{k+1}(t) - \xi_k(t)|^2 \leqslant \frac{(Mt)^{k+1}}{(k+1)!} \qquad \text{for} \quad 0 \leqslant k \leqslant m - 1, \quad (1.7)$$

where M is some positive constant (depending only on K, K_*, T).

Since $\xi_0 \in \mathcal{F}_0$, ξ_{m+1} is well defined if $m = 0$. Further

$$|\xi_1(t) - \xi_0|^2 \leqslant 2\left|\int_0^t b(\xi_0, s)\, ds\right|^2 + 2\left|\int_0^t \sigma(\xi_0, s)\, dw(s)\right|^2.$$

Taking the expectation and using (1.3), we get

$$E|\xi_1(t) - \xi_0|^2 \leqslant 2K^2 t^2(1 + E|\xi_0|^2) + 2K^2 t(1 + E|\xi_0|^2) \leqslant Mt$$

if $M \geqslant 2K^2(T+1)(1+E|\xi_0|^2)$. This implies that $\xi_1 \in M_w^2[0, T]$ and (1.7) holds for $m = 0$.

We now make the inductive assumption for any $m \geqslant 0$ and prove it for $m+1$.

Since $\xi_m \in M_w^2[0, T]$ it follows, using (1.3), that $b(\xi_m(t), t)$ and $\sigma(\xi_m(t), t)$ belong to $M_w^2[0, T]$. Thus the integrals on the right-hand side of (1.6) are well defined.

Next,

$$|\xi_{m+1}(t) - \xi_m(t)|^2 \leqslant 2\left|\int_0^t [b(\xi_m(s), s) - b(\xi_{m-1}(s), s)]\, ds\right|^2 + 2\left|\int_0^t [\sigma(\xi_m(s), s) - \sigma(\xi_{m-1}(s), s)]\, dw(s)\right|^2. \tag{1.8}$$

Taking the expectation and using (1.3),

$$E|\xi_{m+1}(s) - \xi_m(s)|^2 \leqslant 2K_*^2 tE\int_0^t |\xi_m(s) - \xi_{m-1}(s)|^2\, ds + 2K_*^2 E\int_0^t |\xi_m(s) - \xi_{m-1}(s)|^2\, ds.$$

Thus,

$$E|\xi_{m+1}(t) - \xi_m(t)|^2 \leqslant M\int_0^t E|\xi_m(s) - \xi_{m-1}(s)|^2\, ds$$

if $M \geqslant 2K_*^2(T+1)$. Substituting (1.7) with $k = m-1$ into the right-hand side, we get

$$E|\xi_{m+1}(t) - \xi_m(t)|^2 \leqslant M\int_0^t \frac{(Ms)^m}{m!}\, ds = \frac{(Mt)^{m+1}}{(m+1)!}.$$

Thus (1.7) holds for $k = m$. Since this implies that $\xi_{m+1} \in M_w^2[0, T]$, the proof of the inductive assumption for $m+1$ is complete.

From (1.8) we also have

$$\sup_{0 \leqslant t \leqslant T} |\xi_{m+1}(t) - \xi_m(t)|^2 \leqslant 2TK_*^2 \int_0^T |\xi_m(s) - \xi_{m-1}(s)|^2\, ds + 2\sup_{0 \leqslant t \leqslant T}\left|\int_0^t [\sigma(\xi_m(s), s) - \sigma(\xi_{m-1}(s), s)]\, dw(s)\right|^2.$$

Taking the expectation and using Theorem 4.3.6 and (1.7), we find that

$$E\sup_{0 \leqslant t \leqslant T} |\xi_{m+1}(t) - \xi_m(t)|^2 \leqslant 2K_*^2 T\int_0^T E|\xi_m(s) - \xi_{m-1}(s)|^2\, ds + 8K_*^2\int_0^T E|\xi_m(s) - \xi_{m-1}(s)|^2\, ds \leqslant C\frac{(MT)^m}{m!}$$

where $C = 2K_*^2T^2 + 8K_*^2T$. Hence

$$P\left\{ \sup_{0 \leq t \leq T} |\xi_{m+1}(t) - \xi_m(t)| > \frac{1}{2^m} \right\} \leq 2^{2m}C \frac{(MT)^m}{m!} .$$

Since $\Sigma[2^m(MT)^m/m!] < \infty$, the Borel–Cantelli lemma implies that

$$P\left\{ \sup_{0 \leq t \leq T} |\xi_{m+1}(t) - \xi_m(t)| > \frac{1}{2^m} \text{ i.o.} \right\} = 0.$$

Thus, for almost any ω there is a positive integer $m_0 = m_0(\omega)$ such that

$$\sup_{0 \leq t \leq T} |\xi_{m+1}(t) - \xi_m(t)| \leq \frac{1}{2^m} \qquad \text{if} \quad m \geq m_0(\omega).$$

It follows that the partial sums

$$\xi_0 + \sum_{m=0}^{k-1} (\xi_{m+1}(t) - \xi_m(t)) = \xi_k(t)$$

are convergent uniformly in $t \in [0, T]$. Denote the limit by $\xi(t)$. Then $\xi(t)$ is a continuous process. It is clearly also a nonanticipative function and it belongs to $L_w^2[0, T]$. Since for a.a. ω,

$$b(\xi_m(t), t) \to b(\xi(t), t) \quad \text{uniformly in} \quad t \in [0, T],$$
$$\sigma(\xi_m(t), t) \to \sigma(\xi(t), t) \quad \text{uniformly in} \quad t \in [0, T],$$

and hence also

$$\int_0^T |\sigma(\xi_m(t), t) - \sigma(\xi(t), t)|^2 \xrightarrow{P} 0,$$

if we take $m \to \infty$ in (1.6) we obtain the relation

$$\xi(t) = \xi_0 + \int_0^t b(\xi(s), s)\, ds + \int_0^t \sigma(\xi(s), s)\, dw(s). \tag{1.9}$$

Thus $\xi(t)$ is a solution of (1.1), (1.2).

From (1.6) we have

$$E|\xi_{m+1}(t)|^2 \leq 3E|\xi_0|^2 + 3E\left| \int_0^t b(\xi_m(s), s)\, ds\right|^2 + 3E\left| \int_0^t \sigma(\xi_m(s), s)\, dw\right|^2$$
$$\leq C\left(1 + E|\xi_0|^2\right) + C\int_0^t E|\xi_m(s)|^2\, ds$$

where C is some constant depending only on K, T. By induction we then get

$$E|\xi_{m+1}(t)|^2 \leq \left[C + C^2t + C^3\frac{t^2}{2!} + \cdots + C^{m+2}\frac{t^{m+1}}{(m+1)!} \right]\left[1 + E|\xi_0|^2\right].$$

Therefore

$$E|\xi_{m+1}(t)|^2 \leq C(1 + E|\xi_0|^2)e^{Ct}.$$

Taking $m \uparrow \infty$ and using Fatou's lemma, we conclude that

$$E|\xi(t)|^2 \leqslant C(1 + E|\xi_0|^2)e^{Ct}. \tag{1.10}$$

This implies that $\xi(t)$ belongs to $M_w^2[0, T]$.

The above method used to prove the existence of a solution $\xi(t)$ is called the *method of successive approximations*; it is modeled after the corresponding proof for ordinary differential equations.

Remark. Very often we shall take the initial value ξ_0 to be a constant function x, i.e., $\xi_0(\omega) = x$ a.s. Notice that this random variable is independent of $\mathcal{F}(w(t), t \geqslant 0)$.

From (1.9) we obtain

$$\sup_{0 \leqslant t \leqslant T} |\xi(t)|^2 \leqslant 3|\xi_0|^2 + 3\left[\int_0^T |b(\xi(s), s)|\, ds\right]^2 + 3 \sup_{0 \leqslant t \leqslant T} \left|\int_0^t \sigma(\xi(s), s)\, dw(s)\right|^2.$$

Taking the expectation and using (1.3) and Theorem 4.3.6, we get

$$E \sup_{0 \leqslant t \leqslant T} |\xi(t)|^2 \leqslant C_0(1 + E|\xi_0|^2) + C_0 \int_0^T E|\xi(s)|^2\, ds$$

where C_0 is a constant depending only on K, T. Making use of (1.10), we obtain

Corollary 1.2. *Under the assumptions of Theorem* 1.1,

$$E\Big[\sup_{0 \leqslant t \leqslant T} |\xi(t)|^2\Big] \leqslant C^*(1 + E|\xi_0|^2) \tag{1.11}$$

where C^* is a constant depending only on K, T.

2. Stronger uniqueness and existence theorems

Theorem 2.1. *Suppose $b_i(x, t)$, $\sigma_i(x, t)$ are measurable functions in $(x, t) \in R^n \times [0,T]$, for $i = 1, 2$, satisfying*

$$|b_i(x, t) - b_i(\bar{x}, t)| \leqslant K_*|x - \bar{x}|, \qquad |\sigma_i(x, t) - \sigma_i(\bar{x}, t)| \leqslant K_*|x - \bar{x}|,$$
$$|b_i(x, t)| \leqslant K(1 + |x|), \qquad |\sigma_i(x, t)| \leqslant K(1 + |x|).$$

Let D be a domain in R^n and suppose that

$$b_1(x, t) = b_2(x, t), \quad \sigma_1(x, t) = \sigma_2(x, t) \qquad \text{if} \quad x \in D,\ 0 \leqslant t \leqslant T. \tag{2.1}$$

Let $\xi_i(t)$ $(i = 1, 2)$ be the solution of

$$d\xi_i(t) = b_i(\xi_i(t), t)\, dt + \sigma_i(\xi_i(t), t), \qquad \xi_i(0) = \xi_{i0}$$

in $M_w^2[0, T]$ (with the same family of σ-fields $\mathcal{F}_t$) where $E|\xi_{i0}|^2 < \infty$. Assume finally that $\xi_{10} = \xi_{20}$ for a.a. ω for which either $\xi_{10}(\omega) \in D$ or $\xi_{20}(\omega) \in D$. Denote by τ_i the first time $\xi_i(t)$ intersects $R^n \setminus D$ if such time $t \leqslant T$ exists, and $\tau_i = T$ otherwise. Then

$$P(\tau_1 = \tau_2) = 1,$$

$$P\Big\{\sup_{0 \leqslant s \leqslant \tau_1} |\xi_1(s) - \xi_2(s)| = 0\Big\} = 1.$$

Thus, if two stochastic equation have the same coefficients in a cylinder $Q = D \times [0, T]$ and if the initial conditions coincide in D, then the corresponding solutions agree until the first time they both leave D; they first leave D at the same time. This is a local uniqueness theorem. It remains true (with similar proof) for the general domains Q.

Proof. Let $\phi_i(t) = 1$ if $\xi_i(s) \in D$ for all $0 \leqslant s \leqslant t$, and $\phi_i(t) = 0$ in all other cases. Then

$$\phi_1(t)(\xi_{10} - \xi_{20}) = 0 \quad \text{a.s.}$$

Hence

$$\begin{aligned}
\phi_1(t)[\xi_1(t) - \xi_2(t)] &= \phi_1(t) \int_0^t [b_1(\xi_1(s), s) - b_2(\xi_1(s), s)]\, ds \\
&\quad + \phi_1(t) \int_0^t [b_2(\xi_1(s), s) - b_2(\xi_2(s), s)]\, ds \\
&\quad + \phi_1(t) \int_0^t [\sigma_1(\xi_1(s), s) - \sigma_2(\xi_1(s), s)]\, dw(s) \\
&\quad + \phi_1(t) \int_0^t [\sigma_2(\xi_1(s), s) - \sigma_2(\xi_2(s), s)]\, dw(s) \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

If $\phi_1(t) = 1$, then $b_1(\xi_1(s), s) = b_2(\xi_1(s), s)$. Hence $I_1 = 0$. If $t \leqslant \tau_1$, then $\sigma_1(\xi_1(s), s) = \sigma_2(\xi_1(s), s)$. Hence, by Theorem 4.4.7, $I_3 = 0$ a.s. if $t \leqslant \tau_1$. Since also $\phi_1(t) = 0$ when $t > \tau_1$, we have $I_3 = 0$ a.s. Thus

$$\begin{aligned}
\phi_1(t)|\xi_1(t) - \xi_2(t)|^2 &\leqslant 2\phi_1(t) \left| \int_0^t [b_2(\xi_1(s), s) - b_2(\xi_2(s), s)]\, ds \right|^2 \\
&\quad + 2\phi_1(t) \left| \int_0^t [\sigma_2(\xi_1(s), s) - \sigma_2(\xi_2(s), s)]\, dw(s) \right|^2. \qquad (2.2)
\end{aligned}$$

Since $\phi_1(t) \leqslant \phi_1(s)$ if $s \leqslant t$,

$$\phi_1(t)\left| \int_0^t [b_2(\xi_1(s), s) - b_2(\xi_2(s), s)]\, ds \right|^2$$
$$\leqslant n\left\{ \int_0^t \phi_1(s)\,|\,b_2(\xi_1(s), s) - b_2(\xi_2(s), s)|\, ds \right\}^2. \tag{2.3}$$

Next, if $f \in L_w^2[0, T]$ and $t \in [0, T]$,

$$\phi_1(t)\left| \int_0^t f(s)\, dw(s) \right|^2 \leqslant \left| \int_0^t \phi_1(s) f(s)\, dw(s) \right|^2. \tag{2.4}$$

Indeed, let $A = \{\phi_1(t) = 1\}$, $B = \{\phi_1(t) = 0\}$. If $\omega \in B$, then the left-hand side of (2.4) vanishes; hence (2.4) is true. If $\omega \in A$, then $\phi_1(s)f(s) = f(s)$ if $s < t$; hence, by Lemma 4.2.11,

$$\int_0^t \phi_1(s) f(s)\, dw(s) = \int_0^t f(s)\, dw(s),$$

and (2.4) again follows.

Using (2.3), (2.4) in (2.2), we find that

$$\phi_1(t)|\xi_1(t) - \xi_2(t)|^2 \leqslant 2n\left[\int_0^t \phi_1(s)|b_2(\xi_1(s), s) - b_2(\xi_2(s), s)|\, ds \right]^2$$
$$+ 2\left| \int_0^t \phi_1(s)[\sigma_2(\xi_1(s), s) - \sigma_2(\xi_2(s), s)]\, dw(s) \right|^2.$$

Using the Lipschitz continuity of $b_2(x, t)$, $\sigma_2(x, t)$ in x, we get

$$E\phi_1(t)|\xi_1(t) - \xi_2(t)|^2 \leqslant C \int_0^t E\phi_1(s)|\xi_1(s) - \xi_2(s)|^2\, ds \qquad (C \text{ constant}).$$

This implies that

$$E\phi_1(t)|\xi_1(t) - \xi_2(t)|^2 = 0.$$

Hence, by the continuity of $\xi_1(t)$, $\xi_2(t)$,

$$\sup_{0 \leqslant t \leqslant T} \phi_1(t)|\xi_1(t) - \xi_2(t)| = 0 \quad \text{a.s.}$$

It follows that $\xi_1(t, \omega) = \xi_2(t, \omega)$ a.s. if $0 \leqslant t \leqslant \tau_1(\omega)$. Consequently, $P(\tau_2 \geqslant \tau_1) = 1$. Similarly $P(\tau_1 \geqslant \tau_2) = 1$, and the proof is complete.

We shall now use Theorem 2.1 in order to improve Theorem 1.1.

Theorem 2.2. *Suppose $b(x, t)$, $\sigma(x, t)$ are measurable in $(x, t) \in R^n \times [0, T]$ and*

$$|b(x, t)| \leqslant K(1 + |x|), \qquad |\sigma(x, t)| \leqslant K(1 + |x|). \tag{2.5}$$

Suppose further that for any $N > 0$ there is a positive constant K_N such that

$$|b(x, t) - b(\bar{x}, t)| \leqslant K_N|x - \bar{x}|, \qquad |\sigma(x, t) - \sigma(\bar{x}, t)| \leqslant K_N|x - \bar{x}| \tag{2.6}$$

if $|x| \leqslant N$, $|\bar{x}| \leqslant N$, $0 \leqslant t \leqslant T$. *Let* ξ_0 *be any n-dimensional random vector independent of* $\mathcal{F}(w(t), 0 \leqslant t \leqslant T)$ *such that* $E|\xi_0|^2 < \infty$. *Then there exists a unique solution* $\xi(t)$ *of* (1.1), (1.2) *in* $M_w^2[0, T]$. *Further*

$$E\Big[\sup_{0\leqslant t\leqslant T} |\xi(t)|^2\Big] \leqslant C^*(1 + E|\xi_0|^2) \tag{2.7}$$

where C^* *is a constant depending only on* T, K.

Thus, Theorem 1.1 and Corollary 1.2 remain true if the condition (1.3) is replaced by (2.5), (2.6).

Proof. Let

$$\xi_{N0}(\omega) = \begin{cases} \xi_0(\omega) & \text{if} \quad |\xi_0(\omega)| \leqslant N, \\ 0 & \text{if} \quad |\xi_0(\omega)| > N, \end{cases}$$

$$b_N(x, t) = \begin{cases} b(x, t) & \text{if} \quad |x| \leqslant N, \\ b(x, t)(2 - (|x|/N)) & \text{if} \quad N < |x| \leqslant 2N, \\ 0 & \text{if} \quad |x| > 2N, \end{cases}$$

$$\sigma_N(x, t) = \begin{cases} \sigma(x, t) & \text{if} \quad |x| \leqslant N, \\ \sigma(x, t)(2 - (|x|/N)) & \text{if} \quad N < |x| \leqslant 2N, \\ 0 & \text{if} \quad |x| > 2N. \end{cases}$$

By Theorem 1.1 there exists a unique solution $\xi_N(t)$ in $M_w^2[0, T]$ of

$$d\xi_N(t) = b_N(\xi_N(t), t)\, dt + \sigma_N(\xi_N(t), t)\, dw(t),$$
$$\xi_N(0) = \xi_{N0}. \tag{2.8}$$

For definiteness we take $\mathcal{F}_t$ to be the σ-field generated by $w(s)$, $0 \leqslant s \leqslant t$ and ξ_0. By Corollary 1.2,

$$E\Big[\sup_{0\leqslant t\leqslant T} |\xi_N(t)|^2\Big] \leqslant C^*(1 + E|\xi_0|^2) \tag{2.9}$$

where C^* is a constant independent of N.

Let τ_N be the first time $|\xi_N(t)| \geqslant N$, if such a time $t \leqslant T$ exists, and $\tau_N = T$ otherwise. If $N' > N$, then by Theorem 2.1 $\xi_N(t) = \xi_{N'}(t)$ a.s. if $0 \leqslant t \leqslant \tau_N$. Therefore,

$$P\Big\{\sup_{0\leqslant t\leqslant T} |\xi_N(t) - \xi_{N'}(t)| > 0\Big\} = P\Big\{\sup_{0\leqslant t\leqslant T} |\xi_N(t)| > N\Big\}$$
$$\leqslant \frac{1}{N^2} E\Big[\sup_{0\leqslant t\leqslant T} |\xi_N(t)|\Big]^2 \leqslant \frac{C^*}{N^2}(1 + E|\xi_0|^2).$$

Notice that the set $\Omega_N = \{\sup_{0\leqslant t\leqslant T} |\xi_N(t)| > N\}$ decreases when N increases, and $P(\Omega_N) \downarrow 0$ if $N \uparrow \infty$. Define

$$\xi(t, \omega) = \lim_{N\to\infty} \xi_N(t, \omega) \qquad \text{if} \quad \omega \notin \lim \Omega_N.$$

Then, $\xi(t, \omega) = \xi_N(t, \omega)$ if $\omega \not\in \Omega_N$. It is clear that $\xi(t)$ is in $L_w^2[0, T]$.

By Lemma 4.2.11,

$$\int_0^t \sigma(\xi(s), s)\, dw(s) = \int_0^t \sigma_N(\xi_N(s), s)\, dw(s) \qquad \text{a.s.} \qquad \text{if} \quad \omega \not\in \Omega_N.$$

From the stochastic differential equation for $\xi_N(t)$ we then obtain

$$\xi(t) = \xi_0 + \int_0^t b(\xi(s), s)\, ds + \int_0^t \sigma(\xi(s), s)\, dw(s)$$

a.s. if $\omega \not\in \Omega_N$. It follows that $\xi(t)$ is a solution of (1.1), (1.2).

Finally, from (2.9) and Fatou's lemma we obtain (2.7).

It remains to prove uniqueness. Let $\xi_1(t)$, $\xi_2(t)$ be two solutions of (1.1), (1.2) in $M_w^2[0, T]$; ξ_i is nonanticipative with respect to $\mathcal{F}_t^i$. Let $\phi(t) = 1$ if $\sup_{0<s<T} |\xi_i(s)| \leqslant N$ for $i = 1, 2$ and $\phi(t) = 0$ in all other cases. Then (cf. the proof of (2.2))

$$\begin{aligned} E\left[\phi(t)|\xi_1(t) - \xi_2(t)|^2\right] &\leqslant 2E\left\{\phi(t)\left|\int_0^t [b(\xi_1(s), s) - b(\xi_2(s), s)]\, ds\right|^2\right\} \\ &\quad + 2E\left\{\phi(t)\left|\int_0^t \sigma(\xi_1(s), s)\, dw(s) - \int_0^t \sigma(\xi_2(s), s)\, dw(s)\right|^2\right\} \\ &\equiv I + J. \end{aligned} \tag{2.10}$$

It is clear that

$$I \leqslant 2t \int_0^t K_N |\xi_1(s) - \xi_2(s)|^2 \phi(s)\, ds. \tag{2.11}$$

Next, if $f_i(s)$ are step functions, nonanticipative with respect to the σ-fields $\mathcal{F}_t^i$ and in $M_w^2[0, T]$, then (cf. the proof of (2.4))

$$E\left\{\phi(t)\left|\int_0^t f_1(s)\, dw(s) - \int_0^t f_2(s)\, dw(s)\right|^2\right\} \leqslant E \int_0^t |f_1(s) - f_2(s)|^2 \phi(s)\, ds. \tag{2.12}$$

By approximation, this inequality holds for any f_1, f_2 in $M_w^2[0, T]$; f_i being nonanticipative with respect to $\mathcal{F}_t^i$. Using (2.12) with $f_i(s) = \sigma(\xi_i(s), s))$ and using (2.11), we obtain from (2.10)

$$E\{\phi(t)|\xi_1(t) - \xi_2(t)|^2\} \leqslant (2T + 2)K_N \int_0^t E\phi(s)|\xi_1(s) - \xi_2(s)|^2\, ds.$$

Consequently

$$E\{\phi(t)|\xi_1(t) - \xi_2(t)|^2\} = 0,$$

i.e.,

$$P\{\xi_1(t) \neq \xi_2(t)\} \leqslant P\Big\{\sup_{0\leqslant s\leqslant T} |\xi_1(s)| > N\Big\} + P\Big\{\sup_{0\leqslant s\leqslant T} |\xi_2(s)| > N\Big\}.$$

Since $\xi_1(s)$ and $\xi_2(s)$ are continuous process, the right-hand side converges to 0 as $N\to\infty$. Therefore,

$$P\{\xi_1(t) \neq \xi_2(t)\} = 0.$$

Since both processes $\xi_1(t)$ and $\xi_2(t)$ are continuous,

$$P\{\xi_1(t) = \xi_2(t) \text{ for } 0 \leqslant t \leqslant T\} = 1.$$

We conclude this section with an estimate on the higher moments of the solution $\xi(t)$.

Theorem 2.3. *Let $b(x, t)$, $\sigma(x, t)$ be as in Theorem 2.2, and let $E|\xi_0|^{2m} < \infty$ for some positive integer m. Then*

$$E|\xi(t)|^{2m} \leqslant (1 + E|\xi_0|^{2m})e^{Ct}, \tag{2.13}$$

$$E\Big\{\sup_{0\leqslant s\leqslant t} |\xi(s) - \xi_0|^{2m}\Big\} \leqslant \bar{C}(1 + E|\xi_0|^{2m})t^m, \tag{2.14}$$

where C, $\bar{C}$ are constants depending only on K, T, m.

Proof. By Itô's formula with $f(x, t) = |x|^{2m}$, $\xi_N(t)$,

$$\begin{aligned} |\xi_N(t)|^{2m} = |\xi_N(0)|^{2m} + \int_0^t \Big\{ & 2m|\xi_N(s)|^{2m-2}\xi_N(s) \cdot b_N(\xi_N(s), s) \\ & + \sum_{i=1}^n 2m|\xi_N(s)|^{2m-2}a_{ii}^N \\ & + \sum_{i,j=1}^n 2m(2m-2)|\xi_N(s)|^{2m-4}a_{ij}^N\xi_N^i(s)\xi_N^j(s)\Big\}\, ds \\ & + \int_0^t 2m|\xi_N(s)|^{2m-2}\xi_N(s) \cdot \sigma_N(\xi(s), s)\, dw(s) \end{aligned} \tag{2.15}$$

where $(a_{ij}^N) = \frac{1}{2}\sigma^N(\xi_N(s), s)(\sigma^N(\xi_N(s), s))^*$ (A^* = transpose of A).

Since

$$\xi_N(s) = \xi_N(0) + \int_0^t b_N(\xi_N(s), s)\, ds + \int_0^t \sigma_N(\xi_N(s), s)\, dw(s)$$

and $E|\xi_N(0)|^2 < \infty$, while $b_N(x, t)$ and $\sigma_N(x, t)$ are bounded, it follows (using Theorem 4.6.3) that $E|\xi_N(s)|^{2m} \leqslant C_N$ for all $0 \leqslant s \leqslant T$, where C_N is a constant. Hence the expectation of the stochastic integral on the right-hand side of (2.15) is equal to 0. We therefore get

$$E|\xi_N(t)|^{2m} \leqslant E|\xi_N(0)|^{2m} + L\int_0^t E\{[1 + |\xi_N(s)|^2]|\xi_N(s)|^{2m-2}\}\, ds,$$

where L is a constant depending only on m, K. Since the integrand on the right is $\leqslant 1 + 2E|\xi_N(s)|^{2m}$, the inequality (2.13) for ξ_N readily follows. Now take $N \uparrow \infty$ to obtain (2.13).

To prove (2.14) notice that for any $\lambda \in [0, T]$

$$\sup_{0 \le t \le \lambda} |\xi(t) - \xi_0|^{2m} \le n 2^{2m} \left[\int_0^\lambda |b(\xi(s), s)|\, ds \right]^2 + 2^m \sup_{0 \le t \le \lambda} \left| \int_0^t \sigma(\xi(s), s)\, dw(s) \right|^{2m}.$$

Taking the expectation and using Corollary 4.6.4 and (2.13) we obtain

$$E \sup_{0 \le t \le \lambda} |\xi(t) - \xi_0|^{2m} \le L\lambda^{m-1} \int_0^\lambda [1 + E|\xi(s)|^{2m}]\, ds \le L_1 \lambda^m (1 + E|\xi_0|^{2m}),$$

where L, L_1 are constants depending only on K, T, m; thus (2.14) holds.

Notice that (2.14) implies that

$$E\left\{ \sup_{0 \le s \le t} |\xi(s)|^{2m} \right\} \le C^*(1 + E|\xi_0|^2), \tag{2.16}$$

where C^* is a constant depending only on K, T, m.

Remark. The results proved in this and the previous sections extend immediately to stochastic differential equations (1.1) with initial condition $\xi(s) = \xi_s$ where $s \in [0, T]$. Here ξ_s is assumed to be independent of the σ-field $\mathcal{F}(w(t + s) - w(s), s \le t \le T - s)$.

3. The solution of a stochastic differential system as a Markov process

We shall assume:

(A) The n-vector $b(x, t)$ and the $n \times n$ matrix $\sigma(x, t)$ are measurable functions for $(x, t) \in R^n \times [0, \infty)$ and, for any $T > 0$,

$$|b(x, t)| \le K(1 + |x|), \qquad |\sigma(x, t)| \le K(1 + |x|), \tag{3.1}$$

$$|b(x, t) - b(\bar{x}, t)| \le K|x - \bar{x}|, \qquad |\sigma(x, t) - \sigma(\bar{x}, t)| \le K|x - \bar{x}| \tag{3.2}$$

if $0 \le t \le T$, $x \in R^n$, $\bar{x} \in R^n$, where K is a constant depending on T.

By Theorem 1.1 there exists a unique solution in $M_w^2[0, \infty)$ of

$$d\xi(t) = b(\xi(t), t)\, dt + \sigma(\xi(t), t)\, dw(t), \tag{3.3}$$

$$\xi(0) = \xi_0, \tag{3.4}$$

provided ξ_0 is independent of $\mathcal{F}(w(\lambda), \lambda \ge 0)$ and $E|\xi_0|^2 < \infty$. Similarly (see the remark at the end of Section 2), for any $s \ge 0$ there exists a unique solution in $M_w^2[s, \infty)$ of (3.3) and

$$\xi(s) = \xi_s \tag{3.5}$$

provided ξ_s is independent of $\mathcal{F}(w(\lambda + s) - w(s), \lambda > 0)$ and $E|\xi_s|^2 < \infty$.

If $\xi_s = x$ a.s., where x is a point in R^n, then we denote the solution of (3.3), (3.5) by $\xi_{x,s}(t)$.

For any Borel set A in R^n and for any $t \geqslant s$, let

$$p(s, x, t, A) = P(\xi_{x,s}(t) \in A). \tag{3.6}$$

Theorem 3.1. *Let* (A) *hold and let* ξ_0 *be independent of* $\mathcal{F}(w(t), t \geqslant 0)$, $E|\xi_0|^2 < \infty$. *Denote by* $\mathcal{F}_t$ *the* σ*-field spanned by* ξ_0 *and* $w(s)$, $0 \leqslant s \leqslant t$. *Then the unique solution* $\xi(t)$ *of* (3.3), (3.4) *satisfies*

$$P(\xi(t) \in A | \mathcal{F}_s) = P(\xi(t) \in A | \xi(s)) = p(s, \xi(s), t, A) \quad \text{a.s.} \tag{3.7}$$

for all $t > s$ *and for any Borel set* A. *Further,* $p(s, x, t, A)$ *is a transition probability function.*

Proof. We can write

$$\xi(t) = \xi(s) + \int_s^t b(\xi(\lambda), \lambda)\, d\lambda + \int_s^t \sigma(\xi(\lambda), \lambda)\, dw(\lambda).$$

By the method of successive approximations, if we set

$$\xi_0(t) = \gamma \qquad (\text{where} \quad \gamma = \xi(s))$$

$$\xi_k(t) = \gamma + \int_s^t b(\xi_{k-1}(\lambda), \lambda)\, d\lambda + \int_s^t \sigma(\xi_{k-1}(\lambda), \lambda)\, dw(\lambda),$$

then $\xi_k(t) \to \xi(t)$ a.s. By induction one can show that each $\xi_k(t)$ is measurable with respect to the σ-field spanned by $\mathcal{F}(w(u+s) - w(s), s \leqslant u \leqslant t)$ and γ. The same is therefore true of $\xi(t)$. More specifically, using Problems 20, 21 of Chapter 4 and Theorem 4.3.4, one can approximate a.s. (and uniformly in t) each $\xi_k(t)$ by a sequence of functions

$$F_m(t, \gamma, w(u_{m,1} + s) - w(s), \ldots, w(u_{m,\mu_m} + s) - w(s))$$

where $0 < u_{m,i} \leqslant t - s$ and $F_m(t, x_0, x_1, \ldots, x_{\mu_m})$ are Borel measurable functions depending only on the functions $b(x, \lambda)$, $\sigma(x, \lambda)$ in the interval $s \leqslant \lambda \leqslant t$. The same therefore holds for $\xi(t)$, i.e., a.s.

$$\xi(t) = \lim_{m\to\infty} F_m(t, \xi(s), w(u_{m,1} + s) - w(s), \ldots, w(u_{m,\mu_m} + s) - w(s)) \tag{3.8}$$

with suitable Borel functions F_m and suitable $u_{m,i}$. In particular,

$$\xi_{x,s}(t) = \lim_{m\to\infty} F_m(t, x, w(u_{m,1} + s) - w(s), \ldots, w(u_{m,\mu_m} + s) - w(s)). \tag{3.9}$$

Now, for any bounded measurable function

$$F(x_0, x_1, \ldots, x_k) = F_0(x_0)F_1(x_1, \ldots, x_k)$$

and for any $u_i \geqslant 0$,

$$\begin{aligned} E[F(\xi(s), w(u_1+s)-w(s), \dots, w(u_k+s)-w(s))|\mathcal{F}_s] \\ = F_0(\xi(s))E[F_1(w(u_1+s)-w(s), \dots, w(u_k+s)-w(s))|\mathcal{F}_s] \\ = F_0(\xi(s))E[F_1(w(u_1+s)-w(s), \dots, w(u_k+s)-w(s))] \end{aligned}$$

since the random variables $w(u_i+s)-w(s)$ are independent of $\mathcal{F}_s$. Thus,

$$\begin{aligned} E[F(\xi(s), w(u_1+s)-w(s), \dots, w(u_k+s)-w(s))|\mathcal{F}_s] \\ = \{EF(x_0, w(u_1+s)-w(s), \dots, w(u_k+s)-w(s))\}_{x_0=\xi(s)}. \end{aligned} \tag{3.10}$$

By approximation, (3.10) remains true for any bounded Borel measurable function $F(x_0, x_1, \dots, x_k)$. Taking, in particular, $F = f(F_m)$ where f is a bounded continuous function and F_m are as in (3.8), (3.9), we conclude that

$$E\{f(\xi(t))|\mathcal{F}_s\} = E\{f(\xi(t))|\xi(s)\} = \phi(x)|_{x=\xi(s)}$$

where $\phi(x) = Ef(\xi_{x,s}(t))$.

Taking a sequence of f's that increase to the indicator function of an open set A in R^n, we obtain the assertions (3.6), (3.7) for any open set A. Since the class of Borel sets for which (3.6), (3.7) hold form a monotone class, and since the open sets belong to this class, this class must contain the σ-field generated by the open sets, i.e., it includes all the Borel sets.

It remains to prove that p is a transition probability function.

From (3.6) it is clear that $p(s, x, t, A)$ is a probability measure in A, for fixed s, x, t. Since $\xi_{x,s}(t)$ is Borel measurable in x (for the same is true of each function in the sequence that converges to $\xi_{x,s}(t)$ by the method of successive approximations), $p(s, x, t, A)$ is also Borel measurable in x, for fixed s, t, A. Thus it remains to verify the Chapman–Kolmogorov equation.

From (3.6),

$$\int p(s, x, t, dy)\psi(y) = \int \psi(\xi_{x,s}(t))\, dP$$

for any bounded Borel measurable function $\psi(y)$. Taking $\psi(y) = p(t, y, \tau, A)$ where $t < \tau$, we get

$$\begin{aligned} \int p(s, x, t, dy)p(t, y, \tau, A) &= \int p(t, \xi_{x,s}(t), \tau, A)\, dP \\ &= Ep(t, \xi_{x,s}(t), \tau, A) = EP[\xi_{x,s}(\tau) \in A | \xi_{x,s}(t)] \end{aligned}$$

where (3.7) has been used in the last equality. Since the expression on the right is equal to

$$\begin{aligned} EE[\chi_A(\xi_{x,s}(\tau))|\xi_{x,s}(t)] &= E\chi_A(\xi_{x,s}(\tau)) \\ &= P(\xi_{x,s}(\tau) \in A) = p(s, x, \tau, A), \end{aligned}$$

the Chapman–Kolmogorov equation holds for all $x \in R^n$, $0 \leqslant s < t < \tau$.

Denote by $\mathcal{C}$ the class of all continuous functions $x(\cdot)$ from $[0, \infty)$ into R^n. Denote by $\mathfrak{M}_t^s$ the smallest σ-field generated by the sets

$$\{x(\cdot);\, x(u) \in A\}, \qquad s \leqslant u \leqslant t$$

where A is any Borel set in R^n. Denote by $\mathfrak{M}_\infty^s$ the smallest σ-field generated by $\{x(\cdot);\, x(u) \in A\}$, $u \geqslant s$.

Define a continuous process $X(t) = X(t, x(\cdot))$ by

$$X(t, x(\cdot)) = x(t). \tag{3.11}$$

Finally, let

$$P_{x,s}\{x(\cdot) \in B\} = P\{\omega;\, \xi_{x,s}(\cdot, \omega) \in B\} \tag{3.12}$$

where B is any set in $\mathfrak{M}_t^s$. Notice that for each $\omega \in \Omega$, $\xi_{x,s}(t) = \xi_{x,s}(t, \omega)$ is a continuous path. This path is the continuous function $\xi_{x,s}(\cdot, \omega)$ appearing on the right-hand side of (3.12).

It is easily seen that $P_{x,s}$ is a probability measure on $\mathfrak{M}_t^s$. We shall now show that

$$P_{x,s}\{X(t+h) \in A \mid \mathfrak{M}_t^s\} = p(t, X(t), t+h, A) \quad \text{a.s.} \tag{3.13}$$

Since

$$P\{\xi_{x,s}(t+h) \in A \mid \mathcal{F}(\xi_{x,s}(\lambda), \lambda \leqslant t)\} = p(t, \xi_{x,s}(t), t+h, A),$$

for any $s \leqslant t_1 < t_2 < \cdots < t_m \leqslant t$ and any Borel sets $A_1, \ldots, A_m$, we have

$$\begin{aligned} &P\{\xi_{x,s}(t+h) \in A,\, \xi_{x,s}(t_1) \in A_1, \ldots, \xi_{x,s}(t_m) \in A_m\} \\ &\quad = \int_{\bigcap_{i=1}^m [\xi_{x,s}(t_i) \in A_i]} p(t, \xi_{x,s}(t), t+h, A)\, dP \\ &\quad = \int \chi_{A_1}(\xi_{x,s}(t_1)) \cdots \chi_{A_m}(\xi_{x,s}(t_m)) p(t, \xi_{x,s}(t), t+h, A)\, dP. \end{aligned}$$

Using (3.11), (3.12), we conclude that

$$\begin{aligned} &P_{x,s}[X(t+h) \in A,\, X(t_1) \in A_1, \ldots, X(t_m) \in A_m] \\ &\quad = \int \chi_{A_1}(X(t_1)) \cdots \chi_{A_m}(X(t_m)) p(t, X(t), t+h, A)\, dP_{x,s} \\ &\quad = \int_{\bigcap_{i=1}^m [X(t_i) \in A_i]} p(t, X(t), t+h, A)\, dP_{x,s}. \end{aligned}$$

This implies (3.13).

We have proved:

Theorem 3.2. *Let* (A) *hold. Then* $\{\mathcal{C}, \mathfrak{M}, \mathfrak{M}_t^s, X(t), P_{x,s}\}$ *is a continuous n-dimensional Markov process with the transition probability function* (3.6).

We shall call this Markov process also the *solution of the stochastic* *differential system* (1.1). The process $X(t)$ will often be denoted also by $\xi(t)$.

Lemma 3.3. *Let* (A) *hold. Then for any* $R > 0$, $T > 0$,

$$E \sup_{\tau \leqslant t \leqslant T} |\xi_{x,s}(t) - \xi_{y,\tau}(t)|^2 \leqslant C(|x - y|^2 + |s - \tau|) \tag{3.14}$$

if $|x| \leqslant R$, $|y| \leqslant R$, $0 \leqslant s \leqslant \tau \leqslant T$; *C is a constant depending on* R, T.

Proof. Clearly, if $\tau \leqslant t$,

$$\xi_{x,s}(t) - \xi_{y,\tau}(t) = \xi_{x,s}(\tau) - y + \int_\tau^t [b(\xi_{x,s}(\lambda), \lambda) - b(\xi_{y,\tau}(\lambda), \lambda)]\, d\lambda$$

$$+ \int_\tau^t [\sigma(\xi_{x,s}(\lambda), \lambda) - \sigma(\xi_{y,\tau}(\lambda), \lambda)]\, dw(\lambda). \tag{3.15}$$

By Theorem 2.3,

$$E|\xi_{x,s}(\tau) - y|^2 \leqslant 2E|\xi_{x,s}(\tau) - x|^2 + 2|x - y|^2 \leqslant C_0|\tau - s| + 2|x - y|^2 \tag{3.16}$$

where C_0 is a constant. Taking the expectation of the supremum of the squares of both side of (3.15) and using (3.16), (3.2), we get

$$E \sup_{\tau \leqslant t \leqslant t'} |\xi_{x,s}(t) - \xi_{y,\tau}(t)|^2 \leqslant C_1(|x - y|^2 + |s - \tau|)$$

$$+ C_1 E \int_\tau^{t'} |\xi_{x,s}(\lambda) - \xi_{y,\tau}(\lambda)|^2\, d\lambda$$

where C_1 is a constant. This easily implies (3.14).

Theorem 3.4. *Let* (A) *hold. Then the solution* $\{\mathcal{C}, \mathfrak{M}, \mathfrak{M}_t^s, X(t), P_{x,s}\}$ *of the stochastic differential system* (1.1) *satisfies the Feller property, and therefore also the strong Markov property.*

Proof. If f is a bounded continuous function, then, by Lemma 3.3 and the Lebesgue bounded convergence theorem,

$$Ef(\xi_{y,\tau}(t + \tau)) - Ef(\xi_{x,s}(t + \tau)) \to 0 \quad \text{if} \quad y \to x, \quad \tau \to s.$$

Also,

$$Ef(\xi_{x,s}(t + \tau)) \to Ef(\xi_{x,s}(t + s)) \quad \text{if} \quad \tau \to s.$$

Thus, the function

$$(s, x) \to \int p(s, x, t + s, dy) f(y)$$

is continuous, i.e., the Markov process $\{\mathcal{C}, \mathfrak{M}, \mathfrak{M}_t^s, X(t), P_{x,s}\}$ satisfies the Feller property. Since it is also a continuous process, Corollary 2.2.6 asserts that it possesses the strong Markov property.

Consider a stochastic differential system

$$d\xi(t) = b(\xi(t), t)\, dt + \sigma(\xi(t), t)\, d\tilde{w}(t) \tag{3.17}$$

where $\tilde{w}(t)$ is another n-dimensional Brownian motion. According to Theorem 3.3, there is a Markov process solution $\{\mathcal{C}, \mathfrak{M}, \mathfrak{M}_t^s, X(t), \tilde{P}_{x,s}\}$.

Definition. If for any Brownian motion $\tilde{w}(t)$, $\tilde{P}_{x,s} = P_{x,s}$ for all x, s, then we say that there is *uniqueness in the sense of probability law*. By contrast, the uniqueness asserted in Theorem 1.1 is called *pathwise uniqueness*.

Theorem 3.5. *Let* (A) *hold. Then there is a unique solution of* (1.1) *in the sense of probability law.*

Proof. From the proof of (3.8) one sees that the F_m do not depend on the particular Brownian motion $w(t)$. Hence, for any $s \leqslant t_1 < t_2 < \cdots < t_k < \infty$ and for any bounded continuous function $\phi(x_1, \ldots, x_k)$, if $\tilde{\xi}(t)$ is a solution of (3.17) with $\tilde{\xi}(s) = \xi(s)$ and if $\xi(s) = x$, then

$$\begin{aligned}
&E\phi(\xi(t_1), \ldots, \xi(t_k))\\
&\quad = \lim_{m\to\infty} E\phi\big(F_m(t_1, x, w(u_{m,1} + s) - w(s), \ldots, w(u_{m,\mu_m} + s) - w(s)),\\
&\qquad \ldots, F_m(t_k, x, w(u_{m,1} + s) - w(s), \ldots, w(u_{m,\mu_m} + s) - w(s))\big)\\
&\quad = \lim_{m\to\infty} E\phi\big(F_m(t_1, x, \tilde{w}(u_{m,1} + s) - w(s), \ldots, \tilde{w}(u_{m,\mu_m} + s) - \tilde{w}(s)),\\
&\qquad \ldots, F_m(t_k, x, \tilde{w}(u_{m,1} + s) - \tilde{w}(s), \ldots, \tilde{w}(u_{m,\mu_m} + s) - \tilde{w}(s))\big)\\
&\quad = E\phi(\tilde{\xi}(t_1), \ldots, \tilde{\xi}(t_k)).
\end{aligned}$$

Since ϕ is arbitrary, it follows that $P_{x,s} = \tilde{P}_{x,s}$ on cylinder sets in $\mathfrak{M}_t^s$. Therefore also $P_{x,s} = \tilde{P}_{x,s}$ on the whole of $\mathfrak{M}_t^s$.

The results of this section can be extended to the case where the condition (A) is replaced by the weaker condition:

(A′) $b(x, t)$ and $\sigma(x, t)$ are measurable; for every $T > 0$ the inequality (3.1) holds for $0 \leqslant t \leqslant T$, $x \in R^n$ with K depending only on T; for every $T > 0$, $R > 0$,

$$|b(x, t) - b(\bar{x}, t)| \leqslant K_R|x - \bar{x}|, \qquad |\sigma(x, t) - \sigma(\bar{x}, t)| \leqslant K_R|x - \bar{x}|$$

if $0 \leqslant t \leqslant T$, $|x| \leqslant R$, $|\bar{x}| \leqslant R$ where K_R is a constant depending on T, R.

Theorem 3.6. *Theorems* 3.1, 3.2, 3.4, 3.5 *remain true if the condition* (A) *is replaced by the weaker condition* (A′).

The proof is left to the reader.

4. Diffusion processes

A continuous n-dimensional Markov process with transition probability function $p(s, x, t, A)$ is called a *diffusion process* if:

(i) for any $\epsilon > 0$, $t \geqslant 0$, $x \in R^n$,

$$\lim_{h \downarrow 0} \frac{1}{h} \int_{|y-x|>\epsilon} p(t, x, t+h, dy) = 0; \tag{4.1}$$

(ii) there exist an n-vector $b(x, t)$ and an $n \times n$ matrix $a(x, t)$ such that for any $\epsilon > 0$, $t \geqslant 0$, $x \in R^n$,

$$\lim_{h \downarrow 0} \frac{1}{h} \int_{|y-x|<\epsilon} (y_i - x_i) p(t, x, t+h, dy) = b_i(x, t) \qquad (1 \leqslant i \leqslant n), \tag{4.2}$$

$$\lim_{h \downarrow 0} \frac{1}{h} \int_{|y-x|<\epsilon} (y_i - x_i)(y_j - x_j) p(t, x, t+h, dy) = a_{ij}(x, t) \qquad (1 \leqslant i, j \leqslant n), \tag{4.3}$$

where $b = (b_1, \ldots, b_n)$, $a = (a_{ij})$.

The vector b is called the *drift coefficient* and the matrix a is called the *diffusion matrix*; when $n = 1$ we call a the *diffusion coefficient.*

Lemma 4.1. *The following conditions imply the conditions* (i), (ii):

(i*) *for some* $\delta > 0$, $t \geqslant 0$, $x \in R^n$,

$$\lim_{h \downarrow 0} \frac{1}{h} \int_{R^n} |x - y|^{2+\delta} p(t, x, t+h, dy) = 0; \tag{4.4}$$

(ii*) *for any* $t \geqslant 0$, $x \in R^n$,

$$\lim_{h \downarrow 0} \frac{1}{h} \int_{R^n} (y_i - x_i) p(t, x, t+h, dy) = b_i(x, t) \qquad (1 \leqslant i \leqslant n), \tag{4.5}$$

$$\lim_{h \downarrow 0} \frac{1}{h} \int_{R^n} (y_i - x_i)(y_j - x_j) p(t, x, t+h, dy) = a_{ij}(x, t) \qquad (1 \leqslant i, j \leqslant n). \tag{4.6}$$

Proof. Using (4.4) we have

$$\int_{|y-x|>\epsilon} p(t, x, t+h, dy) \leqslant \frac{1}{\epsilon^{2+\delta}} \int_{R^n} |y - x|^{2+\delta} p(t, x, t+h, dy) \to 0 \quad \text{if} \quad h \to 0,$$

so that (4.1) holds. By (4.4) we also have, for $j = 1, 2$,

$$\frac{1}{h}\int_{|y-x|>\epsilon} |y - x|^j p(t, x, t + h, dy)$$
$$\leqslant \frac{1}{\epsilon^{2+\delta-j}} \cdot \frac{1}{h}\int_{R^n} |y - x|^{2+\delta} p(t, x, t + h, dy) \to 0$$

if $h \to 0$. Consequently (4.2), (4.3) hold if and only if (4.5), (4.6) hold.

Theorem 4.2. *Let* (A′) *hold and let* $b(x, t)$, $\sigma(x, t)$ *be continuous in* $(x, t) \in R^n \times [0, \infty)$. *Then the solution of* (1.1) *is a diffusion process with drift* $b(x, t)$ *and diffusion matrix* $a(x, t) = \sigma(x, t)\sigma^*(x, t)$.

Here σ^* is the transpose of σ. Thus, if $a = (a_{ij})$, then $a_{ij} = \sum_{k=1}^n \sigma_{ik}\sigma_{jk}$.

Proof. Since $p(t, x, t + h, A)$ is the probability distribution of the random variable $\xi_{x,t}(t + h)$,

$$Ef(\xi_{x,t}(t + h) - x) = \int_{R^n} f(y - x)p(t, x, t + h, dy) \tag{4.7}$$

for any continuous function $f(z)$ with $|f(z)| \leqslant K(1 + |z|^\alpha)$ for some $K > 0$, $\alpha > 0$; we use here the fact that $E|\xi_{x,t}(t + h)|^\alpha < \infty$ for any $\alpha > 0$ (see Theorem 2.3).

In view of Lemma 4.1 it is therefore sufficient to prove that

$$\frac{1}{h} E|\xi_{x,t}(t + h) - x|^4 \to 0 \quad \text{if} \quad h \to 0, \tag{4.8}$$

$$\frac{1}{h} E[\xi_{x,t}(t + h) - x] \to b(x, t) \quad \text{if} \quad h \to 0, \tag{4.9}$$

$$\frac{1}{h} E(\xi^i_{x,t}(t + h) - x)(\xi^j_{x,t}(t + h) - x) \to a_{ij}(x, t) \quad \text{if} \quad h \to 0, \tag{4.10}$$

where $\xi^i_{x,t}$ is the ith component of $\xi_{x,t}$.

From Theorem 2.3,

$$E|\xi_{x,t}(t + h) - x|^4 \leqslant Kh^2(1 + |x|^4) \qquad (K \text{ constant});$$

this gives (4.8).

Next,

$$\frac{1}{h} E[\xi_{x,t}(t + h) - x] = \frac{1}{h} E \int_t^{t+h} b(\xi_{x,t}(\lambda), \lambda)\, d\lambda$$
$$= \int_0^1 Eb(\xi_{x,t}(t + hs), t + hs)\, ds.$$

Consider the random variables $X_h(s, \omega) = b(\xi_{x,t}(t + hs), t + hs)$. Since

$$\int_0^1 E|X_h|^2 \, ds \leqslant C \int_0^1 E[1 + |\xi_{x,t}(t + hs)|^2] \, ds \leqslant C_1$$

where C, C_1 are constants independent of h, the random variables X_h satisfy the condition of uniform integrability, i.e.,

$$\sup_h \int_{|X_h| > \lambda} |X_h| \, dP \to 0 \qquad \text{if} \quad \lambda \to \infty. \tag{4.11}$$

We also have that $X_h(s, \omega) \xrightarrow{P} b(x, t)$ uniformly with respect to s. Hence, by Lemma 1.3.6,

$$\int_0^1 EX_h(s) \, ds \to b(x, t) \qquad \text{as} \quad h \to 0, \tag{4.12}$$

and the proof of (4.9) is complete.

To prove (4.10) we use Itô's formula (4.7.11) with $u(z, t) = z_i z_j$:

$$\begin{aligned} &\frac{1}{h} \{E\xi^i_{x,t}(t + h)\xi^j_{x,t}(t + h) - x_i x_j\} \\ &\quad = \frac{1}{h} E \int_t^{t+h} [\xi^i_{x,t}(\lambda) b_j(\xi_{x,t}(\lambda), \lambda) \\ &\qquad + \xi^j_{x,t}(\lambda) b_i(\xi_{x,t}(\lambda), \lambda) + a_{ij}(\xi_{x,t}(\lambda), \lambda)] \, d\lambda, \end{aligned} \tag{4.13}$$

where $x = (x_1, \ldots, x_n)$.

Noting that

$$\int_0^1 E[1 + |\xi_{x,t}(t + hs)|^4] \, ds \leqslant C_2$$

where C_2 is a constant independent of h, we can apply the argument used to prove (4.12) in order to deduce from (4.13) that, as $h \to 0$,

$$\frac{1}{h} E[\xi^i_{x,t}(t + h)\xi^j_{x,t}(t + h) - x_i x_j] \to x_i b_j(x, t) + x_j b_i(x, t) + a_{ij}(x, t).$$

It follows that

$$\begin{aligned} &\lim_{h \to 0} \left\{ \frac{1}{h} E(\xi^i_{x,t}(t + h) - x_i)(\xi^j_{x,t}(t + h) - x_j) \right\} \\ &\quad = x_i b_j(x, t) + x_j b_i(x, t) + a_{ij}(x, t) - x_i \lim_{h \to 0} \frac{1}{h} E[\xi^j_{x,t}(t + h) - x_j] \\ &\qquad - x_j \lim_{h \to 0} \frac{1}{h} E[\xi^i_{x,t}(t + h) - x_i] \\ &\quad = a_{ij}(x, t), \end{aligned}$$

where (4.9) has been used in the last equality.

We shall use Theorem 4.2 in order to compute the infinitesimal generators $\mathcal{a}_s$ of the Markov process solution of (1.1). By definition (see (2.3.7)),

$$(\mathcal{a}_s f)(x) = \lim_{h \downarrow 0} \frac{(T_{s, s+h} f)(x) - f(x)}{h}$$

where

$$(T_{s, s+h} f)(x) = \int p(s, x, s+h, dy) f(y).$$

Thus,

$$(\mathcal{a}_s f)(x) = \lim_{h \downarrow 0} \frac{1}{h} \int [f(y) - f(x)] p(s, x, s+h, dy). \tag{4.14}$$

Suppose f is bounded and twice continuously differentiable. Then

$$f(y) - f(x) = \sum_{i=1}^{n} (y_i - x_i) f_{x_i}(x) + \frac{1}{2} \sum_{i,j=1}^{n} (y_i - x_i)(y_j - x_j) f_{x_i x_j}(x) + o(|y - x|^2).$$

Substituting this in the integral in (4.14) for y in a small neighborhood of x, then taking $h \downarrow 0$ and using (4.1)–(4.3), we find that

$$(\mathcal{a}_s f)(x) = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x, s) \frac{\partial^2 f}{\partial x_i \, \partial x_j} + \sum_{i,j=1}^{n} b_i(x, s) \frac{\partial f}{\partial x_i} \equiv (L_s f)(x). \tag{4.15}$$

The operator $L_s f$ is a partial differential operator. This same operator arises also in Itô's formula:

$$u(\xi(t), t) - u(\xi(s), s) = \int_s^t \left(\frac{\partial u}{\partial \lambda} + L_\lambda u \right)(\xi(\lambda), \lambda) \, d\lambda + \int_s^t u_x(\xi(\lambda), \lambda) \sigma(\xi(\lambda), \lambda) \, dw(\lambda) \tag{4.16}$$

where $\xi(t)$ is any solution of (1.1). We refer to L_s as the differential operator corresponding to the stochastic system (1.1). Using (4.16) with $t = \tau$, one can easily derive Dynkin's formula (2.3.6) when f is twice continuously differen tiable.

5. Equations depending on a parameter

Let $A(x, t)$, $B(x, t)$ be a random n-vector and $n \times n$ matrix respectively, defined for $(x, t) \in R^n \times [s, \infty)$ for some $s \geqslant 0$ and assume:

(i) $A(x, t)$, $B(x, t)$ are continuous in (x, t), for each ω;
(ii) $A(x, t)$, $B(x, t)$ are measurable in (x, t, ω);
(iii) $A(x, t)$, $B(x, t)$ are $\mathcal{F}_t$ measurable for each (x, t), where $\mathcal{F}_t$ is an

increasing family of σ-fields such that $w(t)$ is $\mathcal{F}_t$ measurable and $\mathcal{F}(w(t+\lambda) - w(t), \lambda \geqslant 0)$ is independent of $\mathcal{F}_t$ for all $t \geqslant 0$;

(iv) there is a constant K such that

$$|A(x, t)| \leqslant K(1 + |x|), \qquad |B(x, t)| \leqslant K(1 + |x|) \quad \text{a.s.,} \tag{5.1}$$

$$|A(x, t) - A(\bar{x}, t)| \leqslant K|x - \bar{x}|, \qquad |B(x, t) - B(\bar{x}, t)| \leqslant K|x - \bar{x}| \quad \text{a.s.} \tag{5.2}$$

Let $\phi(t)$ be a function in $M_w^\infty[s, T]$. Consider the equation

$$\xi(t) = \phi(t) + \int_s^t A(\xi(\lambda), \lambda)\, d\lambda + \int_s^t B(\xi(\lambda), \lambda)\, dw(\lambda). \tag{5.3}$$

We refer to it as a *system of stochastic differential equations with random coefficients.*

Theorem 5.1. *If* (i)–(iv) *hold and* $\phi \in M_w^\infty[s, T]$, *then there exists a unique solution* $\xi(t)$ *in* $M_w^2[s, T]$; *further,* $\xi(t)$ *belongs to* $M_w^\infty[s, T]$.

We outline the proof leaving the details to the reader.

One defines by induction

$$\xi_m(t) = \phi(t) + \int_s^t A(\xi_{m-1}(\lambda), \lambda)\, d\lambda + \int_s^t B(\xi_{m-1}(\lambda), \lambda)\, dw(\lambda),$$

$$\xi_0(t) = \phi(t),$$

and shows that the ξ_m are well defined and

$$E|\xi_m(t) - \xi_{m-1}(t)|^2 \leqslant K_0 \int_s^t E|\xi_{m-1}(s) - \xi_{m-2}(s)|^2\, ds.$$

Since

$$E|\xi_1(t) - \xi_0(t)|^2 \leqslant K_1 \int_s^t E(1 + |\phi(\lambda)|^2)\, d\lambda,$$

$$E|\xi_m(t) - \xi_{m-1}(t)|^2 \leqslant \frac{L^m(t-s)^m}{m!} \qquad (L \text{ constant}).$$

Now we proceed as in the proof of Theorem 1.1 to show that $\lim \xi_m(t)$ is a solution. The proof of uniqueness is the same as for Theorem 1.1.

Theorem 5.2. *Let* $A_\alpha(x, t)$, $B_\alpha(x, t)$, $\phi_\alpha(t)$ *satisfy the assumptions of Theorem* 5.1 *for any* $0 \leqslant \alpha \leqslant 1$, *with the constant* K (*in* (5.1), (5.2)) *independent of* α, *and with* $\sup_{s \leqslant t \leqslant T} E|\phi_\alpha(t)|^2 \leqslant C$, *C a constant indepen-*

dent of α. *Suppose that for any* $N > 0$, $t \in [s, T]$, $\epsilon > 0$,

$$\lim_{\alpha \downarrow 0} P\left\{ \sup_{|x| \leq N} |A_\alpha(x, t) - A_0(x, t)| > \epsilon \right\} = 0, \tag{5.4}$$

$$\lim_{\alpha \downarrow 0} P\left\{ \sup_{|x| \leq N} |B_\alpha(x, t) - B_0(x, t)| > \epsilon \right\} = 0. \tag{5.5}$$

Suppose also that

$$\lim_{\alpha \downarrow 0} \sup_{s \leq t \leq T} E|\phi_\alpha(t) - \phi_0(t)|^2 = 0. \tag{5.6}$$

Consider the solutions $\xi_\alpha(t)$ *of the equations*

$$\xi_\alpha(t) = \phi_\alpha(t) + \int_s^t A_\alpha(\xi_\alpha(\lambda), \lambda)\, d\lambda + \int_s^t B_\alpha(\xi_\alpha(\lambda), \lambda)\, dw(\lambda). \tag{5.7}$$

Then,

$$\sup_{s \leq t \leq T} E|\xi_\alpha(t) - \xi_0(t)|^2 \to 0 \qquad \textit{if} \quad \alpha \downarrow 0. \tag{5.8}$$

Proof. Take for simplicity $s = 0$. We can write

$$\xi_\alpha(t) - \xi_0(t) = \eta_\alpha(t) + \int_0^t [A_\alpha(\xi_\alpha(s), s) - A_\alpha(\xi_0(s), s)]\, ds + \int_0^t [B_\alpha(\xi_\alpha(s), s) - B_\alpha(\xi_0(s), s)]\, dw(s)$$

where

$$\eta_\alpha(t) = \phi_\alpha(t) - \phi_0(t) + \int_0^t [A_\alpha(\xi_0(s), s) - A_0(\xi_0(s), s)]\, ds + \int_0^t [B_\alpha(\xi_0(s), s) - B_0(\xi_0(s), s)]\, dw(s).$$

Using (5.1), (5.2) for A_α, B_α, we easily find that

$$E|\xi_\alpha(t) - \xi_0(t)|^2 \leq 3E|\eta_\alpha(t)|^2 + C \int_0^t E|\xi_\alpha(s) - \xi_0(s)|^2\, ds. \tag{5.9}$$

If we prove that

$$\sup_{0 \leq t \leq T} E|\eta_\alpha(t)|^2 \to 0 \qquad \text{as} \quad \alpha \to 0, \tag{5.10}$$

then the assertion (5.8) follows from (5.9).

Now,

$$I_\alpha \equiv E\left| \int_0^t [A_\alpha(\xi_0(s), s) - A_0(\xi_0(s), s)]\, ds \right|^2 \leq tE \int_0^t |A_\alpha(\xi_0(s), s) - A_0(\xi_0(s), s)|^2\, ds.$$

The last integrand is bounded by $2K(1 + |\xi_0(s)|^2)$, which is an integrable function. In view of (5.4), this integrand also converges to 0, in probability, as $\alpha \to 0$. Hence, by the Lebesgue bounded convergence theorem, $I_\alpha \to 0$ if $\alpha \to 0$.

Next,

$$J_\alpha \equiv E\left| \int_0^t [B_\alpha(\xi_0(s), s) - B_0(\xi_0(s), s)]\, dw \right|^2$$
$$= E \int_0^t |B_\alpha(\xi_0(s), s) - B_0(\xi_0(s), s)|^2\, ds.$$

By the previous argument, $J_\alpha \to 0$ if $\alpha \to 0$.

Finally, making use also of (5.6), the assertion (5.10) follows.

We shall apply Theorems 5.1, 5.2 in studying the behavior of the solution $\xi_{x,s}(t)$ in the parameters x, s. Recall that

$$\xi_{x,s}(t) = x + \int_s^t b(\xi_{x,s}(\lambda), \lambda)\, d\lambda + \int_s^t \sigma(\xi_{x,s}(\lambda), \lambda)\, dw(\lambda). \tag{5.11}$$

We shall use the notation

$$D_x^\alpha = \frac{\partial^\alpha}{\partial x^\alpha} = \frac{\partial^{\alpha_1 + \cdots + \alpha_n}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \qquad |\alpha| = \alpha_1 + \cdots + \alpha_n.$$

If $f = f(x, t)$, then the vector $(\partial f/\partial x_1, \ldots, \partial f/\partial x_n)$ is denoted briefly by $D_x f$ or by f_x.

Definition. Let $g(x) = g(x_1, \ldots, x_n)$, $f(x) = f(x_1, \ldots, x_n)$ be random functions for x in some open set. If

$$\int \left| \frac{1}{h} [g(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - g(x_1, \ldots, x_n)] - f(x_1, \ldots, x_n) \right|^2 dP \to 0$$

as $h \to 0$, then we say that $g(x)$ has a derivative with respect to x_i in the $L^2(\Omega)$ sense, and the derivative is equal to $f(x)$. We write $(\partial/\partial x_i)g(x) = f(x)$. Similarly one defines the derivative $D^\alpha g(x)$ in the $L^2(\Omega)$ sense, for any $\alpha = (\alpha_1, \ldots, \alpha_n)$.

We shall need the following condition:

$$\partial b/\partial x_i, \quad \partial\sigma/\partial x_i \quad \text{exist and are continuous} \qquad (1 \leqslant i \leqslant n). \tag{5.12}$$

Theorem 5.3. *If* (A) *and* (5.12) *hold, then the derivatives* $\partial \xi_{x,s}(t)/\partial x_i$ *exist in the* $L^2(\Omega)$ *sense and the functions* $\zeta_i(t) = \partial \xi_{x,s}(t)/\partial x_i$ *satisfy the* *stochastic differential system with random coefficients*

$$\zeta_i(t) = e_i + \int_s^t \zeta_i(\lambda) \cdot b_x(\xi_{x,s}(\lambda), \lambda)\, d\lambda + \int_s \zeta_i(\lambda) \cdot \sigma_x(\xi_{x,s}(\lambda), \lambda)\, dw(\lambda) \tag{5.13}$$

where e_i is the vector with components δ_{ij}.

Proof. Take for simplicity $i = 1$, and write $h = (h_1, 0, \ldots, 0)$ where $h_1 \neq 0$. Then

$$\frac{1}{h_1}[\xi_{x+h,s}(t) - \xi_{x,s}(t)] = e_1 + \int_s^t [b(\xi_{x+h,s}(\lambda), \lambda) - b(\xi_{x,s}(\lambda), \lambda)]\, d\lambda$$
$$+ \int_s^t [\sigma(\xi_{x+h,s}(\lambda), \lambda) - \sigma(\xi_{x,s}(\lambda), \lambda)]\, dw(\lambda). \tag{5.14}$$

By Theorem 5.1 there exists a unique solution ζ_1 of (5.13) with $i = 1$. We shall now complete the proof by invoking Theorem 5.2 with $\alpha = h_1$ if $h_1 > 0$ and $\alpha = -h_1$ if $h_1 < 0$. First we note that

$$\frac{1}{h_1} \int_s^t [b(\xi_{x+h,s}(\lambda), \lambda) - b(\xi_{x,s}(\lambda), \lambda)]\, d\lambda$$
$$= \frac{1}{h_1} \int_s^t d\lambda \int_0^1 \frac{d}{d\mu} b(\xi_{x,s}(\lambda) + \mu(\xi_{x+h,s}(\lambda) - \xi_{x,s}(\lambda)), \lambda)\, d\mu$$
$$= \int_s^t \left[\int_0^1 b_x(\xi_{x,s}(\lambda) + \mu(\xi_{x+h,s}(\lambda) - \xi_{x,s}(\lambda)),\lambda)\, d\mu \right]$$
$$\cdot \frac{\xi_{x+h,s}(\lambda) - \xi_{x,s}(\lambda)}{h_1}\, d\lambda.$$

A similar formula holds for the stochastic integral on the right-hand side of (5.14). Therefore, the function

$$\xi_\alpha(t) = \frac{1}{h_1}[\xi_{x+h,s}(t) - \xi_{x,s}(t)]$$

satisfies (5.7) with

$$A_\alpha(z, t) = z \cdot \int_0^1 b_x(\xi_{x,s}(t) + \mu(\xi_{x+h,s}(t) - \xi_{x,s}(t)), t)\, d\mu,$$

$$B_\alpha(z, t) = z \cdot \int_0^1 \sigma_x(\xi_{x,s}(t) + \mu(\xi_{x+h,s}(t) - \xi_{x,s}(t)), t)\, d\mu$$

if $\alpha \neq 0$, and $\xi_0(t) \equiv \zeta_1(t)$ satisfies the same equation with

$$A_0(z, t) = z \cdot b_x(\xi_{x,s}(t), t), \qquad B_0(z, t) = z \cdot \sigma_x(\xi_{x,s}(t), t).$$

Since, by Lemma 3.3,

$$E \sup_t |\xi_{x+h,s}(t) - \xi_{x,s}(t)| \to 0 \qquad \text{if} \quad h \to 0,$$

we conclude, upon using (5.12), that (5.4), (5.5) are satisfied. The assertion of the theorem now follows from Theorem 5.2.

Remark. Notice that $\partial \xi_{x,s}(t)/\partial x_i$ satisfies the stochastic differential system with random coefficients obtained by differentiating formally the stochastic differential system of $\xi_{x,s}(t)$ with respect to x_i.

Theorem 5.4. *Let* (A) *hold and assume that* $D_x^\alpha b(x, t)$, $D_x^\alpha \sigma(x, t)$ *exist and are continuous if* $|\alpha| \leq 2$, *and*

$$|D_x^\alpha b(x, t)| + |D_x^\alpha \sigma(x, t)| \leq K_0(1 + |x|^\beta) \qquad (|\alpha| \leq 2) \tag{5.15}$$

where K_0, β *are positive constants. Then the second derivatives* $\partial^2 \xi_{x,s}(t)/\partial x_i\, \partial x_j$ *exist in the* $L^2(\Omega)$ *sense, and they satisfy the stochastic differential system with random coefficients obtained by applying formally* $\partial^2/\partial x_i\, \partial x_j$ *to* (5.11).

The proof is left to the reader.

Theorem 5.5. *Let* $f(x)$ *be a function with two continuous derivatives, satisfying*

$$|D_x^\alpha f(x)| \leq C(1 + |x|^\beta) \qquad (|\alpha| \leq 2) \tag{5.16}$$

where C, β *are positive constants. Let the conditions of Theorem* 5.4 *hold, and set*

$$\phi(x) = Ef(\xi_{x,s}(t)). \tag{5.17}$$

Then $\phi(x)$ *has two continuous derivatives; these derivatives can be computed by differentiating the right-hand side of* (5.17) *under the integral sign. Finally,*

$$|D^\alpha \phi(x)| \leq C_0(1 + |x|^\gamma) \qquad \text{if} \quad |\alpha| \leq 2, \tag{5.18}$$

where C_0, γ *are positive constants.*

Proof. We shall prove that

$$\frac{\partial \phi}{\partial x_i} = Ef_x(\xi_{x,s}(t)) \cdot \frac{\partial}{\partial x_i} \xi_{x,s}(t). \tag{5.19}$$

Take for simplicity $i = 1$ and set $h = (h_1, 0, \ldots, 0)$. Then

$$\begin{aligned}
\phi(x + h) - \phi(x) &= E[f(\xi_{x+h,s}(t)) - f(\xi_{x,s}(t))] \\
&= E \int_0^1 \frac{d}{d\mu} f(\xi_{x,s}(t) + \mu(\xi_{x+h,s}(t) - \xi_{x,s}(t)))\, d\mu \\
&= E \int_0^1 f_x(\xi_{x,s}(t) + \mu(\xi_{x+h,s}(t) - \xi_{x,s}(t))) \cdot (\xi_{x+h,s}(t) - \xi_{x,s}(t))\, d\mu.
\end{aligned}$$

Hence

$$\frac{\phi(x+h)-\phi(x)}{h_1} = E\int_0^1 f_x(\xi_{x,s}(t) + \mu(\xi_{x+h,s}(t)-\xi_{x,s}(t)))\,d\mu \cdot \frac{\xi_{x+h,s}(t)-\xi_{x,s}(t)}{h_1}. \tag{5.20}$$

As $h \to 0$

$$\frac{\xi_{x+h,s}(t)-\xi_{x,s}(t)}{h_1} \to \frac{\partial}{\partial x_1}\xi_{x,s}(t) \quad \text{in} \quad L^2(\Omega). \tag{5.21}$$

We claim that also

$$\psi_h \equiv \int_0^1 f_x(\xi_{x,s}(t) + \mu(\xi_{x+h,s}(t)-\xi_{x,s}(t)))\,d\mu \to f_x(\xi_{x,s}(t)) \quad \text{in} \quad L^2(\Omega) \tag{5.22}$$

as $h \to 0$. Indeed, this follows from Lemma 1.3.6 if we observe that $|\psi_h - f_x(\xi_{x,s}(t)|^2 \to 0$ in probability and

$$E|\psi_h - f_x(\xi_{x,s}(t))|^4 \leqslant C_1$$

where C_1 is a constant independent of h; the assumption (5.16) is used here.

From (5.20), (5.21), and (5.22) we see that

$$\frac{\phi(x+h)-\phi(x)}{h_1} \to Ef_x(\xi_{x,s}(t)) \cdot \frac{\partial}{\partial x_1}\xi_{x,s}(t) \qquad \text{as} \quad h \to 0.$$

Having proved (5.19), one can now check that the right-hand side of (5.19) is a continuous function in x and is bounded by $C_0(1+|x|^\gamma)$ for some positive constants C_0, γ. The same is then true of $\partial\phi/\partial x_i$. The proof that $\phi(x)$ has second derivatives, that these derivatives are obtained by differentiating under the integral sign of (5.17), and that they are continuous and satisfy (5.18), is similar to the proof for the first derivatives; this is left to the reader.

Theorems 5.4, 5.5 extend without difficulty to higher derivatives.

6. The Kolmogorov equation

Consider the function

$$u(x,t) = Ef(\xi_{x,t}(T)).$$

Notice that

$$u(x,t) = E_{x,t}f(\xi(T))$$

where $\{\mathcal{C}, \mathfrak{M}, \mathfrak{M}_t^s, \xi(t), P_{x,s}\}$ is the Markov process that solves the system of stochastic differential equations (1.1).

Theorem 6.1. *If f and b, σ satisfy the conditions of Theorem 5.5, then u_t, u_x, u_{xx} are continuous functions in $(x, t) \in R^n \times [0, T)$ and satisfy*

$$\frac{\partial u}{\partial t} + \sum_{i=1}^{n} b_i \frac{\partial u}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 u}{\partial x_i \, \partial x_j} = 0 \qquad \text{in} \quad R^n \times [0, T) \tag{6.1}$$

$$u(x, t) \to f(x) \qquad \text{if} \quad t \uparrow T. \tag{6.2}$$

Equation (6.2) is called the *Kolmogorov equation*, or the *backward parabolic equation*.

Proof. Let $g(x)$ be a twice continuously differentiable function satisfying $|D^\alpha g(x)| \leqslant C(1 + |x|^\beta)$ if $|\alpha| \leqslant 2$, where C, β are positive constants. By Itô's formula (see (4.16))

$$Eg(\xi_{x,t-h}(t)) - g(x) = E \int_{t-h}^{t} (L_s g)(\xi_{x,t-h}(s), s) \, ds \qquad (h > 0) \tag{6.3}$$

where L_s is the partial differential operator defined in (4.15). The argument used to derive (4.12) can be applied also here to deduce that

$$\frac{1}{h} E \int_{t-h}^{t} (L_s g)(\xi_{x,t-h}(s), s) \, ds \to (L_t g)(x, t) \qquad \text{as} \quad h \to 0.$$

Hence, if we divide both sides of (6.3) by h and let $h \to 0$, we get the formula

$$\lim_{h \downarrow 0} \frac{1}{h} E[g(\xi_{x,t-h}(t)) - g(x)] = (L_t g)(x, t). \tag{6.4}$$

Next, by the Markov property,

$$\begin{aligned} u(x, t-h) &= E_{x,t-h} f(\xi(T)) = E_{x,t-h} E_{x,t-h}(f(\xi(T)) \mid \mathfrak{M}_t^{t-h}) \\ &= E_{x,t-h} E_{\xi(t),t} f(\xi(T)) = E_{x,t-h} u(\xi(t), t) = Eu(\xi_{x,t-h}(t), t), \end{aligned}$$

so that

$$\frac{u(x, t) - u(x, t-h)}{h} = - \frac{Eu(\xi_{x,t-h}(t), t) - u(x, t)}{h}. \tag{6.5}$$

In view of Theorem 5.5, the function $g(x) = u(x, t)$ has two continuous x-derivatives, bounded by $C_0(1 + |x|^\gamma)$, for some positive constants C_0, γ. Consequently the formula (6.4) can be applied. From (6.5) we then conclude that

$$\lim_{h \downarrow 0} \frac{u(x, t) - u(x, t-h)}{h} = -L_t u(x, t). \tag{6.6}$$

From (6.3) one easily deduces that

$$\frac{1}{h} |Eg(\xi_{x,t-h}(t)) - g(x)| \leqslant C$$

where C is a constant independent of h. Applying this to the function

$g(x) = u(x, t)$ and then using (6.5), we get

$$|u(x, t) - u(x, t - h)| \leqslant Ch.$$

It follows that, for fixed x, $u(x, t)$ is absolutely continuous in t. Therefore, $\partial u(x, s)/\partial s$ exists for almost all s and

$$u(x, t) = u(x, 0) + \int_0^t \frac{\partial u(x, s)}{\partial s}\, ds.$$

In view of (6.6),

$$u(x, t) = u(x, 0) - \int_0^t L_s u(x, s)\, ds.$$

But since $L_s u(x, s)$ is a continuous function, it follows that $\partial u(x, t)/\partial t$ exists everywhere and (6.1) holds.

Next,

$$u(x, t) - f(x) = Ef(\xi_{x,t}(T)) - Ef(\xi_{x,T}(T)).$$

Using Lemma 3.3 and Lemma 1.3.6, we find that the right-hand side converges to 0 if $t \uparrow T$. This proves (6.2).

PROBLEMS

1. In Theorem 2.2, replace (2.5) by the inequalities

$$x \cdot b(x, t) \leqslant K(1 + |x|^2), \qquad |\sigma(x)| \leqslant K,$$

and prove that the assertions of Theorem 2.2 are still valid. [*Hint*: Apply Itô's formula to $\xi_N^2(t)$. Prove that

$$E \sup_{0 \leqslant t \leqslant T} |\xi_N(t)|^2 \leqslant C,$$

C independent of N.]

2. Prove that if (A′) holds, then for every $x \in R^n$, $s \geqslant 0$,

$$E|\xi_{y,0}(t) - \xi_{x,0}(s)|^2 \leqslant \eta(|y - x|^2 + |t - s|)$$

where $\eta(r) \to 0$ if $r \to 0$. [*Hint*: Use (2.7).]

3. Prove Theorem 3.6.

4. If $f \in L_w^2[0, T]$ and τ is a stopping time, $0 \leqslant \tau \leqslant T$, then

$$\int_\tau^T f(t)\, dw(t) = \int_0^{T-\tau} f(\tau + s)\, d\hat{w}(s)$$

where $\hat{w}(s) = w(s + \tau) - w(\tau)$.

5. If $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, \xi(t), P_{x,s}\}$ is a Markov process with transition probability $p(s, x, t, A)$ and if $g(x, t)$ is a continuous function in $(x, t) \in R^1 \times [0, \infty)$, strictly monotone in x, then $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, g(\xi(t), t), P_{x,s}\}$ is a Markov process with the transition probability $\tilde{p}(s, x, t, A) = p(s, g^{-1}(x, s), t, g^{-1}(A, t))$

where $g^{-1}(x, t)$ is the inverse function to $g(x, t)$.

6. If the Markov process in the preceding problem is a diffusion process with drift $b(x, t)$ and diffusion coefficient $a(x, t)$ and if g_t, g_x, g_{xx} are continuous and $g_x(x, t) \neq 0$ for all (x, t), then the process $\{\Omega, \mathcal{F}, \mathcal{F}_t^s, g(\xi(t), t), P_{x,s}\}$ is also a diffusion process with drift and diffusion coefficients given by:

$$\tilde{b}(x, t) = g_t(y, t) + b(y, t) g_x(y, t) + \tfrac{1}{2} a(y, t) g_{xx}(y, t),$$

$$\tilde{a}(x, t) = a(y, t)(g_x(y, t))^2, \qquad y = g^{-1}(x, t).$$

7. Complete the details of the proof of Theorem 5.1.

8. Prove Theorem 5.4.

9. Complete the proof of Theorem 5.5.

10. Consider a linear stochastic differential equation

$$d\xi(t) = [\alpha(t) + \beta(t)\xi(t)]\, dt + [\gamma(t) + \delta(t)\xi(t)]\, dw(t) \tag{6.7}$$

where the functions α, β, γ, δ are bounded and measurable. Prove:

(i) If $\alpha \equiv 0$, $\gamma \equiv 0$, then the solution $\xi = \xi_0(t)$ is given by

$$\xi_0(t) = \xi_0(0) \exp\left\{ \int_0^t [\beta(s) - \tfrac{1}{2}\delta^2(s)]\, ds + \int_0^t \delta(s)\, dw(s) \right\}.$$

(ii) Setting $\xi(t) = \xi_0(t)\zeta(t)$ show that $\xi(t)$ is a solution of (6.7) if and only if

$$\zeta(t) = \zeta(0) + \int_0^t [\xi_0(s)\alpha(s) - \gamma(s)\delta(s)]\, ds + \int_0^t \gamma(s)\xi_0(s)\, ds.$$

Thus the solution of (6.7) is $\xi_0(t)\zeta(t)$ with $\xi(0) = \xi_0(0)\zeta(0)$.

11. If an equation of the form $d\xi = b(x, t)\, dt + \sigma(x, t)\, dw(t)$ can be reduced by a transformation $\eta(t) = f(\xi(t), t)$ to an equation of the form $d\eta(t) = \beta(t)\, dt + \delta(t)\, dw(t)$, then

$$\frac{\partial}{\partial x}\left\{ \sigma\left[\frac{\sigma_t}{\sigma^2} - \frac{\partial}{\partial x}\left(\frac{b}{\sigma}\right) + \tfrac{1}{2}\sigma_{xx} \right] \right\} = 0.$$

12. Let $g(x, t)$, $h(x, t)$ be continuous functions in $(x, t) \in R^1 \times [0, T]$ together with their first two x-derivatives, and assume that

$$|D_x^\alpha g(x, t)| + |D_x^\alpha h(x, t)| \leqslant K_0(1 + |x|^\beta) \qquad \text{if} \quad |\alpha| \leqslant 2,$$

where K_0, β are positive constants. Let the conditions of Theorem 5.5 also hold, and consider the function

$$u(x, t) = E \exp\left\{ i\lambda \int_t^T g(\xi_{x,t}(s), s)\, ds + i\mu \int_t^T h(\xi_{x,t}(s), s)\, dw(s) \right\} f(\xi_{x,t}(T)).$$

Prove that

$$\frac{\partial u}{\partial t} + Lu + i\mu h u_x + \left(i\lambda g - \tfrac{1}{2}\mu^2 h^2\right)u = 0 \qquad \text{if} \quad x \in R^1, \quad 0 \leqslant t < T,$$

$$u(x, t) \to f(x) \qquad \text{if} \quad t \uparrow T.$$

13. Verify that all the results of this chapter remain true for a system of n stochastic differential equations

$$d\xi_i = \sum_{j=1}^{l} \sigma_{ij}(\xi, t)\, dw_j + b_i(\xi, t)\, dt \qquad (1 \leqslant i \leqslant n)$$

where $(w_1, \ldots, w_l)$ is an l-dimensional Brownian motion.

6

Elliptic and Parabolic Partial Differential Equations and Their Relations to Stochastic Differential Equations

1. Square root of a nonnegative definite matrix

As is well known, if a is an $n \times n$ nonnegative definite matrix, then it has a unique nonnegative square root σ, i.e., there is a unique nonnegative definite matrix σ such that $\sigma\sigma = a$. If a depends on a parameter x, say $a = a(x)$, then also $\sigma = \sigma(x)$. We shall be concerned here with the smoothness of $\sigma(x)$ in x, assuming that $a(x)$ varies smoothly with x.

Set $a(x) = (a_{ij}(x))$, $\sigma(x) = (\sigma_{ij}(x))$. The point x varies in an open set G of R^ν.

We shall denote by $C^m(G)$ the class of all functions $f(x)$ having continuous derivatives up to order m in G. If the mth derivatives of $f(x)$ satisfy a Hölder condition with exponent α in compact subsets of G, then we say that f belongs to $C^{m+\alpha}(G)$; here $0 < \alpha \leqslant 1$.

Lemma 1.1. *If $a(x)$ is positive definite for all $x \in G$ and if $a_{ij} \in C^{m+\alpha}(G)$ for all i, j* (or *if $a_{ij} \in C^m(G)$ for all i, j*), *then the elements of $\sigma(x)$ belong to $C^{m+\alpha}(G)$* (or *to $C^m(G)$*).

Proof. Let G_0 be an open bounded set with closure in G. Let Γ be a smooth contour lying in the half-plane $\operatorname{Re} z > 0$ of the complex plane and containing all the eigenvalues of $a(x)$, $x \in G_0$. We claim that if $x \in G_0$, then

$$\sigma(x) = \frac{1}{2\pi} \int_\Gamma \sqrt{z}\,(a(x) - zI)^{-1}\,dz \qquad (I = \text{identity matrix}). \quad (1.1)$$

To prove it, let Γ' be another smooth contour in $\operatorname{Re} z > 0$ whose interior contains Γ. Then

$$\begin{aligned}\sigma^2(x) &= \frac{1}{4\pi^2}\int_{\Gamma'}\int_{\Gamma}\sqrt{z}\sqrt{\zeta}\,(a(x)-\zeta I)^{-1}(a(x)-zI)^{-1}\,dz\,d\zeta\\ &= \frac{1}{4\pi^2}\int_{\Gamma'}\left\{\int_{\Gamma}\sqrt{z}\sqrt{\zeta}\,\frac{(a(x)-\zeta I)^{-1}-(a(x)-zI)^{-1}}{\zeta-z}\,dz\right\}d\zeta\\ &= -\frac{1}{4\pi^2}\int_{\Gamma'}\int_{\Gamma}\frac{\sqrt{z}\sqrt{\zeta}\,(a(x)-zI)^{-1}}{\zeta-z}\,dz\,d\zeta\end{aligned}$$

by Cauchy's theorem. Changing the order of integration and using Cauchy's theorem, we get

$$\begin{aligned}\sigma^2(x) &= -\frac{1}{4\pi^2}\int_{\Gamma}(a(x)-zI)^{-1}\sqrt{z}\int_{\Gamma'}\frac{\sqrt{\zeta}}{\zeta-z}\,d\zeta\,dz\\ &= \frac{1}{2\pi i}\int_{\Gamma}(a(x)-zI)^{-1}z\,dz = \frac{1}{2\pi i}\int_{\Gamma}(a(x)-zI)^{-1}a(x)\,dz.\end{aligned}$$

Now modify Γ into the disk $\Gamma_R = \{z;\ |z| = R\}$, R large. Since

$$\|(a(x)-zI)^{-1}-(zI)^{-1}\| \leq \frac{C}{1+|z|^2} \qquad \text{if} \quad |z| = R,$$

$$\left|\int_{\Gamma_R}\left[(a(x)-zI)^{-1}-(zI)^{-1}\right]dz\right| \leq \frac{2\pi CR}{1+R^2} \to 0 \qquad \text{if} \quad R\to\infty.$$

It follows that

$$\sigma^2(x) = \lim_{R\to\infty}\left\{\frac{1}{2\pi i}\int_{\Gamma_R}a(x)\,\frac{dz}{z}\right\} = a(x).$$

From (1.1) it is obvious that the elements $\sigma(x)$ are as smooth in G_0 as the elements of $a(x)$. Since G_0 is arbitrary, the proof of the lemma is complete.

Theorem 1.2. *If $a(x)$ is nonnegative definite for all $x \in R^\nu$ and if the $a_{ij}(x)$ belong to $C^2(R^\nu)$, then the $\sigma_{ij}(x)$ are Lipschitz continuous in compact subsets. If, further,*

$$\sup_{x\in R^\nu}\ \sup_{i,j,k,l}\left|\frac{\partial^2 a_{ij}(x)}{\partial x_k\,\partial x_l}\right| \leq M, \tag{1.2}$$

then

$$|\sigma_{ij}(x)-\sigma_{ij}(y)| \leq 2\nu\sqrt{\nu}\,\sqrt{M}\,|x-y|. \tag{1.3}$$

Proof. Suppose first that $a(x)$ is positive definite and (1.2) holds. Consider the function $\langle a(x)\xi, \xi\rangle$ where $\langle\ ,\ \rangle$ denotes the scalar product. By Taylor's formula,

$$0 \leqslant \langle a(x_1, \ldots, x_{k-1}, x_k + h, x_{k+1}, \ldots, x_\nu)\xi, \xi\rangle$$
$$= \langle a(x)\xi, \xi\rangle + h\left\langle \frac{\partial a(x)}{\partial x_k}\xi, \xi\right\rangle + \frac{h^2}{2}\left\langle \frac{\partial^2 a(\tilde{x})}{\partial x_k^2}\xi, \xi\right\rangle$$

for a suitable point $\tilde{x}$. Using (1.2) we get

$$0 \leqslant \langle a(x)\xi, \xi\rangle + h\left\langle \frac{\partial a(x)}{\partial x_k}\xi, \xi\right\rangle + \frac{h^2}{2}\nu M|\xi|^2.$$

Since the right-hand side is a quadratic polynomial in h taking only nonnegative values,

$$\left\langle \frac{\partial a(x)}{\partial x_k}\xi, \xi\right\rangle^2 \leqslant 2\nu M|\xi|^2\langle a(x)\xi, \xi\rangle. \tag{1.4}$$

Denoting by $'$ the derivative with respect to x_k and differentiating $\sigma^2(x) = a(x)$ with respect to x_k, we obtain

$$a'(x) = \sigma(x)\sigma'(x) + \sigma'(x)\sigma(x). \tag{1.5}$$

Let $T(x)$ be an orthogonal matrix such that $T(x)\sigma(x)T^{-1}(x)$ is a diagonal matrix $\Lambda(x)$. Multiplying (1.5) by $T(x)$ on the left and by $T^{-1}(x)$ on the right, we get

$$\bar{a}'(x) \equiv T(x)a'(x)T^{-1}(x) = Y(x)\Lambda(x) + \Lambda(x)Y(x) \tag{1.6}$$

where $Y(x) = (y_{ij}(x)) = T(x)\sigma'(x)T^{-1}(x)$. Denoting by $\lambda_i(x)$ the eigenvalues of $\sigma(x)$, we obtain from (1.6),

$$y_{ij} = \frac{\bar{a}'_{ij}}{\lambda_i + \lambda_j} \qquad \text{where} \quad \bar{a}'(x) = (\bar{a}'_{ij}(x)). \tag{1.7}$$

Using the estimate (1.4) we find that

$$\langle \bar{a}'(x)\xi, \xi\rangle^2 = \langle T(x)a'(x)T^{-1}(x)\xi, \xi\rangle^2 = \langle a'(x)T^{-1}(x)\xi, T^{-1}(x)\xi\rangle^2$$
$$\leqslant 2\nu M|T^{-1}(x)\xi|^2\langle a(x)T^{-1}(x)\xi, T^{-1}(x)\xi\rangle = 2\nu M|\xi|^2\langle \Lambda^2\xi, \xi\rangle. \tag{1.8}$$

Taking $\xi = (\xi_1, \ldots, \xi_\nu)$ with $\xi_j = 0$ if $j \neq i$, $\xi_i = 1$, we get

$$|\bar{a}'_{ii}(x)|^2 \leqslant 2\nu M\lambda_i^2(x). \tag{1.9}$$

Next, taking in (1.8) $\xi = (\xi_1, \ldots, \xi_\nu)$ with $\xi_k = 0$ if $k \neq i$, $k \neq j$ (where $i \neq j$) and $\xi_i = \xi_j = 1$, we find that

$$|\bar{a}'_{ii}(x) + 2\bar{a}'_{ij}(x) + \bar{a}'_{jj}(x)|^2 \leqslant 4\nu M(\lambda_i^2(x) + \lambda_j^2(x)).$$

Hence

$$\left[\frac{\bar{a}'_{ii}(x)}{\lambda_i + \lambda_j} + \frac{2\bar{a}'_{ij}(x)}{\lambda_i + \lambda_j} + \frac{\bar{a}'_{jj}(x)}{\lambda_i + \lambda_j} \right]^2 \leqslant 4\nu M.$$

Recalling (1.9) we get

$$2\left| \frac{\bar{a}'_{ij}(x)}{\lambda_i + \lambda_j} \right| \leqslant \sqrt{2\nu M} + 2\sqrt{\nu M} . \tag{1.10}$$

Substituting (1.9), (1.10) into (1.7), we obtain

$$| y_{il}(x)| < 2\sqrt{\nu M} \qquad \text{if} \quad 1 \leqslant i, \quad l \leqslant \nu. \tag{1.11}$$

Since $(\sigma'_{ij}(x)) = T^{-1}(x)Y(x)T(x)$ and $T(x) = (T_{ij}(x))$ is orthogonal,

$$|\sigma'_{ij}(x)| \leqslant \left\{ \max_{k,l} | y_{kl}(x)| \right\} \sum_{k,l} |T_{ik}(x)T_{lj}(x)| \leqslant 2\nu\sqrt{\nu M} . \tag{1.12}$$

Suppose now that $a(x)$ is nonnegative definite and (1.2) holds. Then the matrix $a(x) + \epsilon I$ $(\epsilon > 0)$ is positive definite and satisfies (1.2) with M replaced by a constant M_ϵ, $M_\epsilon \downarrow M$ if $\epsilon \downarrow 0$. By what we have already proved, the square root $\sigma^\epsilon(x)$ of $a(x) + \epsilon I$ satisfies

$$|\sigma^\epsilon_{ij}(x) - \sigma^\epsilon_{ij}(y)| \leqslant 2\nu\sqrt{\nu M_\epsilon}\, |x - y|.$$

We can apply the Ascoli–Arzela lemma to conclude that, for some sequence $\epsilon_m \downarrow 0$, $\sigma^{\epsilon_m}(x) \to \sigma(x)$ uniformly on compact sets. This gives (1.3). We have thus completed the proof of the theorem in case (1.2) holds.

If the a_{ij} belong to $C^2(R^\nu)$ but (1.2) is not assumed to hold, then we can slightly modify the previous proof, making use of the fact that for any $R > 0$

$$\sup_{|x|<R} \sup_{i,j,k,l} \left| \frac{\partial^2 a_{ij}(x)}{\partial x_k\, \partial x_l} \right| \leqslant M(R) \qquad (M(R) \text{ constant}).$$

Remark 1. If $a(x)$ is nonnegative definite in an open set G and $a_{ij} \in C^2(G)$, then the square root $\sigma(x)$ is Lipschitz continuous in compact subsets of G. Indeed, this follows from the proof of Theorem 1.2.

Remark 2. If $n = 1$ and $a_{11}(x) = |x|^\lambda$, then $\sigma_{11}(x) = |x|^{\lambda/2}$. If $\lambda = 2$, then a_{11} is in C^2 and σ_{11} is Lipschitz continuous, but if $\lambda = 2 - \epsilon$ for any small $\epsilon > 0$, then a_{11} is in $C^{2-\epsilon}$ and σ_{11} is not Lipschitz continuous. This example shows that assumption that $a_{ij} \in C^2$ in Theorem 1.2 is sharp.

***Remark* 3.** If $a(x)$ is nonnegative definite in a closed domain G with smooth boundary, and if $a_{ij} \in C^\infty(G)$, there does not exist in general a nonnegative definite matrix σ in G that is Lipschitz continuous and satisfies $\sigma^2 = a$ in G. For example, if $n = 1$ and $G = \{x;\ 0 \leqslant x \leqslant 1\}$, $a_{11}(x) = x$, then $\sigma_{11}(x) = \sqrt{x}$ is not Lipschitz continuous.

2. The maximum principle for elliptic equations

Consider the linear partial differential operator

$$Lu \equiv \tfrac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i\, \partial x_j} + \sum_{i=1}^{n} b_i(x) \frac{\partial u}{\partial x_i} + c(x)u \tag{2.1}$$

with real coefficients defined in an n-dimensional domain D. L is said to be of *elliptic type* (or *elliptic*) at a point x^0 if the matrix $(a_{ij}(x^0))$ is positive definite, i.e., for any real vector $\xi \neq 0$, $\Sigma a_{ij}(x^0)\xi_i\xi_j > 0$.

Lemma 2.1. *Let $(a_{ij}(x))$ be a nonnegative definite matrix for $x \in D$, and let $c(x) \leqslant 0$. If $u(x)$ is in $C^2(D)$ and if it attains a positive maximum in D at a point x^0, then $Lu(x^0) \leqslant 0$.*

Proof. If A and B are nonnegative definite matrices, then $\operatorname{tr}(AB) \geqslant 0$. Indeed, if λ is an eigenvalue of AB, then $ABx = \lambda x$ for some $x \neq 0$. Taking the scalar product with Bx we see that $\lambda \geqslant 0$. Since $\operatorname{tr}(AB)$ is equal to the sum of the eigenvalues of AB, it follows that $\operatorname{tr}(AB) \geqslant 0$.

Applying the previous result to $A = (a_{ij}(x^0))$, $B = -(u_{x_i x_j}(x^0))$ we conclude that $\Sigma a_{ij}(x^0)u_{x_i x_j}(x^0) \leqslant 0$. Since also $u_{x_i}(x^0) = 0$, $c(x^0)u(x^0) \leqslant 0$, the assertion $Lu(x^0) \leqslant 0$ follows.

The *weak maximum principle* is the following theorem:

Theorem 2.2. *Let $(a_{ij}(x))$ be nonnegative definite matrix for all x in a bounded domain D, and assume that $c(x) \leqslant 0$ and*

$$\tfrac{1}{2} a_{11}(x)\lambda^2 + b_1(x)\lambda > 0 \qquad \textit{for some} \quad \lambda > 0 \quad \textit{and for all} \quad x \in D.$$

If $u \in C^2(D) \cap C^0(\bar{D})$ and $Lu \geqslant 0$ in D, and if $\max_{\bar{D}} u(x) > 0$, then

$$\operatorname*{l.u.b.}_{x \in D} u(x) \leqslant \max_{x \in \partial D} u(x)$$

where ∂D is the boundary of D.

Thus, if u takes positive values in $\bar{D}$, then the maximum of u in $\bar{D}$ is attained on ∂D; it may also be attained at points of D.

Proof. If the assertion is not true, then there is a point $\bar{x} \in D$ such that

$u(\bar{x}) > \max_{\partial D} u(x)$. Consider the function

$$k(x) = \exp(\lambda x_1^0) - \exp(\lambda x_1)$$

where x_1^0 is a real number such that $x_1 < x_1^0$ for all $x = (x_1, \ldots, x_n)$ in $\bar{D}$. It is clear that $k(x) > 0$ and $Lk(x) < 0$ in D. If ϵ is positive and sufficiently small, then the function $v = u - \epsilon k$ satisfies: $v(\bar{x}) > \max_{\partial D} v(x)$. Hence v assumes a positive maximum in $\bar{D}$ at some point $x^0 \in D$. By Lemma 2.1, $Lv(x^0) \leqslant 0$, i.e., $Lu(x^0) - \epsilon Lk(x^0) \leqslant 0$. Since $Lk(x^0) < 0$, we arrive at the inequality $Lu(x^0) < 0$; this contradicts the assumption that $Lu \geqslant 0$ in D.

Remark. If we apply the weak maximum principle to $-u$, we obtain the following assertion: If the coefficients of L satisfy the conditions imposed in Theorem 2.2, and if $u \in C^2(D) \cap C^0(\bar{D})$, $Lu \leqslant 0$ in D and u assumes a negative minimum in $\bar{D}$, then

$$\operatorname*{g.l.b.}_{x \in D} u(x) \geqslant \min_{x \in \partial D} u(x).$$

Definition. If there is a positive constant μ such that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geqslant \mu|\xi|^2$$

for all $x \in D$, $\xi \in R^n$, then we say that L is *uniformly elliptic* in D.

The *strong maximum principle* for elliptic operators is the following theorem.

Theorem 2.3. *Let L be a uniformly elliptic operator with bounded coefficients in compact subsets of D, and assume that $c(x) \leqslant 0$ in D. If $u \in C^2(D)$ and $Lu \geqslant 0$ in D, and if $u(x) \not\equiv$ const, then u cannot attain a positive maximum in D.*

Applying this result to $-u$ we conclude that if $Lu \leqslant 0$ in D and $u(x) \not\equiv$ const, then u cannot attain a negative minimum in D.

Proof. We shall assume that u takes a positive maximum M at some point $x^0 \in D$, and derive a contradiction. Since $u \not\equiv$ const, there is a point $x^1 \in D$ such that $u(x^1) < M$. Connect x^1 to x^0 by a continuous curve γ lying in D. Then there is a first point x^2 along γ such that $u(x) < M$ if x lies on γ between x^1 and x^2, and $u(x^2) = M$ (x^2 may coincide with x^0). Take a point x^* on γ between x^1 and x^2 such that the ball $\{x;\ |x - x^*| \leqslant |x^2 - x^*|\}$ lies in D. Let x^{**} be the nearest point of the set $\{x \in D;\ u(x) = M\}$ to x^*. Clearly, the closure of the ball $B^* = \{x;\ |x - x^*| < |x^{**} - x^*|\}$ lies in D, $u(x) < M$ for all $x \in B^*$, and $u(x^{**}) = M$. Take a ball B contained in B^* whose boundary ∂B touches the boundary of B^* at x^{**}. Then $u(x) < M$ for all $x \in B \cup \partial B$, $x \neq x^{**}$. Denote the center of B by $\bar{x}$.

Denote by E a closed ball with center x^{**} and radius $< |x^{**} - \bar{x}|$ lying in D. Its boundary consists of a part E_1 lying in B and a part E_2 lying outside B. On E_1, $u \leqslant M - \delta$ for some $\delta > 0$.

Set $R = |x^{**} - \bar{x}|$, and consider the function

$$h(x) = \exp[-\alpha|x - \bar{x}|^2] - \exp[-\alpha R^2].$$

It satisfies: $h \geqslant 0$ in B, $h < 0$ outside B, $h = 0$ on ∂B, and $Lh > 0$ in E if α is sufficiently large. Consider the function $v = u + \epsilon h$ $(\epsilon > 0)$. If ϵ is sufficiently small, then $v < M$ on E_1, $v < M$ on E_2, and $Lv > 0$ in E. By Lemma 2.1, v cannot take a positive maximum in the interior of E. Since however $v(x^{**}) = M > 0$, we get a contradiction.

Consider the problem of finding a solution u of

$$Lu(x) = f(x) \quad \text{in} \quad D, \tag{2.2}$$

$$u(x) = \phi(x) \quad \text{on} \quad \partial D. \tag{2.3}$$

This is called the *Dirichlet problem* or the *first boundary value problem*.

A *barrier* $w_y(x)$ at the point $y \in \partial D$ is a continuous nonnegative function in $\bar{D}$ that vanishes only at the point y and for which $Lw_y(x) \leqslant -1$.

If there exists a closed ball K such that $K \cap D = \emptyset$, $K \cap \bar{D} = \{y\}$, then

$$w_y(x) = k\{|x^0 - y|^{-p} - |x - y|^{-p}\} \qquad (x^0 = \text{center of } K) \tag{2.4}$$

is a barrier at y provided k and p are sufficiently large.

Theorem 2.4. *Assume that L is uniformly elliptic in D, that $c(x) \leqslant 0$, and that a_{ij}, b_i, c, f are uniformly Hölder continuous (exponent α) in $\bar{D}$. If every point of ∂D has a barrier and if ϕ is a continuous function on ∂D, then there exists a unique solution u in $C^2(D) \cap C^0(\bar{D})$ of the Dirichlet problem* (2.2), (2.3).

For the proof of existence the reader is referred to Friedman [1, 2]. The uniqueness follows from the maximum principle.

3. The maximum principle for parabolic equations

Consider the partial differential operator

$$Mu \equiv \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x, t) \frac{\partial^2 u}{\partial x_i \, \partial x_j} + \sum_{i=1}^{n} b_i(x, t) \frac{\partial u}{\partial x_i} + c(x, t)u - \frac{\partial u}{\partial t} \tag{3.1}$$

with real coefficients defined in an $(n + 1)$-dimensional domain Q. If $\sum a_{ij}(x^0, t^0)\xi_i\xi_j > 0$ for all $\xi \in R^n$, $\xi \neq 0$, then we say that M is of *parabolic type* at (x^0, t^0). M is *uniformly parabolic* in Q if there is a positive constant μ

such that

$$\sum a_{ij}(x, t)\xi_i\xi_j \geqslant \mu|\xi|^2 \qquad \text{for all} \quad (x, t) \in Q, \quad \xi \in R^n.$$

Suppose Q lies in the strip $0 < t < T$, and define

$$D_T = \{(x, T);\ (x, T-\delta) \in Q \text{ for all } 0 < \delta < \delta_0,\ \delta_0 \textit{ depending on } x\},$$

$$Q_0 = Q \cup D_T, \qquad \partial Q = \text{boundary of } Q, \qquad \partial_0 Q = \partial Q \setminus (\text{int } D_T).$$

The *weak maximum principle* is the following theorem:

Theorem 3.1. *Let $(a_{ij}(x, t))$ be nonnegative definite matrix for all (x, t) in a bounded domain Q, and assume that $c(x, t) \leqslant 0$. If $u \in C^0(\bar{Q})$ and u_{x_i}, $u_{x_ix_j}$, u_t belong to $C^0(Q_0)$, and if $Mu \geqslant 0$ in Q_0 and $\max_{\bar{Q}} u > 0$, then*

$$\underset{(x,t) \in Q_0}{\text{l.u.b.}}\ u(x, t) \leqslant \max_{(x,t) \in \partial_0 Q} u(x, t).$$

Proof. If the assertion is false, then, for some $\epsilon > 0$ sufficiently small, the function $v = u - \epsilon t$ takes positive maximum in $\bar{Q}$ at a point (x^0, t^0) in $Q \cup D_T$. If $t^0 = T$, then $v_t(x^0, t^0) \leqslant 0$; and if $t^0 < T$, then $v_t(x^0, t^0) = 0$. Employing Lemma 2.1 we conclude that $Mv(x^0, t^0) \leqslant 0$. Since however $Mu \geqslant 0$ in Q_0 and $M(-\epsilon t) = -c\epsilon t + \epsilon > 0$, we have $Mv(x^0, t^0) > 0$, a contradiction.

For any point $P^0 = (x^0, t^0)$ in Q, denote by $S(P^0)$ the set of all points P in Q that can be connected to P^0 by a simple continuous curve in Q along which the t-coordinate is nondecreasing from P to P^0. We denote by $C(P^0)$ the component, in $t = t^0$, of $Q \cap \{t = t^0\}$ that contains P^0. Notice that $C(P^0) \subset S(P^0)$.

The *strong maximum principle* for parabolic equations states the following.

Theorem 3.2. *Let M be uniformly parabolic in a domain Q, with bounded coefficients, and assume that $c(x, t) \leqslant 0$. If u, u_{x_i}, $u_{x_ix_j}$, u_t are continuous in Q and $Mu \geqslant 0$ in Q, and if u takes a positive maximum in Q at a point $P^0 = (x^0, t^0)$, then $u(P) = u(P^0)$ for all $P \in S(P^0)$.*

We need a few lemmas.

Lemma 3.3. *Let $Mu \geqslant 0$ in Q and suppose that u takes a positive maximum K in Q. Suppose that Q contains a closed solid ellipsoid E:*

$$\sum_{i=1}^{n} \lambda_i(x_i - x_i^*)^2 + \lambda_0(t - t^*)^2 \leqslant R^2 \qquad (\lambda_j > 0, \quad R > 0),$$

$u < K$ in the interior of E, and $u(\bar{x}, \bar{t}) = K$ for some point $\bar{P} = (\bar{x}, \bar{t})$ on the boundary ∂E of E. Then $\bar{x} = x^*$ where $x^* = (x_1^*, \ldots, x_n^*)$.

Proof. We may suppose that $\bar{P}$ is the only point on ∂E where $u = K$ since otherwise we confine ourselves to a smaller ellipsoid lying in E and touching ∂E at $\bar{P}$ only. Suppose $\bar{x} \neq x^*$ and let C be an $(n + 1)$-dimensional closed ball contained in Q with center $\bar{P}$ and radius $< |\bar{x} - x^*|$. Then

$$|x - x^*| \geqslant \text{const} > 0 \qquad \text{for all} \quad (x, t) \in C. \tag{3.2}$$

The boundary of C is composed of a part C_1 lying in E and a part C_2 lying outside E. Clearly

$$u \leqslant K - \delta \qquad \text{on } C_1, \qquad \text{for some constant} \quad \delta > 0.$$

Consider the function

$$h(x, t) = \exp\left\{ -\alpha\left[\sum_{i=1}^{n} \lambda_i (x_i - x_i^*)^2 + \lambda_0 (t - t^*)^2 \right] \right\} - \exp\{-\alpha R^2\}$$

where α is a positive constant. Clearly $h > 0$ in the interior of E, $h = 0$ on ∂E, and $h < 0$ outside E.

Using (3.2) one easily verifies that $Mh > 0$ in C if α is sufficiently large. Consider now the function $v = u + \epsilon h$ in C, where $\epsilon > 0$. If ϵ is sufficiently small, then $v < K$ on both C_1 and C_2. Since $v(\bar{P}) = K$, it follows that v takes its positive maximum in C at an interior point $\tilde{P} = (\tilde{x}, \tilde{t})$. By Lemma 2.1, $Mv \leqslant 0$ at $\tilde{P}$. Since, however, $Mv = Mu + \epsilon Mh > 0$ in C, we get a contradiction.

Lemma 3.4. *If $Mu \geqslant 0$ in Q and if u takes a positive maximum in Q at a point $P^0 = (x^0, t^0)$, then $u(P) = u(P^0)$ for all $P \in C(P^0)$.*

Proof. If the assertion is not true, then there exists a point $P^1 = (x^1, t^0)$ in $C(P^0)$ such that $u(P^1) < u(P^0)$. Connect P^1 to P^0 by a simple continuous curve γ lying in $C(P^0)$. Then there exists a point $P^* = (x^*, t^0)$ on γ such that $u(P^*) = u(P^0)$ and $u(\bar{P}) < u(P^0)$ for all $\bar{P}$ on γ between P^1 and P^*. Take one such point $\bar{P} = (\bar{x}, t^0)$ whose distance to the boundary of Q is $\geqslant 2|\bar{P} - P^*|$.

Since $u(\bar{P}) < u(P^*)$, there exists a sufficiently small interval $\sigma = \{(\bar{x}, t); t^0 - \epsilon \leqslant t \leqslant t^0 + \epsilon\}$ such that

$$u(P) < u(P^*) \qquad \text{for all} \quad P \in \sigma. \tag{3.3}$$

Consider the family of ellipsoids E_λ:

$$|x - \bar{x}|^2 + \lambda(t - t^0)^2 \leqslant R^2 \qquad (\lambda > 0).$$

Take $R^2 = \lambda\epsilon^2$ so that the end points of σ lie on the boundary of E_λ. If $\lambda \to 0$, $E_\lambda \to \sigma$. As λ increases, $E_\lambda \cap \{t = t^0\}$ increases; for sufficiently large λ it will contain the point P^*. Hence, because of (3.3), there is a first value of λ, say $\lambda = \lambda_0$, such that $u < u(P^*)$ in the interior of E_{λ_0} and $u = u(P^*)$ at some boundary point $\hat{P} = (y, \hat{t})$ of E_{λ_0}. Since $u < u(P^*)$ on σ, it is clear that $y \neq \bar{x}$. This contradicts Lemma 3.3.

Lemma 3.5. *Let R be a rectangle*

$$x_i^0 - a_i \leqslant x_i \leqslant x_i^0 + a_i \qquad (i = 1, \ldots, n), \qquad t^0 - a_0 \leqslant t \leqslant t^0$$

contained in Q, and let $Mu \geqslant 0$ in Q. If u takes a positive maximum in R at the point $P^0 = (x^0, t^0)$, where $x^0 = (x_1^0, \ldots, x_n^0)$, then $u(P) = u(P^0)$ for all $P \in R$.

Proof. If the assertion is not true, then there is a point $\hat{P}$ in R such that $u(\hat{P}) < u(P^0)$. Since then $u < u(P^0)$ in a neighborhood of $\hat{P}$, we may assume that $\hat{P}$ does not lie on $t = t^0$. On the straight segment γ connecting $\hat{P}$ to P^0 there exists a point P^1 such that $u(P^1) = u(P^0)$ and $u(\bar{P}) < u(P^1)$ for all $\bar{P}$ on γ between $\hat{P}$ and P^1. Without loss of generality we may assume that $P^1 = P^0$ and that $\hat{P}$ lies on $t = t^0 - a_0$, for otherwise we can confine ourselves to a smaller rectangle.

Denote by R_0 the subset of R consisting of all points with $t < t^0$. Since, for every P^1 in R_0, $C(P^1)$ contains a point of γ, and since $u < u(P^0)$ on γ, Lemma 3.4 implies that $u(P^1) < u(P^0)$.

Consider the function

$$h(x, t) = t^0 - t - A|x - x^0|^2 \qquad (A > 0).$$

Clearly $h = 0$ on the paraboloid Π: $t^0 - t = A|x - x^0|^2$, $h < 0$ above Π, and $h > 0$ below Π. Also, as easily seen, $Mh > 0$ in R if $4A\Sigma a_{ii} \leqslant 1$ in R and if the dimensions of R are sufficiently small.

Π divides R into two regions. Denote by R' the lower region. Its upper boundary B' touches $t = t^0$ at the point P^0 only. Hence on the complementary boundary B'' of R' we have $u \leqslant u(P^0) - \delta$ for some $\delta > 0$. It follows that if $\epsilon > 0$ is sufficiently small, then the function $v = u + \epsilon h$ satisfies $v < u(P^0)$ on B''. Next, $v = u < u(P^0)$ at all the points of B' other than P^0. Since $Mv = Mu + \epsilon Mh > 0$ in R', v cannot take a positive maximum in R' at an interior point (by Lemma 2.1). We thus conclude that the maximum of v in the closure of R' is attained at just one point, namely, at P^0. This implies that $\partial v(P^0)/\partial t \geqslant 0$. Since $\partial h(P^0)/\partial t = -1$, we conclude that

$$\partial u(P^0)/\partial t > 0. \tag{3.4}$$

Since however u takes a positive maximum at P^0, Lemma 2.1 implies that $Mu \leqslant -\partial u/\partial t$ at P^0. But since we have assumed that $Mu \geqslant 0$ in Q, $\partial u(P^0)/\partial t \leqslant 0$. This contradicts (3.4).

Proof of Theorem 3.2. Suppose $u \not\equiv u(P^0)$ in $S(P^0)$. Then there exists a point P in $S(P^0)$ such that $u(P) < u(P^0)$. Connect P to P^0 by a simple continuous curve γ lying in $S(P^0)$ such that the t-coordinate is nondecreasing from P to P^0. There exists a point P^1 on γ such that $u(P^1) = u(P^0)$ and $u(\bar{P}) < u(P^1)$ for all $\bar{P}$ on γ lying between P and P^1. Construct a rectangle R:

$$x_i^1 - \alpha \leqslant x_i \leqslant x_i^1 + \alpha \qquad (i = 1, \ldots, n), \qquad t^1 - \alpha \leqslant t \leqslant t^1$$

where $P^1 = (x_1^1, \ldots, x_n^1, t^1)$ with a sufficiently small α so that R is contained in Q. Applying Lemma 3.5 we conclude that $u \equiv u(P^1)$ in R. This contradicts the definition of P^1.

Let Q be a bounded domain in the $(n + 1)$-dimensional space of variables (x, t). Assume that Q lies in the strip $0 < t < T$ and that $\tilde{B} \equiv \bar{Q} \cap \{t = 0\}$, $\tilde{B}_T \equiv \bar{Q} \cap \{t = T\}$ are nonempty. Let B_T = interior of $\tilde{B}_T$, B = interior of $\tilde{B}$. Denote by S_0 the boundary of Q lying in the strip $0 < t \leqslant T$, and let $S = S_0 \setminus B_T$. The set $\partial_0 Q = B \cup S$ is called the *normal* (or *parabolic*) *boundary* of Q.

The *first initial-boundary value problem* consists of finding a solution u of

$$Mu(x, t) = f(x, t) \qquad \text{in} \quad Q \cup B_T \tag{3.5}$$

$$u(x, 0) = \phi(x) \qquad \text{on } B, \tag{3.6}$$

$$u(x, t) = g(x, t) \qquad \text{on } S, \tag{3.7}$$

where f, ϕ, g are given functions. We refer to (3.6) as the *initial condition* and to (3.7) as the *boundary condition*. If $g = \phi$ on $\bar{B} \cap \bar{S}$, then the solution u is always understood to be continuous in $\bar{Q}$.

A function $w_R(P)$ $(R \in \bar{B} \cup S)$ is called a *barrier* at the point R if $w_R(P)$ is continuous in $P \in \bar{Q}$, $w_R(P) > 0$ if $P \in \bar{Q}$, $P \neq R$, $w_R(R) = 0$, and $Mw_R \leqslant -1$ in $Q \cup B_T$.

Suppose Q is a cylinder $B \times (0, T)$, and assume that there exists an n-dimensional closed ball K with center $\bar{x}$ such that $K \cap B = \emptyset$, $K \cap \bar{B} = \{x^0\}$. Then there exists a barrier at each point $R = (x^0, t^0)$ of S $(0 < t^0 \leqslant T)$, namely,

$$w_R(x, t) = ke^{\gamma t}\left(\frac{1}{R_0^p} - \frac{1}{R^p}\right)$$

where $\gamma \geqslant c(x, t)$, $R_0 = |x^0 - \bar{x}|$, $R = [|x - \bar{x}|^2 + (t - t^0)^2]^{1/2}$, and k, p are appropriate positive numbers.

Theorem 3.6. *Assume that M is uniformly parabolic in Q, that a_{ij}, b_i, c, f are uniformly Hölder continuous in $\bar{Q}$, and that g, ϕ are continuous functions on $\bar{B}$, $\bar{S}$ respectively and $g = \phi$ on $\bar{B} \cap \bar{S}$. Assume also that there exists a barrier at every point of S. Then there exists a unique solution u of the initial-boundary value problem* (3.5)–(3.7).

The solution u has Hölder continuous derivatives u_{x_i}, $u_{x_i x_j}$, u_t.

For the proof of existence we refer the reader to Friedman [1]. The uniqueness follows from the inequality

$$\max_{Q} |u| \leqslant e^{\alpha T} \max_{B \cup S} |u| \quad \text{if } Mu = 0 \text{ and } c(x, t) \leqslant \alpha \quad \text{in } Q. \tag{3.8}$$

This inequality follows from the weak maximum principle applied to $v = ue^{-\gamma t}$.

4. The Cauchy problem and fundamental solutions for parabolic equations

Let

$$Lu \equiv \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x, t) \frac{\partial^2 u}{\partial x_i \, \partial x_j} + \sum_{i=1}^{n} b_i(x, t) \frac{\partial u}{\partial x_i} + c(x, t)u \tag{4.1}$$

be an elliptic operator in R^n for each $t \in [0, T]$. Consider the parabolic equation

$$Mu \equiv Lu(x, t) - \frac{\partial u(x, t)}{\partial t} = f(x, t) \quad \text{in} \quad R^n \times (0, T] \tag{4.2}$$

with the *initial condition*

$$u(x, 0) = \phi(x) \qquad \text{on } R^n. \tag{4.3}$$

The problem of solving (4.2), (4.3), for given f, ϕ, is called the *Cauchy problem*. The solution is understood to be continuous in $R^n \times [0, T]$ and to have continuous derivatives u_{x_i}, $u_{x_i x_j}$, u_t in $R^n \times (0, T]$.

We first prove uniqueness.

Theorem 4.1. *Let $(a_{ij}(x, t))$ be nonnegative definite and let*

$$|a_{ij}(x, t)| \leqslant C, \qquad |b_i(x, t)| \leqslant C(|x| + 1), \qquad c(x, t) \leqslant C(|x|^2 + 1) \tag{4.4}$$

(C constant). If $Mu \leqslant 0$ in $R^n \times (0, T]$ and

$$u(x, t) \geqslant -B \exp[\beta |x|^2] \quad \text{in} \quad R^n \times [0, T] \tag{4.5}$$

for some positive constants B, β, and if $u(x, 0) \geqslant 0$ on R^n, then $u(x, t) \geqslant 0$ in $R^n \times [0, T]$.

Proof. Consider the function

$$H(x, t) = \exp\left\{ \frac{k|x|^2}{1 - \mu t} + \nu t \right\} \qquad \left(0 < t < \frac{1}{2\mu}\right).$$

One easily checks that, for every $k > 0$, if μ and ν are sufficiently large positive constants, then $MH \leqslant 0$. Consider the function $v = u/H$ and take $k > \beta$ and μ, ν such that $MH \leqslant 0$. From (4.5) it follows that

$$\liminf_{|x|\to\infty} v(x, t) \geqslant 0 \tag{4.6}$$

uniformly in t, $0 \leqslant t \leqslant 1/2\mu$. Further, v satisfies

$$\tilde{M}v \equiv \frac{1}{2} \sum a_{ij} \frac{\partial^2 v}{\partial x_i\, \partial x_j} + \sum \tilde{b}_i \frac{\partial v}{\partial x_i} + \tilde{c}v - \frac{\partial v}{\partial t} = \tilde{f} \qquad \left(0 < t < \frac{1}{2\mu}\right)$$

where $\tilde{f} = (Mu)/H \leqslant 0$ and

$$\tilde{b}_i = b_i + \sum a_{ij} \frac{\partial H/\partial x_j}{H}, \qquad \tilde{c} = \frac{MH}{H} \leqslant 0.$$

By (4.6), for any $\epsilon > 0$, $v(x, t) + \epsilon > 0$ if $|x| = R$, $0 \leqslant t \leqslant 1/(2\mu)$ provided R is sufficiently large. Also $\tilde{M}(v + \epsilon) \leqslant \tilde{c}\epsilon \leqslant 0$ if $0 < t < 1/(2\mu)$ and $v(x, 0) + \epsilon > 0$ if $|x| \leqslant R$. By the maximum principle, $v(x, t) + \epsilon > 0$ if $|x| \leqslant R$, $0 < t < 1/2\mu$. Taking $R \to \infty$ and then $\epsilon \downarrow 0$ it follows that $v(x, t) \geqslant 0$ if $0 \leqslant t \leqslant 1/2\mu$. Hence $u(x, t) \geqslant 0$ if $0 \leqslant t \leqslant 1/2\mu$. We can now proceed step-by-step to prove the nonnegativity of u in the strip $0 \leqslant t \leqslant T$.

Corollary 4.2. *If $(a_{ij}(x, t))$ is nonnegative definite and* (4.4) *holds, then there exists at most one solution u of the Cauchy problem* (4.2), (4.3) *satisfying*

$$|u(x, t)| \leqslant B \exp[\beta |x|^2]$$

where B, β are some positive constants.

The next result on uniqueness has different growth conditions on the coefficients of L.

Theorem 4.3. *Let $(a_{ij}(x, t))$ be nonnegative definite and let*

$$|a_{ij}(x, t)| \leqslant C(|x|^2 + 1), \quad |b_i(x, t)| \leqslant C(|x| + 1),$$
$$c(x, t) \leqslant C \quad (C \text{ constant}). \tag{4.7}$$

If $Mu \leqslant 0$ in $R^n \times (0, T]$ and

$$u(x, t) \geqslant -N(|x|^q + 1) \quad \text{in} \quad R^n \times [0, T] \tag{4.8}$$

where N, q are positive constants, and if $u(x, 0) \geqslant 0$ on R^n, then $u(x, t) \geqslant 0$ in $R^n \times [0, T]$.

Proof. For any $p > 0$, the function

$$w(x, t) = (|x|^2 + Kt)^p e^{\alpha t}$$

satisfies $Mw < 0$ in $R^n \times [0, T]$ provided K, α are appropriate positive constants. Take $2p > q$ and consider the function $v = u + \epsilon w$ for any $\epsilon > 0$. Then $Mv < 0$. Since $v(x, 0) \geqslant 0$ and (by (4.8)) $v(x, t) > 0$ if $|x| = R$ (R large), $0 < t < T$, the maximum principle can be applied (to ve^{-Ct}) to yield the inequality $v(x, t) \geqslant 0$ if $|x| \leqslant R$, $0 \leqslant t \leqslant T$. Taking first $R \to \infty$ and then $\epsilon \to 0$, the assertion follows.

Corollary 4.4. *Let $(a_{ij}(x, t))$ be nonnegative definite and let* (4.7) *hold. Then there exists at most one solution of the Cauchy problem* (4.2), (4.3) *satisfying*

$$|u(x, t)| \leqslant N(1 + |x|^q)$$

where N, q are some positive constants.

Definition. A *fundamental solution* of the parabolic operator $L - \partial/\partial t$ in $R^n \times [0, T]$ is a function $\Gamma(x, t; \xi, \tau)$ defined for all (x, t) and (ξ, τ) in $R^n \times [0, T]$, $t > \tau$, satisfying the following condition:

For any continuous function $f(x)$ with compact support, the function

$$u(x, t) = \int_{R^n} \Gamma(x, t; \xi, \tau) f(\xi)\, d\xi \tag{4.9}$$

satisfies

$$Lu - \partial u/\partial t = 0 \quad \text{if} \quad x \in R^n, \quad \tau < t \leqslant T, \tag{4.10}$$

$$u(x, t) \to f(x) \quad \text{if} \quad t \downarrow \tau. \tag{4.11}$$

We shall need the conditions:

(A_1) There is a positive constant μ such that

$$\sum a_{ij}(x, t)\xi_i\xi_j \geqslant \mu|\xi|^2 \qquad \text{for all} \quad (x, t) \in R^n \times [0, T], \quad \xi \in R^n.$$

(A_2) The coefficients of L are bounded continuous functions in $R^n \times [0, T]$, and the coefficients $a_{ij}(x, t)$ are continuous in t, uniformly with respect to (x, t) in $R^n \times [0, T]$.

(A_3) The coefficients of L are Hölder continuous (exponent α) in x, uniformly with respect to (x, t) in compact subsets of $R^n \times [0, T]$; furthermore, the $a_{ij}(x, t)$ are Hölder continuous (exponent α) in x uniformly with respect to (x, t) in $R^n \times [0, T]$.

Theorem 4.5. *Let* (A_1)–(A_3) *hold. Then there exists a fundamental solution $\Gamma(x, t, \xi, \tau)$ for $L - \partial/\partial t$ satisfying the inequalities*

$$|D_x^m \Gamma(x, t; \xi, \tau)| \leqslant C(t - \tau)^{-(n+|m|)/2} \exp\left[-c\frac{|x - \xi|^2}{t - \tau}\right] \tag{4.12}$$

for $|m| = 0, 1$, where C, c are positive constants. The functions $D_x^m \Gamma(x, t; \xi, \tau)$ $(0 \leqslant |m| \leqslant 2)$ and $D_t \Gamma(x, t; \xi, \tau)$ are continuous in

$(x, t, \xi, \tau) \in R^n \times [0, T] \times R^n \times [0, T]$, $t > \tau$, *and* $L\Gamma - \partial\Gamma/\partial t = 0$ *as a function in* (x, t). *Finally, for any continuous bounded function* $f(x)$,

$$\int_{R^n} \Gamma(x, t; \xi, \tau) f(x)\, dx \to f(\xi) \qquad \text{if} \quad t \downarrow \tau. \tag{4.13}$$

The construction of Γ can be given by the parametrix method; for details regarding the construction of this Γ and the proof of the other assertions of Theorem 4.5, see Friedman [1; Chapter 9]. (The assertion (4.13) follows from Friedman [1, p. 247, formula (2.29); p. 252, formula (4.4), and the estimate on $\Gamma - Z$].)

The following theorem is also proved in Friedman [1]:

Theorem 4.6. *Let* (A_1)–(A_3) *hold. Let* $f(x, t)$ *be a continuous function in* $R^n \times [0, T])$, *Hölder continuous in* x *uniformly with respect to* (x, t) *in compact subsets, and let* $\phi(x)$ *be a continuous function in* R^n. *Assume also that*

$$|f(x, t)| \leqslant A \exp(a|x|^2) \quad \text{in} \quad R^n \times [0, T], \tag{4.14}$$

$$|\phi(x)| \leqslant A \exp(a|x|^2) \quad \text{in} \quad R^n \tag{4.15}$$

where A, a *are positive constants. Then there exists a solution of the Cauchy problem* (4.2), (4.3) *in the strip* $0 \leqslant t \leqslant T^*$ *where* $T^* = \min\{T, \bar{c}/a\}$ *and* $\bar{c}$ *is a constant depending only on the coefficients of* L, *and*

$$|u(x, t)| \leqslant A' \exp(a'|x|^2) \quad \text{in} \quad R^n \times [0, T^*] \tag{4.16}$$

for some positive constants A', a'. *The solution is given by*

$$u(x, t) = \int_{R^n} \Gamma(x, t; \xi, 0)\phi(\xi)\, d\xi - \int_0^t \int_{R^n} \Gamma(x, t; \xi, \tau) f(\xi, \tau)\, d\xi\, d\tau. \tag{4.17}$$

The *formal adjoint operator* M^* of $M = L - \partial/\partial t$, where L is given by (4.1), is given by

$$M^* v = L^* v + \partial v/\partial t,$$

$$L^* v = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x, t) \frac{\partial^2 v}{\partial x_i\, \partial x_j} + \sum_{i,j=1}^n b_i^*(x, t) \frac{\partial v}{\partial x_i} + c^*(x, t) v \tag{4.18}$$

where

$$b_i^* = -b_i + \frac{1}{2} \sum_{j=1}^n \frac{\partial a_{ij}}{\partial x_j},$$

$$c^* = c - \sum_{i=1}^n \frac{\partial b_i}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 a_{ij}}{\partial x_i\, \partial x_j}. \tag{4.19}$$

It is assumed here that $\partial a_{ij}/\partial x_j$, $\partial^2 a_{ij}/\partial x_i\, \partial x_j$, $\partial b_i/\partial x_i$ exist and are bounded functions.

Notice that

$$M^*v = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i\,\partial x_j}(a_{ij}v) - \sum_{i=1}^{n}\frac{\partial}{\partial x_i}(b_i v) + cv + \frac{\partial v}{\partial t}. \tag{4.20}$$

Using this form of M^*v, one can easily verify *Green's identity*

$$vMu - uM^*v = \sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[\frac{1}{2}\sum_{j=1}^{n}\left(va_{ij}\frac{\partial u}{\partial x_j} - ua_{ij}\frac{\partial v}{\partial x_j} - uv\frac{\partial a_{ij}}{\partial x_j}\right) + b_i uv\right] - \frac{\partial}{\partial t}(uv). \tag{4.21}$$

Thus, if u, v have compact support in a domain G, then

$$\int\int_G (vMu - uM^*v)\,dx\,dt = 0. \tag{4.22}$$

Definition. A *fundamental solution* of the operator $L^* + \partial/\partial t$ in $R^n \times [0, T]$ is a function $\Gamma^*(x, t; \xi, \tau)$ defined for all (x, t) and (ξ, τ) in $R^n \times [0, T]$, $t < \tau$, satisfying the following condition:

For any continuous function $g(x)$ with compact support, the function

$$v(x, t) = \int_{R^n} \Gamma^*(x, t; \xi, \tau)g(\xi)\,d\xi$$

satisfies

$$L^*v + \partial v/\partial t = 0 \quad \text{if} \quad x \in R^n, \quad 0 \leqslant t < \tau,$$
$$v(x, t) \to g(x) \quad \text{if} \quad t \uparrow \tau.$$

We shall need the following condition:

(A_4) The functions

$$a_{ij}, \qquad \frac{\partial}{\partial x_i}a_{ij}, \qquad \frac{\partial^2}{\partial x_i\,\partial x_j}a_{ij}, \qquad b_i, \qquad \frac{\partial}{\partial x_i}b_i, \qquad c$$

are bounded functions, and the coefficients of L^* satisfy the conditions (A_2), (A_3).

Theorem 4.7. *If* (A_1)–(A_4) *hold, then there exists a fundamental solution* $\Gamma^*(x, t; \xi, \tau)$ *of* $L^* + \partial/\partial t$, *and*

$$\Gamma(x, t; \xi, \tau) = \Gamma^*(\xi, \tau; x, t) \qquad (t > \tau). \tag{4.23}$$

Proof. The construction of Γ^* can be carried out in the same way as for Γ. Further, Γ^* satisfies inequalities similar to (4.12) and $L^*\Gamma^* + \partial\Gamma^*/\partial t = 0$ as a function of (ξ, τ). Thus it remains to prove (4.23). Consider the functions

$$u(y, \sigma) = \Gamma(y, \sigma; \xi, \tau), \qquad v(y, \sigma) = \Gamma^*(y, \sigma; x, t)$$

for $y \in R^n$, $\tau < \sigma < t$. Integrating Green's identity (4.21) over the domain $|y| < R$, $\tau + \epsilon < \sigma < t - \epsilon$ $(R > 0)$ and using the relations $Mu = 0$, $M^*v = 0$, we obtain

$$\int_{|y|<R} u(y, t-\epsilon)v(y, t-\epsilon)\, dy - \int_{|y|<R} u(y, \tau+\epsilon)v(y, \tau+\epsilon)\, dy = I_{\epsilon, R}$$

where

$$I_{\epsilon, R} = \sum_{i=1}^{n} \int_{\tau+\epsilon}^{t-\epsilon} \int_{|y|=R} \left[\frac{1}{2} \sum_{j=1}^{n} \left(va_{ij} \frac{\partial u}{\partial y_j} - ua_{ij} \frac{\partial v}{\partial y_j} - uv \frac{\partial a_{ij}}{\partial y_j} \right) + b_i uv \right] \cos(\nu, y_i)\, dS_y\, d\sigma;$$

ν is the outwardly directed normal to $|y| = R$ and dS_y is the surface element on $|y| = R$. Using (4.12) and the corresponding inequalities for Γ^* we find that $I_{\epsilon, R} \to 0$ if $R \to \infty$. Hence

$$\int_{R^n} u(y, t-\epsilon)\Gamma^*(y, t-\epsilon; x, t)\, dy = \int_{R^n} v(y, \tau+\epsilon)\Gamma(y, \tau+\epsilon; \xi, \tau)\, dy. \tag{4.24}$$

Taking $\epsilon \downarrow 0$ the assertion (4.23) follows; cf. Problem 10.

5. Stochastic representation of solutions of partial differential equations

Let L be an elliptic operator in a bounded domain D given by (2.1). Assume that L is uniformly elliptic in D, i.e.,

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geqslant \mu|\xi|^2 \qquad \text{if} \quad x \in D, \quad \xi \in R^n \qquad (\mu > 0). \tag{5.1}$$

Assume also that

$$a_{ij},\ b_i \quad \text{are uniformly Lipschitz continuous in } \bar{D}, \tag{5.2}$$

$$c \leqslant 0,\ c \quad \text{uniformly Hölder continuous in } \bar{D}. \tag{5.3}$$

Assume finally that the boundary ∂D of D is in C^2, so that barriers exist at all the points of ∂D (see Problem 2).

Then, by Theorem 2.4, the Dirichlet problem (2.2), (2.3) has a unique solution u for any given functions f, ϕ satisfying:

$$f \quad \text{is uniformly Hölder continuous in } \bar{D}, \tag{5.4}$$

$$\phi \quad \text{is continuous on } \partial D. \tag{5.5}$$

We shall now represent u in terms of a solution of a stochastic differential system.

By Lemma 1.1 and (1.1) it is clear that the nonnegative definite square root $\sigma(x) = (\sigma_{ij}(x))$ of the matrix $a(x) = (a_{ij}(x))$ is Lipschitz continuous in $\bar{D}$. Extend $\sigma(x)$ and $b(x) = (b_1(x), \ldots, b_n(x))$ into the whole space R^n so that

$$|\sigma(x) - \sigma(y)| \leqslant C|x - y|, \qquad |b(x) - b(y)| \leqslant C|x - y|. \tag{5.6}$$

Consider the system of stochastic differential equations

$$d\xi(t) = \sigma(\xi(t))\, dw(t) + b(\xi(t))\, dt. \tag{5.7}$$

Denote by V_ϵ the closed ϵ-neighborhood of ∂D and let $D_\epsilon = D \setminus V_\epsilon$. Let v be a function in $C^2(R^n)$ that coincides with the solution u of (2.2), (2.3) in $D_{\epsilon/2}$, and let τ be any Markov time with respect to the time-homogeneous Markov process solution of (5.7). By Itô's formula,

$$E_x v(\xi(\tau)) \exp\left[\int_0^\tau c(\xi(s))\, ds\right] - v(x)$$
$$= E_x \int_0^\tau [Lv(\xi(t))] \exp\left[\int_0^t c(\xi(s))\, ds\right] dt. \tag{5.8}$$

Take $x \in D_\epsilon$ and $\tau = \tau_\epsilon \wedge T$ where τ_ϵ is the hitting time of V_ϵ. Then $v(\xi(t)) = u(\xi(t))$ for all $0 \leqslant t \leqslant \tau_\epsilon \wedge T$. Hence (5.8) holds for $v = u$. Taking $\epsilon \to 0$ and using the Lebesgue bounded convergence theorem, we get

$$u(x) = E_x u(\xi(\tau \wedge T)) \exp\left[\int_0^{\tau\wedge T} c(\xi(s))\, ds\right]$$
$$- E_x \int_0^{\tau\wedge T} f(\xi(t)) \exp\left[\int_0^t c(\xi(s))\, ds\right] dt \tag{5.9}$$

where τ is the exit time from D.

Theorem 5.1. *Let* (5.1)–(5.5) *hold and let* ∂D *belong to* C^2. *Then the unique solution* u *of the Dirichlet problem* (2.2), (2.3) *is given by*

$$u(x) = E_x \phi(\xi(\tau)) \exp\left[\int_0^\tau c(\xi(s))\, ds\right]$$
$$- E_x \int_0^\tau f(\xi(t)) \exp\left[\int_0^t c(\xi(s))\, ds\right] dt \tag{5.10}$$

where τ *is the exit time from* D.

Proof. If we prove that

$$E_x \tau < \infty \tag{5.11}$$

then, by taking $T \uparrow \infty$ in (5.9) and using the Lebesgue bounded convergence theorem, we get the assertion (5.9).

To prove (5.11) consider the function

$$h(x) = -A e^{\lambda x_1}$$

for all $x = (x_1, \ldots, x_n)$ in D. If A, λ are sufficiently large (A depending on λ), then

$$\sum a_{ij} h_{x_i x_j} + \sum b_i h_{x_i} \leqslant -1 \qquad \text{in } D.$$

By Itô's formula,

$$E_x h(\xi(\tau \wedge T)) - h(x) \leqslant -E_x(\tau \wedge T).$$

Since $|h(x)| \leqslant K$ in D (K constant), we then have $E_x(\tau \wedge T) \leqslant 2K$. Taking $T \uparrow \infty$ and using the monotone convergence theorem we get $E_x \tau \leqslant 2K$; this proves (5.11).

Remark. The proof of (5.11) requires only that $a_{11}(x) > 0$ in $\bar{D}$. Thus the right-hand side of (5.10) makes sense even if ∂D is not in C^2, L is not elliptic but has bounded coefficients with $c(x) \leqslant 0$, $a_{11}(x) > 0$ in $\bar{D}$, $a = \sigma\sigma^*$, and σ, b are Lipschitz continuous in $\bar{D}$. Recall that, by Theorem 1.2 and Remark 1 following it, if (a_{ij}) is nonnegative definite and in C^2 in some neighborhood of $\bar{D}$ then σ is Lipschitz continuous in $\bar{D}$.

Consider next the initial-boundary value problem (written for t replaced by $T - t$)

$$\begin{aligned} Lu + \partial u/\partial t &= f(x, t) \qquad \text{in} \quad Q = B \times [0, T), \\ u(x, T) &= \phi(x) \qquad \text{on } B, \\ u(x, t) &= g(x, t) \qquad \text{on } S, \end{aligned} \tag{5.12}$$

where B is a bounded domain with C^2 boundary ∂B, $S = \partial B \times [0, T)$, and L is defined by (4.1).

Consider also the system of stochastic differential equations

$$d\xi(t) = \sigma(\xi(t), t)\, dw(t) + b(\xi(t), t)\, dt \tag{5.13}$$

where $(\sigma(x, t))^2 = a(x, t)$ in $\bar{Q}$. The coefficients σ, b here are extensions of $\sigma(x, t)$, $b(x, t)$, originally defined in $\bar{Q}$, such that

$$\begin{aligned} |\sigma(x, t) - \sigma(y, s)| &\leqslant C(|x - y| + |t - s|), \\ |b(x, t) - b(y, s)| &\leqslant C(|x - y| + |t - s|). \end{aligned}$$

We shall assume:

$$\sum a_{ij}(x, t)\xi_i\xi_j \geqslant \mu|\xi|^2 \qquad \text{if} \quad (x, t) \in Q, \quad \xi \in R^n \qquad (\mu > 0),$$

$$\begin{aligned} &a_{ij}, b_i \quad \text{are uniformly Lipschitz continuous in } (x, t) \in \bar{Q}, \\ &c \quad \text{is uniformly Hölder continuous in } (x, t) \in \bar{Q}, \\ &f \quad \text{is uniformly Hölder continuous in } (x, t) \in \bar{Q}, \\ &g \quad \text{is continuous on } \bar{S}, \ \phi \text{ is continuous on } \bar{B} \text{ and} \\ &\qquad g(x, T) = \phi(x) \qquad \text{if} \quad x \in \partial B. \end{aligned} \tag{5.14}$$

By Theorem 3.6 there exists a unique solution u of (5.12).

Theorem 5.2. *Let ∂B belong to C^2 and let* (5.14) *hold. Then the unique solution u of the initial-boundary value problem* (5.12) *is given by*

$$\begin{aligned} u(x,t) &= E_{x,t} g(\xi(\tau),\tau) \exp\left[\int_t^\tau c(\xi(s),s)\,ds\right]\chi_{\tau<T} \\ &\quad + E_{x,t}\phi(\xi(T)) \exp\left[\int_t^T c(\xi(s),s)\,ds\right]\chi_{\tau=T} \\ &\quad - E_{x,t}\int_t^\tau f(\xi(s),s)\exp\left[\int_t^s c(\xi(\lambda),\lambda)\,d\lambda\right]ds \end{aligned} \tag{5.15}$$

where τ is the first time $\lambda \in [t, T)$ that $\xi(\lambda)$ leaves B if such a time exists and $\tau = T$ otherwise.

The proof of Theorem 5.2 is similar to the proof of Theorem 5.1. Here one applies Itô's formula to

$$u(\xi(\lambda),\lambda)\exp\left[\int_t^\lambda c(\xi(s),s)\,ds\right]. \tag{5.16}$$

Consider next the Cauchy problem

$$\begin{aligned} Lu + \partial u/\partial t &= f(x,t) \quad &&\text{in} \quad R^n \times [0,T), \\ u(x,T) &= \phi(x) &&\text{in } R^n \end{aligned} \tag{5.17}$$

where L is given by (4.1).

We shall assume:

(B_1) (i) The functions a_{ij}, b_i are bounded in $R^n \times [0, T]$ and uniformly Lipschitz continuous in (x, t) in compact subsets of $R^n \times [0, T]$.

(ii) The functions a_{ij} are Hölder continuous in x, uniformly with respect to (x, t) in $R^n \times [0, T]$.

(B_2) The function c is bounded in $R^n \times [0, T]$ and uniformly Hölder continuous in (x, t) in compact subsets of $R^n \times [0, T]$.

We shall also assume:

$f(x, t)$ is continuous in $R^n \times [0, T]$, Hölder continuous in x uniformly with respect to $(x, t) \in R^n \times [0, T]$, and

$$|f(x,t)| \le A(1+|x|^a) \quad \text{in} \quad R^n \times [0,T], \tag{5.18}$$

$$\phi(x) \quad \text{is continuous in } R^n \quad \text{and} \quad |\phi(x)| \le A(1+|x|^a) \tag{5.19}$$

where A, a are positive constants.

Under the conditions (A_1), (B_1), (B_2), (5.18), and (5.19), there exists a unique solution u of the Cauchy problem (5.17) satisfying

$$|u(x,t)| \le \text{const}(1+|x|^a). \tag{5.20}$$

Indeed, uniqueness follows from Corollary 4.2. The existence of u follows from Theorem 4.6; the estimate (5.20) on $u(x, t)$ is an easy consequence of the estimate (4.12) (with $m = 0$) on Γ and the assumptions (5.18), (5.19). Using (4.12) with $|m| = 1$ we also get

$$|u_x(x, t)| \leqslant \text{const}(1 + |x|^a). \tag{5.21}$$

By Lemma 1.1 and (1.1), the nonnegative definite square root $\sigma(x, t)$ of $a(x, t)$ is Lipschitz continuous in (x, t), uniformly in compact subsets of $R^n \times [0, T]$. It is also clear that $\sigma(x, t)$ is uniformly bounded in $R^n \times [0, T]$. Thus (by results in Chapter 5) the stochastic system (5.13) has a unique solution for any initial value $\xi(0) = x$.

Theorem 5.3. *If* (A_1), (B_1), (B_2) *and* (5.18), (5.19) *hold, then the solution of the Cauchy problem* (5.17), (5.20) *is given by*

$$u(x, t) = E_{x,t}\phi(\xi(T)) \exp\left[\int_t^T c(\xi(s), s)\, ds\right] - E_{x,t}\int_t^T f(\xi(s), s) \exp\left[\int_t^s c(\xi(\lambda), \lambda)\, d\lambda\right] ds. \tag{5.22}$$

The proof follows by applying Itô's formula to the function in (5.16) and the process (5.13), and then taking the expectation. Since, by (5.21),

$$|u_x(x, t)\sigma(x, t)| \leqslant \text{const}(1 + |x|^a)$$

and since

$$\sup_{t \leqslant s \leqslant T} E_{x,t}|\xi(s)|^m < \infty \qquad \text{for any} \quad m > 0,$$

the stochastic integral occurring in Itô's formula has zero expectation. The assumptions (5.18), (5.19) ensure that the expectations on the right-hand side of (5.22) exist.

Let

$$L_0 = \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x, t)\frac{\partial^2}{\partial x_i\, \partial x_j} + \sum_{i=1}^n b_i(x, t)\frac{\partial}{\partial x_i} \tag{5.23}$$

and denote by $\Gamma_0^*(x, s; y, t)$ the fundamental solution of $L_0 + \partial/\partial s$ $(s < t)$. Taking $f \equiv 0$, $c \equiv 0$ in (5.22), we get

$$u(x, t) = E_{x,t}\phi(\xi(T)). \tag{5.24}$$

By Theorem 4.6 (with t replaced by $T - t$) we can also represent u in the form

$$u(x, t) = \int_{R^n} \Gamma_0^*(x, t; y, T)\phi(y)\, dy.$$

Comparing this with (5.24) we conclude that

$$\int_{R^n} \Gamma_0^*(x, t; y, T)\phi(y)\, dy = E_{x,t}\phi(\xi(T)).$$

Notice that this relation holds for any function ϕ satisfying (5.19).

Similarly, for any $0 \leqslant s < t \leqslant T$,

$$\int_{R^n} \Gamma_0^*(x, s; y, t)\phi(y)\, dy = E_{x,s}\phi(\xi(t)) = \int_{R^n} \phi(y) P_{x,s}(\xi(t) \in dy). \quad (5.25)$$

Since ϕ is arbitrary function satisfying (5.19), we conclude that the transition probability function

$$p(s, x, t, A) = P_{x,s}(\xi(t) \in A) = P(\xi_{x,s}(t) \in A)$$

of the Markov process solution of (5.13) has density, and that the density function is $\Gamma_0^*(s, x, t, y)$. We sum up:

Theorem 5.4. *If* (A_1), (B_1) *hold, then the transition probability function of the solution of the stochastic differential system* (5.13) *has density, i.e.,*

$$P_{x,s}(\xi(t) \in A) = \int_A \Gamma_0^*(x, s; y, t)\, dy \qquad (s < t) \quad (5.26)$$

for any Borel set A, *and* $\Gamma_0^*(x, s; y, t)$ *is the fundamental solution of* $L_0 + \partial/\partial t$ *constructed in Section* 4.

The density function of the transition probability function is called the *transition density function.* From the results of Section 4 we conclude that the transition density function $\Gamma_0^*(x, s; y, t)$ of the solution $\xi(t)$ of (5.13) satisfies in (x, s) the *backward parabolic equation*

$$\frac{\partial}{\partial s} \Gamma_0^*(x, s; y, t) + \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x, s) \frac{\partial^2}{\partial x_i\, \partial x_j} \Gamma_0^*(x, s; y, t) + \sum_{i=1}^n b_i(x, s) \frac{\partial}{\partial x_i} \Gamma_0^*(x, s; y, t) = 0. \quad (5.27)$$

Under the assumptions of Theorem 4.7, $\Gamma_0^*(x, s; y, t)$ also satisfies in (y, t) the *forward parabolic equation*

$$-\frac{\partial}{\partial t} \Gamma_0^*(x, s; y, t) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial y_i\, \partial y_j} [a_{ij}(y, t)\Gamma_0^*(x, s; y, t)] - \sum_{i=1}^n \frac{\partial}{\partial y_i} [b_i(y, t)\Gamma_0^*(x, s; y, t)] = 0. \quad (5.28)$$

Notice that under the conditions (A_1), (B_1) the backward equation (5.27) is equivalent to the Kolmogorov equation (5.6.1).

If the coefficients a_{ij}, b_i are independent of t, then $\Gamma_0^*(x, s; y, t) = \Gamma_0^*(x, 0; y, t - s) \equiv \Gamma(x, t - s, y)$; $\Gamma(x, t, y)$ satisfies in (x, t) the parabolic equation

$$-\frac{\partial \Gamma(x, t, y)}{\partial t} + \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 \Gamma(x, t, y)}{\partial x_i\, \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial \Gamma(x, t, y)}{\partial x_i} = 0. \quad (5.29)$$

Under the conditions of Theorem 4.7, it also satisfies in (y, t) the parabolic equation

$$-\frac{\partial \Gamma(x, t, y)}{\partial t} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial y_i \partial y_j} [a_{ij}(y)\Gamma(x, t, y)]$$
$$- \sum_{i=1}^{n} \left[\frac{\partial}{\partial y_i} b_i(y)\Gamma(x, t, y) \right]. \tag{5.30}$$

PROBLEMS

1. Verify that the function in (2.4) is a barrier for L.

2. If D is a bounded domain with C^2 boundary ∂D, then for any $y \in \partial D$, there exists a closed ball K such that $K \cap D = \emptyset$ and $K \cap \bar{D} = \{y\}$; thus, a barrier exists.

3. Let L be as in Theorem 2.2 and let u be a solution of (2.2), (2.3) where D is a bounded domain. Prove that

$$\max_{\bar{D}} |u| \leqslant \max_{D} |\phi| + A \max_{\bar{D}} |f|$$

where A is a constant independent of f, ϕ.

4. Let $Lu = y^2 u_{xx} - 2xyu_{xy} + x^2 u_{yy} - xu_x - yu_y$. Verify that any C^2 function $u = f(r)$ (where $r = \sqrt{x^2 + y^2}$) satisfies $Lu = 0$. Hence the weak maximum principle does not hold in any domain $\epsilon^2 < x^2 + y^2 < R^2$. Notice that (a_{ij}) is nonnegative definite, $a_{11} + a_{22} \geqslant \epsilon^2$, but neither a_{11} nor a_{22} are positive throughout the domain.

5. Let $(a_{ij}(x, t))$ be nonnegative definite, $c(x, t) \leqslant \alpha$ and $a_{11}(x, t)\lambda^2 + b_1(x)\lambda \geqslant 1$ in a cylinder $Q = D \times [0, T]$. Denote the diameter of D by d. Show that if u is continuous in $\bar{Q}$ and Mu is bounded in Q (M given by (3.1)), then

$$\max_{\bar{Q}} |u| \leqslant e^{\alpha T}\left\{ \max_{\partial_0 Q} |u| + (e^{\lambda d} - 1) \max_{\bar{Q}} |Mu| \right\}$$

where $\partial_0 Q$ is the normal boundary of Q.

6. Let the coefficients of M satisfy the conditions of either Theorem 4.1 or 4.3. Prove that a fundamental solution Γ is uniquely determined by the requirements:

(i) $\Gamma(x, t; \xi, \tau)$ is continuous in ξ, for fixed x, t, τ;

(ii) for any continuous function f with compact support, the function $u(x, t)$ given by (4.9) satisfies: $|u(x, t)| \leqslant C(1 + |x|^q)$ in $R^n \times [0, T]$, where C, q are some positive constants (depending on f).

7. Let the coefficients of M satisfy the conditions of either Theorem 4.1 or 4.3 and let Γ be a fundamental solution satisfying (i), (ii) of the preceding problem. Prove that $\Gamma(x, t; \xi, \tau) > 0$.

8. The operator $\Delta - \partial/\partial t$ (where Δ is the Laplacian operator $\sum_{i=1}^n \partial^2/\partial x_i^2$) is called the *heat operator*. Prove that

$$\Gamma(x, t; \xi, \tau) = \frac{1}{(2\sqrt{\pi})^n (t-\tau)^{n/2}} \exp\left\{ \frac{|x-\xi|^2}{4(t-\tau)} \right\}$$

is a fundamental solution for the heat operator.

9. Let Γ be the fundamental solution constructed in Theorem 4.5. Prove that, for any $\epsilon > 0$,

$$\int_{|x-\xi|>\epsilon} \Gamma(x, t; \xi, \tau)\, d\xi \to 0 \qquad \text{if} \quad t \downarrow \tau,$$

$$\int_{|x-\xi|>\epsilon} \Gamma(x, t; \xi, \tau)\, dx \to 0 \qquad \text{if} \quad t \downarrow \tau.$$

10. Let $f(x, t)$ be a continuous bounded function in $R^n \times [0, T]$, and let Γ be the fundamental solution constructed in Theorem 4.5. Prove that

$$\int_{R^n} \Gamma(x, t; \xi, \tau) f(\xi, \tau)\, d\xi \to f(x, \tau) \qquad \text{if} \quad t \downarrow \tau,$$

$$\int_{R^n} \Gamma(x, t; \xi, \tau) f(x, \tau)\, dx \to f(\xi, \tau) \qquad \text{if} \quad t \downarrow \tau.$$

[*Hint*: Use the preceding problem.]

11. Let $\Gamma(x, t; \xi, \tau)$ be the fundamental solution constructed in Theorem 4.5. Prove that

$$\Gamma(x, t; \xi, \tau) = \int_{R^n} \Gamma(x, t; y, \sigma) \Gamma(y, \sigma; \xi, \tau)\, dy \qquad (\tau < \sigma < t).$$

12. Give another proof of Theorem 4.1, by applying the maximum principle to $u + \epsilon v$ in $0 \leqslant t \leqslant 1/\alpha$, where

$$v(x, t) = \exp\{ \beta(|x|^2 + 1) e^{\alpha t} \}$$

with suitable constants α, β.

13. Let M^* be an operator of the form $M^* v = \sum \alpha_{ij} v_{x_i x_j} + \sum \beta_i v_{x_i} + \gamma v + \delta v_t$ where α_{ij}, β_i, γ, δ are continuous functions. Prove that if (4.22) holds for any u, v in C^∞ with compact support in G, then M^* is given by (4.20), i.e., M^* is the formal adjoint of M.

14. Prove Theorem 5.2.

15. Consider two stochastic differential systems of n equations

$$d\xi = \sigma(\xi, t)\, dw + b(\xi, t)\, dt, \qquad d\xi' = \sigma'(\xi', t)\, dw' + b'(\xi', t)\, dt$$

where w and w' are n-dimensional Brownian motions. Suppose that the coefficients are uniformly Lipschitz continuous and that $\sigma\sigma^*$ is uniformly positive definite. Prove that if $\sigma\sigma^* \equiv \sigma'(\sigma')^*$, $b \equiv b'$, then the two processes have the same joint distribution functions, given that $\xi(0)$, $\xi'(0)$ have the same distribution functions.

7

The Cameron–Martin–Girsanov Theorem

1. A class of absolutely continuous probabilities

Any nonnegative random variable g defines an absolutely continuous measure P_g whose Radon–Nikodym derivative is g, i.e., $P_g(A) = \int_A g\, dP$ for any measurable set A. In this section we show that the random variable

$$g = \exp\left\{ \int_0^T f(s)\, dw(s) - \tfrac{1}{2} \int_0^T |f(s)|^2\, ds \right\}$$

defines a *probability* P_g, i.e., $Eg = 1$, provided f belongs to $L^2_w[0, T]$ and satisfies a certain growth condition. More precisely:

Theorem 1.1. *Let $f = (f_1, \ldots, f_n)$ belong to $L^2_w[0, T]$ and assume that there exist positive numbers μ, C such that*

$$E \exp\left[\mu |f(t)|^2 \right] \leqslant C \qquad \text{for} \quad 0 \leqslant t \leqslant T. \tag{1.1}$$

Then

$$E \exp\left\{ \int_{t_1}^{t_2} f(s)\, dw(s) - \tfrac{1}{2} \int_{t_1}^{t_2} |f(s)|^2\, ds \right\} = 1 \qquad \text{if} \quad 0 \leqslant t_1 < t_2 \leqslant T. \tag{1.2}$$

We first prove two lemmas.

Lemma 1.2. *If in Theorem* 1.1 *the condition* (1.1) *is replaced by the condition*

$$E \exp\left\{ \lambda \int_0^T |f(s)|^2\, ds \right\} < \infty \qquad \text{for some} \quad \lambda > 1, \tag{1.3}$$

then the assertion (1.2) *is valid.*

Proof. By Lemma 4.1.1 (see (4.1.5)) there exists a sequence of continuous functions $f_m(t) = (f_{m1}(t), \ldots, f_{mn}(t))$ in $L_w^2[0, T]$ such that

$$\int_0^T |f_m(t) - f(t)|^2 \, dt \xrightarrow{P} 0 \quad \text{as} \quad m \to \infty, \tag{1.4}$$

$$\int_0^T |f_m(t)|^2 \, dt \leqslant \int_0^T |f(t)|^2 \, dt. \tag{1.5}$$

We may assume that each f_m is bounded, i.e., $|f_m(t, \omega)| \leqslant C_m$ a.s. in (t, ω), for otherwise we replace f_m by f_{m, N_m} with components $\phi_{N_m}(f_{mi})$, where $\phi_N(r) = r$ if $|r| \leqslant N$ and $\phi_N(r) = Nr/|r|$ if $|r| > N$, and $N_m \uparrow \infty$ if $m \uparrow \infty$. We claim that

$$E \exp\left\{2\int_{t_1}^{t_2} f_m(s) \, dw(s)\right\} \leqslant C_m' \qquad \text{for all} \quad 0 \leqslant t_1 < t_2 \leqslant T, \tag{1.6}$$

where C_m' is a constant depending on m.

To prove it notice that

$$Ee^{\gamma|w(t)|} = \int_{R^n} e^{\gamma|x|} \frac{e^{-|x|^2/2t}}{\sqrt{2\pi t}} \, dx = e^{\gamma^2 t/2}. \tag{1.7}$$

Let $\Pi_l = \{s_1^l, \ldots, s_{k_l}^l\}$ be a sequence of partitions of $[t_1, t_2]$ with mesh $|\Pi_l| \to 0$. Then

$$\left|\int_{t_1}^{t_2} f_m(s) \, dw(s)\right| = \left|\lim_{l\to\infty} \sum f_m(s_i^l)\left[w(s_{i+1}^l) - w(s_i^l)\right]\right|$$
$$\leqslant C \varliminf_{l\to\infty} \sum |w(s_{i+1}^l) - w(s_i^l)|$$

where C is a constant (depending on m). By Fatou's lemma,

$$E \exp\left|2\int_{t_1}^{t_2} f_m(s) \, dw(s)\right| \leqslant \varliminf_{l\to\infty} E \exp\left[2C\sum_i |w(s_{i+1}^l) - w(s_i^l)|\right]$$
$$= \varliminf_{l\to\infty} \prod_i E \exp\left[2C|w(s_{i+1}^l) - w(s_i^l)|\right].$$

By (1.7), the right-hand side is equal to

$$= \varliminf_{l\to\infty} \prod_i \exp\left[2C^2(s_{i+1}^l - s_i^l)\right] = \exp\left[2C^2(t_2 - t_1)\right].$$

Thus (1.6) follows.

Applying Itô's formula to $u = e^x$ and the process

$$\int_{t_1}^t f_m(s) \, dw(s) - \tfrac{1}{2}\int_{t_1}^t |f_m(s)|^2 \, ds,$$

we obtain

$$\exp\left\{\int_{t_1}^{t_2} f_m(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t_2} |f_m(s)|^2\,ds\right\} - 1$$
$$= \int_{t_1}^{t} \exp\left\{\int_{t_1}^{t_2} f_m(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t} |f_m(s)|^2\,ds\right\} f_m(t)\,dw(t). \qquad (1.8)$$

Denote the last integral by $\int_{t_1}^{t_2} h(s)\,dw(s)$. By (1.6) and the fact that $|f_m(t)| \leqslant C_m$ it follows that

$$E\int_{t_1}^{t_2} |h(t)|^2\,dt < \infty.$$

Hence $E\int_{t_1}^{t_2} h(t)\,dw(t) = 0$. Taking the expectation in (1.8) we then get

$$E\exp\left\{\int_{t_1}^{t_2} f_m(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t_2} |f_m(s)|^2\,ds\right\} = 1. \qquad (1.9)$$

We shall prove that

$$I \equiv E\left[\exp\left\{\int_{t_1}^{t_2} f_m(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t_2} |f_m(s)|^2\,ds\right\}\right]^{1+\epsilon} \leqslant K \qquad (1.10)$$

for some $\epsilon > 0$, where K is a constant independent of m. This would imply that the sequence of random variables

$$X_m \equiv \exp\left\{\int_{t_1}^{t_2} f_m(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t_2} |f_m(s)|^2\,ds\right\}$$

is uniformly integrable. Since, by (1.4),

$$\int_{t_1}^{t_2} f_m(s)\,dw(s) \xrightarrow{P} \int_{t_1}^{t_2} f(s)\,dw(s) \qquad \text{as} \quad m\to\infty,$$

$$X_m \xrightarrow{P} \exp\left\{\int_{t_1}^{t_2} f(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t_2} |f(s)|^2\,ds\right\}.$$

Hence, taking $m\to\infty$ in (1.19), the assertion (1.2) follows.

It remains to prove (1.10). By Hölder's inequality,

$$I = E\left\{\exp\left[(1+\epsilon)\int_{t_1}^{t_2} f_m(s)\,dw(s) - \frac{(1+\epsilon)^3}{2}\int_{t_1}^{t_2} |f_m(s)|^2\,ds\right]\right.$$
$$\left.\cdot\exp\left[\frac{(1+\epsilon)(2\epsilon+\epsilon^2)}{2}\int_{t_1}^{t_2} |f_m(s)|^2\,ds\right]\right\}$$
$$\leqslant \left\{E\exp\left[(1+\epsilon)^2\int_{t_1}^{t_2} f_m(s)\,dw(s) - \frac{(1+\epsilon)^4}{2}\int_{t_1}^{t_2} |f_m(s)|^2\,ds\right]\right\}^{1/(1+\epsilon)}$$
$$\cdot\left\{E\exp\left[\frac{(1+\epsilon)^2(2+\epsilon)}{2}\int_{t_1}^{t_2} |f_m(s)|^2\,ds\right]\right\}^{\epsilon/(1+\epsilon)}.$$

By (1.9) the first factor on the right is equal to 1. Using (1.5) we get

$$I \leqslant \left\{ E \exp\left[\frac{(1+\epsilon)^2(2+\epsilon)}{2} \int_{t_1}^{t_2} |f(s)|^2 \, ds \right] \right\}^{\epsilon/(1+\epsilon)} \equiv K.$$

If ϵ is sufficiently small so that $(1 + \epsilon)^2(2 + \epsilon) < 2\lambda$, then, by (1.3), K is a finite number. This completes the proof of (1.10).

Lemma 1.3. *Under the assumptions of Lemma* 1.2,

$$E\left\{ \exp\left[\int_{t_1}^{t_2} f(s)\, dw(s) - \tfrac{1}{2} \int_{t_1}^{t_2} |f(s)|^2 \, ds \right] \Big| \mathcal{F}_{t_1} \right\} = 1 \quad a.s. \tag{1.11}$$

for any $0 \leqslant t_1 < t_2 \leqslant T$.

Proof. Taking the conditional expectation of both sides of (1.8) with respect to $\mathcal{F}_{t_1}$, we obtain the assertion (1.11) for $f = f_m$. Using (1.10) one can easily justify passage to the limit $m \to \infty$ in the formula (1.11) for $f = f_m$.

Proof of Theorem 1.1. Let $\lambda > 1$. Since $e^{\lambda x}$ is a convex function, Jensen's inequality (see Problem 7, Chapter 1) gives

$$\exp\left[\lambda \int_{t'}^{t''} |f(s)|^2 \, ds \right] = \exp\left[\frac{1}{t'' - t'} \int_{t'}^{t''} \lambda(t'' - t')|f(s)|^2 \, ds \right]$$
$$\leqslant \frac{1}{t'' - t'} \int_{t'}^{t''} \exp\left[\lambda(t'' - t')|f(s)|^2 \right] ds \qquad (t' < t'').$$

Hence, if $\lambda(t'' - t') < \mu$, then (1.1) implies that

$$E \exp\left[\lambda \int_{t'}^{t''} |f(s)|^2 \, ds \right] < \infty.$$

Consequently, by Lemma 1.3,

$$E\left[\exp\left\{ \int_{t'}^{t''} f(s)\, dw(s) - \tfrac{1}{2} \int_{t'}^{t''} |f(s)|^2 \, ds \right\} \Big| \mathcal{F}_{t'} \right] = 1 \quad \text{a.s.} \tag{1.12}$$

Now let $t_1 = s_1 < s_2 < \cdots < s_m = t_2$ where $\lambda(s_{i+1} - s_i) < \mu$. Then

$$E\left\{ \exp\left[\int_{t_1}^{t_2} f \, dw - \tfrac{1}{2} \int_{t_1}^{t_2} |f|^2 \, ds| \right] \mathcal{F}_{t_1} \right\}$$
$$= E\left\{ \exp\left[\int_{t_1}^{s_{m-1}} f \, dw - \tfrac{1}{2} \int_{t_1}^{s_{m-1}} |f|^2 \, ds \right] \right.$$
$$\left. \cdot\, E \exp\left[\int_{s_{m-1}}^{s_m} f \, dw - \tfrac{1}{2} \int_{s_{m-1}}^{s_m} |f|^2 \, ds \right] \Big| \mathcal{F}_{s_{m-1}} \right\} \Big| \mathcal{F}_{t_1}$$
$$= E \exp\left[\int_{t_1}^{s_{m-1}} f \, dw - \tfrac{1}{2} \int_{t_1}^{s_{m-1}} |f|^2 \, ds \right] \Big| \mathcal{F}_{t_1}$$

by (1.12). Proceeding similarly step-by-step, we arrive at the formula

$$E\left\{\exp\left[\int_{t_1}^{t_2} f(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t_2} |f(s)|^2\,ds\right] \mid \mathcal{F}_{t_1}\right\} = 1. \quad \text{a.s.} \tag{1.13}$$

Taking the expectation, (1.2) follows.

Corollary 1.4. *Under the assumptions of Theorem* 1.1, *the relation* (1.13) *holds for any* $0 \leqslant t_1 < t_2 \leqslant T$.

Example. If $f(t) = w(t)$, then condition (1.1) is satisfied with $\mu < 1/2T$.

Corollary 1.5. *For any* $f = (f_1, \ldots, f_m)$ *in* $L_w^2[0, T]$,

$$E\exp\left[\int_{t_1}^{t_2} f(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t_2} |f(s)|^2\,ds\right] \leqslant 1, \tag{1.14}$$

$$E\left\{\exp\left[\int_{t_1}^{t_2} f(s)\,dw(s) - \tfrac{1}{2}\int_{t_1}^{t_2} |f(s)|^2\,ds\right] \mid \mathcal{F}_{t_1}\right\} \leqslant 1 \quad a.s. \tag{1.15}$$

for any $0 \leqslant t_1 < t_2 \leqslant T$.

To prove (1.14) we take $m \to \infty$ in (1.9) and use Fatou's lemma. To prove (1.15) we first note (by taking the conditional expectation in (1.8)) that (1.15) holds (with "=") for $f = f_m$. Now take $m \to \infty$ and use Fatou's lemma.

2. Transformation of Brownian motion

A process $\{w(t),\, 0 \leqslant t \leqslant T\}$ that satisfies all the conditions imposed on an n-dimensional Brownian motion (including continuity) in the interval $0 \leqslant t \leqslant T$ is called *an n-dimensional Brownian motion in the interval* $[0, T]$.

Let $w(t)$ be an n-dimensional Brownian motion in an interval $[0, T]$. Let $\mathcal{F}_t$ be an increasing family of σ-fields such that $\mathcal{F}(w(\lambda), \lambda \leqslant t)$ is a subset of $\mathcal{F}_t$ and $\mathcal{F}\{w(\lambda + t) - w(t),\, 0 \leqslant \lambda \leqslant T - t\}$ is independent of $\mathcal{F}_t$, for all $t \in [0, T]$. As in Chapter 3 (where $w(t)$ and $\mathcal{F}_t$ were defined for all $t \geqslant 0$), we can define $L_w^2[0, T]$ and stochastic integrals $\int_0^t f\,dw$ $(0 < t \leqslant T)$ with respect to the present family $\mathcal{F}_t$.

If $\tilde{P}$ is a measure on $(\Omega, \mathcal{F})$ given by

$$\tilde{P}(A) = \int_A f\,dP \qquad (A \in \mathcal{F}),$$

then we write $d\tilde{P}(\omega) = f(\omega)\,dP(\omega)$.

Theorem 2.1. *Let* $w(t)$ *be an n-dimensional Brownian motion in* $[0, T]$ *and* *let $\phi = (\phi_1, \ldots, \phi_n)$ be a function in $L^2_w[0, T]$. Define*

$$\zeta_s^t(\phi) = \int_s^t \phi(u)\, dw(u) - \tfrac{1}{2}\int_s^t |\phi(u)|^2\, du, \tag{2.1}$$

$$\tilde{w}(t) = w(t) - \int_0^t \phi(s)\, ds, \tag{2.2}$$

$$d\tilde{P}(\omega) = \exp\left[\zeta_0^T(\phi)\right] dP(\omega). \tag{2.3}$$

If

$$\tilde{P}(\Omega) = 1, \tag{2.4}$$

then $\tilde{w}(t)$, $0 \leqslant t \leqslant T$ is an n-dimensional Brownian motion in the probability space $(\Omega, \mathcal{F}, \tilde{P})$.

Recall that in Section 1 we have established sufficient conditions for (2.4) to hold. We have also proved (Corollary 1.5) that $\tilde{P}(\Omega) \leqslant 1$ for any $\phi \in L^2_w[0, T]$.

Theorem 2.1 is due to Girsanov [1]. Cameron and Martin [1] have previously established results of the same nature on nonlinear transformations of Brownian motion. We shall refer to Theorem 2.1 as the *Cameron–Martin–Girsanov theorem.*

We begin with some lemmas.

Lemma 2.2. *If $\phi \in L^2_w[0, T]$ and if $|\phi(t)| \leqslant c$ a.s., then, for any $\alpha > 1$,*

$$E \exp[\alpha\zeta_s^t(\phi)] \leqslant \exp\left[\frac{\alpha^2 - \alpha}{2}(t - s)c^2\right]. \tag{2.5}$$

Proof. By Corollary 1.5, $E \exp[\zeta_s^t(\alpha\phi)] \leqslant 1$. Hence

$$E \exp[\alpha\zeta_s^t(\phi)] = E \exp[\zeta_s^t(\alpha\phi)] \exp\left[\frac{\alpha^2 - \alpha}{2}\int_s^t |\phi(u)|^2\, du\right]$$

$$\leqslant E \exp[\zeta_s^t(\alpha\phi)] \exp\left[\frac{\alpha^2 - \alpha}{2}(t - s)c^2\right] \leqslant \exp\left[\frac{\alpha^2 - \alpha}{2}(t - s)c^2\right].$$

Denote by $\tilde{E}$ the expectation with respect to $\tilde{P}$.

Lemma 2.3. *Let $\phi \in L^2_w[0, T]$. Then for any nonnegative and $\mathcal{F}_t$ measurable random variable η,*

$$\tilde{E}\eta \leqslant E\{\eta \exp[\zeta_0^t(\phi)]\}, \qquad 0 \leqslant t \leqslant T. \tag{2.6}$$

If ϕ satisfies (2.4), then

$$E \exp[\zeta_s^t(\phi)] = 1, \tag{2.7}$$

$$E\{\exp[\zeta_s^t(\phi)] \mid \mathcal{F}_s\} = 1 \quad a.s. \tag{2.8}$$

for all $0 \leqslant s < t \leqslant T$, *and*

$$\tilde{E}\eta = E\{\eta \exp[\zeta_0^t(\phi)]\} \tag{2.9}$$

for any $\mathcal{F}_t$ *measurable random variable* η *for which* $\tilde{E}\eta$ *exists.*

Proof. The assertion (2.6) follows from

$$\begin{aligned}\tilde{E}\eta &= E\eta \exp[\zeta_0^T(\phi)] = EE\{\eta \exp[\zeta_0^T(\phi)] | \mathcal{F}_t\} \\ &= E\eta \exp[\zeta_0^t(\phi)] E\{\exp[\zeta_t^T(\phi)] | \mathcal{F}_t\} \\ &\leqslant E\eta \exp[\zeta_0^t(\phi)],\end{aligned}$$

where (1.15) has been used.

Assume now that (2.4) holds. Using (1.15) we have

$$\begin{aligned}1 = E\exp[\zeta_0^T(\phi)] &= E\exp[\zeta_0^s(\phi)] E\{\exp[\zeta_s^t(\phi)] E[\exp[\zeta_t^T(\phi)] | \mathcal{F}_t] | \mathcal{F}_s\} \\ &\leqslant E\exp[\zeta_0^s(\phi)] E\{\exp[\zeta_s^t(\phi)] | \mathcal{F}_s\}.\end{aligned} \tag{2.10}$$

Denote by B_m $(m > 0)$ the set on which

$$E\exp[\zeta_s^t(\phi)] | \mathcal{F}_s \leqslant 1 - 1/m.$$

Since, by (1.15), $E\{\exp[\zeta_s^t(\phi)] | \mathcal{F}_s\} \leqslant 1$ a.s., we conclude from (2.10) that

$$1 \leqslant \int_{\Omega \setminus B_m} \exp[\zeta_0^s(\phi)]\, dP + \left(1 - \frac{1}{m}\right) \int_{B_m} \exp[\zeta_0^s(\phi)]\, dP.$$

But since

$$\int_\Omega \exp[\zeta_0^s(\phi)]\, dP \leqslant 1 \qquad \text{and} \qquad \exp \zeta_0^s(\phi) > 0 \quad \text{a.s.},$$

it follows that $P(B_m) = 0$. Since m is arbitrary, the assertion (2.8) follows. The validity of (2.7) is now obvious, and (2.9) follows from

$$\tilde{E}\eta = E\eta \exp[\zeta_0^t(\phi)] E\{\exp[\zeta_t^T(\phi)] | \mathcal{F}_t\} = E\eta \exp[\zeta_0^t(\phi)].$$

Lemma 2.4. *Let* ϕ *be as in Theorem* 2.1 *and let* η *be any* $\mathcal{F}_t$ *measurable random variable for which* $\tilde{E}\eta$ *exists. Then*

$$\tilde{E}[\eta | \mathcal{F}_s] = E\{\eta \exp[\zeta_s^t(\phi)] | \mathcal{F}_s\}. \tag{2.11}$$

Proof. Set $\gamma_1 = \exp \zeta_0^s(\phi)$, $\gamma_2 = \exp \zeta_s^t(\phi)$. Let λ be any bounded and $\mathcal{F}_s$ measurable random variable. Then

$$\tilde{E}(\lambda\eta) = \tilde{E}\lambda\tilde{E}(\eta | \mathcal{F}_s). \tag{2.12}$$

We also have

$$\tilde{E}(\lambda\eta) = E(\lambda\eta\gamma_1\gamma_2) = E\lambda\gamma_1 E(\eta\gamma_2 | \mathcal{F}_s). \tag{2.13}$$

If γ is $\mathcal{F}_s$ measurable, then, by (2.8),

$$E\gamma = E\gamma E(\gamma_2 | \mathcal{F}_s) = EE[(\gamma\gamma_2) | \mathcal{F}_s] = E\gamma\gamma_2.$$

Applying this to $\gamma = \lambda\gamma_1 E[(\eta\gamma_2)|\mathcal{F}_s]$, we get from (2.13)

$$\tilde{E}(\lambda\eta) = E\lambda\gamma_1\gamma_2 E[(\eta\gamma_2)|\mathcal{F}_s] = \tilde{E}\lambda E(\eta\gamma_2)|\mathcal{F}_s.$$

Comparing this with (2.12), we get

$$\tilde{E}\lambda\tilde{E}(\eta|\mathcal{F}_s) = \tilde{E}\lambda E(\eta\gamma_2)|\mathcal{F}_s.$$

Since λ is arbitrary bounded random variable that is $\mathcal{F}_s$ measurable, (2.11) follows.

Lemma 2.5. *Let ϕ be as in Theorem 2.1 and let $\{\phi^N\}$ be a sequence of bounded functions in $L_w^2[0, T]$ such that $\zeta_s^t(\phi^N) \to \zeta_s^t(\phi)$ in probability as $N \to \infty$. Then*

$$\int_\Omega |\exp \zeta_s^t(\phi^N) - \exp \zeta_s^t(\phi)|\, dP \to 0 \qquad \textit{if} \quad N \to \infty. \tag{2.14}$$

Proof. By Theorem 1.1,

$$\int_\Omega \exp \zeta_s^t(\phi^N)\, dP = 1. \tag{2.15}$$

Denote by $\Omega_{N,\epsilon}$ the set where

$$|\exp \zeta_s^t(\phi) - \exp \zeta_s^t(\phi^N)| < \epsilon.$$

By assumption, $P(\Omega_{N,\epsilon}) > 1 - \delta_N(\epsilon)$, $\delta_N(\epsilon) \to 0$ if $N \to \infty$. Since $\exp \zeta_s^t(\phi)$ is integrable,

$$\int_{\Omega_{N,\epsilon}} \exp[\zeta_s^t(\phi)]\, dP = 1 - \int_{\Omega \setminus \Omega_{N,\epsilon}} \exp[\zeta_s^t(\phi)]\, dP > 1 - \epsilon \tag{2.16}$$

if N is sufficiently large. Hence

$$\int_{\Omega_{N,\epsilon}} \exp\big[\zeta_s^t(\phi^N)\big]\, dP > 1 - 2\epsilon.$$

From (2.15) it then follows that

$$\int_{\Omega \setminus \Omega_{N,\epsilon}} \exp\big[\zeta_s^t(\phi^N)\big]\, dP < 2\epsilon. \tag{2.17}$$

Since ϕ satisfies (2.7), (2.16) implies that

$$\int_{\Omega \setminus \Omega_{N,\epsilon}} \exp[\zeta_s^t(\phi)]\, dP < \epsilon. \tag{2.18}$$

Making use of (2.17), (2.18), we find that

$$\begin{aligned}
&\int_\Omega |\exp \zeta_s^t(\phi) - \exp \zeta_s^t(\phi^N)|\, dP \\
&\quad < \int_{\Omega_{N,\epsilon}} |\exp \zeta_s^t(\phi) - \exp \zeta_s^t(\phi^N)|\, dP \\
&\quad + \int_{\Omega \setminus \Omega_{N,\epsilon}} \big[\exp \zeta_s^t(\phi) + \exp \zeta_s^t(\phi^N)\big]\, dP > 4\epsilon
\end{aligned}$$

if N is sufficiently large, and the proof of (2.14) is complete.

In proving Theorem 2.1 we shall rely upon Theorem 3.6.2. Since $\tilde{w}(t)$ is a continuous process, it will suffice to prove that, for any $0 \leqslant s < t \leqslant T$,

$$\tilde{E}[\tilde{w}_i(t) - \tilde{w}_i(s)] | \mathcal{F}_s = 0 \quad \text{a.s.,} \tag{2.19}$$

$$\tilde{E}[\tilde{w}_i(t) - \tilde{w}_i(s)][\tilde{w}_j(t) - \tilde{w}_j(s)] | \mathcal{F}_s = \delta_{ij}(t - s) \quad \text{a.s.} \tag{2.20}$$

Lemma 2.6. *The assertion of Theorem* 2.1 *is true if* $\phi(t)$ *is bounded.*

Proof. We shall verify (2.19), (2.20).

We need the following special cases of the integrated form of Itô's formula $d(\zeta_1\zeta_2) = \zeta_1\, d\zeta_2 + \zeta_2\, d\zeta_1 + d\zeta_1\, d\zeta_2$:

$$\int_{t_0}^{t} f(s)\, dw(s) \int_{t_0}^{t} g(s)\, dw(s) = \int_{t_0}^{t} f(s) \cdot g(s)\, ds + \int_{t_0}^{t} \left(\int_{t_0}^{s} g(u)\, dw(u) \right) f(s)\, dw(s) + \int_{t_0}^{t} \left(\int_{t_0}^{s} f(u)\, dw(u) \right) g(s)\, dw(s), \tag{2.21}$$

$$\int_{t_0}^{t} f(s)\, dw(s) \int_{t_0}^{t} h(s)\, ds = \int_{t_0}^{t} \left(\int_{t_0}^{s} h(u)\, du \right) f(s)\, dw(s) + \int_{t_0}^{t} \left(\int_{t_0}^{s} f(u)\, dw(u) \right) h(s)\, ds. \tag{2.22}$$

Here f and g are n-dimensional functions in $L_w^2[t_0, t]$ and h is a scalar function in $L_w^1[t_0, t]$.

We shall also need (1.8) with $f_m = \phi$, i.e.,

$$\exp \zeta_{t_1}^{t_2}(\phi) = 1 + \int_{t_1}^{t_2} \exp[\zeta_{t_1}^{u}(\phi)]\phi(u)\, dw(u).$$

By Lemma 2.4,

$$\begin{aligned}
\tilde{E}[\tilde{w}_i(t) - \tilde{w}_i(s)] | \mathcal{F}_s &= E[w_i(t) - w_i(s)] \exp[\zeta_s^t(\phi)] | \mathcal{F}_s \\
&\quad - E\left(\int_s^t \phi_i(u)\, du \right) \exp[\zeta_s^t(\phi)] | \mathcal{F}_s \\
&= E[w_i(t) - w_i(s)]\left[1 + \int_s^t \exp[\zeta_s^u(\phi)]\phi(u)\, dw(u) \right] | \mathcal{F}_s \\
&\quad - E\left(\int_s^t \phi_i(u)\, du \right) \exp[\zeta_s^t(\phi)] | \mathcal{F}_s \\
&= E\left\{ \int_s^t \phi_i(u) \exp[\zeta_s^u(\phi)]\, du | \mathcal{F}_s \right\} - E\left(\int_s^t \phi_i(u)\, du \right) \exp[\zeta_s^t(\phi)] | \mathcal{F}_s.
\end{aligned} \tag{2.23}$$

In the last equation we have used (2.21) and also the fact that

$$E\left[\int_s^t h(u)\, dw(u) \right] | \mathcal{F}_s = 0$$

if

$$E\int_s^t |h(u)|^2\, du < \infty \qquad (h = (h_1, \ldots, h_n)) \tag{2.24}$$

where, for one stochastic integral,

$$h_i(u) = \int_s^u \exp[\zeta_s^v(\phi)]\phi(v)\, dw(v), \qquad h_j = 0 \quad \text{if} \quad j \neq i$$

and, for another stochastic integral,

$$h(u) = (w_i(u) - w_i(s)) \exp[\zeta_s^u(\phi)]\phi(u).$$

Using the fact that $|\phi| \leqslant c$ (c constant) and the estimate (2.5), one can easily see that these two functions $h(u)$ indeed satisfy (2.24).

In view of (2.8) the right-hand side of (2.23) is equal to zero. Thus (2.19) holds.

In order to prove (2.20), set

$$\xi(t) = w(t) - w(s), \qquad \eta(t) = \exp \zeta_0^t(\phi).$$

Then

$$\begin{aligned}
&[\tilde{w}_i(t) - \tilde{w}_i(s)][\tilde{w}_j(t) - \tilde{w}_j(s)]\, \eta(t) \\
&= \left[\xi_i(t) - \int_s^t \phi_i(u)\, du\right]\left[\xi_j(t) - \int_s^t \phi_j(u)\, du\right]\eta(t) \\
&= \xi_i(t)\xi_j(t)\left[1 + \int_s^t \eta(u)\phi(u)\, dw(u)\right] - \left[\int_s^t \phi_i(u)\, du\right]\xi_j(t)\eta(t) \\
&\quad - \left[\int_s^t \phi_j(u)\, du\right]\xi_i(t)\eta(t) + \left[\int_s^t \phi_i(u)\, du\right]\left[\int_s^t \phi_j(u)\, du\right]\eta(t) \\
&\equiv \xi_i(t)\xi_j(t) + J_1 + J_2 + J_3 + J_4,
\end{aligned}$$

where J_i ($i = 2, 3, 4$) denotes the $(i + 1)$th term on the right.

Denote by $\cong$ equality moduli stochastic integrals, i.e., $A \cong B$ if $A = B + \Sigma\int_s^t h_j(u)\, dw_j(u)$. Then, using (2.21), (2.22), we find that

$$\begin{aligned}
J_1 &= \xi_i(t)\int_s^t \left[\int_s^u \eta(v)\phi(v)\, dw(v)\right] dw_j(u) + \xi_i(t)\int_s^t \xi_j(u)\eta(u)\phi(u)\, dw(u) \\
&\quad + \xi_i(t)\int_s^t \eta(u)\phi_j(u)\, du \\
&\cong \delta_{ij}\int_s^t \left[\int_s^u \eta(v)\phi(v)\, dw(v)\right] du \\
&\quad + \int_s^t \xi_j(u)\eta(u)\phi_i(u)\, du + \int_s^t \xi_i(u)\eta(u)\phi_j(u)\, du.
\end{aligned}$$

Next,

$$J_2 = -\left[\int_s^t \phi_i(u)\, du\right]\xi_j(t)\eta(t) \cong -\left[\int_s^t \xi_j(u)\phi_i(u)\, du\right]\eta(t),$$

$$J_3 = -\left[\int_s^t \phi_j(u)\, du\right]\xi_i(t)\eta(t) \cong -\left[\int_s^t \xi_i(u)\phi_j(u)\, du\right]\eta(t).$$

Notice that due to the boundedness of ϕ and Lemma 2.2, all the stochastic integrals $\int_s^t h(u)\,dw(u)$ which were dropped in the final expressions for J_1, J_2, J_3 are such that (2.24) holds. Consequently,

$$\begin{aligned}
&\tilde{E}[\tilde{w}_i(t) - \tilde{w}_i(s)][\tilde{w}_j(t) - \tilde{w}_j(s)]|\mathcal{F}_s - \delta_{ij}(t-s)\\
&\quad = -E\left[\int_s^t [w_i(t) - w_i(u)]\phi_j(u)\,du\right]\eta(t)|\mathcal{F}_s\\
&\qquad -E\left[\int_s^t [w_j(t) - w_j(u)]\phi_i(u)\,du\right]\eta(t)|\mathcal{F}_s + EJ_4|\mathcal{F}_s. \qquad (2.25)
\end{aligned}$$

The first term on the right-hand side is equal to

$$\begin{aligned}
&-E\left[\int_s^t [\tilde{w}_i(t) - \tilde{w}_i(u)]\phi_j(u)\eta(t)\,du\right]|\mathcal{F}_s\\
&\quad -E\left[\int_s^t\left[\int_u^t \phi_i(v)\,dv\right]\phi_j(u)\eta(t)\,du\right]|\mathcal{F}_s. \qquad (2.26)
\end{aligned}$$

The first term in (2.26) is equal to

$$\begin{aligned}
&-E\left\{\int_s^t \{E[\tilde{w}_i(t) - \tilde{w}_i(u)\ \phi_j(u)]\,\eta(t)|\mathcal{F}_u\}\,du\right\}|\mathcal{F}_s\\
&\quad = -E\left\{\int_s^t \{\tilde{E}[\tilde{w}_i(t) - \tilde{w}_i(u)]\ \phi_j(u)|\mathcal{F}_u\}\,du\right\}|\mathcal{F}_s\\
&\quad = -E\left\{\int_s^t \{\phi_j(u)\tilde{E}[\tilde{w}_i(t) - \tilde{w}_i(u)]\ |\mathcal{F}_u\}\,du\right\}|\mathcal{F}_s = 0
\end{aligned}$$

by (2.8) and (2.19). Hence the first term on the right-hand side of (2.25) is equal to the second term in (2.26).

Treating the second term on the right-hand side of (2.25) similarly, we conclude that

$$\begin{aligned}
&\tilde{E}[\tilde{w}_i(t) - \tilde{w}_i(s)][\tilde{w}_j(t) - \tilde{w}_j(s)]|\mathcal{F}_s - \delta_{ij}(t-s)\\
&\quad = E\left\{\int_s^t \frac{d}{du}\left\{\left[\int_u^t \phi_i(v)\,dv\right]\left[\int_u^t \phi_j(v)\,dv\right]\right\}\eta(t)\,du\right\}|\mathcal{F}_s + EJ_4|\mathcal{F}_s\\
&\quad = -E\left\{\left[\int_s^t \phi_i(u)\,du\right]\left[\int_s^t \phi_j(u)\,du\right]\eta(t)\right\}|\mathcal{F}_s + EJ_4|\mathcal{F}_s = 0
\end{aligned}$$

by the definition of J_4. This completes the proof of (2.20).

Proof of Theorem 2.1. Let $\phi^N = (\phi_1^N, \ldots, \phi_n^N)$ be bounded functions in $L_w^2[0, T]$ such that

$$\int_0^T |\phi^N(t) - \phi(t)|^2\,dt \to 0 \quad \text{a.s.} \qquad \text{as} \quad N\to\infty.$$

Define $\tilde{w}^N(t) = (\tilde{w}_1^N(t), \ldots, \tilde{w}_n^N(t))$ by

$$\tilde{w}^N(t) = w(t) - \int_s^t \phi^N(u)\, du.$$

Notice that

$$\tilde{w}^N(t) \to \tilde{w}(t) \quad \text{a.s.} \tag{2.27}$$

In view of Lemma 2.6, $\tilde{w}^N(t)$ is a Brownian motion in the interval $[0, T]$, with respect to the probability space $(\Omega, \mathcal{F}, \tilde{P}_N)$ where $\tilde{P}_N$ is defined by (2.3) with $\phi = \phi^N$. Consequently, for any $0 \leqslant t_0 < t_1 < \cdots < t_k \leqslant T$ and $\lambda_1, \ldots, \lambda_k$ real,

$$E \exp\left\{ \sqrt{-1} \sum_{j=1}^{k} \lambda_j \left[\tilde{w}_i^N(t_j) - \tilde{w}_i^N(t_{j-1}) \right] \right\} \exp \zeta_0^T(\phi^N)$$

$$= \exp\left\{ - \sum_{j=1}^{k} \tfrac{1}{2}\lambda_j^2 (t_j - t_{j-1}) \right\}, \tag{2.28}$$

and for any $0 \leqslant t, s \leqslant T$ and λ, μ real,

$$E \exp\left\{ \sqrt{-1}\, [\lambda \tilde{w}_i^N(t) + \mu \tilde{w}_j^N(s)] \right\} \exp \zeta_0^T(\phi^N)$$

$$= \exp[-(\lambda^2/2)t] \exp[-(\mu^2/2)s] \qquad \text{if} \quad i \neq j. \tag{2.29}$$

We shall need the following fact:

$$\text{If } \alpha_N \to \alpha \text{ a.s., } |\alpha_N| \leqslant c \ (c \text{ const}), \text{ and } E|\beta_N - \beta| \to 0,$$
$$\text{then } E|\alpha_N \beta_N - \alpha\beta| \to 0. \tag{2.30}$$

To prove (2.30) write

$$E|\alpha_N\beta_N - \alpha\beta| \leqslant E|\alpha_N|\,|\beta_N - \beta| + E|\alpha_N - \alpha|\,|\beta|$$
$$\leqslant cE|\beta_N - \beta| + E|\alpha_N - \alpha|\,|\beta|.$$

The first integral on the right converges to zero by assumption, and the second integral converges to zero by the Lebesgue bounded convergence theorem. Thus (2.30) follows.

Taking α_N to be the first exponent on the left-hand side of (2.28) and taking $\beta_N = \exp \zeta_0^T(\phi^N)$, and recalling (2.27) and Lemma 2.5, we see that the assumptions in (2.30) are satisfied. Hence, letting $N \to \infty$ in (2.28) we obtain

$$\tilde{E} \exp\left\{ \sqrt{-1} \sum_{j=1}^{k} \lambda_j [\tilde{w}_i(t_j) - \tilde{w}_i(t_{j-1})] \right\} = \exp\left\{ - \sum_{j=1}^{k} \tfrac{1}{2}\lambda_j^2 (t_j - t_{j-1}) \right\}.$$

This implies (see Problem 2, Chapter 3) that $\tilde{w}_i(t)$, $0 \leqslant t \leqslant T$ is a Brownian motion in the space $(\Omega, \mathcal{F}, \tilde{P})$.

If we let $N \to \infty$ in (2.29) and apply (2.30) (again using (2.27) and Lemma 2.5), we obtain

$$\tilde{E} \exp\left\{ \sqrt{-1}\, [\lambda \tilde{w}_i(t) + \mu \tilde{w}_j(s)] \right\} = \exp[-(\lambda^2/2)t] \exp[-(\mu^2/2)s].$$

This proves that $\mathcal{F}\{\tilde{w}_i(t), 0 \leqslant t \leqslant T\}$ is independent of $\mathcal{F}\{\tilde{w}_j(t), 0 \leqslant t \leqslant T\}$ if $i \neq j$. Thus $\tilde{w}(t)$ is an n-dimensional Brownian motion in the space $(\Omega, \mathcal{F}, \tilde{P})$.

Remark. Since $\tilde{w}^N$ satisfies (2.19), (2.20), the remark at the end of Section 3.5 implies that, if $0 \leqslant s < t \leqslant T$,

$$\tilde{E}\exp\left[i\lambda \cdot \left(\tilde{w}^N(t) - \tilde{w}^N(s)\right)\right]|\mathcal{F}_s = \exp\left[-\tfrac{1}{2}|\lambda|^2(t-s)\right] \quad \text{a.s.} \tag{2.31}$$

If α_N, β_N are as in (2.30) then $E(\alpha_N\beta_N - \alpha\beta)|\mathcal{F}_s \to 0$ in probability, as $N \to \infty$. Using this fact we get, as $N \to \infty$ in (2.31),

$$\tilde{E}\exp[i\lambda \cdot (\tilde{w}(t) - \tilde{w}(s))]|\mathcal{F}_s = \exp\left[-\tfrac{1}{2}|\lambda|^2(t-s)\right] \quad \text{a.s.} \tag{2.32}$$

Now, if X has n-dimensional normal distribution, then $E|X|^m \leqslant c^{m+1}m!$ for some $c > 0$. Hence the series

$$\sum_{m=0}^{\infty} \frac{i^m(\lambda \cdot X)^m}{m!}$$

is absolutely convergent in $L^1(\Omega)$ if $|\lambda| < 1/c$. It follows that

$$E\left[e^{i\lambda \cdot X} - \sum_{m=1}^{l} \frac{i^m(\lambda \cdot X)^m}{m!}\right]|\mathcal{F}_s \xrightarrow{P} 0 \qquad \text{if} \quad l \to \infty. \tag{2.33}$$

(For if $E|\alpha_N - \alpha| \to 0$, then $E(\alpha_N - \alpha)|\mathcal{F}_s \xrightarrow{P} 0$.) If

$$Ee^{i\lambda \cdot X}|\mathcal{F}_s = \exp\left[-\tfrac{1}{2}|\lambda|^2(t-s)\right],$$

then (2.33) implies that

$$\sum_{m=0}^{\infty} \frac{i^m}{m!} E(\lambda \cdot X)^m|\mathcal{F}_s = \sum_{m=0}^{\infty} \frac{|\lambda|^{2m}(t-s)^m}{2^m m!} \quad \text{a.s.} \quad \left(|\lambda| < \frac{1}{c}\right). \tag{2.34}$$

In view of (2.32), (2.34) can be applied to $X = \tilde{w}(t) - \tilde{w}(s)$. Comparing coefficients for $m = 1, 2$ we deduce:

$$\tilde{E}(\tilde{w}(t) - \tilde{w}(s))|\mathcal{F}_s = 0 \qquad \text{if} \quad 0 \leqslant s < t \leqslant T, \tag{2.35}$$

$$\tilde{E}[\tilde{w}_i(t) - \tilde{w}_i(s)][\tilde{w}_j(t) - \tilde{w}_j(s)]|\mathcal{F}_s = \delta_{ij}(t-s) \qquad \text{if} \quad 0 \leqslant s < t \leqslant T. \tag{2.36}$$

3. Girsanov's formula

Let $w(t)$ be an n-dimensional Brownian motion in the interval $0 \leqslant t \leqslant T$. In Chapter 4 we have defined the concept of the stochastic integral $\int_0^T f(s)\,dw(s)$ for functions f in $L_w^2[0, T]$, where $L_w^2[0, T]$ is defined with respect to an increasing family of σ-fields $\mathcal{F}_t$ satisfying:

(i) $\mathcal{F}(w(\lambda), \lambda \leq t)$ is in $\mathcal{F}_t$, for any $0 \leq t \leq T$;

(ii) $\mathcal{F}(w(t+\lambda) - w(t), 0 \leq \lambda \leq T - t)$ is independent of $\mathcal{F}_t$, for any $0 \leq t \leq T$.

(More precisely, we have assumed that the $\mathcal{F}_t$ are defined for all $t \geq 0$ and satisfy (i) for all $t \geq 0$, and (ii) with $0 \leq \lambda < \infty$ and for any $t \geq 0$. However we have actually made use only of (i), (ii).)

Suppose we replace (ii) by the weaker condition:

(ii′) For all $0 \leq s < t \leq T$,

$$E[w(t) - w(s)]|\mathcal{F}_s = 0,$$

$$E[w_i(t) - w_i(s)][w_j(t) - w_j(s)]|\mathcal{F}_s = \delta_{ij}(t - s).$$

Then we can still define the stochastic integral and derive all the formulas and estimates as in Chapter 4. In fact we first define the stochastic integral for any step function, next derive Lemmas 4.2.2, 4.2.3 (see Problem 1), and then approximate any f in $L_w^2[0, T]$ by bounded step functions and use the procedure of Section 4.2. All the results of Chapters 4,5 and of Sections 1, 2 of this chapter remain valid, without any change, for this slightly more general notion of the stochastic integral.

Definition. From now on we shall assume only that the family $\mathcal{F}_t$ satisfies (i) and (ii′). We shall say that $\mathcal{F}_t$ is *adapted* to $w(t)$.

Let $\phi(t)$ be as in Theorem 2.1. In view of (2.35), (2.36), $\mathcal{F}_t$ is adapted to $\tilde{w}(t)$. If $f \in L_w^2[0, T]$, then

$$P\left\{ \int_0^T |f(s, \omega)|^2 \, ds < \infty \right\} = 1.$$

Since $\tilde{P}$ is absolutely continuous with respect to P,

$$\tilde{P}\left\{ \int_0^T |f(s, \omega)|^2 \, ds < \infty \right\} = 1.$$

Hence, if $f \in L_w^2[0, T]$, with any $\mathcal{F}_t$ adapted to $w(t)$, then $f \in L_{\tilde{w}}^2[0, T]$ with the same $\mathcal{F}_t$.

Let f_k be a sequence of step functions in $L_w^2[0, T]$ such that

$$\int_0^T |f_k(s) - f(s)|^2 \, ds \to 0 \quad \text{a.s.} \qquad \text{in } P.$$

Then $f_k \in L_{\tilde{w}}^2[0, T]$ and

$$\int_0^T |f_k(s) - f(s)|^2 \, ds \to 0 \quad \text{a.s.} \qquad \text{in } \tilde{P}.$$

Consequently

$$\int_0^t f_k(s)\,dw(s) \xrightarrow{P} \int_0^t f(s)\,dw(s),$$
$$\int_0^t f_k(s)\,d\tilde{w}(s) \xrightarrow{\tilde{P}} \int_0^t f(s)\,d\tilde{w}(s).$$

Hence, for a subsequence $\{k'\}$,

$$\begin{aligned} &\int_0^t f_{k'}(s)\,dw(s) \to \int_0^t f(s)\,dw(s) \quad \text{a.s.} \quad \text{in } P,\\ &\int_0^t f_{k'}(s)\,d\tilde{w}(s) \to \int_0^t f(s)\,d\tilde{w}(s) \quad \text{a.s.} \quad \text{in } \tilde{P}. \end{aligned} \tag{3.1}$$

From (2.2) we easily see that

$$\int_0^t f_{k'}(s)\,dw(s) = \int_0^t f_{k'}(s)\,d\tilde{w}(s) + \int_0^t f_{k'}(s)\phi(s)\,ds. \tag{3.2}$$

It is also clear that

$$\int_0^t f_{k'}(s)\phi(s)\,ds \to \int_0^t f(s)\phi(s)\,ds \quad \text{a.s.} \quad \text{in } P.$$

Hence, taking $k' \to \infty$ in (3.2) and using (3.1), we get

$$\int_0^t f(s)\,dw(s) = \int_0^t f(s)\,d\tilde{w}(s) + \int_0^t f(s)\phi(s)\,ds \qquad \text{if} \quad f \in L_w^2[0, T], \tag{3.3}$$

a.s. in P (or in $\tilde{P}$).

Definition. A stochastic process $x(t)$ $(0 \leqslant t \leqslant T)$ is called an *Itô process with respect to* $\{w(t), P, \mathcal{F}_t\}$ (where $\mathcal{F}_t$ is adapted to $w(t)$) *relative to the pair* $\sigma(t)$, $b(t)$ if

$$x(t) = x(0) + \int_0^t b(s)\,ds + \int_0^t \sigma(s)\,dw(s) \qquad \text{for} \quad 0 \leqslant t \leqslant T \tag{3.4}$$

where $b = (b_1, \ldots, b_n)$ is in $L_w^1[0, T]$ and $\sigma = (\sigma_{ij})_{i,j=1}^n$ is in $L_w^2[0, T]$.

Theorem 3.1. *Let $x(t)$ $(0 \leqslant t \leqslant T)$ be an Itô process with respect to $\{w(t), P, \mathcal{F}_t\}$ relative to the matrix $\sigma(t)$ and the vector $b(t)$. Let $\phi(t) = (\phi_1(t), \ldots, \phi_n(t))$ belong to $L_w^2[0, T]$ and set*

$$\zeta_s^t(\phi) = \int_s^t \phi(u)\,dw(u) - \tfrac{1}{2}\int_s^t |\phi(u)|^2\,du, \tag{3.5}$$

$$\tilde{w}(t) = w(t) - \int_0^t \phi(s)\,ds, \tag{3.6}$$

$$d\tilde{P}(\omega) = \exp\left[\zeta_0^T(\phi)\right] dP(\omega). \tag{3.7}$$

Assume that

$$\tilde{P}(\Omega) = 1. \tag{3.8}$$

Then $\tilde{w}(t)$ is an n-dimensional Brownian motion in the probability space $(\Omega, \mathcal{F}, \tilde{P})$, $\mathcal{F}_t$ is adapted to $\tilde{w}(t)$, and the process $x(t)$ is an Itô process with respect to $\{\tilde{w}(t), \tilde{P}, \mathcal{F}_t\}$ relative to the matrix $\sigma(t)$ and the vector

$$\tilde{b}(t) = b(t) + \sigma(t)\phi(t). \tag{3.9}$$

This follows by combining Theorem 2.1 with (2.35), (2.36), and (3.3).

Denote by $\mathcal{C}_T^n$ the space of continuous functions $x(t)$ from $0 \leqslant t \leqslant T$ into R^n. Denote by $\mathfrak{M}_T$ the σ-algebra generated by the sets $\{x(t) \in A\}$ where $0 \leqslant t \leqslant T$ and A is any Borel set in R^n.

A continuous n-dimensional process $\xi(t)$ $(0 \leqslant t \leqslant T)$ in $(\Omega, \mathcal{F}, P)$ induces a probability μ_p on $(\mathcal{C}_T^n, \mathfrak{M}_T)$ as follows:

Denote by ξ_ω the sample path $t \to \xi(t, \omega)$ $(0 \leqslant t \leqslant T)$, and denote by $\mathcal{C}_T^n(\xi)$ the set of all elements ξ_ω, $\omega \in \Omega$. For any set $B \in \mathfrak{M}_T$ define

$$\mu_P(B) = P\{\omega; \xi_\omega \in B\}. \tag{3.10}$$

It is clear that μ_P is a probability, and $\mu_P(B) = \mu_P[B \cap \mathcal{C}_T^n(\xi)]$. In particular

$$\mu_P\{x; x(t_1) \in A_1, \ldots, x(t_k) \in A_k\} = P\{\omega; \xi(t_1, \omega) \in A_1, \ldots, \xi(t_k, \omega) \in A_k\}. \tag{3.11}$$

Suppose now that Q is another probability on $(\Omega, \mathcal{F})$ given by

$$dQ(\omega) = \rho(\omega)\, dP(\omega); \tag{3.12}$$

thus $\int \rho(\omega)\, dP(\omega) = 1$. Introduce the probability μ_Q on $(\mathcal{C}_T^n, \mathfrak{M}_T)$ corresponding to Q:

$$\mu_Q(B) = Q\{\omega; \xi_\omega \in B\}.$$

Lemma 3.2 *Under the foregoing assumptions,*

$$\frac{d\mu_Q}{d\mu_P}(\xi_\omega) = \rho(\omega), \tag{3.13}$$

i.e., for any $B \in \mathfrak{M}_T$, $B \subset \mathcal{C}_T^n(\xi)$,

$$\mu_Q(B) = \int_B \rho(\omega)\, d\mu_P(\xi_\omega). \tag{3.14}$$

Proof. It suffices to verify (3.14) for sets of the form

$$B = \{\xi_\omega; \xi(t_1, \omega) \in A_1, \ldots, \xi(t_k, \omega) \in A_k\}.$$

Let

$$\tilde{B} = \{\omega; \xi(t_1, \omega) \in A_1, \ldots, \xi(t_k, \omega) \in A_k\}.$$

$|f_i| > k$, then

$$\int_0^T |f - g_k|^2 \, dt \xrightarrow{P} 0, \qquad \int_0^T g_k \, dw \xrightarrow{P} \int_0^T f \, dw \qquad \text{if } f \in L_w^2[0, T],$$

$$E\int_0^T |g_k - f|^2 \, dt \to 0, \qquad E\left|\int_0^T g_k \, dw - \int_0^T f \, dw\right|^2 \to 0 \qquad \text{if } f \in M_w^2[0, T].$$

(c) Use (a), (b) in order to prove Lemmas 4.2.2, 4.2.3 for any step function f in $L_w^2[0, T]$.

2. Consider a system of n *stochastic functional differential equations*

$$x(t, \omega) = x(0, \omega) + \int_0^t \sigma(x(\,\cdot\,, \omega), s) \, dw(s, \omega) + \int_0^t b(x(\,\cdot\,, \omega), s) \, ds \quad (3.23)$$

where $b = (b_1, \dots, b_n)$, $\sigma = (\sigma_{ij})_{i,j=1}^n$, and $b_i(x(\,\cdot\,), t)$, $\sigma_{ij}(x(\,\cdot\,), t)$ are measurable functions on $\mathcal{C}_T^n \times [0, T]$, measurable with respect to $\mathfrak{M}_t$, for each t (where $\mathfrak{M}_t$ is the σ-algebra generated by the sets $\{x(s) \in A\}$, $0 \leqslant s \leqslant t$, A a Borel set in R^n). Assume that

$$|f(x(\,\cdot\,), t)| \leqslant K(1 + \|x(\,\cdot\,)\|),$$
$$|f(x(\,\cdot\,), t) - f(\hat{x}(\,\cdot\,), t)| \leqslant K\|x(\,\cdot\,) - \hat{x}(\,\cdot\,)\|$$

for $f = \sigma_{ij}$ and $f = b_i$, where $\|x(\,\cdot\,)\| = \max_{0 \leqslant t \leqslant T}|x(t)|$.

(i) Prove by the method of successive approximations that for any $x(0, \omega)$ in $L^2(\Omega)$ which is independent of $\mathcal{F}(w(s), 0 \leqslant s \leqslant T)$ there exists a unique solution of (3.23) satisfying $E|x(t, \omega)|^2 \leqslant$ const $(0 \leqslant t \leqslant T)$.

(ii) Prove uniqueness in the sense of probability law.

3. Extend Girsanov's formula to stochastic functional differential equations.

4. Let $b(x)$ be uniformly Lipschitz continuous in $x \in R^1$. Consider the stochastic differential equation $d\xi(t) = dw(t) + b(\xi(t)) \, dt$. Prove that

$$P(s, x, t, A) = \int_A \frac{\Phi(s, x, t, y)}{\sqrt{2\pi(t - s)}} \exp\left[-\frac{(y - x)^2}{2(t - s)}\right] dy$$

where

$$\Phi(s, x, t, y) = E \exp\left\{\int_s^t b(\bar{\xi}_{x,s}(u)) \, dw(u) - \tfrac{1}{2}\int_s^t b^2(\bar{\xi}_{x,s}(u)) \, du\right\} | \bar{\xi}_{x,s}(t)$$

and $\bar{\xi}_{x,s}(t) = x + w(t) - w(s)$.

5. Let $w(t)$ be an n-dimensional Brownian motion. The processes $w(t)$ and $2w(t)$ induce probability measures in $(\mathcal{C}_T^n, \mathfrak{M}_T)$, μ_w and μ_{2w} respectively. Prove that μ_w and μ_{2w} are mutually singular.

6. Let $y = f(x)$ be a local diffeomorphism in R^n. A given stochastic system $dx = \sigma(x) \, dw + b(x) \, dt$ is then transformed, by means of Itô's formula, into $dy = \hat{\sigma}(y) \, dw + \hat{b}(y) \, dt$. Denote by L and $\hat{L}$ the differential operators

corresponding to the first and second stochastic systems, respectively. If $u(x) = \hat{u}(y)$ where $y = f(x)$, prove that $Lu(x) = \hat{L}\hat{u}(y)$; thus the relation between a stochastic differential system and its differential operator is invariant under diffeomorphism.

8

Asymptotic Estimates for Solutions

In this chapter (except for Section 1) we consider the behavior, as $t \to \infty$, of solutions of a stochastic differential system in case the diffusion matrix is nondegenerate for all $(x, t) \in R^n \times [0, \infty)$.

1. Unboundedness of solutions

Consider a system of n stochastic differential equations

$$d\xi(t) = \sigma(\xi(t), t)\, dw(t) + b(\xi(t), t)\, dt \tag{1.1}$$

with initial condition

$$\xi(0) = \xi_0 \tag{1.2}$$

where ξ_0 is independent of $\mathcal{F}\{w(t), t \geq 0\}$ and $E|\xi_0|^2 < \infty$. Set

$$b = (b_1, \ldots, b_n), \qquad \sigma = (\sigma_{ij})_{i,j=1}^{n}, \qquad a = (a_{ij}), \qquad a_{ij} = \sum_{k=1}^{n} \sigma_{ik}\sigma_{jk}.$$

If C is the matrix (c_{ij}), we write

$$|C| = \left\{ \sum_{i,j} (c_{ij})^2 \right\}^{1/2}.$$

We shall need the following conditions:

(A_0) $\sigma(x, t)$ and $b(x, t)$ are measurable functions in $R^n \times [0, \infty)$ and, for any $T > 0$, $R > 0$, there are positive constants C_T, $C_{T,R}$ such that

$$|\sigma(x, t)| + |b(x, t)| \leq C_T(1 + |x|) \quad \text{if } x \in R^n, \quad 0 \leq t \leq T,$$

$$|\sigma(x, t) - \sigma(\bar{x}, t)| + |b(x, t) - b(\bar{x}, t)| \leq C_{T,R}|x - \bar{x}| \quad \text{if } |x| \leq R, \quad |\bar{x}| \leq R, \quad 0 \leq t \leq T.$$

(A_1) For any $R > 0$ there exist positive constants γ_R, Γ_R such that

$$a_{11}(x, t)\Gamma_R + b_1(x, t) \geqslant \gamma_R \quad \text{if} \quad |x| \leqslant R, \quad t \geqslant 0.$$

This condition is satisfied, for example, if $a_{11}(x, t) \geqslant \lambda$, $b_1(x, t) \leqslant C$ where λ, C are positive constants.

Theorem 1.1. *Suppose* (A_0) *and* (A_1) *hold. Let* $\xi(t)$ *be the solution of* (1.1), (1.2). *Then*

$$\limsup_{t\to\infty} |\xi(t)| = \infty \quad a.s. \tag{1.3}$$

Proof. Since $\xi(t)$ is a continuous process, the assertion (1.3) is equivalent to the assertion that

$$\sup_{t>0} |\xi(t)| = \infty \quad \text{a.s.} \tag{1.4}$$

Consider first the case where the range of ξ_0 lies in a bounded set K. Let B be any open ball containing K and denote by $\tau(B)$ the exit time from B, i.e.,

$\tau(B)$ = first $\bar{t}$ such that $\xi(\bar{t}) \not\in B$ if such $\bar{t}$ exists,
$\tau(B) = \infty$ if no such $\bar{t}$ exist.

For any $T > 0$ let $\tau_T = \tau(B) \wedge T$. Set $B_T = B \times \{t = T\}$, $S_T = \partial B \times \{0 < t < T\}$ where ∂B is the boundary of B.

Suppose $\phi(x, t)$ is a smooth function satisfying

$$L\phi \equiv \frac{\partial \phi}{\partial t} + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x, t) \frac{\partial^2 \phi}{\partial x_i\, \partial x_j} + \sum_{i=1}^{n} b_i(x, t) \frac{\partial \phi}{\partial t} \leqslant -1$$
$$\text{in} \quad B \times [0, T], \qquad \phi \geqslant 0 \quad \text{on} \quad B_T \cup S_T. \tag{1.5}$$

If we apply Itô's formula to $\phi(\xi(t), t)$ and substitute $t = \tau_T$, we get after taking the expectation (cf. Theorem 6.5.2)

$$E\phi(\xi(0), 0) \geqslant E\tau_T, \tag{1.6}$$

provided the range of $\xi(0)$ is in K.

If $1 + x_1 \leqslant x_1^0$ for all $(x_1, x_2, \ldots, x_n) \in B$, then we can take

$$\phi(x, t) = A[\exp(\alpha x_1^0) - \exp(\alpha x_1)],$$

where α, A are suitable positive constants. The condition (A_1) is used here in verifying (1.5). Since $\phi \leqslant C$ where C is a constant independent of T, (1.6) yields $E\tau_T \leqslant C$.

Now, $\tau_T \uparrow \tau(B)$ if $T \uparrow \infty$. Using the monotone convergence theorem we conclude that $\tau(B)$ satisfies

$$E\tau(B) \leqslant C. \tag{1.7}$$

In particular,

$$\tau(B) < \infty \quad \text{a.s.} \tag{1.8}$$

Take a strictly increasing sequence of open balls B_m with center 0 and radius $R_m \to \infty$. Let

$$\Omega_m = \Big\{ \sup_{t>0} |\xi(t)| \geq R_m \Big\}, \qquad \Omega^* = \bigcap_{m=1}^{\infty} \Omega_m.$$

Denote by χ_m the indicator function of B_m and let $\xi_m(t)$ be the solution of (1.1) with the initial condition $\xi_m(0) = \chi_m \xi(0)$. Denote by $\tau_m(B_m)$ the exit time from B_m of the solution $\xi_m(t)$. By Theorem 5.2.1, $\tau_m(B_m) = \tau(B_m)$ a.s. and $\xi_m(t) = \xi(t)$ if $0 \leq t \leq \tau(B_m)$. Hence,

$$P\Big\{ \sup_{t>0} |\xi(t)| \geq R_m \Big\} = P\Big\{ \sup_{t>0} |\xi_m(t)| \geq R_m \Big\}$$

and the right-hand side is equal to 1, by (1.8). We conclude that $P(\Omega_m) = 1$. Since the sequence Ω_m is monotone decreasing to $\Omega^* = \bigcap_{m=1}^{\infty} \Omega_m$,

$$P(\Omega^*) = \lim_{m\to\infty} P(\Omega_m) = 1.$$

Now, if $\omega \in \Omega^*$, then $\omega \in \Omega_m$ for all m, so that

$$\sup_{t>0} |\xi(t)| \geq R_m \quad \text{for all} \quad m, \qquad \text{i.e.,} \qquad \sup_{t>0} |\xi(t)| = \infty.$$

This completes the proof.

From the proof of (1.7) we have:

Corollary 1.2. *Let $\xi(t)$ be a solution of* (1.1) *with* $|\xi(0)| < R$ *a.s. Suppose* $a_{11}(x, t) \geq \lambda$, $b_1(x, t) \leq C$ *if* $|x| \leq R$, $t \geq 0$, *where* λ, C *are positive constants. Then*

$$E\tau^R < \infty$$

where τ^R *is the exit time from the ball* $|x| < R$.

2. Auxiliary estimates

In the sequel we shall assume that $\sigma(x, t)$, $b(x, t)$ are measurable and $\xi(t)$ is any solution of (1.1), (1.2) such that

$$\sup_{0 \leq t \leq T} E|\xi(t)|^2 < \infty \qquad \text{for any} \quad T > 0.$$

If the condition (A_0) holds, then, of course, there exists a unique solution.

Set

$$Lu \equiv \tfrac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x, t) \frac{\partial^2 u}{\partial x_i \, \partial x_j} + \sum_{i=1}^{n} b_i(x, t) \frac{\partial u}{\partial x_i} + \frac{\partial u}{\partial t}.$$

Theorem 2.1. *Assume that*

$$\sum_{i=1}^{n} a_{ii}(x, t) \leqslant C, \qquad \sum_{i=1}^{n} b_i(x, t) \leqslant C$$

where C is a positive constant. Then

$$E|\xi(t)|^2 \leqslant Kt + K' \qquad \text{for all} \quad t \geqslant 0, \tag{2.1}$$

where K, K′ are positive constants.

Proof. Using Itô's formula with $u(x) = |x|^2$ and $\xi(t)$, and taking the expectation, we get

$$E|\xi(t)|^2 = E|\xi_0|^2 + E \int_0^t g(\xi(s), s)\, ds$$

where

$$g(x, t) = L|x|^2 = \sum a_{ii}(x, t) + 2\sum x_i b_i(x, t).$$

Since, by our assumptions, $g(x, t) \leqslant C_1$ where C_1 is a positive constant, (2.1) follows.

Notation. We shall denote the eigenvalues of $a(x, t)$ by $\lambda_i(x, t)$ where

$$\lambda_1(x, t) \leqslant \lambda_2(x, t) \leqslant \cdots \leqslant \lambda_n(x, t).$$

Lemma 2.2. *Assume that*

$$\sum_{i,j} |a_{ij}(x, t)| \leqslant C \qquad \text{for all} \quad x \in R^n, \quad t \geqslant 0 \qquad (C\ const); \tag{2.2}$$

for any $R > 0$ there is a positive constant $\mu(R)$ such that

$$\sum_{i,j} a_{ij}(x, t)\xi_i\xi_j \geqslant \mu(R)|\xi|^2 \qquad if \quad |x| \leqslant R, \quad t \geqslant 0, \quad \xi \in R^n, \tag{2.3}$$

$$(1 + |x|)\sum_i |b_i(x, t)| \leqslant \epsilon(|x|) \qquad for \quad x \in R^n, \quad t \geqslant 0,$$

$$where \quad \epsilon(r) \to 0 \quad if \ r \to \infty; \tag{2.4}$$

$$\lambda_n(x, t) \geqslant \gamma > 0 \qquad for \quad x \in R^n, \quad t \geqslant 0 \qquad (\gamma\ const). \tag{2.5}$$

Let $\phi(x, t)$ be a bounded measurable function such that $\phi(x, t) = 0$ if $x \notin G$ where G is a compact set. Then, for any $\eta > 0$,

$$E\left|\int_0^t \phi(\xi(s), s)\, ds\right| \leqslant K_1 + K_2 t^{(1+\eta)/2} \tag{2.6}$$

where K_1, K_2 are positive constants. If further,

$$\lambda_{n-1}(x, t) \geqslant \gamma' > 0 \qquad for \quad x \in R^n, \quad t \geqslant 0 \qquad (\gamma'\ const), \tag{2.7}$$

then (2.6) holds for some $-1 < \eta < 0$.

Proof. Let $r = |x - x^0|$ where x^0 is a point lying outside G. For simplicity we shall assume in what follows that $x^0 = 0$. It is clear that there exists a continuous function $\psi(r)$ such that

$$|\phi(x, t)| \leqslant \psi(r) \tag{2.8}$$

and $\psi(r) = 0$ if $r \leqslant r_0$ or if $r \geqslant r_1$ for some $0 < r_0 < r_1 < \infty$. We shall construct a function $F(r)$ such that the function $f(x) = F(|x|)$ is in $C^2(R^n)$ and

$$(Lf)(x, t) \geqslant \psi(|x|). \tag{2.9}$$

One can easily verify that

$$2Lf(r) = A(x, t)F''(r) + \frac{F'(r)}{r}[B(x, t) - A(x, t) + C(x, t)] \tag{2.10}$$

where

$$A(x, t) = \frac{1}{r^2}\sum_{i,j=1}^{n} a_{ij}(x, t)x_i x_j, \qquad B(x, t) = \sum_{i=1}^{n} a_{ii}(x, t),$$

$$C(x, t) = 2\sum_{i=1}^{n} x_i b_i(x, t).$$

Let $\theta(r)$ be a continuous function, vanishing for $r < r_0/2$ and satisfying, for $r > r_0$,

$$(1 + \theta(r))A(x, t) \leqslant B(x, t) + C(x, t) \tag{2.11}$$

Since, for any $R > 0$, $A(x, t) \geqslant \mu(R) > 0$ if $t \geqslant 0$, $|x| \leqslant R$, and since $B(x, t) \geqslant 0$ and $|C(x, t)|$ is bounded, one can certainly construct such a function $\theta(r)$. Noting that

$$\lambda_1(x, t) \leqslant A(x, t) \leqslant \lambda_n(x, t), \qquad B(x, t) = \lambda_1(x, t) + \cdots + \lambda_n(x, t)$$

and using (2.5) and (2.4), we can construct $\theta(r)$ such that

$$\theta(r) = -\eta \quad \text{if } r \text{ is sufficiently large (for any } \eta > 0). \tag{2.12}$$

If (2.7) holds, then we can take

$$\theta(r) = -\eta \quad \text{if } r \text{ is sufficiently large (for some } -1 < \eta < 0); \tag{2.13}$$

in fact, we can take any η such that

$$-\eta < \gamma'/\delta' \qquad \text{where} \quad \delta' = \limsup_{|x|\to\infty} \Big\{ \sup_{t>0} \lambda_n(x, t) \Big\}.$$

Introduce the functions

$$\begin{aligned} I(s) &= \int_{r_0}^{s} \frac{\theta(u)}{u}\, du, \\ f(x) = F(r) &= \frac{1}{\lambda}\int_0^r e^{-I(s)}\, ds \cdot \int_0^s e^{I(u)}\psi(u)\, du \end{aligned} \tag{2.14}$$

for some $\lambda > 0$, $r = |x|$. Since F is in C^2 and $F(r) = 0$ if $r \leqslant r_0$, $f(x)$ is in $C^2(R^n)$. One can easily check that

$$F''(r) + \frac{\theta(r)}{r} F'(r) = \frac{\psi(r)}{\lambda} \tag{2.15}$$

and that $F'(r) \geqslant 0$. Using (2.10), (2.11), we then get

$$2Lf(x) \geqslant \left[F''(r) + \frac{\theta(r)}{r} F'(r) \right] A(x, t) \geqslant A(x, t) \frac{\psi(r)}{\lambda} \geqslant 2\psi(r)$$

if $2\lambda = \mu(r_1)$ (since $A(x, t) \geqslant \mu(r_1)$ if $|x| \leqslant r_1$, $t \geqslant 0$).

Having proved (2.9), we now use Itô's formula to get

$$E\left| \int_0^t \phi(\xi(s), s)\, ds \right| \leqslant E \int_0^t \psi(\xi(s))\, ds \leqslant E \int_0^t (Lf)(\xi(s), s)\, ds$$
$$= EF(|\xi(t)|) - EF(|\xi_0|) \leqslant EF(|\xi(t)|).$$

Noting that $F(r) \leqslant Cr^{1+\eta}$, we obtain

$$E\left| \int_0^t \phi(\xi(s), s)\, ds \right| \leqslant CE|\xi(t)|^{1+\eta} + C \leqslant C\left\{ E|\xi(t)|^2 \right\}^{(1+\eta)/2} + C$$

for suitable constants C. Using Theorem 2.1, the assertion (2.6) follows.

We shall now extend Lemma 2.2 to the case where $\phi(x, t)$ does not necessarily vanish if $|x|$ is large, but either

$$|\phi(x, t) - \Phi| \leqslant \frac{C}{(1 + |x|)^\alpha} \qquad (\alpha > 0), \tag{2.16}$$

or

$$|\phi(x, t) - \Phi| \leqslant \frac{\epsilon(|x|)}{(1 + |x|)^\alpha} \qquad (\alpha \geqslant 0, \quad \epsilon(r) \to 0 \quad \text{if } r \to 0), \tag{2.17}$$

where Φ is a constant.

Lemma 2.3. *Let the conditions* (2.2)–(2.5) *hold and let* $\phi(x, t)$ *be a bounded measurable function satisfying* (2.16). *Then*

$$E\left| \int_0^t \phi(\xi(s), s)\, ds - \Phi t \right| = o(t^{(1+\eta)/2}) + O(t^{1-\alpha/2}) \tag{2.18}$$

for any $\eta > 0$; *if* (2.7) *also holds, then* (2.18) *holds for some* $-1 < \eta < 0$. *If the function* $\phi(x, t)$ *satisfies* (2.17) *instead of* (2.16), *then*

$$E\left| \int_0^t \phi(\xi(s), s)\, ds - \Phi t \right| = O(t^{(1+\eta)/2}) + o(t^{1-\alpha/2}) \tag{2.19}$$

with the same η *as before.*

Proof. Suppose first that (2.17) holds. Then, for any $\epsilon > 0$, there exists an

R_ϵ such that

$$|\phi(x, t) - \Phi| < \frac{\epsilon}{r^\alpha} \qquad \text{if} \quad |x| > R_\epsilon, \quad t \geq 0.$$

Let $\chi_\epsilon(x) = 1$ if $|x| < R_\epsilon$, $\chi_\epsilon(x) = 0$ if $|x| \geq R_\epsilon$, and write

$$\begin{aligned}\phi(x, t) - \Phi &= [\phi(x, t) - \Phi]\chi_\epsilon(x) + [\phi(x, t) - \Phi][1 - \chi_\epsilon(x)] \\ &= \phi_1(x, t) + \phi_2(x, t).\end{aligned} \tag{2.20}$$

By Lemma 2.2,

$$E\left|\int_0^t \phi_1(\xi(s), s)\, ds\right| = O(t^{(1+\eta)/2}). \tag{2.21}$$

Next, let $\psi(r)$ be a continuous function such that

$$\psi(r) = \begin{cases} 0 & \text{if} \quad 0 < r < r_0/2 \\ \epsilon/r^\alpha & \text{if} \quad r > r_0 \end{cases} \tag{2.22}$$

for some $0 < r_0 < R_\epsilon$. We shall slightly modify the definition of $\theta(r)$ by replacing the condition (2.11) by the stricter condition

$$(1 + \theta(r))A(x, t) \leq \rho B(x, t) + C(x, t) \tag{2.23}$$

where ρ is any positive number < 1. We can still satisfy (2.12) for any $\eta = \eta(\rho) > 0$ provided ρ is sufficiently close to 1; if (2.7) holds, then we can even satisfy (2.13) if ρ is sufficiently close to 1.

Let $f(x)$ be defined by (2.14) with $\lambda > 0$ to be determined later. Then, by (2.10), (2.23),.

$$\begin{aligned}2Lf(x) &\geq \left[F''(r) + \frac{\theta(r)}{r} F'(r)\right]A(x, t) + \frac{F'(r)}{r}(1 - \rho)B(x, t) \\ &\geq \frac{F'(r)}{r}(1 - \rho)B(x, t) \geq \frac{c}{\lambda}\,\frac{1}{r}\,e^{-I(r)} \int_{r_0}^{r} e^{I(s)} \frac{\epsilon}{s^\alpha}\, ds \\ &\geq \frac{c}{\lambda}\,\frac{\epsilon}{r^{1+\alpha}}\,e^{-I(r)} \int_{r_0}^{r} e^{I(s)}\, ds\end{aligned} \tag{2.24}$$

where c is a positive constant $\leq (1 - \rho)B(x, t)$. Since $e^{I(r)} \sim r^{-\eta}$ if $r \to \infty$, we find that

$$Lf(x) \geq \frac{c_1}{2\lambda}\,\frac{\epsilon}{r^\alpha} \geq |\phi_2(x, t)| \qquad \text{if} \quad |x| \geq R_\epsilon \tag{2.25}$$

where c_1 is a positive constant independent of ϵ, provided $2\lambda \leq c_1$. If $|x| < R_\epsilon$, then clearly $Lf(x) \geq 0 = |\phi_2(x, t)|$.

From Itô's formula and (2.25) we obtain

$$\begin{aligned} E\left|\int_0^t \phi_2(\xi(s), s)\, ds\right| &\leqslant E\int_0^t |\phi_2(\xi(s), s)|\, ds \\ &\leqslant E\int_0^t (Lf)(\xi(s), s)\, ds \\ &= EF(|\xi(t)|) - EF(|\xi_0|) \leqslant EF(|\xi(t)|). \end{aligned} \tag{2.26}$$

Since

$$F(r) \leqslant \begin{cases} C\epsilon r^{2-\alpha} + C & \text{if } 2-\alpha > 1+\eta, \\ C\epsilon r^{1+\eta} + C & \text{if } 2-\alpha < 1+\eta, \end{cases}$$

and since we may assume, without loss of generality, that $\alpha + \eta \neq 1$, we deduce, after using Hölder's inequality and Theorem 2.1,

$$EF(|\xi(t)|) \leqslant \begin{cases} C\epsilon t^{1-\alpha/2} + C & \text{if } \alpha+\eta < 1, \\ cc\epsilon t^{(1+\eta)/2} + C & \text{if } \alpha+\eta \geqslant 1. \end{cases}$$

Substituting this into (2.26) and combining the resulting inequality with (2.21), (2.20), we get

$$E\left|\int_0^t \phi(\xi(s), s)\, ds - \Phi t\right| \leqslant C_0 t^{(1+\eta)/2} + C_1\epsilon t^{1-\alpha/2} \tag{2.27}$$

where C_1 is a constant independent of ϵ. Since ϵ is arbitrary, (2.27) implies the assertion (2.19) with η replaced by η', for any $\eta' > \eta$. But this clearly completes the proof of (2.19). The proof of (2.18) (for ϕ satisfying (2.16)) is similar.

We proceed to evaluate (under additional conditions) more precisely the number η occurring in Lemma 2.3. We shall assume that, for $1 \leqslant i, j \leqslant n$,

$$a_{ij}(x, t) \to \bar{a}_{ij} \quad \text{as } |x| \to \infty, \text{ uniformly with respect to } t, \tag{2.28}$$

where $\bar{a}_{ij}$ are constants. Let

$$d = \text{number of positive eigenvalues of } \bar{a} = (\bar{a}_{ij}). \tag{2.29}$$

Observe that if (2.3) holds, then (2.5) holds if and only if $d \geqslant 1$, and (2.7) holds if and only if $d \geqslant 2$.

Note that the assumptions and assertions of Lemma 2.3 remain unchanged if we perform a nonsingular linear transformation $x \to Ax$. Such a transformation changes $a(x, t)$ into $Aa(x, t)A^*$. We can therefore choose A such that

$$\begin{aligned} \bar{a}_{ij} &= 0 && \text{if } i \neq j, \\ \bar{a}_{ii} &= 1 && \text{if } i = 1, 2, \ldots, d, \\ \bar{a}_{ii} &= 0 && \text{if } i = d+1, \ldots, n; \end{aligned}$$

$d = 0$ means that $\bar{a}_{ii} = 0$ for all i, and $d = n$ means that $\bar{a}_{ii} = 1$ for all i.

If $d \geqslant 2$, then we can take, in the proof of Lemmas 2.2, 2.3, $\theta(r) = \nu$ for any $\nu < 1$ provided r is sufficiently large. This leads to the assertions of Lemmas 2.2, 2.3 with any $-1 < \eta < 0$.

If $d \geqslant 3$, then we can take $\theta(r) = \frac{3}{2}$ provided r is sufficiently large. This leads to the assertion of Lemma 2.2 with $\eta = -1$. If ϕ satisfies (2.16), then, instead of (2.18), we have

$$E\left|\int_0^t \phi(\xi(s), s)\, ds - \Phi t\right| = O(1) + O(t^{1-\alpha/2}) \qquad \text{if} \quad \alpha \neq 2. \tag{2.30}$$

We sum up:

Lemma 2.4. (a) *Let the conditions* (2.2)–(2.4), (2.28) *hold, and let* $d \geqslant 2$. *Let* $\phi(x, t)$ *be a bounded measurable function. If* ϕ *satisfies* (2.16), *then* (2.18) *holds for any* $-1 < \eta < 0$, *and if* ϕ *satisfies* (2.17), *then* (2.19) *holds for any* $-1 < \eta < 0$.

(b) *Let the conditions* (2.2)–(2.4), (2.28) *hold and let* $d \geqslant 3$. *If* $\phi(x, t)$ *is a bounded measurable function satisfying* (2.16), *then* (2.30) *holds.*

3. Asymptotic estimates

Theorem 3.1. *Let the conditions* (2.2)–(2.5) *hold. Then*

$$E|\xi(t)|^2 \geqslant Kt - K' \qquad \textit{for all} \quad t \geqslant 0 \tag{3.1}$$

where K, K' *are positive constants.*

Proof.

$$L|x|^2 = \sum a_{ii}(x, t) + 2\sum x_i b_i(x, t) \geqslant \gamma - \phi(x)$$

where $\phi(x)$ is a bounded measurable and nonnegative function having compact support. By Itô's formula,

$$E|\xi(t)|^2 - E|\xi_0|^2 = E\int_0^t (L|x|^2)(\xi(s), s)\, ds \geqslant \gamma t - E\int_0^t \phi(\xi(s))\, ds.$$

Since, by Lemma 2.2,

$$E\int_0^t \phi(\xi(s))\, ds = o(t)$$

the inequality (3.1) follows.

Remark. Lemma 2.2 and Theorem 3.1 remain true if the condition (2.4) is replaced by the weaker condition

$$(1 + |x|)\sum_i |b_i(x, t)| \leqslant \epsilon_0 \tag{3.2}$$

provided ϵ_0 is a sufficiently small positive constant.

We shall need the following condition:

$$(1 + |x|)^{1+\delta} \sum_{i=1}^{n} |b_i(x, t)| \leqslant \epsilon(|x|), \quad \epsilon(r) \to 0 \quad \text{if} \quad r \to 0, \quad \delta \geqslant 0. \tag{3.3}$$

Theorem 3.2. *Let* (2.2), (2.3), (2.28), (3.3) *hold and let* $d \geqslant 1$. *Then*

$$E|\xi(t)|^2 = (\operatorname{tr} \bar{a})t + O(t^{(1+\eta)/2}) + o(t^{1-\delta/2}) \tag{3.4}$$

where $\operatorname{tr} \bar{a} = \sum_i \bar{a}_{ii}$ *and* η *is any positive number; if* $d \geqslant 2$, *then* η *is any number* > -1.

Proof. Notice that

$$|L|x|^2 - \operatorname{tr} \bar{a}| \leqslant \epsilon'(|x|)/(1 + |x|^\delta)$$

where $\epsilon'(r) \to 0$ if $r \to \infty$. Hence, by Lemmas 2.3 and 2.4(a) with $\phi(x, t) = L|x|^2$, $\Phi = \operatorname{tr} \bar{a}$,

$$\begin{aligned} E|\xi(t)|^2 - E|\xi_0|^2 &= E \int_0^t (L|x|^2)(\xi(s), s)\, ds \\ &= (\operatorname{tr} \bar{a})t + O(t^{(1+\eta)/2}) + o(t^{1-\delta/2}) \end{aligned}$$

where η is any positive number if $d \geqslant 1$ and any number > -1 if $d \geqslant 2$. This yields (3.4).

The special case $\delta = 0$ gives

Corollary 3.3. *Let* (2.2)–(2.4), (2.28) *hold and let* $d \geqslant 1$. *Then*

$$E|\xi(t)|^2 = 2(\operatorname{tr} \bar{a})t + o(t). \tag{3.5}$$

We can now state a key lemma needed for deriving the main results of this section.

Lemma 3.4. *Let the conditions* (2.2), (2.3), *and* (3.3) *hold with* $0 \leqslant \delta < 1$, *and let* $d \geqslant 2$. *Then, for any* $j = 1, 2, \ldots, n$,

$$E\left| \int_0^t b_j(\xi(s), s)\, ds \right|^2 = o(t^{1-\delta}). \tag{3.6}$$

Proof. For any $\epsilon > 0$ we can write

$$|b_j(x, t)| \leqslant g_1(x) + g_2(|x|)$$

where g_i are bounded measurable functions; $g_1(x)$ has compact support,

$$g_2(r) = \epsilon/(1 + r)^{1+\delta} \quad \text{if} \quad r > r_1, \tag{3.7}$$

and $g_2(r) = 0$ if $r \leqslant r_1$ for some $0 < r_1 < \infty$. Clearly

$$E\left| \int_0^t b_j(\xi(s), s)\, ds \right|^2 \leqslant 2E\left| \int_0^t g_1(\xi(s), s)\, ds \right|^2 + 2E\left| \int_0^t g_2(|\xi(s)|)\, ds \right|^2. \tag{3.8}$$

Let x^0 be a point outside the support of g_1 and let $\rho = |x - x^0|$.

Let $F(\rho)$ be the function constructed in the proof of Lemma 2.2 for $\phi = g_1$. Then, for any $-1 < \eta < 0$,

$$F(r) \leqslant cr^{1+\eta}, \qquad F'(r) \leqslant cr^{\eta} \qquad (c \text{ positive constant}) \tag{3.9}$$

provided r is sufficiently large. By Itô's formula,

$$\begin{aligned}\left|\int_0^t g_1(\xi(s))\,ds\right| &\leqslant \int_0^t (LF)(\xi(s), s)\,ds \\ &\leqslant F(|\xi(t) - x^0|) - \int_0^t D_x F \cdot \sigma(\xi(s), s)\,dw(s).\end{aligned}$$

Consequently

$$E\left|\int_0^t g_1(\xi(s))\,ds\right|^2 \leqslant Ct^{1+\eta} + C + 2E\int_0^t |D_x F \cdot \sigma|^2\,ds. \tag{3.10}$$

Next, for large r,

$$|D_x F \cdot \sigma|^2 \leqslant Cr^{2\eta} \qquad \text{for any} \quad -1 < \eta < 0. \tag{3.11}$$

Hence, by Lemma 2.4(a) with $\alpha = -2\eta$,

$$E\int_0^t |D_x F \cdot \sigma|^2\,ds \leqslant Ct^{1+\eta} + C, \tag{3.12}$$

and (3.10) then yields

$$E\left|\int_0^t g_1(\xi(s))\,ds\right|^2 \leqslant Ct^{1+\eta} + C. \tag{3.13}$$

To evaluate the integral corresponding to g_2, let $F_0(r)$ be the function $F(r)$ constructed in Lemma 2.3 when $\phi_2 = g_2$. We can take $-1 < \eta < -\delta$. Then

$$F_0(r) \leqslant C\epsilon r^{1-\delta} + C\epsilon, \qquad F_0'(r) \leqslant C\epsilon r^{-\delta} + C\epsilon \tag{3.14}$$

where C is a positive constant independent of ϵ. Analogously to (3.10), (3.11) we get

$$E\left|\int_0^t g_2(|\xi(s)|)\,ds\right|^2 \leqslant C\epsilon t^{1-\delta} + C + 2E\int_0^t |D_x F_0 \cdot \sigma|^2\,ds \tag{3.15}$$

$$|D_x F_0 \cdot \sigma| \leqslant C\epsilon^2/r^{2\delta}. \tag{3.16}$$

By Lemma 2.4(a),

$$E\int_0^t |D_x F_0 \cdot \sigma|^2\,ds \leqslant Ct^{(1+\eta)/2} + C\epsilon^2 t^{1-\delta} + C. \tag{3.17}$$

Hence

$$E\left|\int_0^t g_2(|\xi(s)|)\,ds\right|^2 \leqslant C\epsilon t^{1-\delta} + Ct^{(1+\eta)/2} + C. \tag{3.18}$$

Combining (3.18) with (3.13) and (3.8), and recalling that $\eta < -\delta$, the assertion (3.6) follows.

The following lemma is a variant of Lemma 3.4 in case $d \geqslant 3$.

Lemma 3.5. *Let the conditions* (2.2), (2.3), (2.28) *hold, let* $d \geqslant 3$ *and let* (3.3) *hold with some* $\delta > 1$. *Then, for any* $j = 1, \ldots, n$,

$$E\left|\int_0^t b_j(\xi(s), s)\,ds\right|^2 = O(1). \tag{3.19}$$

Proof. The function $F(r)$ occurring in (3.9) is now given by

$$F(r) = \int_{\rho_0}^r e^{-I(s)}\,ds \int_{\rho_0}^s e^{I(u)}\bar{\gamma}(u)\,du \qquad \text{if} \quad r > \rho_0,$$

where $\bar{\gamma}(r) = 0$ if $r > \rho_1$, for some $0 < \rho_0 < \rho_1 < \infty$. We can take $\eta(r) = \frac{3}{2}$ for large r, so that $I(r) \sim \log r^{3/2}$. Therefore

$$F(r) \leqslant C, \qquad F'(r) \leqslant C/r^{3/2}. \tag{3.20}$$

Hence (3.11) is replaced by

$$|D_xF \cdot \sigma|^2 \leqslant C/r^3. \tag{3.21}$$

Lemma 2.4(b) with $\alpha = 3$ gives

$$E\int_0^t |D_xF \cdot \sigma|^2\,ds = O(1), \tag{3.22}$$

and this estimate is to replace (3.12). We conclude that (instead of (3.13))

$$E\left|\int_0^t g_1(\xi(s))\,ds\right|^2 \leqslant C. \tag{3.23}$$

Let $0 < r_0 < r_1$ and let $\bar{g}_2(r)$ be a continuous function satisfying

$$\bar{g}_2(r) = \begin{cases} \epsilon/\,(1+r)^{1+\delta} & \text{if} \quad r > r_0, \\ 0 & \text{if} \quad r < r_0/2. \end{cases}$$

Take

$$F_0(r) = C\int_0^r e^{-I(s)} \int_0^s e^{I(u)}\bar{g}_2(u)\,du \qquad (C > 0).$$

Since $\delta > 1$, F_0 satisfies (instead of (3.14))

$$F_0(r) \leqslant C, \qquad F_0'(r) \leqslant C/r^k \qquad k = \min(\delta, \tfrac{3}{2}). \tag{3.24}$$

Using Itô's formula we find that

$$E\left|\int_0^t g_2(|\xi(s)|)\,ds\right|^2 \leqslant C + E\int_0^t |D_xF_0 \cdot \sigma|^2\,ds. \tag{3.25}$$

Since, by Lemma 2.4(b) with $\alpha = 2k$,

$$E\int_0^t |D_xF_0 \cdot \sigma|^2\,ds \leqslant C,$$

the right-hand side of (3.25) is bounded by a constant. Combining this with (3.23), (3.8), the assertion (3.19) follows.

Remark. If in Lemma 3.4 we replace the condition (3.3) by

$$\sum_i |b_i(x, t)| \leqslant C/(1 + |x|)^{1+\delta} \qquad (0 < \delta < 1), \tag{3.26}$$

then

$$E\left|\int_0^t b_j(\xi(s), s)\, ds\right|^2 = O(t^{1-\delta}). \tag{3.27}$$

We shall need the following conditions:

$$\sum_{i,j=1}^n |\sigma_{ij}(x, t) - \bar{\sigma}_{ij}| \leqslant \frac{\epsilon(|x|)}{(1 + |x|)^\delta}, \qquad \delta \geqslant 0, \quad \epsilon(r) \to 0 \quad \text{if} \quad r \to \infty, \tag{3.28}$$

where $\bar{\sigma}_{ij}$ are constants. Note that

$$\bar{a} = \bar{\sigma}\bar{\sigma}^*, \qquad \operatorname{tr} \bar{a} = |\bar{\sigma}|^2.$$

We shall also need the condition

$$\sum_{i,j=1}^n |\sigma_{ij}(x, t) - \bar{\sigma}_{ij}| \leqslant C/(1 + |x|)^\delta, \qquad \delta > 0. \tag{3.29}$$

We shall now state the main results of this section.

Theorem 3.6. (a) *Let* (2.2), (2.3), (3.3), (3.28) *hold with* $0 \leqslant \delta < 1$, *and let* $d \geqslant 2$. *Then*

$$E|\xi(t) - \bar{\sigma}w(t)|^2 = o(t^{1-\delta}). \tag{3.30}$$

If (3.3), (3.28) *are replaced by* (3.26), (3.29) *and* $0 < \delta < 1$, *then*

$$E|\xi(t) - \bar{\sigma}w(t)|^2 = O(t^{1-\delta}). \tag{3.31}$$

(b) *Let* (2.2), (2.3), (3.26), (3.29) *hold for some* $\delta > 1$, *and let* $d \geqslant 3$. *Then*

$$E|\xi(t) - \bar{\sigma}w(t)|^2 = O(1). \tag{3.32}$$

Proof. Let (2.2), (2.3), (3.3), (3.28) hold with $0 \leqslant \delta < 1$, and let $d \geqslant 2$. Consider the expression

$$E\left|\int_0^t (\sigma - \bar{\sigma})\, dw(s)\right|^2 = E\int_0^t |\sigma - \bar{\sigma}|^2\, ds.$$

Since $|\sigma(x, t) - \bar{\sigma}| \leqslant \epsilon^2(|x|)/(1 + |x|)^{2\delta}$, Lemma 2.4(a) implies that

$$E\int_0^t |\sigma(\xi(s), s) - \bar{\sigma}|^2\, ds = o(t^{1-\delta}).$$

Writing

$$\xi(t) - \bar{\sigma}w(t) = \xi_0 + \int_0^t [\sigma(\xi(s), s) - \bar{\sigma}]\, dw(s) + \int_0^t b(\xi(s), s)\, ds$$

and using the previous estimate and Lemma 3.4, the assertion (3.30) follows.

The proofs of (3.31), (3.32) are similar. In proving (3.31) we make use of the remark following Lemma 3.5. In proving (3.32) we employ Lemma 3.5.

4. Applications of the asymptotic estimates

We shall need the following conditions:

$$\lim_{|x|\to\infty} \sigma_{ij}(x, t) = \bar{\sigma}_{ij} \quad \text{uniformly with respect to} \quad t, \qquad 1 \leqslant i, j \leqslant n, \tag{4.1}$$

$$\begin{gathered} \sum |b_i(x, t)| \leqslant C \qquad \text{for all} \quad x \in R^n, \quad t \geqslant 0 \quad (C \text{ const}), \\ \lim_{|x|\to\infty} (1 + |x|) \sum_{i=1}^{n} |b_i(x, t)| = 0 \quad \text{uniformly with respect to } t. \end{gathered} \tag{4.2}$$

Theorem 3.6(a) with $\delta = 0$ states:

If (2.2), (2.3), (4.1), (4.2) hold and if $d \geqslant 2$, then

$$E|\xi(t) - \bar{\sigma}w(t)|^2 = o(t). \tag{4.3}$$

From (4.3) we see that, as $t\to\infty$,

$$\frac{\xi(t)}{\sqrt{t}} - \bar{\sigma}\frac{w(t)}{\sqrt{t}} \to 0 \qquad \text{in } L^2;$$

consequently also in probability. This immediately yields the following theorem on convergence in distribution of $\xi(t)$:

Theorem 4.1. *Let the conditions* (2.2), (2.3), (4.1), (4.2) *hold, let* $\bar{\sigma}$ *be nonsingular matrix, and let* $n \geqslant 2$. *Then*

$$\begin{aligned} &\lim_{t\to\infty} P\{\xi(t) < x\sqrt{t}\,\} \\ &= \frac{1}{(2\pi)^{n/2}\det\bar{\sigma}} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \exp\Big\{-\tfrac{1}{2}\sum \hat{a}_{ij}y_iy_j\Big\}\, dy_n \cdots dy_1 \end{aligned} \tag{4.4}$$

where $\hat{a}$ *is the inverse matrix to* $\bar{a}$.

Suppose next that (2.2), (2.3) hold and that

$$\sum_{i,j} |\sigma_{ij}(x, t) - \bar{\sigma}_{ij}| \leqslant C/(1 + |x|)^{\delta}, \tag{4.5}$$

$$\sum_{i} |b_i(x, t)| \leqslant C/(1 + |x|)^{1+\delta} \tag{4.6}$$

for some $0 < \delta < 1$. Suppose $\bar{\sigma}$ is nonsingular and $n \geqslant 2$. Denote by $\hat{\sigma}$ the inverse of $\bar{\sigma}$. Then

$$\begin{aligned} \frac{\hat{\sigma}\xi(t) - w(t)}{\sqrt{2t\log\log t}} &= \frac{\hat{\sigma}}{\sqrt{2t\log\log t}} \int_0^t b(\xi(s), s)\, ds \\ &+ \frac{\hat{\sigma}}{\sqrt{2t\log\log t}} \int_0^t [\sigma(\xi(s), s) - \bar{\sigma}]\, dw(s) \\ &+ \frac{\hat{\sigma}\xi_0}{\sqrt{2t\log\log t}} \equiv J_1(t) + J_2(t) + J_3(t) \equiv J(t). \end{aligned} \tag{4.7}$$

We shall denote various positive constants by the same symbol C. Let $t_m = m^\lambda$, $\lambda = 4/\delta$, m a positive integer. By the proof of Theorem 3.6(a),

$$P\left\{ \sup_{t_m \leqslant t \leqslant t_{m+1}} |J_1(t)| > \frac{1}{m} \right\} \leqslant P\left\{ \frac{C}{\sqrt{t_m}} \int_0^{t_{m+1}} |b(\xi(s), s)|\, ds > \frac{1}{m} \right\}$$

$$\leqslant \frac{Cm^2}{t_m} E\left[\int_0^{t_{m+1}} |b(\xi(s), s)|\, ds \right]^2$$

$$\leqslant \frac{Cm^2}{t_m} (t_{m+1})^{1-\delta} \leqslant \frac{C}{m^2}\, .$$

Next, by the martingale inequality and the proof of Theorem 3.6(a),

$$P\left\{ \sup_{t_m \leqslant t \leqslant t_{m+1}} |J_2(t)| > \frac{1}{m} \right\}$$

$$\leqslant P\left\{ \frac{C}{\sqrt{t_m}} \sup_{t_m \leqslant t \leqslant t_{m+1}} \left| \int_0^t [\sigma(\xi(s), s) - \bar{\sigma}]\, dw(s) \right| > \frac{1}{m} \right\}$$

$$\leqslant \frac{Cm^2}{t_m} (t_{m+1})^{1-\delta} \leqslant \frac{C}{m^2}\, .$$

Since, finally,

$$P\left\{ \sup_{t_m \leqslant t \leqslant t_{m+1}} |J_3(t)| > \frac{1}{m} \right\} \leqslant \frac{Cm^2}{t_m} E|\xi_0|^2 \leqslant \frac{C}{m^2}\, ,$$

we conclude that

$$P\left\{ \sup_{t_m \leqslant t \leqslant t_{m+1}} |J(t)| > \frac{3}{m} \right\} \leqslant \frac{C}{m^2}\, .$$

Applying the Borel–Cantelli lemma we deduce

$$P\left\{ \sup_{t_m \leqslant t \leqslant t_{m+1}} |J(t)| > \frac{3}{m} \text{ i.o.} \right\} = 0;$$

consequently

$$P\left\{ \lim_{t \to \infty} |J(t)| = 0 \right\} = 1. \tag{4.8}$$

From (4.7) and the law of the iterated logarithm for $w(t)$ (Theorem 3.3.1, Corollary 3.3.2 and Theorem 3.6.1) we obtain:

Theorem 4.2. *Let the conditions* (2.2), (2.3), (4.5), (4.6) *hold for some* $\delta > 0$, *let* $\bar{\sigma}$ *be nonsingular, and let* $n \geqslant 2$. *Then, for any* $i = 1, \ldots, n$,

$$\overline{\lim_{t \to \infty}} \frac{\sum_{j=1}^n \hat{\sigma}_{ij} \xi_j(t)}{\sqrt{2t \log \log t}} = 1 \quad a.s., \tag{4.9}$$

$$\underline{\lim_{t \to \infty}} \frac{\sum_{j=1}^n \hat{\sigma}_{ij} \xi_j(t)}{\sqrt{2t \log \log t}} = 1 \quad a.s. \tag{4.10}$$

Further,

$$\overline{\lim_{t\to\infty}} \frac{|\hat{\sigma}\xi(t)|}{\sqrt{2t \log\log t}} = 1 \quad a.s. \tag{4.11}$$

The next application is to the Cauchy problem:

$$\tfrac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i\, \partial x_j} + \sum_{i=1}^{n} b_i(x) \frac{\partial u}{\partial x_i} - \frac{\partial u}{\partial t} = 0 \qquad \text{if} \quad x \in R^n, \quad t > 0, \tag{4.12}$$

$$u(x, 0) = f(x) \qquad \text{if} \quad x \in R^n. \tag{4.13}$$

We shall assume:

(i) $\sigma_{ij}(x)$, $b_i(x)$ are bounded functions in R^n, uniformly Lipschitz continuous, and the matrix $(a_{ij}(x))$ is uniformly positive definite.

(ii) For all x, $\bar{x}$ in R^n,

$$|f(x) - f(\bar{x})| \leqslant C|x - \bar{x}|^\gamma \qquad (0 < \gamma \leqslant 2, \quad C > 0).$$

By Sections 6.4, 6.5, there is a unique solution $u(x, t)$ of (4.12), (4.13) bounded by $O(|x|^\gamma)$ uniformly in t in bounded intervals, and it is given by

$$u(x, t) = Ef(\xi_x(t)) \tag{4.14}$$

where $\xi_x(t)$ is the solution of (1.1) with $\xi(0) = x$.

Let us further assume that

$$\sum_{i,j} |\sigma_{ij}(x) - \delta_{ij}| \leqslant C/(1 + |x|)^\delta, \tag{4.15}$$

$$\sum_{i} |b_i(x)| \leqslant C/(1 + |x|)^{1+\delta} \tag{4.16}$$

for some $\delta > 0$.

If $a_{ij} = \delta_{ij}$ and $b_i \equiv 0$, then the solution $\tilde{u}$ is given either in the form $Ef(w(t) + x)$, or in terms of the fundamental solution for the heat equation, namely

$$\tilde{u}(x, t) = \frac{1}{(2\pi)^{n/2} t^{n/2}} \int_{R^n} \exp\left\{ \frac{|x - y|^2}{2t} \right\} f(y)\, dy. \tag{4.17}$$

From (4.14) and (ii) we get

$$|u(x, t) - \tilde{u}(x, t)| \leqslant C\{E|\xi_x(t) - w(t) - x|^2\}^{\gamma/2}.$$

We can now apply Theorems 3.6(a), 3.6(b) to estimate the right-hand side. A careful review of the proof of (3.31) and (3.32) for $\xi(t) = \xi_x(t)$ shows that

$$|O(t^{1-\delta})| \leqslant C[t^{1-\delta} + |x|^{2(1-\delta)} + 1], \qquad |O(1)| \leqslant C$$

where C is a constant independent of the initial condition $\xi(0) = x$. Hence:

$0 < \delta < 1$, *then for all* $x \in R^n$, $t \geqslant 0$,

$$|u(x, t) - \tilde{u}(x, t)| \leqslant C(1 + t + |x|^2)^{(1-\delta)\gamma/2} \qquad (C \text{ const}). \qquad (4.18)$$

If $n \geqslant 3$ *and* $\delta > 1$, *then for all* $x \in R^n$, $t \geqslant 0$,

$$|u(x, t) - \tilde{u}(x, t)| \leqslant C \qquad (C \text{ const}). \qquad (4.19)$$

5. The one-dimensional case

In proving Theorem 3.6(a) or even in deriving (4.3), we have assumed that the matrix $\bar{a}$ has at least two positive eigenvalues. To see what may happen when $\bar{a}$ has only one positive eigenvalue, we resort to the case $n = 1$. For simplicity, consider first the equation

$$d\xi(t) = dw(t) + b(\xi(t))\, dt. \qquad (5.1)$$

We assume that $b(x)$ is measurable and

$$\int_{-\infty}^{\infty} |b(x)|\, dx < \infty. \qquad (5.2)$$

One can construct comparison functions explicitly by solving differential equations of the form

$$Lf = \tfrac{1}{2} f''(x) + b(x) f'(x) = \psi(x). \qquad (5.3)$$

Setting $A(x) = 2\int_{-\infty}^{x} b(y)\, dy$, the general solution of (5.3) is given by

$$f(x) = \int_0^x e^{-A(z)}\, dz \left[2 \int_0^t e^{A(y)} \psi(y)\, dy + C_1 \right] + C_2$$

where C_1, C_2 are constants. Using this solution in the proof of Theorem 3.6, one can derive the estimate

$$E|\xi(t) - w(t)|^2 = o(t) \qquad (5.4)$$

provided $A(+\infty) = 0$, i.e., provided

$$\int_{-\infty}^{\infty} b(x)\, dx = 0; \qquad (5.5)$$

if further

$$|b(x)| \leqslant C/(1 + |x|)^{1+\delta} \qquad (0 < \delta < 1) \qquad (5.6)$$

then

$$E|\xi(t) - w(t)|^2 = O(t^{1-\delta}); \qquad (5.7)$$

the details are left to the reader.

Consider next the more general stochastic differential equation

$$d\xi(t) = \sigma(\xi(t))\, dt + b(\xi(t))\, dt. \qquad (5.8)$$

We assume that $\sigma(x) > 0$ for all x, and set

$$h(z) = \int_0^z \frac{dy}{\sigma(y)}, \qquad g \text{ inverse to } h.$$

As easily verified, the process $\eta(t) = h(\xi(t))$ satisfies the equation

$$d\eta(t) = dw(t) + \tilde{b}(\eta(t))\,dt \tag{5.9}$$

where

$$\tilde{b}(x) = \frac{b(g(x))}{\sigma(g(x))} - \tfrac{1}{2}\sigma'(g(x)). \tag{5.10}$$

The condition (5.5) for the equation (5.9) becomes

$$\int_{-\infty}^{\infty} \frac{b(x) - \frac{1}{2}\sigma(x)\sigma'(x)}{\sigma^2(x)}\,dx = 0. \tag{5.11}$$

If

$$\left| \frac{b(g(x))}{\sigma(g(x))} - \tfrac{1}{2}\sigma'(g(x)) \right| \leqslant \frac{C}{(1+|x|)^{1+\delta}} \qquad (\delta > 0) \tag{5.12}$$

and if (5.11) holds, then, by the assertion (5.7) applied to the solution $\eta(t)$ of (5.9),

$$E|\eta(t) - w(t)|^2 \leqslant Ct^{1-\delta} + C.$$

If we further assume that

$$|\bar{\sigma}h(x) - x| \leqslant C(1 + |x|^{1-\mu}) \qquad (0 \leqslant \mu < 1, \bar{\sigma} > 0) \tag{5.13}$$

for some constant $\bar{\sigma}$, then we easily conclude that

$$E|\xi(t) - \bar{\sigma}w(t)|^2 \leqslant Kt^{1-\nu} + K' \qquad \text{for all} \quad t \geqslant 0 \tag{5.14}$$

where $\nu = \min(\delta, \mu)$ and K, K' are positive constants.

Observe that if

$$|\sigma(x) - \bar{\sigma}| \leqslant \frac{C}{(1+|x|)^{\mu}} \tag{5.15}$$

then (5.13) holds and, further, (5.12) is equivalent to

$$\left| \frac{b(x)}{\sigma(x)} - \tfrac{1}{2}\sigma'(x) \right| \leqslant \frac{C}{(1+|x|)^{1+\delta}} \tag{5.16}$$

We can therefore state:

Theorem 5.1. *If* (5.15), (5.16), *and* (5.11) *hold, then for any solution* $\xi(t)$ *of* (5.8) *the estimate* (5.14) *holds with* $\nu = \min(\delta, \mu)$.

If (5.15) is replaced by

$$\bar{\sigma} = \lim_{|x|\to\infty} \sigma(x) \quad \text{exists}, \qquad \bar{\sigma} \neq 0,$$

if the left-hand side of (5.12) is in $L^1(-\infty, \infty)$, and if (5.11) holds, then, by applying the assertion (5.4) to the solution $\eta(t)$ of (5.9), we easily obtain

$$E|\xi(t) - \bar{\sigma}w(t)|^2 = o(t). \tag{5.17}$$

The condition (5.11) is essential for the validity of Theorem 5.1. Taking for simplicity the case $\sigma \equiv 1$, where the condition (5.11) reduces to the condition (5.5), we shall prove:

Theorem 5.2. *Let* $b(x)$ *satisfy* (5.2). *If* $\xi(t)$ *is a solution of* (5.1), *then the estimate* (5.4) *holds if and only if* (5.5) *holds.*

Proof. We only have to prove that if

$$\int_{-\infty}^{\infty} b(x)\, dx \neq 0 \tag{5.18}$$

then (5.4) does not hold. We shall assume that (5.18) and (5.4) hold, and derive a contradiction.

Clearly

$$E\xi(t) = E\xi_0 + E\int_0^t b(\xi(s))\, ds. \tag{5.19}$$

A particular solution of (5.3) for $\psi = b$ is

$$f(x) = \int_0^x e^{-A(z)}[e^{A(z)} - 1]\, dz.$$

Since $A(-\infty) = 0$ and (by (5.18)) $A(+\infty) \neq 0$,

$$f(x) = \lambda x^+ + o(|x|) \qquad \text{as} \quad |x|\to\infty \tag{5.20}$$

where λ is a nonvanishing constant.

By Itô's formula

$$E\int_0^t b(\xi(s))\, ds = Ef(\xi(t)) - Ef(\xi_0).$$

Substituting this into (5.19), we get

$$E\xi(t) = E\xi_0 - Ef(\xi_0) + Ef(\xi(t)). \tag{5.21}$$

Since $f(x)$ satisfies a uniform Lipschitz condition,

$$|Ef(\xi(t)) - Ef(w(t))| \leq CE|\xi(t) - w(t)| \qquad (C \text{ const}),$$

and, in view of (5.4),

$$Ef(\xi(t)) = Ef(w(t)) + o(t^{1/2}). \tag{5.22}$$

From (5.20) we have, for any $\epsilon > 0$,

$$|f(x) - \lambda x^+| \leqslant \epsilon|x| + C(\epsilon) \qquad (C(\epsilon) \text{ const}).$$

Combining this (when $x = w(t)$) with (5.22), we find that

$$|Ef(\xi(t)) - E[\lambda w^+(t)| \leqslant \epsilon E|w(t)| + C(\epsilon) + o(t^{1/2}) \leqslant 2\epsilon t^{1/2}$$

if t is sufficiently large. Upon substituting this into (5.21), we obtain

$$E\xi(t) = \lambda E w^+(t) + \theta\epsilon t^{1/2} + C' \qquad (C' \text{ const}, |\theta| \leqslant 2).$$

But, by (5.4),

$$E\xi(t) = Ew(t) + o(t^{1/2}) = o(t^{1/2}).$$

Consequently,

$$|\lambda| E w^+(t) \leqslant 3\epsilon t^{1/2} \qquad \text{if} \quad t \text{ is sufficiently large.}$$

Since, however, $Ew^+(t) = (t/2\pi)^{1/2}$, we get a contradiction if $18\pi\epsilon^2 < \lambda^2$.

There is an intuitive reason why for $n = 1$, $\sigma \equiv 1$ the assertion (5.4) cannot hold unless (5.5) is satisfied. In order for the distribution of $\xi(t)/\sqrt{t}$ to approximate the normal distribution as $t \to \infty$ the particles represented by $\xi(t)$, or $\xi(t, \omega)$, must be able to move without significant "resistance" from intervals (α, β) near $+\infty$ to intervals $(-\beta, -\alpha)$. Since in performing this move they must cross the interval $(-\alpha, \alpha)$, they are subject to the influence of the drift term $b(x)$. This drift coefficient will resist the motion if $\int_{-\infty}^{\infty} b(x)\, dx > 0$; thus this inequality cannot take place. Similarly, the reverse inequality cannot take place.

If $n \geqslant 2$ and $\bar{\sigma}$ is nonsingular, then $\xi(t)$ may move from one n-dimensional interval in a neighborhood G of infinity to another without leaving G. If the drift coefficient $b(x, t)$ is "small" in G (i.e., if $|b(x, t)| \leqslant C(1 + |x|)^{-1-\mu}$, $\mu > 0$), then there will be negligible resistance by the drift coefficient to the movement of $\xi(t)$ in G. Thus, no condition analogous to (5.5) is required in the case $n \geqslant 2$.

6. Counterexample

We shall give an example of a system of n equations $(n \geqslant 1)$

$$d\xi(t) = dw(t) + b(\xi(t))\, dt \tag{6.1}$$

for which

$$b(x) = O\left(\frac{\log |x|}{|x|} \right) \qquad \text{if} \quad |x| \to \infty \tag{6.2}$$

such that the estimate

$$E|\xi(t)|^2 \leqslant Kt + K' \qquad \text{for all} \quad t > 0 \qquad (K, K' \text{ positive constants}) \tag{6.3}$$

is false.

Let $f(x)$ be a function in $C^2(R^n)$ such that

$$f(x) = r^2/\log r \qquad \text{if} \quad r = |x| > 2. \tag{6.4}$$

If $b_i(x) = x_i\beta(r)$ for $|x| > 2$, then

$$\tfrac{1}{2}\,\Delta f(x) + \sum_{i=1}^{n} b_i(x)\,\frac{\partial f(x)}{\partial x_i}$$
$$= \frac{1}{\log r}\left[1 - \frac{m+2}{2}\,\frac{1}{\log r} + \frac{1}{(\log r)^2}\right] + \frac{2r^2}{\log r}\left[1 - \frac{1}{2\log r}\right]\beta(r).$$

Take

$$\beta(r) = \frac{\log r}{2r^2}\,[1 + \gamma(r)]$$

where $\gamma(r)$ is defined so that

$$\tfrac{1}{2}\,\Delta f(x) + \sum_{i=1}^{n} b_i(x)\,\frac{\partial f(x)}{\partial x_i} = 1 \qquad \text{if} \quad |x| > 2. \tag{6.5}$$

It is easily seen that $\gamma(r) = O(1)$. We now define $b_i(x)$ in R^n such that

$$b_i(x) = x_i[\log|x| + \gamma(|x|)]/2|x|^2 \qquad \text{if} \quad |x| > 2; \tag{6.6}$$

we can make the $b_i(x)$ as smooth in R^n as we wish.

By Itô's formula and (6.5),

$$Ef(\xi(t)) = Ef(\xi_0) + E\int_0^t c(\xi(s))\,ds \tag{6.7}$$

where $c(x) - 1 = 0$ if $|x| > 2$.

We wish to apply the proof of Lemma 2.2 to $\phi(x, t) = c(x) - 1$. Here the $b_i(x, t) = b_i(x)$ are not bounded by $\epsilon(|x|)/(1 + |x|)$ where $\epsilon(r) \to 0$ if $r \to \infty$. Nevertheless (2.11) takes the form

$$\tfrac{1}{2}[1 + \theta(r)] \leqslant \tfrac{1}{2}n + \sum_i (x_i - e_i)b_i(x) \tag{6.8}$$

where $r = |x - e|$ and $e = (e_1, \ldots, e_n)$ is a point outside the support of $c(x) - 1$, i.e., $|e| > 2$. In view of (6.6) and the boundedness of $\gamma(r)$ as $r \to \infty$, we can actually construct a continuous function $\theta(r)$ satisfying (6.8) such that, for any $A > 0$, $\lim_{r\to\infty}\theta(r) = A$. Hence, by the proof of Lemma 2.2,

$$E\int_0^t [c(\xi(s)) - 1]\,ds = O(1) \qquad \text{as} \quad t \to \infty.$$

Combining this with (6.7) we get

$$E|f(\xi(t)) - t| \leqslant C_0 \qquad \text{for} \quad t \geqslant 0 \qquad (C_0 \text{ const}). \tag{6.9}$$

For any $\epsilon > 0$ there is a constant B such that

$$|x|^2/\log|x| < \epsilon|x|^2 + B \qquad \text{if} \quad |x| > 2.$$

Hence

$$f(x) \leqslant \epsilon|x|^2 + C \qquad \text{for all} \quad x \in R^n,$$

where C is a constant depending on ϵ. This implies that

$$Ef(\xi(t)) \leqslant \epsilon E|\xi(t)|^2 + C. \tag{6.10}$$

Now, if (6.3) holds, then from (6.9), (6.10) we obtain

$$t \leqslant \epsilon Kt + \epsilon K' + C + C_0 \qquad \text{for all} \quad t \geqslant 0.$$

But this is impossible if $\epsilon < 1/K$.

If the function $\log |x|$ occurring in (6.2) is replaced by other functions that increase to infinity more mildly, as $|x| \to \infty$, such as $\log \log |x|$, then one can again show by the above method that (6.3) cannot hold. If however $b(x) = O(1/|x|)$, then (6.3) does hold (by Theorem 2.1). Recall that if (3.2) holds, then the estimate (3.1) is also valid.

PROBLEMS

1. All the results of Sections 2, 3 remain valid if the condition (2.4) is replaced by the condition that

$$(1 + |x|) \sum_i |b_i(x, t)| \leqslant \epsilon \qquad \text{for all} \quad x \in R^n, \quad t \geqslant 0$$

provided ϵ is sufficiently small. Check this in the case of Theorem 3.6(a).

2. The methods of Sections 2, 3 can be used to estimate $E|\xi(t) - \hat{\sigma}w(t)|^4$, provided $0 < \delta < \frac{1}{2}$. Show that under the assumptions (2.2), (2.3), (3.26), (3.29) with $0 < \delta < \frac{1}{2}$,

$$E|\xi(t) - \bar{\sigma}w(t)|^4 = O(t^{2(1-\delta)}).$$

3. Prove that if (5.2) and (5.5) hold, then any solution of (5.1) satisfies the estimate (5.4).

4. Prove that if (5.6) and (5.5) hold, then the estimate (5.7) is valid for the solution of (5.1).

5. If

$$\sum a_{ii}(x, t) \leqslant C(1 + |x|^2)^{\mu}, \qquad |\sum x_i b_i(x, t)| \leqslant C(1 + |x|^2)^{\mu}$$

for some constants $C > 0$, $0 < \mu < 1$, then

$$E|\xi(t)|^{2-2\mu} \leqslant Kt + K' \qquad \text{for all} \quad t \geqslant 0,$$

where K, K' are positive constants.

6. Let $\tilde{b}(x, t) = b(x, t)/(1 + |x|^2)^{\mu}$, $\tilde{\sigma}(x, t) = \sigma(x, t)/(1 + |x|^2)^{\mu/2}$, $\tilde{a} = \tilde{\sigma}\tilde{\sigma}^*$. Let the assumptions of Lemma 2.2 hold for a, b replaced by $\tilde{a}$, $\tilde{b}$. Let $\phi(x)$ be any bounded measurable function with compact support. Prove that if $\xi(t)$ is a solution of (1.1), then

$$E\left| \int_0^t \phi(\xi(s))\, ds \right| \leqslant K_1(1 + t^{(1+\eta)/(2-2\mu)}) \qquad \text{for all} \quad t \geqslant 0,$$

where K_1 is a constant and η is as in Lemma 2.2. [*Hint*: Construct $f(x) = F(r)$ satisfying $Lf(x) \geqslant (1 + |x|^2)^\eta \psi(|x|)$.]

7. Let (2.2)–(2.5) hold with a, b replaced by $\tilde{a}$, $\tilde{b}$ (as in Problem 6), and let $0 \leqslant \mu < 1$. Prove that

$$E|\xi(t)|^{2-2\mu} \geqslant Kt - K' \qquad \text{for all} \quad t \geqslant 0,$$

where K, K' are positive constants.

8. Extend Theorem 5.2 to the case of Eq. (5.8).

9. Let $\xi(t)$ be a solution of (1.1) and assume that $|\sigma(x, t)| \leqslant C$, $|b(x, t)| \leqslant C$, $|b(x, t) - \bar{b}| \to 0$ if $t \to \infty$, uniformly with respect to x. Prove that

$$\lim_{t\to\infty} \frac{\xi(t)}{t} = \bar{b} \quad \text{a.s.}$$

10. Consider one-dimensional equations

$$d\xi_i(t) = dw(t) + b_i(\xi_i(t), t) \qquad (i = 1, 2).$$

Assume that $b_1(x, t) < b_2(x, t)$ for all $x \in R^1$, $t \geqslant 0$. Prove that if $\xi_1(0) = \xi_2(0)$ then $\xi_1(t) < \xi_2(t)$ for all $t > 0$.

11. Let $\xi(t)$ be a solution of a one-dimensional equation

$$d\xi(t) = dw(t) + b(\xi(t), t)\, dt, \qquad \xi(0) = x$$

and assume that

$$b(x, t) \geqslant \beta(t) \qquad \text{for all} \quad x \in R^1,$$

$$\lim_{t\to\infty} \frac{1}{\sqrt{2t \log\log t}} \int_0^t \beta(s)\, ds > 1.$$

Prove that $\lim_{t\to\infty} \xi(t) = \infty$ a.s.

12. Consider a stochastic differential equation

$$d\xi(t) = \sigma(\xi(t))\, dw(t) + b(\xi(t))\, dt$$

and assume that $\sigma(x) > 0$ in the interval $a \leqslant x \leqslant b$. Denote by $\tau_x(a, b)$ the exit time from the interval (a, b), given $\xi(0) = x$. Denote by $p_a(x; b)$ the probability that $\xi(\tau_x(a, b)) = a$. Prove:

(a) If $\frac{1}{2}\sigma(x)u''(x) + b(x)u'(x) = -1$ in $a \leqslant x \leqslant b$, and $u(a) = u(b) = 0$ then $E\tau_x(a, b) = u(x)$.

(b) If $\Phi(x) = \exp\{-\int_a^x (2b(z)/\sigma^2(z))\, dz\}$, then

$$u(x) = -\int_a^x 2\Phi(y) \int_a^y \frac{dz}{\sigma^2(z)\Phi(z)}\, dy + \gamma \int_a^b 2\Phi(y) \int_a^y \frac{dz}{\sigma^2(z)\Phi(z)}\, dy,$$

$$\gamma = \int_a^x \Phi(z)\, dz \Big/ \int_a^b \Phi(z)\, dz.$$

(c) If $\frac{1}{2}\sigma^2(x)w''(x) + b(x)w'(x) = 0$ in $a \leqslant x \leqslant b$, then

$$p_a(x; b) = [w(x) - w(b)]/[w(a) - w(b)].$$

13. If (2.2), (2.3), (3.3) hold with $\delta > 0$, and if (2.5) is replaced by $\lambda_n(x, t) \geqslant c/(1 + |x|)^\delta$ $(c > 0)$, then the assertion (2.6) of Lemma 2.2 is valid for any $\eta > 0$.

14. If (2.2), (2.3), (2.5), and (3.3) hold with $\delta > 0$, then (2.6) holds with $\eta = 0$. [*Hint*: Take $\theta(r) = -D/r^\beta$ for large r, where β is any positive constant $\leqslant \delta$ and D is a positive constant.]

15. Suppose, in Lemma 2.3, ϕ satisfies (2.16) and the condition (2.4) is replaced by the stricter condition (3.3) with $\delta > 0$. Prove that

$$E\left|\int_0^t \phi(\xi(s), s)\, ds - \Phi t\right| = O(t^{1/2}) + O(t^{1-(\alpha-\beta)/2})$$

for any $\beta > 0$. [*Hint*: Take $1 - \rho = D/(2r^\beta)$, $\theta(r) = -D/r^\beta$. In (2.14) replace $\psi(u)$ by $\psi(u)u^\beta$; this gives (2.24) with α replaced by $\alpha - \beta$.]

16. Under the same conditions as in Problem 13, if ϕ satisfies (2.16), then

$$E\left|\int_0^t \phi(\xi(s), s)\, ds - \Phi t\right| = O(t^{(1+\eta)/2}) + O(t^{1-(\alpha-\delta)/2})$$

for any $\eta > 0$. [*Hint*: Replace, in (2.14), $\psi(u)$ by $\psi(u)u^\delta$.]

17. Under the assumptions of Theorem 3.2,

$$E[\langle e, \xi(t)\rangle^2] = \langle e, \bar{a}e\rangle + O(t^{(1+\eta)/2}) + o(t^{1-\delta/2})$$

where e is any unit vector and $\langle\ ,\ \rangle$ denotes the scalar product.

9

Recurrent and Transient Solutions

Consider a stochastic differential system of n equations

$$d\xi(t) = \sigma(\xi(t), t)\, dw(t) + b(\xi(t), t)\, dt \tag{0.1}$$

Let G be an open set in R^n. Suppose that for any $x \in G$ and for any open subset V of G,

$$P_x\{\xi(t_m) \in V \quad \text{for a sequence of finite random times } t_m \text{ increasing to } \infty\} = 1. \tag{0.2}$$

Then we say that the solution of (0.1) is a *recurrent process* in G. If $G = R^n$, then we say simply that the solution of (0.1) is a recurrent process.

If, for any $x \in G$,

$$P_x\{\lim_{t\to\infty} |\xi(t)| = \infty\} = 1, \tag{0.3}$$

then we say that the solution of (0.1) is a *transient process* in G. If $G = R^n$, we simply say that the solution of (0.1) is a transient process.

The condition (0.2) says that, given $\xi(0) = x$, the solution $\xi(t)$ "visits" any open subset V of G at a sequence of times increasing to infinity. The condition (0.3) says that, given $\xi(0) = x$, $\xi(t)$ "wanders out to infinity" with probability one.

It is well known (see Itô and McKean [1]) that an n-dimensional Brownian motion is recurrent if $n \leqslant 2$, and transient if $n \geqslant 3$. This fact will follow as a very special case of the results of this chapter.

1. Transient solutions

For simplicity we shall consider only temporally homogeneous processes, i.e., solutions of systems of n stochastic differential equations with time-independent coefficients,

$$d\xi(t) = \sigma(\xi(t))\, dw(t) + b(\xi(t))\, dt. \tag{1.1}$$

Set

$$\sigma = (\sigma_{ij})_{i,j=1}^{n}, \qquad b = (b_1, \ldots, b_n), \qquad a_{ij} = \sum_{k=1}^{n} \sigma_{ik}\sigma_{jk},$$

$$Lu = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \, \partial x_j} + \sum_{i=1}^{n} b_i(x) \frac{\partial u}{\partial x_i} . \tag{1.2}$$

We shall assume:

(A_1) For all $x \in R^n$,

$$\sum_{i=1}^{n} |b_i(x)| + \sum_{i,j=1}^{n} |\sigma_{ij}(x)| \leqslant C(1 + |x|) \qquad (C \text{ const});$$

for any $R > 0$ there is a positive constant C_R such that

$$\sum_{i=1}^{n} |b_i(x) - b_i(y)| + \sum_{i,j=1}^{n} |\sigma_{ij}(x) - \sigma_{ij}(y)| \leqslant C_R|x - y|$$

if $|x| < R$, $|y| < R$.

(A_2) The matrix $(a_{ij}(x))$ is positive definite for each $x \in R^n$.

Let

$$A(x, \xi) = \frac{1}{|\xi|^2} \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j,$$

$$B(x) = \sum_{i=1}^{n} a_{ii}(x),$$

$$C(x, \xi) = 2 \sum_{i=1}^{n} \xi_i b_i(x),$$

and set

$$S(x, \xi) = \frac{B(x) - A(x, \xi) + C(x, \xi)}{A(x, \xi)}, \qquad S(x) = S(x, x). \tag{1.3}$$

We shall need the assumption:

(A_3) There is a positive constant R_0 such that

$$S(x) \geqslant 1 + \epsilon(|x|) \qquad \text{if} \quad |x| \geqslant R_0, \tag{1.4}$$

where $\epsilon(r)$ is a continuous function satisfying:

$$\int_{R_0}^{\infty} \frac{1}{t} \exp\left[- \int_{R_0}^{t} \frac{\epsilon(s)}{s} \, ds \right] dt < \infty. \tag{1.5}$$

Notice that (1.5) holds for any of the functions

$$\epsilon(s) = d, \qquad \epsilon(s) = A/s, \qquad \epsilon(s) = B/\log s$$

where d, A, B are positive constants and $B > 1$.

Theorem 1.1. *Let* (A_1)–(A_3) *hold. Then the solution of* (1.1) *is transient, i.e., for any* $x \in R^n$,

$$P_x\{\lim_{t\to\infty} |\xi(t)| = \infty\} = 1. \tag{1.6}$$

Proof. Let $\alpha > 0$. By (A_3), there is a continuous function $\theta(r)$ defined for $r \geqslant \alpha$ such that

$$S(x) \geqslant \theta(|x|) \quad \text{if} \quad |x| \geqslant \alpha, \tag{1.7}$$

$$\theta(r) = 1 + \epsilon(r) \quad \text{if} \quad r \geqslant R_0. \tag{1.8}$$

We shall construct a function $f(x) = F(r)$, where $r = |x|$, such that $Lf(x) \leqslant 0$ if $|x| \geqslant \alpha$. As easily verified,

$$2Lf(x) = A(x, x)F''(r) + \frac{F'(r)}{r}[B(x) - A(x, x) + C(x, x)]. \tag{1.9}$$

Hence, if

$$F'(r) \leqslant 0 \quad \text{for} \quad r \geqslant \alpha, \tag{1.10}$$

$$F''(r) + \frac{\theta(r)}{r} F'(r) = 0 \quad \text{for} \quad r \geqslant \alpha, \tag{1.11}$$

then, by (1.7)–(1.9),

$$Lf(x) \leqslant 0 \quad \text{if} \quad |x| \geqslant \alpha. \tag{1.12}$$

Set

$$I(r) = \int_\alpha^r \frac{\theta(s)}{s}\, ds. \tag{1.13}$$

Then a solution of (1.11) is given by

$$F(r) = \int_r^\infty e^{-I(s)}\, ds; \tag{1.14}$$

the integral is convergent, by (1.5), (1.8). Notice that (1.10) is also satisfied. Hence (1.12) holds when F is given by (1.14).

By Itô's formula and (1.12), if $|x| > \alpha$, then

$$E_x F(|\xi(\tau)|) - E_x F(|x|) = E_x \int_0^\tau Lf(\xi(s))\, ds \leqslant 0 \tag{1.15}$$

where τ is any bounded stopping time such that $|\xi(s)| \geqslant \alpha$ if $0 \leqslant s \leqslant \tau$. Let $\beta > \alpha$, and let $\tau_{\alpha\beta}$ denote the exit time from the shell $\{y;\ \alpha < |y| < \beta\}$. Denote by $P_x(\alpha)$ the probability that $|\xi(\tau_{\alpha\beta})| = \alpha$ (given $\xi(0) = x$) and by

$P_x(\beta)$ the probability that $|\xi(\tau_{\alpha\beta})| = \beta$ (given $\xi(0) = x$). Setting $\tau = T \wedge \tau_{\alpha\beta}$ in (1.15) and taking $T \to \infty$, we get (since $\tau_{\alpha\beta} < \infty$ a.s.; see Theorem 8.1.1),

$$F(\alpha)P_x(\alpha) + F(\beta)P_x(\beta) \leqslant F(|x|).$$

Taking $\beta \to \infty$ and using the fact that $F(\beta) \to 0$ if $\beta \to \infty$, we get

$$\lim_{\beta \to \infty} P_x(\alpha) \leqslant \frac{F(|x|)}{F(\alpha)} . \tag{1.16}$$

Introduce the balls

$$B_\rho = \{ y; |y| \leqslant \rho \} \qquad (0 < \rho < \infty)$$

and the event

$$\Omega(\alpha) = \{\xi(t) \text{ hits the ball } B_\alpha \text{ for some } t \geqslant 0\}. \tag{1.17}$$

Then we can write (1.16) in the form:

$$P_x(\Omega(\alpha)) \leqslant \frac{F(|x|)}{F(\alpha)} . \tag{1.18}$$

Denote by t_R the hitting time of the boundary of the ball B_R by $\xi(t)$. By Theorem 8.1.1, if $|x| < R$, then $P_x\{t_R < \infty\} = 1$. Introduce the event

$$\Omega^*(\alpha) = \{\xi(t) \text{ hits the ball } B_\alpha \text{ at a sequence of times increasing to } \infty\}. \tag{1.19}$$

Thus $\Omega^*(\alpha)$ is a subset of $\Omega(\alpha)$.

Let $\alpha < |x| < R$. Since $P_x\{t_R < \infty\} = 1$,

$$P_x(\Omega^*(\alpha)) = P_x\{\xi(t + t_R) \text{ hits } B_\alpha \text{ at a sequence of times increasing to } \infty\}.$$

Using the strong Markov property (see Problem 3) we get

$$P_x(\Omega^*(\alpha)) = E_x P_{\xi(t_R)}(\Omega^*(\alpha)) \leqslant E_x P_{\xi(t_R)}(\Omega(\alpha)) \leqslant E_x \frac{F(R)}{F(\alpha)} = \frac{F(R)}{F(\alpha)} ,$$

where (1.18) has been used. Taking $R \to \infty$, we get $P_x(\Omega^*(\alpha)) = 0$. This means that

$$P_x\Big\{ \varliminf_{t \to \infty} |\xi(t)| \geqslant \alpha \Big\} = 1.$$

Since α is arbitrary, we get

$$P_x\Big\{ \lim_{t \to \infty} |\xi(t)| = \infty \Big\} = 1,$$

i.e., (1.6) holds.

We shall now replace the condition (A_3) by:

(A_3') As $|x|\to\infty$,

$$a_{ij}(x)\to a_{ij}^0, \tag{1.20}$$

$$\sum_{i=1}^{n} x_i b_i(x)\to 0, \tag{1.21}$$

where the matrix (a_{ij}^0) has at least three positive eigenvalues.

Theorem 1.2. *Let* (A_1), (A_2), (A_3') *hold. Then the solution of* (1.1) *is transient.*

Proof. We can perform an orthogonal transformation $x\to x'$ in R^n which takes (a_{ij}^0) into (a_{ij}^*), where $a_{ij}^* = 0$ if $i \neq j$, $a_{ii}^* = 1$ if $i = 1, 2, 3$ and $a_{ii}^* = 0$ or 1 if $4 \leqslant i \leqslant n$. In the new coordinates the condition (A_3) holds with $\epsilon(s) = d$ where d is any positive constant < 1. Now apply Theorem 1.1.

2. Recurrent solutions

We shall replace the condition (A_3) by:

(A_4) For any $z\in R^n$ there is a positive constant R_z such that

$$S(x, x - z) \leqslant 1 + \epsilon(|x - z|) \qquad \text{if} \quad |x - z| \geqslant R_z, \tag{2.1}$$

where $\epsilon(r)$ is a continuous function satisfying

$$\int_{R^*}^{\infty} \frac{1}{t} \exp\left[-\int_{R^*}^{t} \frac{\epsilon(s)}{s}\, ds\right] dt = \infty \qquad \text{for some} \quad R^* > 0. \tag{2.2}$$

For simplicity we have taken $\epsilon(r)$ to be independent of z; but the subsequent results are unaffected if $\epsilon(r)$ is allowed to depend on z.

Notice that the function $\epsilon(r) = 1/(\log r)$ satisfies (2.2).

Theorem 2.1. *Let* (A_1), (A_2), (A_4) *hold. Then the solution of* (1.1) *is recurrent, i.e., for any* $x\in R^n$ *and for any ball* $B_\alpha(z) = \{y; |y - z| \leqslant \alpha\}$, $\alpha > 0$

$$P_x\{\xi(t) \text{ hits } B_\alpha(z) \text{ at a sequence of times increasing to } \infty\} = 1. \tag{2.3}$$

Proof. We take, for simplicity, $z = 0$ and write $B_\alpha = B_\alpha(0)$. We shall first construct a function $f(x) = F(r)$ for $r = |x| \geqslant \alpha$ such that

$$Lf(x) \geqslant 0 \qquad \text{if} \quad |x| \geqslant \alpha. \tag{2.4}$$

Let $\theta(r)$ be a continuous function such that

$$S(x) \leqslant \theta(|x|) \quad \text{if} \quad |x| \geqslant \alpha, \tag{2.5}$$

$$\theta(r) = 1 + \epsilon(r) \quad \text{if} \quad r \geqslant R_0. \tag{2.6}$$

In view of (1.9), if $F(r)$ satisfies (1.10), (1.11), then (2.4) holds. With the definition (1.13), the function

$$F(r) = -\int_\alpha^r e^{-I(s)}\, ds \tag{2.7}$$

satisfies both (1.11) and (1.10). In view of (2.6), (2.2),

$$F(r) \to -\infty \quad \text{if} \quad r \to \infty. \tag{2.8}$$

We shall now apply the equality in (1.15) to the present function $F(r)$. Making use of (2.4) and taking $\tau = \tau_{\alpha\beta} \wedge T$, we get, after letting $T \to \infty$,

$$F(\alpha)P_x(\alpha) + F(\beta)P_x(\beta) - F(|x|) \geqslant 0. \tag{2.9}$$

Taking $\beta \to \infty$ in (2.9) and using (2.8), we conclude that $P_x(\beta) \to 0$ if $\beta \to \infty$. Hence, $P_x(\alpha) = 1 - P_x(\beta) \to 1$ if $\beta \to \infty$. This means that

$$P_x(\Omega(\alpha)) = 1 \tag{2.10}$$

where $\Omega(\alpha)$ is defined in (1.17).

For any $\rho > 0$, let $\partial B_\rho = \{y; |y| = \rho\}$. Let

$$\alpha < R_1 < R_2 < \cdots < R_m < \cdots, \qquad R_m \to \infty \quad \text{if} \quad m \to \infty.$$

Introduce Markov times:

τ_1 = first time $\xi(t)$ hits B_α;

σ_1 = first time $> \tau_1$ such that $\xi(t)$ hits ∂B_{R_1};

in general,

τ_m = first time $> \sigma_{m-1}$ such that $\xi(t)$ hits B_α;

σ_m = first time $> \tau_m$ such that $\xi(t)$ hits ∂B_{R_m}.

By Theorem 8.1.1, on the set where $\tau_m < \infty$ also $\sigma_m < \infty$. By (2.10), $P_x(\tau_1 < \infty) = 1$. Hence $P_x(\sigma_1 < \infty) = 1$, and by the strong Markov property (see Problem 4),

$$P_x(\tau_2 < \infty) = E_x E_{\xi(\sigma_1)} \chi_{\tau_1 < \infty} = 1. \tag{2.11}$$

where (2.10) has been used in the last equality.

We now proceed by induction. Assuming that $\tau_m < \infty$ a.s., we get

$$P_x(\tau_{m+1} < \infty) = E_x E_{\xi(\sigma_m)} \chi_{\tau_1 < \infty} = 1.$$

Now, at each time $t = \tau_m$, $\xi(t)$ hits B_α. Further, since $|\xi(\sigma_m)| = R_m \to \infty$ as $m \to \infty$, $\lim_{m\to\infty} \sigma_m = \infty$; hence also $\lim_{m\to\infty} \tau_m = \infty$. This completes the proof of (2.3) (in case $z = 0$).

We shall now replace the condition (A_4) by:

(A_4') As $|x|\to\infty$,

$$a_{ij}(x) - a_{ij}^0 = o\left(\frac{1}{\log|x|}\right), \tag{2.12}$$

$$\sum|b_i(x)| = o\left(\frac{1}{|x|\log|x|}\right), \tag{2.13}$$

and the matrix (a_{ij}^0) has precisely two positive eigenvalues.

Theorem 2.2. *Let* (A_1), (A_2), (A_4') *hold. Then, the solution of* (1.1) *is recurrent.*

Proof. We perform an orthogonal transformation $x\to x'$ that takes (a_{ij}^0) into a new matrix (a_{ij}^*) with $a_{ij}^* = 0$ if $i \neq j$ or if $i = j \geqslant 3$, and $a_{ii}^* = 1$ if $i = 1, 2$. In the new coordinates, the condition (A_4) is satisfied with $\epsilon(r) = 1/(\log r)$. Now apply Theorem 2.1.

Remark. Suppose (A_4') is replaced by

$$a_{ij}(x)\to a_{ij}^0 \quad \text{as} \quad |x|\to\infty, \qquad \sum|b_i(x)| = o\left(\frac{1}{|x|}\right) \quad \text{as} \quad |x|\to\infty$$

where the matrix (a_{ij}^0) has precisely one positive eigenvalue. Then the assertion of Theorem 2.1 remains valid, with the same proof; here $\epsilon(r) = -d$ where d is any positive constant < 1.

Example. Consider the case where $n \geqslant 2$, $b_i \equiv 0$, and σ is such that

$$\sigma\sigma^* = (a_{ij}), \qquad a_{ij} = \delta_{ij} + \frac{g(r)}{r^2}x_ix_j \qquad (r = |x|);$$

$g(r)$ is a Lipschitz continuous function vanishing near $r = 0$, and

$$\mu \leqslant g(r) \leqslant M \quad \text{where} \quad \mu > -1, \quad M < \infty \qquad (\mu, M \text{ const}).$$

The eigenvalues of $(a_{ij}(x))$ are 1 (with multiplicity $n - 1$) and $1 + g$. Hence $(a_{ij}(x))$ is positive definite for all $x\in R^n$. Clearly,

$$S(x) - 1 = \frac{n - 2 - g(r)}{1 + g(r)} \equiv \epsilon(r).$$

Hence, if $g(r)$ is such that $\epsilon(r) = A/(\log r)$ for some $A > 1$ (and all large r), then the assertion of Theorem 1.1 holds. If $g(r)$ is such that $\epsilon(r) = 1/(\log r)$ for all large r, then (A_4) holds with $\epsilon(s) = 1/(\log s) + C/s$ for some positive constant C; consequently the assertion of Theorem 2.1 holds. This example shows that conditions (A_3), (A_4) made in Theorems 1.1, 2.1 are rather sharp.

This example also shows that the behavior asserted in Theorems 1.1 and 2.1 does not depend exclusively on the dimension n. In fact, given any $\epsilon(r)$ which converges to 0 as $r \to \infty$, take

$$g = \frac{n-2-\epsilon}{1+\epsilon} \qquad \text{(for all large } r\text{)}$$

in the above example. Then the behavior of $\xi(t)$ does not depend on n; if $\epsilon(r) \geqslant A/\log r$ for some $A > 1$, then $\xi(t)$ is transient, whereas if $\epsilon(r) \leqslant 1/\log r$ then $\xi(t)$ is recurrent.

3. Rate of wandering out to infinity

In this section we return to the situation of Theorem 1.1. We shall assume:

(A_5) $a_{ij}(x)$, $b_i(x)$ are bounded functions in R^n, the $a_{ij}(x)$ are uniformly Hölder continuous in R^n, and, for all $x \in R^n$, $\xi \in R^n$,

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geqslant \alpha_0|\xi|^2 \qquad (\alpha_0 \text{ positive constant}).$$

We shall also assume that the function $\epsilon(r)$ occurring in the condition (A_3) satisfies, for some r_0 sufficiently large,

$$\epsilon(r) = d \qquad \text{if} \quad r \geqslant r_0 \qquad (d \text{ positive constant}). \tag{3.1}$$

Theorem 3.1. *Let* (A_1), (A_5) *hold and let* (A_3) *hold with* $\epsilon(r)$ *satisfying* (3.1). *Then, for any* $0 < \theta < \frac{1}{2}$, $x \in R^n$,

$$P_x\left\{ \frac{|\xi(t)|}{t^\theta} \to \infty \text{ if } t \to \infty \right\} = 1. \tag{3.2}$$

We first prove two lemmas.

Lemma 3.2. *Let* (A_1), (A_5), (A_3), *and* (3.1) *hold. Then there exists a positive constant C such that for any* $\alpha > r_0$, $x \in R^n$,

$$P_x\{|\xi(t)| \leqslant \alpha \text{ for some } t \geqslant 0\} \leqslant C\left(\frac{\alpha}{|x|}\right)^d. \tag{3.3}$$

Proof. Let $R > \alpha$. Consider the Dirichlet problem:

$$Lu_R(x) = 0 \qquad \text{if} \quad \alpha < |x| < R,$$
$$u_R = 1 \qquad \text{if} \quad |x| = \alpha, \qquad u_R = 0 \qquad \text{if} \quad |x| = R.$$

Let $f(x) = F(r)$ $(r = |x|)$ be the function constructed in the proof of Theorem 1.1. Then

$$v(x) = \frac{F(r)}{F(\alpha)} \qquad (r = |x|)$$

satisfies:

$$\begin{aligned} Lv \leqslant 0 \quad &\text{if} \quad \alpha < |x| < R, \\ v = 1 \quad &\text{if} \quad |x| = \alpha, \qquad v \geqslant 0 \quad \text{if} \quad |x| = R. \end{aligned}$$

Hence, by the maximum principle,

$$0 \leqslant u_R(x) \leqslant v(x). \tag{3.4}$$

By Itô's formula, if $\alpha < |x| < R$,

$$u_R(x) = P_x\{\xi(t) \text{ hits } \partial B_\alpha \text{ before it hits } \partial B_R\}. \tag{3.5}$$

Hence $u_R(x) \uparrow u(x)$ as $R \uparrow \infty$, where

$$u(x) = P_x\{\xi(t) \text{ hits } \partial B_\alpha \text{ for some } t \geqslant 0\}. \tag{3.6}$$

From (3.4) we get

$$0 \leqslant u(x) \leqslant v(x) = \frac{\int_r^\infty \exp[-I(s)]\,ds}{\int_\alpha^\infty \exp[-I(s)]\,ds}\,.$$

Using (3.1) to estimate the right-hand side, we obtain

$$0 \leqslant u(x) \leqslant C\frac{\alpha^d}{r^d} \tag{3.7}$$

where C is a constant independent of α, x.

Now, if $|x| \leqslant \alpha$, then the assertion (3.3) is trivially true (with $C = 1$). If, on the other hand, $|x| > \alpha$, then the assertion (3.3) follows from (3.6), (3.7).

Lemma 3.3. *Let* (A_1), (A_5), *and* (A_3), (3.1) *hold, with* $d < n$. *Then there is a positive constant* C' *such that, if* $\alpha > r_0$, $|x| < \alpha/4$, $T > 0$,

$$P_x\{|\xi(t)| \leqslant \alpha \textit{ for some } t \geqslant T\} \leqslant C'\left(\frac{\alpha}{T^{1/2}}\right)^d. \tag{3.8}$$

Proof. By the Markov property and Lemma 3.2,

$$\begin{aligned} P_x\{|\xi(t)| \leqslant \alpha \text{ for some } t \geqslant T\} &= E_x P_{\xi(T)}\{|\xi(t)| \leqslant \alpha \text{ for some } t \geqslant 0\} \\ &= P_x\{|\xi(T)| \leqslant \alpha\} + E_x\left\{\chi_{|\xi(T)|>\alpha}\left[C\frac{\alpha^d}{|\xi(T)|^d}\right]\right\} \\ &\equiv I + J. \end{aligned} \tag{3.9}$$

Denote by $\Gamma(x, t, y)$ the fundamental solution of the parabolic operator

$L - \partial/\partial t$. By Theorem 6.4.5 and Problem 7, Chapter 6,

$$0 \leqslant \Gamma(x, t, y) \leqslant \frac{M}{t^{n/2}} \exp\left[-\frac{\mu|x-y|^2}{t}\right] \tag{3.10}$$

where M, μ are positive constants.

We can write

$$I = \int_{|y| \leqslant \alpha} \Gamma(x, T, y)\, dy,$$

$$J = C\alpha^d \int_{|y| > \alpha} \frac{1}{|y|^d} \Gamma(x, T, y)\, dy.$$

We shall subsequently denote various positive constants by the same symbol C. Substituting $|y - x| = \rho\sqrt{T}$ in the integral I and noting that

$$\rho\sqrt{T} = |y - x| \leqslant 2\alpha \qquad (\text{since } |y| \leqslant \alpha,\ |x| < \alpha),$$

we get

$$I \leqslant C\int_{\rho\sqrt{T} < 2\alpha} \rho^{n-1} e^{-\mu\rho^2}\, d\rho \leqslant C\left(\frac{\alpha}{\sqrt{T}}\right)^n.$$

Substituting $|y - x| = \rho\sqrt{T}$ in the integral J and noting that

$$\rho\sqrt{T} = |y - x| \geqslant \alpha/2 \qquad (\text{since } |y| \geqslant \alpha,\ |x| \leqslant \alpha/2),$$

$$|y| \geqslant |y - x| - |x| \geqslant \rho\sqrt{T}/2 \qquad (\text{since } |x| \leqslant \alpha/4 \leqslant |y - x|/2),$$

we get

$$J \leqslant C\alpha^d \int_{\rho\sqrt{T} > \alpha/2} \frac{\rho^{n-1}}{\left(\rho\sqrt{T}\right)^d} e^{-\mu\rho^2} d\rho$$

$$\leqslant C\left(\frac{\alpha}{\sqrt{T}}\right)^d \int_{\rho > 0} \rho^{n-1-d} e^{-\mu\rho^2}\, d\rho \leqslant C\left(\frac{\alpha}{\sqrt{T}}\right)^d,$$

since $n - 1 - d > -1$. Substituting the estimates for I, J into (3.9), the assertion (3.8) follows.

Proof of Theorem 3.1. Without loss of generality we may assume that $d < n$. We apply Lemma 3.3 with $T = 2^m$, $\alpha = 2^{(m+1)\theta}$ where m is a positive integer such that $|x| < 2^{(m+1)\theta}/4$. We get

$$P_x\{|\xi(t)| \leqslant t^\theta \text{ for some } t,\ 2^m \leqslant t \leqslant 2^{m+1}\}$$

$$\leqslant P_x\{|\xi(t)| \leqslant 2^{(m+1)\theta} \text{ for some } t \geqslant 2^m\} \leqslant C\left[\frac{2^{(m+1)\theta}}{2^{m/2}}\right]^d \leqslant C2^{m(\theta-1/2)d}.$$

Since $\sum 2^{m(\theta - 1/2)d} < \infty$, the Borel–Cantelli lemma implies that, with probability 1, the sequence of events

$$\{|\xi(t)| \leqslant t^{\theta} \text{ for some } t,\ 2^m \leqslant t \leqslant 2^{m+1}\}$$

(where $\xi(0) = x$) occurs only finitely often. Hence

$$P_x\{|\xi(t)| > t^{\theta} \text{ for all } t \text{ sufficiently large}\} = 1.$$

Since this is true also for θ replaced by any θ', $\theta < \theta' < \frac{1}{2}$, the assertion of the theorem follows.

Corollary 3.4. *Let* (A_1), (A_5), *and* (A_3') *hold. Then, for any* $0 < \theta < \frac{1}{2}$, $x \in R^n$, *the assertion* (3.2) *holds.*

Indeed, perform an orthogonal transformation as in the proof of Theorem 1.2. In the new coordinates the conditions (A_3), (3.1) hold.

Remark 1. By Theorem 3.6.1,

$$P_x\left\{ \overline{\lim_{t\to\infty}} \frac{|w(t)|}{\sqrt{t \log\log t}} = 1 \right\} = 1.$$

More generally, under the conditions of Theorem 8.4.2,

$$P_x\left\{ \overline{\lim_{t\to\infty}} \frac{|\hat{\sigma}\xi(t)|}{\sqrt{t \log\log t}} = 1 \right\} = 1 \tag{3.11}$$

where $\hat{\sigma}$ is the inverse of $\lim_{|x|\to\infty}\sigma(x)$. From (3.11) it follows that

$$P_x\left\{ \lim_{t\to\infty} \frac{|\xi(t)|}{t^{\eta}} = 0 \right\} = 1 \quad \text{if} \quad \eta > \tfrac{1}{2}.$$

Thus the assertion (3.2) (for any $\theta < \frac{1}{2}$) is rather sharp. However, this assertion can still be strengthened as follows.

Let $g(t)$ be a positive and monotone decreasing function for $t > 0$, satisfying

$$\sum_{m=1}^{\infty} (g(2^m))^d < \infty \tag{3.12}$$

where d is any constant $< n$ such that (3.1) holds. If in the proof of Theorem 3.1 we replace t^{θ} by $t^{1/2}g(t)$, then we conclude that the events

$$\{|\xi(t)| < t^{1/2}g(t) \text{ for some } t,\ 2^m \leqslant t \leqslant 2^{m+1}\}$$

occur only finitely often. Consequently

$$P_x\left\{ \underline{\lim_{t\to\infty}} \frac{|\xi(t)|}{t^{1/2}g(t)} \geqslant 1 \right\} = 1. \tag{3.13}$$

It can be shown (see Dvoretsky and Erdös [1]) that if the function $g(t)$ satisfies

$$\sum_{m=1}^{\infty} (g(2^m))^d = \infty \qquad (d = n-2, n \geqslant 3), \tag{3.14}$$

instead of (3.12), then

$$P_x\left\{ \overline{\lim_{t\to\infty}} \frac{|w(t)|}{t^{1/2}g(t)} \leqslant 1 \right\} = 1. \tag{3.15}$$

***Remark* 2.** Consider the example at the end of Section 2. If $n = 2$ and $g = -1/(1 + 1/d)$ $(d > 0)$, then the assertion of Theorem 3.1 holds. Thus, even when $n = 2$, a diffusion process $\xi(t)$ may wander out to ∞ at a rate $\geqslant t^\theta$, for any $0 < \theta < \frac{1}{2}$.

4. Obstacles

We shall maintain the condition (A_1) but relax the condition (A_2).

Let G be a closed bounded domain with C^3 connected boundary ∂G, and let $\hat{G} = R^n \backslash G$.

Suppose the diffusion matrix $(a_{ij}(x))$ degenerates on ∂G in such a way that

$$\sum_{i,j=1}^{n} a_{ij}(x)\nu_i\nu_j = 0 \qquad \text{if} \quad x \in \partial G, \tag{4.1}$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is the outward normal at x to ∂G. This condition is equivalent to

$$\sum_{k=1}^{n}\left(\sum_{i=1}^{n} \sigma_{ik}(x)\nu_i\right)^2 = 0,$$

i.e.,

$$\sum_{i=1}^{n} \sigma_{ik}(x)\nu_i = 0 \qquad \text{if} \quad x \in \partial G, \quad 1 \leqslant k \leqslant n. \tag{4.2}$$

The expression $\sum a_{ij}(x)\nu_i\nu_j$ is called the *normal diffusion* to ∂G at x.

Let $\rho(x) = \text{dist}\,(x, \partial G)$ for $x \in \hat{G} \cup \partial G$, and let

$$\hat{G}_\epsilon = \{x \in \hat{G};\ \rho(x) \leqslant \epsilon\}.$$

Since ∂G is in C^3, $\rho(x)$ is in $C^2(\hat{G}_{\epsilon_0} \cup \partial G)$ for some $\epsilon_0 > 0$ sufficiently small. Noting that

$$\nu_i(x) = \partial\rho(x)/\partial x_i \qquad x \in \partial G, \quad 1 \leqslant i \leqslant n,$$

we deduce from (4.2) that

$$\sum_{i=1}^{n} \sigma_{ik}(x) \frac{\partial \rho(x)}{\partial x_i} = O(\rho(x)) \qquad \text{as} \quad \rho(x) \to 0, \quad x \in \hat{G}_{\epsilon_0}.$$

Taking the squares, we get

$$\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial \rho(x)}{\partial x_i} \frac{\partial \rho(x)}{\partial x_j} \leqslant C(\rho(x))^2 \qquad \text{if} \quad x \in \hat{G}_{\epsilon_0} \tag{4.3}$$

where C is a positive constant.

Let (4.1) hold and suppose $\sigma_{ij} \in C^1(\hat{G}_{\epsilon_0})$. By (4.2), the vector $T_k = (\sigma_{1k}, \ldots, \sigma_{nk})$ is tangent to ∂G. Since the function $\sum \sigma_{jk}\, \partial\rho/\partial x_j$ vanishes on ∂G, its derivative with respect to the tangent vector of T_k also vanishes on ∂G. Consequently

$$\sum_k \sum_j \sigma_{jk} \frac{\partial}{\partial x_j} \sum_i \sigma_{ik} \frac{\partial \rho}{\partial x_i} = 0.$$

This gives, after using (4.2) once more,

$$\sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 \rho}{\partial x_i\, \partial x_j} = - \sum_{i,j=1}^{n} \frac{\partial a_{ij}}{\partial x_j} \nu_i \qquad \text{on } \partial G. \tag{4.4}$$

The vector with components

$$b_i - \frac{1}{2} \sum_{j=1}^{n} \frac{\partial a_{ij}}{\partial x_j}$$

is called the *Fichera drift*. We shall impose the condition:

$$\sum_{i=1}^{n} b_i \nu_i + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 \rho}{\partial x_i\, \partial x_j} \geqslant 0 \qquad \text{on } \partial G. \tag{4.5}$$

If $\sigma_{ij} \in C^1(\hat{G}_{\epsilon_0})$ then, by (4.4), this condition is equivalent to

$$\sum_{i=1}^{n} \left(b_i - \frac{1}{2} \sum_{j=1}^{n} \frac{\partial a_{ij}}{\partial x_j} \right) \nu_i \geqslant 0 \qquad \text{on } \partial G, \tag{4.6}$$

i.e., the Fichera drift at ∂G points into the exterior of G.

Notice that (4.5) implies that

$$\sum_{i=1}^{n} b_i \frac{\partial \rho}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 \rho}{\partial x_i\, \partial x_j} \geqslant -c\rho \qquad \text{in } \hat{G}_{\epsilon_0} \tag{4.7}$$

where c is a positive constant.

Theorem 4.1. *Let* (A_1) *and* (4.1), (4.5) *hold. Then*

$$P_x\{\xi(t) \in \hat{G} \text{ for all } t > 0\} = 1 \qquad \text{if} \quad x \in \hat{G}. \tag{4.8}$$

Proof. Let $R(x)$ be a $C^2(\hat{G})$ function such that

$$R(x) = \begin{cases} \rho(x) & \text{if } x \in \hat{G}_{\epsilon_1} \quad \text{(for some } 0 < \epsilon_1 < \epsilon_0), \\ |x| & \text{if } |x| > M, \end{cases}$$

and $\epsilon_1 < R(x) < M$ elsewhere; M is chosen so large that $\hat{G}_{\epsilon_0} \subset \{x;\ |x| < M\}$. (To construct $R(x)$, let

$$\hat{R}(x) = \begin{cases} \rho(x) & \text{if } x \in \hat{G}_{\epsilon_0}, \\ M & \text{if } x \in R^n \backslash \hat{G}_{\epsilon_0} \end{cases}$$

and take $R(x)$ to be a mollifier of $\hat{R}(x)$; see Problem 4, Chapter 4.)

Let $V(x) = 1/(R(x))^\epsilon$ for some $\epsilon > 0$. We have

$$LV = -\epsilon R^{-\epsilon-1} \sum_i b_i \frac{\partial R}{\partial x_i}$$

$$+ \frac{1}{2} \sum_{i,j} a_{ij} \left\{ \epsilon(\epsilon+1) R^{-\epsilon-2} \frac{\partial R}{\partial x_i} \frac{\partial R}{\partial x_j} - \epsilon R^{-\epsilon-1} \frac{\partial^2 R}{\partial x_i \partial x_j} \right\}$$

$$= V \left\{ -\frac{\epsilon}{R} \sum_i b_i \frac{\partial R}{\partial x_i} + \frac{1}{2} \sum_{i,j} \frac{a_{ij}}{R^2} \left[\epsilon(\epsilon+1) \frac{\partial R}{\partial x_i} \frac{\partial R}{\partial x_j} - \epsilon R \frac{\partial^2 R}{\partial x_i \partial x_j} \right] \right\}. \tag{4.9}$$

In view of (4.3) and (4.7), the right-hand side of (4.9) is bounded by μV if $x \in \hat{G}_{\epsilon_1}$, where μ is a positive constant. In view of the condition (A_1), the same estimate is valid also if $|x| > M$. It follows that

$$LV \leqslant \mu V \qquad \text{for all} \quad x \in \hat{G}. \tag{4.10}$$

Further,

$$V(x) \to \infty \qquad \text{if} \quad \rho(x) \to 0, \quad x \in \hat{G}. \tag{4.11}$$

We shall now use (4.10), (4.11) in order to prove (4.8).

Denote by τ_p the hitting time of the set $\hat{G}_{1/p}$ $(p = 1, 2, \ldots)$. By Itô's formula,

$$e^{-\mu(\tau_p \wedge T)} V(\xi(\tau_p \wedge T)) = V(x) + \int_0^{\tau_p \wedge T} e^{-\mu s} V_x(\xi(s)) \sigma(\xi(s))\, dw(s)$$

$$+ \int_0^{\tau_p \wedge T} (LV - \mu V)(\xi(s)) e^{-\mu s}\, ds. \tag{4.12}$$

Taking the expectation and using (4.10), we get

$$E_x\left[e^{-\mu(\tau_p \wedge T)} \chi_{\rho(\xi(\tau_p \wedge T)) = 1/p} \right] p^\epsilon \leqslant V(x),$$

and, as $T \uparrow \infty$,

$$E_x e^{-\mu \tau_p} \chi_{\tau_p < \infty} \leqslant p^{-\epsilon} V(x). \tag{4.13}$$

Now, if (4.8) is not true then the exit time τ from $\hat{G}$ is finite on a set B of positive probability. Since $\tau_p \uparrow \tau$ on B, the monotone convergence theorem gives

$$E_x e^{-\mu\tau} \chi_B \leqslant \lim_{p\to\infty} [p^{-\epsilon} V(x)] = 0.$$

Since $\tau < \infty$ on B, we must then have $P_x(B) = 0$, but this contradicts the definition of B.

Theorem 4.1 motivates the following definition.

Definition. If the conditions (4.1), (4.5) are satisfied, then we say that ∂G is an *obstacle from the outside*.

Suppose (4.5) is replaced by

$$\sum_{i=1}^{n} b_i \nu_i + \frac{1}{2} \sum a_{ij} \frac{\partial^2 \rho}{\partial x_i\, \partial x_j} \leqslant 0 \qquad \text{on } \partial G \tag{4.14}$$

with the same function $\rho(x)$ as before. Define $\hat{\rho}(x) = \text{dist}\,(x, \partial G)$ for $x \in G$, and denote by $\hat{\nu} = (\hat{\nu}_1, \ldots, \hat{\nu}_n)$ the inward normal to ∂G. Then (4.14) holds if and only if

$$\sum_{i=1}^{n} b_i \hat{\nu}_i + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 \hat{\rho}}{\partial x_i\, \partial x_j} \geqslant 0 \qquad \text{on } \partial G \tag{4.15}$$

where the left-hand side is evaluated by taking the limit as x tends to the boundary from inside G. Using (4.1), (4.15), we can now duplicate the proof of Theorem 4.1, taking $R(x)$ as a $C^2(G)$ function which coincides with $\hat{\rho}(x)$ near ∂G, and taking $V(x) = 1/(R(x))^\epsilon$. We thus conclude:

Corollary 4.2. *If* (A_1) *and* (4.1), (4.14) *hold, then*

$$P_x\{\xi(t) \in \text{int}\, G \textit{ for all } t > 0\} = 1 \qquad \textit{if} \quad x \in \text{int}\, G. \tag{4.16}$$

Definitions. If the conditions (4.1), (4.14) are satisfied, then we say that ∂G is an *obstacle from the inside*. If (4.1) holds and if

$$\sum_{i=1}^{n} b_i \nu_i + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 \rho}{\partial x_i\, \partial x_j} = 0 \qquad \text{on } \partial G, \tag{4.17}$$

then we say that ∂G is a *two-sided obstacle*. In case (4.17) is replaced by either (4.5) or (4.14), we speak of a *one-sided obstacle*.

We shall now consider the case of degeneracy at one point z, and assume:

$$a_{ij}(z) = 0, \quad b_i(z) = 0 \qquad \text{for} \quad 1 \leqslant i, j \leqslant n. \tag{4.18}$$

Notice that $a_{ij}(z) = 0$ for $1 \leqslant i, j \leqslant n$ if and only if $\sigma_{ij}(z) = 0$ for $1 \leqslant i, j \leqslant n$.

Using (4.18) one can show that

$$P_x\{\xi(t) \neq z \quad \text{for all} \quad t > 0\} = 1 \qquad \text{if} \quad x \neq z. \tag{4.19}$$

Indeed, the proof is similar to the proof of Theorem 4.1. Here we take $V(x) = 1/(R(x))^\epsilon$ where $R(x) = |x - z|$.

A point z for which (4.18) holds is called a *point obstacle* or a *point of total degeneracy*.

We shall extend the previous considerations to the case where there are several obstacles.

Let $G_1, \ldots, G_k$ be mutually disjoint sets in R^n; if $1 \leqslant j \leqslant k_0$, G_j consists of one point z_j, and, if $k_0 + 1 \leqslant j \leqslant k$, G_j is a closed bounded domain with C^3 connected boundary ∂G_j. Let

$$G = \bigcup_{j=1}^{k} G_j, \qquad \hat{G} = R^n \backslash G$$

$$\rho_j(x) = \text{dist}\,(x, \partial G_j) \qquad \text{if} \quad x \not\in \text{int}\, G_j.$$

We shall assume:

$$a_{ij}(z_h) = 0, \quad b_i(z_h) = 0 \qquad \text{if} \quad 1 \leqslant i, j \leqslant n, \quad 1 \leqslant h \leqslant k_0, \tag{4.20}$$

$$\sum_{i,j=1}^{n} a_{ij}\nu_i\nu_j = 0 \qquad \text{on} \quad \partial G_h, \qquad \text{for} \quad k_0 + 1 \leqslant h \leqslant k, \tag{4.21}$$

$$\sum_{i=1}^{n} b_i\nu_i + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\frac{\partial^2\rho_h}{\partial x_i\,\partial x_j} \geqslant 0 \qquad \text{on} \quad \partial G_h, \qquad \text{for} \quad k_0 + 1 \leqslant h \leqslant k. \tag{4.22}$$

Theorem 4.3. *If* (A_1) *and* (4.20)–(4.22) *hold, then*

$$P_x\{\xi(t) \in \hat{G} \text{ for all } t \geqslant 0\} = 1 \qquad \text{if} \quad x \in \hat{G}. \tag{4.23}$$

The proof is similar to the proof of Theorem 4.1. Here one takes $V = 1/R^\epsilon$ where $R(x)$ is a C^2 function in $\hat{G}$, satisfying:

$$R(x) = \begin{cases} \rho_h(x) & \text{if} \quad \rho_h(x) < \epsilon_1 \quad (1 \leqslant h \leqslant k), \\ |x| & \text{if} \quad |x| > M \end{cases}$$

and $\epsilon_1 \leqslant R(x) \leqslant M$ elsewhere, where ϵ_1 is sufficiently small and M is sufficiently large.

In the following two sections we shall replace the condition (A_2) by the weaker condition:

(A_2') The matrix $(a_{ij}(x))$ is positive definite for all $x \in \hat{G}$.

We shall also assume that (4.20)–(4.22) hold and

$$\partial G_h \text{ is } C^3 \text{ diffeomorphic to a sphere, for } k_0 + 1 \leqslant h \leqslant k. \tag{4.24}$$

We shall then extend the results of Sections 1, 2. The following lemma will be needed.

Lemma 4.4. *Let* (A_1), (A_2'), *and* (4.24) *hold. Then there exists a continuous function* $R(x)$ *in* R^n *having the following properties*:

(i) $R(x)$ *is in* $C^2(\hat{G})$;

(ii) $R(x) > 0$ *in* $\hat{G}$;

(iii) $R(x) = \rho_h(x)$ *if* $\rho_h(x) \leq \epsilon_0$, $R(x) > \epsilon_0$ *if* $\min_h \rho_h(x) > \epsilon_0$, *for some* ϵ_0 *sufficiently small*;

(iv) $R(x) \to \infty$ *if* $|x| \to \infty$;

(v) $\sum_{i,j=1}^{n} a_{ij}(x) \dfrac{\partial^2 R(x)}{\partial x_i \, \partial x_j} < 0$ *if* $x \in \hat{G}$, $D_x R(x) = 0$; *there exist precisely* $k - 1$ *points* x *in* $\hat{G}$ *where* $D_x R(x) = 0$.

The number ϵ_0 will be smaller than

$$\min_{i \neq j} \text{dist}\,(G_i, G_j).$$

Notice that the conditions (4.20)–(4.22) are not assumed in this lemma.

Proof. By the proof of the Schoenflies theorem (see Morse [1]) there is a diffeomorphism $y = f(x)$ of the exterior of $\bigcup_{j=1}^{k} G_j$ onto the exterior of $\bigcup_{j=1}^{k} G_j'$ in R^n, where $G_1', \ldots, G_{k_0}'$ are points situated on the y_1 axis and $G_{k_0+1}', \ldots, G_k'$ are balls with centers on the y_1 axis; the center of G_j' lies to the left of the center of G_{j+1}'. Furthermore, this diffeomorphism preserves the distance functions (to $\bigcup_j G_j$ and to $\bigcup_j G_j'$) as long as the distance is sufficiently small. (The condition (4.24) is needed in order to apply the Schoenflies theorem.) Suppose for simplicity that $k_0 = 0$, $k = 2$. Denote by $(c_1, 0, \ldots, 0)$ the midpoint of the segment connecting the center $(\alpha_1, 0, \ldots, 0)$ of G_1' to the center $(\alpha_2, 0, \ldots, 0)$ of G_2'; it is assumed that $\alpha_1 < \alpha_2$. Construct a positive C^2 function $\phi(y')$ (where $y' = (y_2, \ldots, y_n)$) on the plane $y_1 = c_1$, which increases radially, with $\text{grad}\,\phi(y') \neq 0$ if $y' \neq 0$, such that $\partial\phi/\partial y_i = 0$, $\partial^2\phi/\partial y_i \, \partial y_j = 0$ $(2 \leq i, j \leq n)$ at $y' = 0$.

Denote by μ_i the radius of G_i'. Construct a C^2 function $\psi(y_1)$, positive for $\alpha_1 + \mu_1 < y_1 < \alpha_2 - \mu_2$, such that

$$\psi(y_1) = \begin{cases} y_1 - \alpha_1 - \mu_1 & \text{for} \quad 0 \leq y_1 - \alpha_1 - \mu_1 < \delta_1, \\ \alpha_2 - \mu_2 - y_1 & \text{for} \quad 0 \leq \alpha_2 - \mu_2 - y_1 < \delta_1 \end{cases}$$

where δ_1 is sufficiently small, and such that $\psi'(y_1) \neq 0$ if $y_1 \neq c_1$ and

$$\psi(c_1) = \phi(0, \ldots, 0), \qquad \psi'(c_1) = 0, \qquad \psi''(c_1) < 0.$$

We now construct a C^2 positive function $\lambda(y)$ for $y \not\in (G_1' \cup G_2')$, which extends the functions ϕ, ψ and the distance function from $G_1' \cup G_2'$ (as long as the distance is sufficiently small). This function is to satisfy:

$$\operatorname{grad}\lambda(y) \neq 0 \quad \text{if} \quad y \neq (c_1, 0, \dots, 0);$$

$$\operatorname{grad}\lambda(y) = 0, \quad \frac{\partial^2\lambda(y)}{\partial y_i\, \partial y_j}, \quad = 0 \quad \text{if} \quad (i, j) \neq (1, 1),$$

$$\frac{\partial^2\lambda(y)}{\partial y_1^2} < 0 \quad \text{at} \quad y = (c_1, 0, \dots, 0);$$

$$\lambda(y) = |y| \quad \text{if} \quad |y| \text{ is sufficiently large.}$$

The construction of such a function $\lambda(y)$ can be accomplished by introducing a family of curves $\gamma_{y'}$ connecting $(\alpha_1, 0, \dots, 0)$ to $(\alpha_2, 0, \dots, 0)$ and intersecting the plane $y_1 = c_1$ orthogonally at (c_1, y'). $\lambda(y)$ is defined along $\gamma_{y'}$ such that its tangential derivative vanishes only at $y_1 = c_1$.

Define $R(x) = \lambda(f(x))$. Clearly $D_x R(x) \neq 0$ if $x \neq x^*$ where $f(x^*)$ is the point $(c_1, 0, \dots, 0)$. Furthermore, as easily seen,

$$\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 R}{\partial x_i\, \partial x_j} < 0 \quad \text{at} \quad x = x^*.$$

This completes the proof if $k_0 = 0$, $k = 2$. The proof for any k_0, k is similar.

Lemma 4.5. *Let* (A_1), (A_2') *hold and suppose that*

$$G_h \textit{ is } C^3 \textit{ diffeomorphic to a closed ball, for } k_0 + 1 \leqslant h \leqslant k. \quad (4.25)$$

Then there exists a function $R(x)$ *satisfying* (i)–(v) *of Lemma* 4.4 *and, in addition,*

(iv′) $R(x) = x$ *if* $|x|$ *is sufficiently large.*

For proof see Problem 8. By Milnor [1], (4.24) implies (4.25) if $n \geqslant 5$; the same is true, by conformal mappings, if $n = 2$.

5. Transient solutions for degenerate diffusion

We shall extend the results of Section 1 to the case the diffusion matrix $(a_{ij}(x))$ degenerates in such a way that (A_2') holds, and G_j, G are as in Lemma 4.4.

For later references we state the following condition:

(B_1) (i) $G_h = \{z_h\}$ if $h = 1, \dots, k_0$, G_h is a bounded closed domain with C^3 connected boundary if $h = k_0 + 1, \dots, k$.

(ii) The conditions (4.20)–(4.22) and (4.24) hold.

(iii) The matrix $(a_{ij}(x))$ is nondegenerate for $x \in \hat{G}$, where $\hat{G} = R^n \setminus G$, $G = \bigcup_{h=1}^{k} G_h$.

Let $R(x)$ be the function constructed in Lemma 4.4 or 4.5, and set

$$\mathcal{A} = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial R}{\partial x_i} \frac{\partial R}{\partial x_j},$$

$$\mathcal{B} = \sum_{i=1}^{n} b_i(x) \frac{\partial R}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 R}{\partial x_i \, \partial x_j},$$

$$Q = \frac{1}{R}\left(\mathcal{B} - \frac{\mathcal{A}}{R}\right).$$

Then, if $g(x) = \Phi(R(x))$,

$$Lg(x) = \mathcal{A}\left[\Phi''(R) + \frac{1}{R}\Phi'(R)\right] + RQ\Phi'(R). \tag{5.1}$$

If $R(x)$ is as in Lemma 4.5, then the function $S(x)$ defined in (1.3) satisfies

$$S(x) = 1 + \frac{R^2(x)Q(x)}{\mathcal{A}(x)} \qquad \text{if} \quad |x| \text{ is sufficiently large.}$$

As will be proved in Chapter 11, if $\overline{\lim}\ Q(x) < 0$ as $R(x) \to 0$ and as $R(x) \to \infty$, then

$$P_x\left\{ \min_{1 \le h \le k} [\rho_h(\xi(t))] \to 0 \text{ if } t \to \infty \right\} = 1 \quad \text{if} \quad x \in \hat{G}.$$

Thus, in this case, $\xi(t)$ neither wanders out to ∞ nor visits any open neighborhood in $\hat{G}$ at a sequence of times increasing to ∞. Hence, in order to obtain the type of behavior asserted in Sections 1, 2, we shall have to impose different conditions on $Q(x)$.

As will be shown, in order to generalize the results of Sections 1, 2 to the present case where (B_1) holds, we do not need to change the conditions (A_3), (A_4) near infinity. We only need to impose a condition on $Q(x)$ near $R(x) = 0$. This condition is:

(B_2) For some $0 < \delta_0 \le \epsilon_0$ there is a continuous function $\epsilon(r)$, defined for $0 < r \le \delta_0$, such that

$$Q(x) \ge \frac{\mathcal{A}(x)}{R^2(x)} \epsilon(R(x)) \qquad \text{if} \quad 0 < R(x) \le \delta_0, \tag{5.2}$$

and

$$\int_r^{\delta_0} \frac{1}{s} \exp\left[\int_s^{\delta_0} \frac{\epsilon(t)}{t}\, dt\right] ds \to \infty \qquad \text{if} \quad r \to 0. \tag{5.3}$$

We can take, for example, $\epsilon(s) = -1/[\log(1/s)]$.

Remark. The condition: $\overline{\lim}\ Q(x) < 0$ as $R(x) \to 0$ is a "stability condition," meaning that G attracts $\xi(t)$ near the boundary. The condition (B_2) can be interpreted as a weak "repelling condition."

Theorem 5.1. *Let* (A_1), (B_1), (B_2), *and* (A_3) *hold. Then the solution of* (1.1) *is transient in* $\hat{G}$, *i.e.*,

$$P_x\left\{\lim_{t\to\infty}|\xi(t)| = \infty\right\} = 1 \qquad \text{if} \quad x\in\hat{G}. \tag{5.4}$$

Let G be contained in a ball $B_{R_*} = \{y; |y| < R_*\}$.

We shall first prove a lemma.

Lemma 5.2. *Let* (A_1), (B_1), (B_2) *hold and let* $\beta > R_*$. *Then*

$$P_x\{\xi(t) \text{ hits the set } \partial B_\beta \text{ for some } t \geqslant 0\} = 1 \qquad \text{if} \quad x\in\hat{G}\cap B_\beta. \tag{5.5}$$

Here B_β is the ball $\{y; |y| \leqslant \beta\}$ and ∂B_β is its boundary.

Proof. We first construct a function $g(x) = \Phi(R(x))$ for $x\in\hat{G}\cap B_\beta$, such that

$$Lg(x) \leqslant 0 \qquad \text{if} \quad x\in\hat{G}\cap B_\beta. \tag{5.6}$$

Denote by $\zeta_1, \ldots, \zeta_{k-1}$ the points in $\hat{G}$ where $D_xR(x) = 0$. By slightly modifying the proof of Lemma 4.4, we obtain a modified function $R(x)$ for which the points $\zeta_1, \ldots, \zeta_{k-1}$ lie outside the ball B_β. We shall work, in the present proof, with this modified function $R(x)$; it coincides with the original $R(x)$ in the ϵ_0-neighborhood of G.

We claim that there is a continuous function $\theta(r)$ satisfying:

$$1 + \frac{R^2(x)Q(x)}{\mathcal{Q}(x)} \geqslant \theta(R(x)) \qquad \text{if} \quad x\in\hat{G}\cap B_\beta \tag{5.7}$$

$$\theta(r) = 1 + \epsilon(r) \qquad \text{if} \quad 0 < r \leqslant \delta_0. \tag{5.8}$$

Indeed, since $\mathcal{Q}(x) \neq 0$ if $x\in\hat{G}$, $D_xR(x)\neq 0$, the left-hand side of (5.7) is a bounded function if $x\in\hat{G}\cap B_\beta$, $\min_h\rho_h(x) \geqslant \delta_0$. Using the assumption (5.2), the existence of $\theta(r)$ (satisfying (5.7), (5.8)) follows.

Let $\Phi(r)$ be a solution of

$$\Phi''(r) + \frac{\theta(r)}{r}\Phi'(r) = 0 \qquad \text{if} \quad 0 < r < r_0, \tag{5.9}$$

$$\Phi'(r) \leqslant 0 \qquad \text{if} \quad 0 < r < r_0 \tag{5.10}$$

where $r_0 = \max_{|x|\leqslant\beta} R(x)$. Then, upon using (5.1), (5.7) we conclude that $g(x) = \Phi(R(x))$ satisfies (5.6).

A solution of (5.9), (5.10) is given by

$$\Phi(r) = \int_r^{\delta_0} \exp\left[\int_s^{\delta_0} \frac{\theta(t)}{t}\, dt\right] ds. \tag{5.11}$$

In view of (5.3),

$$\Phi(r) \to \infty \quad \text{if} \quad r \to 0. \tag{5.12}$$

By Itô's formula and (5.6),

$$E_x\Phi(|\xi(t)|) - E_x\Phi(|x|) = E_x\int_0^{\tau} Lg(\xi(s))\, ds \leqslant 0, \tag{5.13}$$

where τ is any bounded stopping time such that $\xi(s) \in \hat{G} \cap B_\beta$ if $0 \leqslant s \leqslant \tau$. Denote by G_ϵ ($\epsilon > 0$) the closed ϵ-neighborhood of G, and denote by $\sigma_{\epsilon\beta}$ the hitting time of the set $G_\epsilon \cup \partial B_\beta$. By Theorem 8.1.1, $P_x(\sigma_{\epsilon\beta} < \infty) = 1$ if $x \not\in G_\epsilon$, $x \in B_\beta$. Denote by $P_x(\epsilon)$ the probability that $\xi(\sigma_\epsilon\beta) \in G_\epsilon$ (given $\xi(0) = x$), and by $P_x(\beta)$ the probability that $\xi(\sigma_{\epsilon\beta}) \in \partial B_\beta$ (given $\xi(0) = x$). Substituting $\tau = \sigma_{\epsilon\beta} \wedge T$ in (5.13), and taking $T \to \infty$, we get

$$\Phi(\epsilon)P_x(\epsilon) + \Phi(\beta)P_x(\beta) \leqslant \Phi(|x|).$$

Taking $\epsilon \to 0$ and using (5.12), we deduce that $P_x(\epsilon) \to 0$ if $\epsilon \to 0$. Hence $P_x(\beta) \to 1$ if $\epsilon \to 0$. But this implies the assertion of the lemma.

Let $R_* < \alpha < R$ and denote by t_R the hitting time of the ball B_R. By Lemma 5.2, $P_x(t_R < \infty) = 1$ if $x \in \hat{G}$. Hence, by the strong Markov property (cf. the proof of Theorem 1.1),

$$P_x(\Omega^*(\alpha)) = E_x P_{\xi(t_R)}(\Omega^*(\alpha)) \leqslant E_x P_{\xi(t_R)}(\Omega(\alpha)),$$

where the notation (1.17), (1.19) is used.

Now, in the domain $\{y; |y| > R_*\}$ the matrix $(a_{ij}(y))$ is nondegenerate. Since the condition (A_3) holds, the estimate (1.16) remains valid for $x = \xi(t_R)$. Hence

$$P_x(\Omega^*(\alpha)) \leqslant \frac{F(R)}{F(\alpha)} \qquad (F(r) \to 0 \text{ if } r \to \infty).$$

We can now complete the proof, as in the case of Theorem 1.1, by taking $R \to \infty$ and noting that α can be arbitrarily large.

Theorem 5.3. *Let* (A_1), (B_1), (B_2), *and* (A_3') *hold. Then for any* $x \in \hat{G}$ *the assertion* (5.4) *holds.*

Proof. Proceeding as in the proof of Theorem 5.1, it remains to establish the estimate (1.16). We now perform an orthogonal transformation as in the proof of Theorem 1.2.

6. Recurrent solutions for degenerate diffusion

Theorem 6.1. *Let* (A_1), (B_1), (B_2), (4.25) *and* (A_4) *hold. Then the solution of* (1.1) *is recurrent in* $\hat{G}$, *i.e., for any ball* $B_\alpha(z) = \{y; |y - z| \leqslant \alpha\}$, $\alpha > 0$, *lying entirely in* $\hat{G}$, *the assertion* (2.3) *holds.*

Proof. For simplicity we take $z = 0$. Let R be any positive number such that $R > \alpha$ and such that G is contained in the interior of the ball B_R. We shall prove: if $x \in \partial B_R$, then

$$P_x\{\xi(t) \text{ hits the ball } B_\alpha \text{ for some } t \geqslant 0\} = 1. \tag{6.1}$$

Let $\delta > 0$ be sufficiently small so that $\delta < \epsilon_0$, the closed δ-neighborhood G_δ of G lies in the interior of G_R, and $G_\delta \cap B_\alpha = \varnothing$.

We shall construct a function $f(x) = F(R(x))$ ($R(x)$ as in Lemma 4.5) in $R^n \setminus G_\delta$ such that

$$Lf(x) \geqslant 0 \qquad \text{if} \quad x \in R^n \setminus G_\delta, \tag{6.2}$$

$$F(r) \to -\infty \qquad \text{if} \quad r \to \infty. \tag{6.3}$$

Notice that at the points ζ_m ($m = 1, \ldots, k-1$) where $D_x R(x) = 0$, $\mathscr{Q} = 0$ and, therefore, by the property (v) of $R(x)$, $Q < 0$. Hence, $Q(x) < 0$ in a neighborhood of each point ζ_m. It follows that there is a continuous function $\theta(r)$, $\delta \leqslant r < \infty$, such that

$$1 + \frac{R^2 Q}{\mathscr{Q}} \leqslant \theta(R) \qquad \text{if} \quad x \in R^n \setminus G_\delta. \tag{6.4}$$

In view of (A_4), we can choose $\theta(r)$ so that, for all r sufficiently large, $\theta(r) = 1 + \epsilon(r)$ where $\epsilon(r)$ satisfies (2.2). If we now define $F(r)$, for $r \geqslant \delta$, by

$$F(r) = -\int_\delta^r e^{-I(s)}\, ds, \qquad I(s) = \int_\delta^s \frac{\theta(t)}{t}\, dt$$

then $F'(r) \leqslant 0$, and (6.2), (6.3) hold. Arguing as in the proof of Theorem 2.1 (following (2.8)) with the present function $f(x) = F(R(x))$ and with the set $\{y; |y| \leqslant \alpha\}$ replaced by G_δ, we conclude that for any $x \in R^n \setminus G_\delta$,

$$P_x\{\xi(t) \text{ hits the set } G_\delta \text{ for some } t \geqslant 0\} = 1. \tag{6.5}$$

Let $\beta > R$ and denote by $\hat{G}_{\alpha\beta\delta}$ the domain bounded by ∂G_δ, ∂B_α, ∂B_β. Denote by $\tau_{\alpha\beta}$ the exit time from this domain. By Theorem 8.1.1, $P_x(\tau_{\alpha\beta} < \infty) = 1$ if $x \in \hat{G}_{\alpha\beta\delta}$. Using the strong Markov property we get, for any $x \in \partial B_R$,

$$\begin{aligned} P_x\{\xi(t) \text{ hits } B_\alpha \text{ for some } t \geqslant 0\} &= P_x\{\xi(t + \tau_{\alpha\beta}) \text{ hits } B_\alpha \text{ for some } t \geqslant 0\} \\ &= E_x P_{\xi(\tau_{\alpha\beta})}\{\xi(t) \text{ hits } B_\alpha \text{ for some } t \geqslant 0\} \\ &\geqslant P_x\{\xi(\tau_{\alpha\beta}) \in \partial B_\alpha\} + P_x\{\xi(\tau_{\alpha\beta}) \in \partial G_\delta\} \\ &\quad \cdot \inf_{y \in \partial G_\delta} P_y\{\xi(t) \text{ hits } B_\alpha \text{ for some } t \geqslant 0\}. \end{aligned}$$

Denote by $\sigma_{\alpha R}$ the hitting time of $B_\alpha \cup \partial B_R$. By Lemma 5.2, if $y \in \partial G_\delta$, then $P_y(\sigma_{\alpha R} < \infty) = 1$. Hence, by the strong Markov property,

$$\begin{aligned} P_y\{\xi(t) \text{ hits } B_\alpha \text{ for some } t \geq 0\} &= P_y\{\xi(t+\sigma_{\alpha R}) \text{ hits } B_\alpha \text{ for some } t \geq 0\} \\ &= E_x P_{\xi(\sigma_{\alpha R})}\{\xi(t) \text{ hits } B_\alpha \text{ for some } t \geq 0\} \\ &= P_y\{\xi(\sigma_{\alpha R}) \in \partial B_\alpha\} + \big(1 - P_y\{\xi(\sigma_{\alpha R}) \in \partial B_\alpha\}\big) \inf_{x \in \partial B_R} P_x\{\xi(t) \text{ hits } B_\alpha \\ &\qquad \text{for some } t \geq 0\}. \end{aligned}$$

Combining this with the previous inequality, and setting

$$\begin{aligned} P_x(\alpha) &= P_x\{\xi(t) \text{ hits } B_\alpha \text{ for some } t \geq 0\}, \\ \gamma_{\alpha\beta}(x) &= P_x\{\xi(\tau_{\alpha\beta}) \in \partial B_\alpha\}, \\ \gamma_\beta(x) &= P_x\{\xi(\tau_{\alpha\beta}) \in \partial G_\delta\}, \\ \mu(y) &= P_y\{\xi(\sigma_{\alpha R}) \in \partial B_\alpha\}, \end{aligned}$$

we arrive at the inequality

$$P_x(\alpha) \geq \gamma_{\alpha\beta}(x) + \gamma_\beta(x) \inf_{y \in \partial G_\delta} \Big\{ \mu(y) + [1 - \mu(y)] \inf_{z \in \partial B_R} P_z(\alpha) \Big\}. \tag{6.6}$$

Note that $\gamma_{\alpha\beta}(x)$ is the solution $u(x)$ of the Dirichlet problem:

$$\begin{aligned} Lu &= 0 && \text{in } \hat{G}_{\alpha\beta\delta}, \\ u &= 1 && \text{on } \partial B_\alpha \\ u &= 0 && \text{on } \partial G_\delta \cup \partial B_\beta. \end{aligned}$$

Hence, by the strong maximum principle, $\gamma_{\alpha\beta}(x)$ is positive on ∂G_R. Further, $\gamma_{\alpha\beta}(x) \uparrow$ if $\beta \uparrow$. Similar assertions are true for $\gamma_\beta(x)$. Since $\gamma_{\alpha\beta}(x) + \gamma_\beta(x) \leq 1$, we conclude that

$$\gamma_\beta(x) \leq \theta < 1 \qquad (x \in \partial G_R) \tag{6.7}$$

where θ is a constant *independent* of β.

Notice that (6.5) implies that

$$\gamma_{\alpha\beta}(x) + \gamma_\beta(x) \to 1 \qquad \text{if } \beta \to \infty \qquad (x \in \partial G_R). \tag{6.8}$$

Let

$$P_\alpha = \inf_{z \in \partial B_R} P_z(\alpha).$$

Let η be a positive number, and choose x_0 in ∂B_R so that

$$P_\alpha \geq P_{x_0}(\alpha) - \eta.$$

Let y_0 be a point in ∂G_δ such that

$$\inf_{y \in \partial G_\delta} \{ \mu(y) + [1 - \mu(y)] P_\alpha \} \geq \{ \mu(y_0) + [1 - \mu(y_0)] \} P_\alpha - \eta.$$

Applying (6.6) with $x = x_0$, we get

$$P_\alpha \geqslant \gamma_{\alpha\beta}(x_0) + \gamma_\beta(x_0)\{\mu(y_0) + [1 - \mu(y_0)]P_\alpha\} - 2\eta.$$

Hence

$$P_\alpha\{1 - \gamma_\beta(x_0)[1 - \mu(y_0)]\} \geqslant \gamma_{\alpha\beta}(x_0) + \gamma_\beta(x_0)\mu(y_0) - 2\eta.$$

Taking β sufficiently large and using (6.8) with $x = x_0$, we get

$$P_\alpha\{1 - \gamma_\beta(x_0)[1 - \mu(y_0)]\} \geqslant \{1 - \gamma_\beta(x_0)[1 - \mu(y_0)]\} - 3\eta.$$

Denote the expression in braces by λ_β. From (6.7) it follows that $\lambda_\beta \geqslant 1 - \theta > 0$. Hence

$$P_\alpha \geqslant 1 - \frac{3\eta}{\lambda_\beta} \geqslant 1 - \frac{3\eta}{1-\theta}.$$

Since η is arbitrary, $P_\alpha = 1$. This implies that $P_\alpha(x) = 1$ for all $x \in \partial G_R$, i.e., (6.1) holds.

Having proved (6.1), we can now easily complete the proof of Theorem 6.1 by the argument given in the proof of Theorem 2.1 (from (2.10) on). Instead of (2.10) we use (6.1), and instead of Theorem 8.1.1 we use Lemma 5.2.

Remark. Theorem 6.1 remains true if the condition (A_4) is replaced by (A_4'), or by the conditions in the remark following Theorem 2.2.

7. The one-dimensional case

Consider the case of one stochastic differential equation

$$d\xi(t) = \sigma(\xi(t))\,dw(t) + b(\xi(t))\,dt. \tag{7.1}$$

We shall assume that the condition (A_1) of Section 1 holds in the present case of $n = 1$, and that $\sigma(x) > 0$ for all x. Let

$$\phi(x) = \int_0^x \exp\left[-\int_0^z \frac{2b(u)}{\sigma^2(u)}\,du\right] dz. \tag{7.2}$$

This is a particular solution of

$$\tfrac{1}{2}\sigma^2 v'' + bv' = 0. \tag{7.3}$$

Notice that $\phi(x)$ is strictly monotone increasing. Set $\phi(\infty) = \lim_{x\to\infty}\phi(x)$, $\phi(-\infty) = \lim_{x\to-\infty}\phi(x)$. These numbers may be finite or infinite.

Theorem 7.1. (a) *If* $\phi(-\infty) = -\infty$, *then*

$$P_x\left\{\sup_{t>0}\xi(t) = \infty\right\} = 1. \tag{7.4}$$

If $\phi(\infty) = \infty$, *then*

$$P_x\left\{ \inf_{t>0} \xi(t) = -\infty \right\} = 1. \tag{7.5}$$

(b) *If* $\phi(\infty) = \infty$, $\phi(-\infty) > -\infty$, *then*

$$P_x\left\{ \sup_{t>0} \xi(t) < \infty \right\} = 1, \tag{7.6}$$

$$P_x\left\{ \lim_{t\to\infty} \xi(t) = -\infty \right\} = 1. \tag{7.7}$$

(c) *If* $\phi(-\infty) = -\infty$, $\phi(\infty) < \infty$, *then*

$$P_x\left\{ \inf_{t>0} \xi(t) > -\infty \right\} = 1, \tag{7.8}$$

$$P_x\left\{ \lim_{t\to\infty} \xi(t) = \infty \right\} = 1. \tag{7.9}$$

(d) *If* $\phi(\infty) < \infty$, $\phi(-\infty) > -\infty$, *then*

$$P_x\left\{ \sup_{t>0} \xi(t) < \infty \right\} = P_x\left\{ \lim_{t\to\infty} \xi(t) = -\infty \right\} = \frac{\phi(\infty) - \phi(x)}{\phi(\infty) - \phi(-\infty)}, \tag{7.10}$$

$$P_x\left\{ \inf_{t>0} \xi(t) > -\infty \right\} = P_x\left\{ \lim_{t\to\infty} \xi(t) = \infty \right\} = \frac{\phi(x) - \phi(-\infty)}{\phi(\infty) - \phi(-\infty)}, \tag{7.11}$$

$$P_x\left\{ \lim_{t\to\infty} |\xi(t)| = \infty \right\} = 1. \tag{7.12}$$

It follows that if $\phi(-\infty) = -\infty$ and $\phi(\infty) = \infty$, then the solution of (7.1) is recurrent; in all other cases the solution of (7.1) is transient.

Proof. Let $x_1 < x < x_2$ and denote by $\tau_x[x_1, x_2]$ the first time $\xi(t)$ leaves (x_1, x_2), given $\xi(0) = x$. By Itô's formula (cf. Problem 12(c), Chapter 8),

$$P_x\{\xi(\tau_x[x_1, x_2]) = x_2\} = \frac{\phi(x) - \phi(x_1)}{\phi(x_2) - \phi(x_1)}. \tag{7.13}$$

Noting that

$$P_x\left\{ \sup_{t>0} \xi(t) \geqslant x_2 \right\} \geqslant P_x\{\xi(\tau_x[x_1, x_2]) = x_2\}, \tag{7.14}$$

and taking $x_1 \to -\infty$ in (7.13) we obtain

$$P_x\left\{ \sup_{t>0} \xi(t) \geqslant x_2 \right\} = 1,$$

provided $\phi(-\infty) = -\infty$. Since x_2 can be arbitrarily large,

$$P\left\{ \sup_{t>0} \xi(t) = \infty \right\} = 1.$$

This proves (7.4). The proof of (7.5), in case $\phi(\infty) = \infty$ is similar.

To prove (b), take $x_1 \to -\infty$ in (7.13) to obtain

$$P_x\left\{\sup_{t>0}\xi(t) \geqslant x_2\right\} = \frac{\phi(x) - \phi(-\infty)}{\phi(x_2) - \phi(-\infty)}.$$

This implies that

$$P\left\{\sup_{t>0}\xi(t) < x_2\right\} = \frac{\phi(x_2) - \phi(x)}{\phi(x_2) - \phi(-\infty)}. \tag{7.15}$$

Taking $x_2 \to \infty$ the assertion (7.6) follows.

To prove (7.7), let $y < x$ and denote by τ_y the first time $\xi(t) = y$. By (a), $P_x(\tau_y < \infty) = 1$. Hence, by the strong Markov property, for any $x_2 > y$,

$$P_x\left\{\sup_{t>0}\xi(t + \tau_y) \geqslant x_2\right\} = E_x P_y\left[\sup_{t>0}\xi(t) \geqslant x_2\right] = \frac{\phi(y) - \phi(-\infty)}{\phi(x_2) - \phi(-\infty)}.$$

But

$$P_x\left\{\sup_{t>0}\xi(t + \tau_y) \geqslant x_2\right\} = P_x\left\{\sup_{t>\tau_y}\xi(t) \geqslant x_2\right\} \geqslant P_x\left\{\overline{\lim_{t\to\infty}}\,\xi(t) \geqslant x_2\right\}.$$

Therefore,

$$P_x\left\{\overline{\lim_{t\to\infty}}\,\xi(t) \geqslant x_2\right\} \leqslant \frac{\phi(y) - \phi(-\infty)}{\phi(x_2) - \phi(-\infty)}.$$

Taking $y \to -\infty$ we conclude that

$$P_x\left\{\overline{\lim_{t\to\infty}}\,\xi(t) \geqslant x_2\right\} = 0.$$

Here x_2 is any real number. Taking $x_2 \to -\infty$, (7.7) follows.

The proof of (c) is similar to the proof of (b), and will be omitted.

We proceed to prove (d). The relation

$$P_x\left\{\sup_{t>0}\xi(t) < \infty\right\} = \frac{\phi(\infty) - \phi(x)}{\phi(\infty) - \phi(-\infty)} \tag{7.16}$$

follows by taking $x_2 \to \infty$ in (7.15). Similarly one proves that

$$P_x\left\{\inf_{t>0}\xi(t) > -\infty\right\} = \frac{\phi(x) - \phi(-\infty)}{\phi(\infty) - \phi(-\infty)}. \tag{7.17}$$

If we prove (7.12), then the second equalities in (7.10), (7.11) follow from the equalities in (7.17), (7.16) respectively. Thus it remains to prove (7.12).

For any small $\epsilon > 0$ and large $M > 0$, choose $x_1 < -M$ and $x_2 > M$ so that $x_1 < x < x_2$,

$$P_{x_1}\left\{\sup_{t>0}\xi(t) < -M\right\} = \frac{\phi(-M) - \phi(x_1)}{\phi(-M) - \phi(-\infty)} > 1 - \epsilon, \tag{7.18}$$

and

$$P_{x_2}\left\{\inf_{t>0}\xi(t) > M\right\} = \frac{\phi(x_2) - \phi(M)}{\phi(\infty) - \phi(M)} \geqslant 1 - \epsilon. \tag{7.19}$$

Let $\tau = \tau_x[x_1, x_2]$. Then, by the strong Markov property,

$$\begin{aligned} &P_x\{\xi(t) \notin [-M, M] \text{ for } t > \tau\} = E_x P_{\xi(\tau)}\{\xi(t) \notin [-M, M] \text{ for } t > 0\} \\ &\quad = P_x[\xi(\tau) = x_2]P_{x_2}\Big[\inf_{t>0}\xi(t) > M\Big] + P_x[\xi(\tau) = x_1]P_{x_1}\Big[\sup_{t>0}\xi(t) < -M\Big] \\ &\quad \geqslant 1 - \epsilon \end{aligned}$$

by (7.18), (7.19). Consequently

$$P_x\left\{\varliminf_{t\to\infty}|\xi(t)| > M\right\} \geqslant 1 - \epsilon.$$

Taking first $\epsilon \to 0$ and then $M \to \infty$, the assertion (7.12) follows.

PROBLEMS

1. The solution of (0.1) is recurrent in G if and only if for every $x \in G$ and for any open set V in G, (0.2) holds with "finite random times t_m" replaced by "finite Markov times t_m." [*Hint*: Let W be a closed domain contained in V. Define t_j = first time $\geqslant t_{j-1} + 1$ such that $\xi(t)$ hits W.]

2. Let $\xi(t)$ be a continuous process for $t \geqslant 0$, with values in R^n. Let $E = \{\xi(t)$ hits a closed set Γ at a sequence of random times t_m increasing to infinity$\}$. Prove that there exists a sequence of bounded, Borel measurable functions $f_k(x_1, \ldots, x_k)$ such that

$$f_k(\xi(t_1), \ldots, \xi(t_k)) \to \chi_E \quad \text{a.s.},$$

and the sequence f_k does not depend on the process $\xi(t)$. [*Hint*: Let $\{y_i\}$ be a dense set in Γ. Show that $E = \{B_m \text{ i.o.}\}$ where

$$B_m = \bigcap_{l=1}^{\infty} \bigcup_{m<r_j<m+1} \bigcup_{i=1}^{\infty} \{|\xi(r_j) - y_i| < 1/l\}$$

where $\{r_j\}$ is the sequence of positive rational numbers; then cf. Problem 9, Chapter 2.]

3. Let Γ be a closed set, τ the hitting time of Γ by the solution $\xi(t)$ of (1.1). Suppose $P_x(\tau < \infty) = 1$. Prove:

$$\begin{aligned} &P_x\{\xi(t + \tau) \text{ hits } \Gamma \text{ at a sequence of times increasing to } \infty\} \\ &\quad = E_x P_{\xi(\tau)}\{\xi(t) \text{ hits } \Gamma \text{ at a sequence of times increasing to } \infty\}. \end{aligned}$$

[*Hint*: Use the preceding problem and the method of solution of Problem 10, Chapter 2.]

4. Prove the first relation in (2.11). [*Hint*: Use the method of proof in Problem 10, Chapter 2.]
5. Let G be a bounded domain with C^3 boundary ∂G. Suppose that (A_1) holds and that the condition (B_2) holds with $R(x)$ replaced by $\rho(x) = \text{dist}(x, \partial G)$, $x \in G$. Suppose further that $(a_{ij}(x))$ is nonsingular for all $x \in \text{int } G$. Prove that $\xi(t)$ is recurrent in int G.
6. Consider (7.1) and set

$$I_1(x) = \int_{-\infty}^{x} \exp\left\{ -\int_0^z \frac{2b(u)}{\sigma^2(u)}\, du \right\} dz,$$

$$I_2(x) = \int_{x}^{\infty} \exp\left\{ -\int_0^z \frac{2b(u)}{\sigma^2(u)}\, du \right\} dz.$$

Prove that if $I_1(x) < \infty$, $I_2(x) < \infty$, then

$$P\left\{ \lim_{t\to\infty} \xi(t) = \infty \right\} = P\left\{ \sup_{t>0} \xi(t) = \infty \right\} = E\,\frac{I_1(\xi(0))}{I_1(\xi(0)) + I_2(\xi(0))}\,,$$

$$P\left\{ \lim_{t\to\infty} \xi(t) = -\infty \right\} = P\left\{ \inf_{t>0} \xi(t) = -\infty \right\} = E\,\frac{I_2(\xi(0))}{I_1(\xi(0)) + I_2(\xi(0))}$$

where $\xi(t)$ is any solution of (7.1).
7. Let $x = 0$ be a two-sided obstacle for (7.1), i.e., $\sigma(0) = b(0) = 0$. Assume that $\sigma(x) > 0$ if $x > 0$ and set

$$\bar{\sigma}(x) = \sigma(e^x)e^{-x}, \qquad \bar{b}(x) = b(e^x)e^{-x} + \tfrac{1}{2}\sigma^2(e^x)e^{-2x}.$$

Let $\psi(x)$ be the solution of

$$\tfrac{1}{2}\bar{\sigma}^2\psi'' + \bar{b}\psi' = 0, \qquad \psi(0) = 0, \quad \psi'(0) = 1.$$

Prove that if $\psi(\infty) = \infty$, $\psi(-\infty) > -\infty$, then

$$P_x\left\{ \xi(t) > 0,\ \lim_{t\to\infty} \xi(t) = 0 \right\} = 1 \qquad \text{if} \quad x > 0.$$

8. Prove Lemma 4.5. [*Hint*: Let $k_0 = 0, k = 1$. By Palais [1] there is a diffeomorphism $y = f(x)$ of $R^n \setminus (\text{int } G_1)$ onto $|y| \geq 1$ such that $f(x) = x$ for x outside a neighborhood of G_1. Take

$$R(x) = \beta\rho_1(x) + (1 - \beta)|f(x)| \qquad (*)$$

with suitable β, $\beta = 0$ outside a neighborhood of G_1, $\beta = 1$ in a smaller neighborhood. For $k_0 = 0, k = 2$ there exists (by applying Palais [1] twice) a diffeomorphism $y = f(x)$ of $R^n \setminus \text{int}(G_1 \cup G_2)$ onto $R^n \setminus \text{int}(G_1' \cup G_2')$ and G_1', G_2' are as in Lemma 4.4. Replace in (*) $|f(x)|$ by $\lambda(f(x))$ where $\lambda(y)$ is constructed as in the proof of Lemma 4.4.]

Bibliographical Remarks

Chapters 1–3. The material of these chapters is standard. We have used, as sources, Breiman [1], Dynkin [1, 2], Doob [1], Varadhan [1], and McKean [1].

Chapters 4–5. The stochastic integral, Itô's formula, and the basic theory of stochastic differential equations are due to *K.* Itô [1, 2] (and the references given there). These results are presented in a book form in Doob [1] and, in more detail, in McKean [1] and Gikhman and Skorokhod [1, 2]. The present treatment is based on Gikhman and Skorokhod [2], but it incorporates some results from McKean [1], such as the exponential martingale inequality. Polynomial mollifiers occur in Courant and Hilbert [1]. For more details on mollifiers, see Friedman [2].

Stroock and Varadhan [1] have developed existence and uniqueness theory, for stochastic differential equations, in which the drift coefficient is any bounded measurable function and the diffusion matrix is continuous, bounded, and uniformly positive definite. In [3] they allow the diffusion matrix to degenerate. They have also extended, in [2], their results from [1] to disfusion processes restricted to a given domain Ω, i.e., when the path hits $\partial\Omega$ it is returned into Ω in a prescribed direction. (Some earlier results in this area are described in McKean [1], Gikhman and Skorokhod [1, 2], and in the references given there.)

Stochastic differential equations with $dw(t)$ replaced by the differential of a Poisson process are studied in Gikhman and Skorokhod [2], Komatsu [1], and Stroock [1]. Some applications in stochastic control are given in Wonham [1].

Chapter 6. Theorem 1.2 is due to Freidlin [1] and to Phillips and Sarason [1]. The material of Sections 2, 3, 4 is based on Friedman [1]. The connection between solutions of partial differential equations and stochastic differential equations is discussed in detail in Gikhman and Skorokhod [2], and in an expository article by Freidlin [2].

Chapter 7. The results of this chapter are based on Girsanov [1]. Earlier work in this direction was done by Cameron and Martin [1].

Chapter 8. The results of this chapter are due to Friedman [3]. Theorem 4.1 for $n = 1$ is due to Kulinič [2] (see also Gikhman and Skorokhod [2]) who also proved, for $n = 1$, other related results [1, 3, 4].

Chapter 9. Sections 1–3, 5–6 are due to Friedman [4]. The concept of obstacle and the results of Section 4 are due to Friedman and Pinsky [1]. The results of Section 3, in the special case of Brownian motion, were first proved by Dvoretzky and Erdös [1]. Theorem 7.1 is given in Gikhman and Skorokhod [2].

References

L. Breiman

[1] *Probability*. Addison-Wesley, Reading, Massachusetts, 1968.

R. H. Cameron and W. T. Martin

[1] The transformation of Wiener integrals by nonlinear transformations, *Trans. Amer. Math. Soc.* **66** (1949), 253–283.

R. Courant and D. Hilbert

[1] *Methods of Mathematical Physics*, Vol. 1. Wiley (Interscience), New York, 1953.

J. L. Doob

[1] *Stochastic Processes*. Wiley, New York, 1967.

A. Dvoretsky and P. Erdös

[1] Some problems on random walk in space, *Proc. 2nd Berkeley. Symp. Math. Statist. Probability*, 1950. Univ. of California Press, Berkeley, California, pp. 353–367.

E. B. Dynkin

[1] *Foundation of the Theory of Markov Processes*. Pergamon Press, New York, 1960.
[2] *Markov Processes*, Vols. 1, 2. Springer-Verlag, Berlin, 1965.

W. Feller

[1] *An Introduction to Probability Theory and Its Applications*, Vol. 2. Wiley, New York, 1971.

M. I. Freidlin

[1] On the factorization of nonnegative definite matrices, *Theor. Probability Appl.* **13** (1968), 354–356 [*Teor. Verojatnost. i Primenen* **13** (1968), 375–378].
[2] Markov processes and differential equations, in *Progress in Mathematics*, Vol. 3. Plenum Press, New York, 1969.

A. Friedman

[1] *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, New Jersey, 1964.
[2] *Partial Differential Equations*. Holt, New York, 1969.
[3] Limit behavior of solutions of stochastic differential equations, *Trans. Amer. Math. Soc.* **170** (1972), 359–384.
[4] Wandering out to infinity of diffusion processes, *Trans. Amer. Math. Soc.* **184** (1973), 185–203.

A. FRIEDMAN and M. A. PINSKY

[1] Asymptotic stability and spiraling properties of solutions of stochastic equations, *Trans. Amer. Math. Soc.* **186** (1973), 331–358.

I. I. GIKHMAN and A. V. SKOROKHOD

[1] *Introduction to the Theory of Random Processes*. Saunders, Philadelphia, 1965.
[2] *Stochastic Differential Equations*. Naukova Dumka, Kiev, 1968. (Springer-Verlag, Vol. 72, New York, 1973).

I. V. GIRSANOV

[1] On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, *Theor. Probability Appl.* **5** (1960), 285–301 [*Teor. Verojatnost. i Primenen.* **5** (1960), 314–330].

K. ITÔ

[1] On stochastic differential equations. *Memoirs Amer. Math. Soc.* No. 4 (1961).
[2] *Lectures on Stochastic Processes*. Tata Inst. Fundamental Res., Bombay, 1961.

K. ITÔ and H. P. MCKEAN

[1] *Diffusion Processes and Their Sample Paths*. Springer-Verlag, Berlin, 1965.

T. KOMATSU

[1] Markov processes associated with certain integro-differential operators, *Osaka J. Math.* **10** (1973), 271–303.

G. L. KULINIČ

[1] On the limit behavior of the distribution of the solution of a stochastic diffusion equation, *Theor. Probability Appl.* **12** (1967), 497–499 [*Teor. Verojatnost. i Primenen* **12** (1967), 548–551].
[2] Limiting behavior of distributions of a solution of the stochastic diffusion equation, *Ukrain. Mat. Ž.* **19** (1967), 119–125.
[3] Asymptotic normality of the distribution of the solution of a stochastic diffusion equation, *Ukrain. Mat. Ž.* **20** (1968), 396–400.
[4] Limit distributions of a solution of the stochastic diffusion equation, *Theor. Probability Appl.* **13** (1968), 478–482 [*Teor. Verojatnost. i Primenen.* 13 (1968), 502–506].

P. LÉVY

[1] *Théorie de l'Addition de Variables Aléatoires*. Gauthier-Villars, Paris, 1937.

H. P. MCKEAN

[1] *Stochastic Integrals*. Academic Press, New York, 1971.

J. MILNOR

[1] *Lectures on the h-Cobordism Theorem*. Princeton Math. Notes, Princeton, New Jersey, 1965.

M. MORSE

[1] A reduction to the Schoenflies extension problem, *Bull. Amer. Math. Soc.* **66** (1960), 113–115.

P. S. PALAIS

[1] Extending diffeomorphisms, *Proc. Amer. Math. Soc.* **11** (1960), 274–277.

R. S. PHILLIPS and L. SARASON

[1] Elliptic-parabolic equations of the second order, *J. Math. Mech.* **17** (1967/8), 891–917.

D. W. STROOCK

[1] Diffusion processes associated with Lévy generators, *to appear*.

D. W. STROOCK and S. R. S. VARADHAN

[1] Diffusion processes with continuous coefficients, *Comm. Pure Appl. Math.* Part I, **22** (1969), 345–400; Part II, *ibid.*, 479–530.

[2] Diffusion processes with boundary conditions, *Comm. Pure Appl. Math.* **24** (1971), 147–225.

[3] On degenerate elliptic-parabolic operators of second order and their associated diffusions, *Comm. Pure Appl. Math.* **24** (1972), 651–713.

S. R. S. VARADHAN

[1] *Stochastic Processes*. Courant Inst. Math. Sci., New York Univ., New York, 1967–68.

W. M. WONHAM

[1] Random differential equations in control theory. *Probabilistic Methods in Applied Mathematics*, Vol. 2, pp. 131–212. Academic Press, New York, 1970.

Index

Probability and Mathematical Statistics

A Series of Monographs and Textbooks

---

1. Thomas Ferguson. Mathematical Statistics: A Decision Theoretic Approach. 1967
2. Howard Tucker. A Graduate Course in Probability. 1967
3. K. R. Parthasarathy. Probahility Measures on Metric Spaces. 1967
4. P. Révész. The Laws of Large Numbers. 1968
5. H. P. McKean, Jr. Stochastic Integrals. 1969
6. B. V. Gnedenko, Yu. K. Belyayev, and A. D. Solovyev. Mathematical Methods of Reliability Theory. 1969
7. Demetrios A. Kappos. Probability Algebras and Stochastic Spaces. 1969
8. Ivan N. Pesin. Classical and Modern Integration Theories. 1970
9. S. Vajda. Probabilistic Programming. 1972
10. Sheldon M. Ross. Introduction to Probability Models. 1972
11. Robert B. Ash. Real Analysis and Probability. 1972
12. V. V. Fedorov. Theory of Optimal Experiments. 1972
13. K. V. Mardia. Statistics of Directional Data. 1972
14. H. Dym and H. P. McKean. Fourier Series and Integrals. 1972
15. Tatsuo Kawata. Fourier Analysis in Probability Theory. 1972
16. Fritz Oberhettinger. Fourier Transforms of Distributions and Their Inverses: A Collection of Tables. 1973
17. Paul Erdös and Joel Spencer. Probabilistic Methods in Combinatorics. 1973
18. K. Sarkadi and I. Vincze. Mathematical Methods of Statistical Quality Control. 1973
19. Michael R. Anderberg. Cluster Analysis for Applications. 1973
20. W. Hengartner and R. Theodorescu. Concentration Functions. 1973
21. Kai Lai Chung. A Course in Probability Theory, Second Edition. 1974
22. L. H. Koopmans. The Spectral Analysis of Time Series. 1974
23. L. E. Maistrov. Probability Theory: A Historical Sketch. 1974
24. William F. Stout. Almost Sure Convergence. 1974

25. E. J. McShane. Stochastic Calculus and Stochastic Models. 1974
26. Z. Govindarajulu. Sequential Statistical Procedures. 1975
27. Robert B. Ash and Melvin F. Gardner. Topics in Stochastic Processes. 1975
28. Avner Friedman. Stochastic Differential Equations and Applications, Volumes 1 and 2. 1975
29. Roger Cuppens. Decomposition of Multivariate Probabilities. 1975

In Preparation

Eugene Lukacs. Stochastic Convergence, Second Edition

Harry Dym and Henry P. McKean. Gaussian Processes: Complex Function Theory and the Inverse Spectral Method

A
B 5
C 6
D 7
E 8
F 9
G 0
H 1
I 2
J 3