Theory and Applications of Stochastic Differential Equations

ZEEV SCHUSS

Theory and Applications of Stochastic Differential Equations

ZEEV SCHUSS

Department of Mathematical Sciences
Tel-Aviv University, Ramat-Aviv
Israel

John Wiley & Sons

New York • Chichester • Brisbane • Toronto

Library of Congress Cataloging in Publication Data:

Schuss, Zeev, 1937–

Theory and applications of stochastic differential equations.

(Wiley series in probability and mathematical statistics)

Includes bibliographical references and index.

1. Stochastic differential equations. I. Title.

QA274.23.S38 519.2 80-14767

ISBN 0-471-04394-X

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Preface

The purpose of this book is to present sources, theory, and applications of stochastic differential equations of Itô's type; that is, differential equations that contain white noise. It gives the basic theory and a wide range of applications. The main theme of the book is the study of first passage problems by modern singular perturbation methods and their role in various fields of science. The material is so presented as to make the theory available to the applied mathematicians, physicists, chemists, and engineers, who are usually well versed in classical analysis but may feel uneasy in the realms of modern probability and measure theory. The prerequisites for this book are therefore a working knowledge of advanced calculus, elementary theory of ordinary and partial differential equations, and of course, elementary probability theory. The professional probabilist will find here some new analytic methods for the computation of first passage times, transition and exit probabilities, and other quantities of interest. Special stress has been put on modeling phenomena in a variety of scientific areas by stochastic differential equations. Thus phenomena in chemical kinetics, solid-state diffusion, genetics, filtering of signals from noise, and more are modeled.

Since Einstein's creation of the mathematical theory of the Brownian motion and molecular diffusion, much scientific work has been done on its applications in such diverse areas as molecular and atomic physics, chemical kinetics, solid-state theory, stability of structures, population genetics, communication, and many other branches of the natural and social sciences and engineering. The most prominent work in the early stages of the theory of stochastic differential equations was done by Einstein, Smoluchowski, Langevin, Ornstein and Uhlenbeck, and Kramers and was summarized in Chandrasekhar's fundamental paper in 1943. The mathematical theory of stochastic differential equations was developed considerably in the last 25 years and several very rigorous mathematical texts on this subject have appeared. Some very important results were discovered by the mathematical researchers in this field; in particular, the equations for first passage times and exit distributions were derived. The Itô and

Stratonovich calculi in particular gave the theory of stochastic differential equations an enormous push forward. Unfortunately, the gap between the mathematical theory and the sources of the problems has widened to such an extent that, generally, physicists, chemists, and engineers remained unaware of the mathematical techniques now available, while the mathematicians remained unaware of the sources and applications of their theories. The complexity of the theory and the mathematical rigor made the mathematical texts virtually unapproachable to the nonspecialist. This book is an attempt to bridge this gap.

Chapters 1 and 2 present the relevant probability theory, construction of the Brownian motion, and the theory of Itô and Stratonovich integration and calculus. The more demanding and mathematically rigorous material has been relegated to separate sections marked by an asterisk. Such sections should be omitted in the first reading of the book. The basic theory of stochastic differential equations is presented in Chapters 3 through 5. Special attention should be given to the exercises because they contain many classical applications of the theory; in particular, Einstein's and Smoluchowski's theories of diffusion and their applications are contained in the exercises. Chapter 4 also establishes the connection between Markov and diffusion processes on the one hand and solutions of stochastic differential equations on the other. In Chapter 5 the relationship between stochastic differential equations and partial differential equations is demonstrated; the basic equations of Fokker–Planck, Kolmogorov, Dynkin, Feynman, and Kac are derived; and boundary behavior is discussed. A method for the treatment of first passage problems by partial differential equations is developed through the Itô calculus. The main contribution of this book are Chapters 6 through 9. Chapter 6 presents the asymptotic theory of stochastic differential equations and its applications in statistical mechanics, transport theory, and mathematical genetics. In Chapter 7 singular perturbation problems that arise from the Smoluchowski–Kramers theory are treated by a new method, developed by B. Matkowsky and myself. Physical applications of the theory are presented in Chapter 8. Mathematical models of chemical reactions, diffusion, and conductivity in ionic crystals are given. Chapter 9 contains elements of filtering theory in state space and the role of first passage times is shown. Finally, Chapter 10 contains Smoluchowski's theory in the context of the kinetic theory of gases, and a short review of some basic notions in the theory of classical mechanics and partial differential equations. The uniform mathematical treatment reduces many of the problems to that of the determination of the expected first exit time by solving singularly perturbed boundary value problems for partial differential equations. The singular perturbation methods presented in this book lead to explicit expressions for probabilistic and consequently physical quantities of interest, such as the steric factor in the Arrhenius law, reaction rates in multistage chemical reactions, the diffusion tensor for atomic migration in crystals, the electric conductivity in ionic crystals, and the "click" frequency in a FM filter. It is my hope that the book will bridge the gap between the mathematical theory of stochastic differential equations and the natural sciences by giving the scientist a new mathematical tool and by giving the mathematicians some insight into the role of stochastic differential equations in the sciences.

My interest in stochastic differential equations was a consequence of a course on this subject given by H. McKean, Jr., at the Weizmann Institute of Science in Rehovot, for which I am grateful. The idea of the book began with a series of lectures in the Applied Mathematics Seminar that I gave at RPI in 1975–1976. It continued with a set of lecture notes from the Applied Mathematics Institute, the University of Delaware in 1976–1977, and completed at Tel-Aviv University and Northwestern University in the summers of 1978 and 1979. I am grateful to these institutions for their hospitality. The book is based on my joint work with B. Matkowsky, E. Larsen, B. Levikson, S. Eliezer and B. Z. Bobrovsky, for whose cooperation I am deeply indebted. The preparation of this manuscript was partially supported by AFOSR grants 77-3372 and 78-3602B in 1977–1980 at the University of Delaware and Northwestern University.

ZEEV SCHUSS

Ramat-Aviv, Israel
September 1980

Contents

CHAPTER 1

Review of Probability Theory

1.1 EVENTS AND SAMPLE SPACES

Consider the experiment of tossing a fair coin three times. The possible outcomes of this experiment are HHH, HHT, HTH, THH, HTT, THT, TTH, and TTT, where H denotes heads and T denotes tails. Each possible outcome of the experiment is called an *elementary event*. Thus there are eight elementary events in the experiment of tossing a coin three times. More complicated events can be expressed as combinations of the elementary events. Thus the event "two or more heads turn up" in the coin-tossing experiment, which we denote by D, consists of the elementary events HHH, HHT, HTH, THH; that is,

$$D = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}\}.$$

The set Ω of all elementary events corresponding to an experiment is called a *sample space* and each elementary event, denoted by ω, is called a *point* in Ω. In the particular example under consideration, we have

$$\Omega = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}, \mathrm{TTH}, \mathrm{THT}, \mathrm{HTT}, \mathrm{TTT}\}.$$

We write $\omega \in \Omega$ to read "ω is a point in (or an element of) Ω." Any event A that consists of elementary events is a subset of Ω. In particular, the impossible event $\varnothing$, that is, an event that contains no elementary events, is called *the empty* set. Let A and B be events in Ω; then we say that *A is a subset of B* if every element of A is also an element of B, and we write $A \subset B$. Obviously, $A \subset \Omega$, $\varnothing \subset A$, and $A \subset A$. For example, the set D in the coin-tossing experiment is a subset of the set (event) E: "at least one head turns up." More specifically,

$$\{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}\} = D \subset E$$
$$= \{\mathrm{HHH}, \mathrm{THH}, \mathrm{HTH}, \mathrm{HHT}, \mathrm{TTH}, \mathrm{THT}, \mathrm{HTT}\}.$$

We say that two subsets A and B of Ω are *equal* if they consist of the same elements; that is, $A = B$ if $A \subset B$ and $B \subset A$. We say that two subsets of Ω (two events), A and B, are *disjoint* if they have no common elements. Thus the set F: "at least two tails turn up" and the set D in the coin-tossing experiment are disjoint sets. However, the sets E and F are not disjoint.

Very often, the sample space contains infinitely many points (elementary events). Consider, for example, the experiment of sampling the velocities of the molecules of a monatomic gas consisting of n molecules of mass m in thermal equilibrium. Let $\mathbf{v}_i = (v_i^1, v_i^2, v_i^3)^T$ $(i = 1, 2, \ldots, n)$ be the velocity vectors of the molecules at the moment the experiment is conducted. We assume that the gas is ideal, thus neglecting the potential energy of intermolecular forces. Denoting the total energy of the gas by E, we have

$$\sum_{i=1}^{n} |\mathbf{v}_i|^2 = \frac{2E}{m},$$

where $|\mathbf{v}_i|^2 = \mathbf{v}_i \cdot \mathbf{v}_i = \sum_{j=1}^{3} v_i^j v_i^j$. Since E is assumed constant, we see that any outcome of the experiment is a point on the surface of the $3n$-dimensional sphere S of radius $(2E/m)^{1/2}$. We may identify the sample space for this experiment with the set of all points on the surface S. A typical event is G: "the first component of $\mathbf{v}_i$ satisfies the inequality $a < \mathbf{v}_i^1 < b$." The set G is therefore a spherical zone on S.

Given two subsets A and B of a sample space Ω, we denote by $A \cup B$ the subset of Ω whose elements are those which belong to A or to B. The set $A \cup B$ is called the *union* of A and B. More generally, given a finite or infinite sequence $\{A_j\}$, $j = 1, 2, \ldots,$ of subsets of Ω, we denote by

$$A = A_1 \cup A_2 \cup \cdots \equiv \bigcup_j A_j$$

the set whose elements are those which belong to *at least* one of the sets A_j. Thus the event $\cup_j A_j$ occurs if at least one of the events A_j $(j = 1, 2, \ldots)$ does. The set $A = \cup_j A_j$ is called the union of the sets A_j. It is easy to see that $A \cup A = A$, $A \cup \Omega = \Omega$ and $A \cup \varnothing = A$. If $A \subset \Omega$ and $B \subset \Omega$, then the set $A - B$ consists of those elements of A which are not elements of B. The set $A - B$ is called the *difference* of A and B. Thus in the coin-tossing example, we have

$$E-D=\{\mathrm{HTT},\mathrm{THT},\mathrm{TTH}\},$$

that is, $E-D$: "exactly one head turns up," so that $E-D$ occurs if E occurs but not D. We also have $D-E=\varnothing$. Obviously, $A-\varnothing=A$, $A-A=\varnothing$, and $A-\Omega=\varnothing$. If A and B are disjoint events, then $A-B=A$. Given a finite or infinite sequence $\{A_j\}$, $j=1,2,\ldots$, of subsets of Ω, we denote by

$$A\equiv A_1\cap A_2\cap\cdots\equiv\bigcap_j A_j$$

the set whose elements belong to all the sets A_j, $j=1,2,\ldots$. The set $A=\cap_j A_j$ is called the *intersection* of the sets A_j. The event $\cap_j A_j$ occurs if *all* the events A_j, $j=1,2,\ldots$, occur. Obviously, $A\cap A=A=A\cap\Omega$, $A\cap\varnothing=\varnothing$; if $A\subset B$, then $A\cap B=A$. Thus in the coin-tossing example, $E\cap D=D$. Clearly, two sets A and B are disjoint if and only if $A\cap B=\varnothing$. If $A\subset\Omega$, we denote by A^c the difference $\Omega-A$. The set A^c is called the *complement* of A in Ω. It contains those elements of Ω that do not belong to A. Thus the set D^c in the coin-tossing example is the event "at most one head turns up in three tosses of a coin," and we have

$$D^c=\{\mathrm{TTT},\mathrm{TTH},\mathrm{THT},\mathrm{HTT}\}.$$

The set G^c in the monatomic gas example consists of all the points on S outside the spherical zone G.

EXERCISE 1.1.1

(i) Construct a sample space Ω corresponding to the experiment of throwing a die.
(ii) How many elementary events are there in Ω?
(iii) How many events consisting of two elementary events are there in Ω?

EXERCISE 1.1.2

Let the sample space consist of n elementary events.

(i) What is the number of events that consist of exactly k elementary events in Ω?
(ii) What is the number of all events in Ω?

EXERCISE 1.1.3

In random sampling of families the event A occurs if the sampled family has only one child and the event B occurs if the family has at least one child. It is known that there are no families having more than n children. Describe the elementary events and express A and B in terms of the elementary events. Describe the events $A \cup B$, A^c, B^c, $B-A$, and $A-B$.

EXERCISE 1.1.4

In the monatomic gas model, let the events A and B be given by $A=\{a<v_i^1<b\}$ and $B=\{c<v_i^2<d\}$. Describe geometrically the events $A \cup B$, $A \cap B$, and $(A-B)\cup(B-A)$.

EXERCISE 1.1.5

Prove de Morgan's law, $(A \cup B)^c = A^c \cap B^c$; $(A \cap B)^c = A^c \cup B^c$.

1.2 PROBABILITY MEASURE

In the experiment of tossing a fair coin three times, we intuitively assign the probability $\frac{1}{8}$ to each of the possible eight outcomes, as all seem to us equally likely. To the event D we assign the probability $\frac{1}{2}$ because it contains one-half of all possible events in Ω. Obviously, we assign the probability zero to the impossible event $\varnothing$ and the probability 1 to the sure event Ω. To make the intuitive notion of probability mathematically precise, we introduce a system of axioms that formalize the basic properties we expect probability to have. We describe first the set $\mathfrak{B}$ of *random events* on which the probability measure is defined. The elements of $\mathfrak{B}$ are subsets of Ω and $\mathfrak{B}$ has the following properties:

(i) $\Omega \in \mathfrak{B}$.

(ii) If $A \in \mathfrak{B}$ and $B \in \mathfrak{B}$, then $A-B \in \mathfrak{B}$.

(iii) If $A_j \in \mathfrak{B}$, $j=1,2,\ldots$, is a sequence of elements of $\mathfrak{B}$, then $\cup_{j=1}^{\infty} A_j \in \mathfrak{B}$.

In particular, complements and intersections of random events are random events. A *probability measure* $P(\cdot)$ is a function defined on the set $\mathfrak{B}$ of random events, which satisfies the following axioms:

Axiom 1. To every element A of $\mathfrak{B}$ there corresponds a number $P(A)$

which satisfies the inequality

$$0 \leqslant P(A) \leqslant 1.$$

Axiom 2. $P(\Omega)=1$.

Axiom 3. If $A_j \in \mathfrak{B}$, $j=1,2,\ldots$, is a finite or infinite sequence of disjoint events, that is, $A_i \cap A_j = \varnothing$ if $i \neq j$, then

$$P\left(\bigcup_j A_j\right)=\sum_j P(A_j).$$

To illustrate the axioms, we again consider the experiment of tossing a fair coin three times. The sample space consists of eight elementary events, and $\mathfrak{B}$ consists of *all* subsets of Ω. It is easy to see that properties (i)–(iii) are satisfied. To each random event in $\mathfrak{B}$ we assign the probability

$$P(A)=\frac{\text{number of elements in } A}{8}.$$

It is easy to verify that Axioms (1)–(3) are satisfied in this case. In the case of the monatomic gas the elementary events have been identified with points of the sphere S of radius $(2E/m)^{1/2}$. We expect the probability of the event G in S to be proportional to the *area* of G; for example, for $G=\{a<v_i^1<b\}$, elementary calculus shows that

$$P(G)=c\int_a^b \left(1-\frac{x^2 m}{2E}\right)^{(3n-3)/2} dx \tag{1.2.1}$$

where c is a proportionality constant. Since by (2) $P(S)=1$, we must have

$$c=\frac{1}{\int_{-(2E/m)^{1/2}}^{(2E/m)^{1/2}} (1-x^2 m/2E)^{(3n-3)/2}\, dx}.$$

The set $\mathfrak{B}$ of random events in S cannot be taken to be the set of *all* subsets of S, since it can be shown that there is no function $P(A)$ defined on the set of all subsets A of S such that (1.2.1) is satisfied and (1)–(3) hold. This is a consequence of the existence of nonmeasurable sets in S (Halmos 1959). The set $\mathfrak{B}$ is defined as the smallest set with properties (i)–(iii) that contains all the events $\{a<v_i^j<b\}$, where $i=1,2,\ldots,n$; $j=1,2,3$; and $-2E/m \leqslant a \leqslant b \leqslant 2E/m$. Thus whenever the area of a subset A

of S is defined, we set

$$P(A)=\frac{\text{area of } A}{\text{area of } S}.$$

The computation of $\lim_{n\to\infty} P(G)$ leads to the well-known result of Maxwell: Assuming that the energy is proportional to the number of particles in the gas, we set $E=\gamma n$, where γ is a constant independent of n; hence

$$P\{a<v_i^1<b\}=\frac{\int_a^b\left(1-\frac{x^2m}{2\gamma n}\right)^{(3n-3)/2}dx}{\int_{-(2\gamma n/m)^{1/2}}^{(2\gamma n/m)^{1/2}}(1-x^2m/2\gamma n)^{(3n-3)/2}\,dx}$$

$$\to\left(\frac{3m}{4\pi\gamma}\right)^{1/2}\int_a^b e^{-3mx^2/4\gamma}\,dx.$$

Setting $\gamma=3kT/2$, we obtain Maxwell's result:

$$\lim_{n\to\infty} P\{a<v_i^1<b\}=\left(\frac{m}{2\pi kT}\right)^{1/2}\int_a^b e^{-mx^2/2kT}\,dx.$$

We call T absolute temperature and k is called Boltzmann's constant. A little more sophisticated example is that of the theory of the game of heads and tails. The possible outcomes of this game are all the infinite sequences of H and T. Thus the sample space Ω is the set of all sequences $A_1, A_2, \ldots$, where each A_j is either the symbol H or the symbol T. There are infinitely many distinct sequences of this kind, and in fact, the elements of Ω cannot even be arranged in a sequence; that is, the set Ω is uncountable (Kamke 1950). If we are to assign a probability measure to each outcome of the game, it would necessarily be $P(\{A_1, A_2,\ldots\})=0$. For, obviously all sequences must have the same probability, and if $P(\{A_1, A_2,\ldots\})=c>0$, then for any sequence of distinct outcomes $A^i=\{A_1^i, A_2^i,\ldots\}$, $i=1,2,\ldots$, we have $A^i\cap A^j=\varnothing$; therefore, by (3),

$$P\left(\bigcup_j A^j\right)=\sum_j P(A^j)=\sum_j c=\infty,$$

which contradicts (1). We therefore take the elementary events to be the sets of sequences, k of whose places ($k=1,2,\ldots$) are fixed. Clearly, the probability of an elementary event, in which k is the number of fixed places, is the probability of an outcome of the experiment of tossing a fair coin k times. Thus the probability measure assigned to such an elementary event is $1/2^k$. To construct a probability measure on Ω that assigns the elementary events the probability $1/2^k$ if k places are fixed, we map Ω onto the unit interval by assigning to each sequence $A_1, A_2, \ldots$ the number

$$t = \sum_{i=1}^{\infty} 2^{-n}\varepsilon_n, \tag{1.2.2}$$

where $\varepsilon_n = 1$ if $A_n = \mathrm{H}$ and $\varepsilon_n = 0$ if $A_n = \mathrm{T}$. Obviously, $0 \leq t \leq 1$. This correspondence is not one-to-one since the sequence H, T, T,... and T, H, H,... are mapped into the same number, namely into the number $\frac{1}{2}$, as

$$2^{-1} = \sum_{n=2}^{\infty} 2^{-n}.$$

The sequences that are mapped into the same numbers correspond to the "dyadic rationals," that is, to numbers of the form $r/2^s$, where r and s are positive integers. It is easy to see that the set of all such numbers can be arranged in a sequence (how?) $A = [A^1, A^2, \ldots]$, say (i.e., the set is countable). Since $P(A^j) = 0$ and $A^i \cap A^j = \varnothing$ if $i \neq j$, we must have, by (3), $P(A) = 0$. Identifying all dyadic rationals with sequences that end with T, T,..., we obtain a one-to-one correspondence between Ω and the interval $[0,1]$. Thus the elementary set $B_1 = \{H, A_1, A_2, \ldots\}$ is mapped onto the interval $\left[\frac{1}{2}, 1\right]$, which is the set of all numbers in the interval $[0,1]$ whose first digit is 1 in their binary expansion. The event B_1 corresponds to the outcome H in a single toss of a fair coin; thus we assign to B_1 the probability $\frac{1}{2}$. This probability is the length of the interval $\left[\frac{1}{2}, 1\right]$ corresponding to B_1 in the one-to-one correspondence between Ω and the interval $[0,1]$. It is now easy to see that any elementary event is mapped onto a finite union of intervals whose end points are dyadic rationals, and the probability of such an elementary event equals the sum of the lengths of the dyadic intervals onto which it is mapped. It is also easy now to construct a probability measure on Ω which is consistent with the probability measure assigned to the elementary events. We simply assign to any interval $[a, b] \subset [0,1]$ its length $b - a$ and extend the definition by properties (1)–(3) to the set $\mathfrak{B}$ of random events (Halmos 1959). The set $\mathfrak{B}$ of random events in Ω can be described by its image in the interval $[0,1]$ under the one-to-one mapping described above. The image of $\mathfrak{B}_\Omega$ in the interval $[0,1]$ is the so-called Borel set $\mathfrak{B}$, which is the smallest set containing all the subintervals of $[0,1]$ such that $\mathfrak{B}$ has the properties (i)–(iii).

Using property (3) of the probability measure, we can derive the formula

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{1.2.3}$$

Indeed,

$$A\cup B=A\cup[B-(A\cap B)]$$

$$B=[A\cap B]\cup[B-(A\cap B)]$$

and obviously

$$\varnothing=A\cap[B-(A\cap B)]$$

and

$$\varnothing=[A\cap B]\cap[B-(A\cap B)].$$

Hence, by (3),

$$P(A\cup B)=P(A)+P[B-(A\cap B)]$$

and

$$P(B)=P(A\cap B)+P[B-(A\cap B)].$$

It follows that (1.2.3) holds.

EXERCISE 1.2.1 (Kac 1959)

Let $H_n(\omega)$ be the number of H's in the sequence $\omega=\{A_1, A_2,\ldots\}\in\Omega$ in the game of heads and tails. Using the identity

$$\lim_{n\to\infty} 2^{-n} \sum_{|k-n/2|<\alpha\sqrt{n}} \binom{n}{k}=(2\pi)^{-1/2}\int_{-2\alpha}^{2\alpha} e^{-x^2/2}\,dx, \tag{1.2.4}$$

show that

$$\lim_{n\to\infty} P\left\{\left|H_n-\frac{n}{2}\right|<\alpha\sqrt{n}\right\}=(2\pi)^{-1/2}\int_{-2\alpha}^{2\alpha} e^{-x^2/2}\,dx.$$

Use this result to devise a test for the fairness of a coin.

EXERCISE 1.2.2

Prove (1.2.4) by using Stirling's formula (Feller 1957).