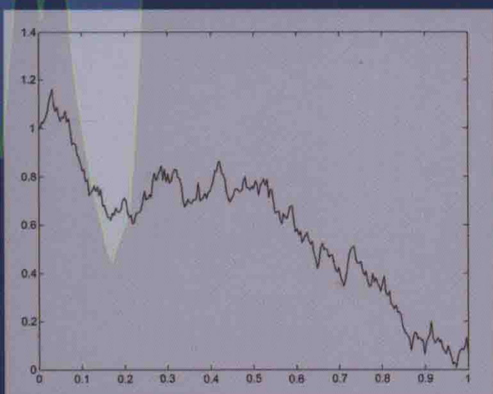
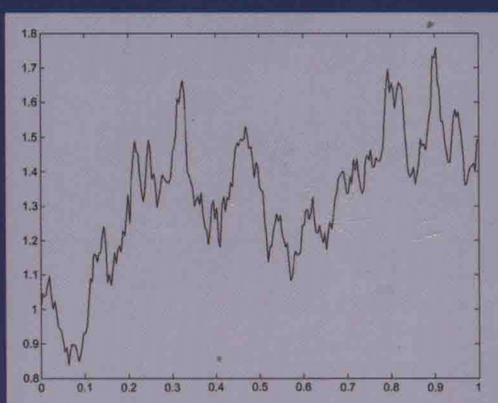
An Introduction to Stochastic Differential Equations

Lawrence C. Evans





AMS
American Mathematical Society

An Introduction to Stochastic Differential Equations

Lawrence C. Evans

Department of Mathematics
University of California, Berkeley

AMERICAN MATHEMATICAL SOCIETY

2010 *Mathematics Subject Classification.* Primary 65C30, 60J65. 60H10, 65N75.

For additional information and updates on this book, visit
www.ams.org/bookpages/mbk-82

Library of Congress Cataloging-in-Publication Data

Evans, Lawrence C., 1949–
An introduction to stochastic differential equations / Lawrence C. Evans.
pages cm
Includes bibliographical references and index.
ISBN 978-1-4704-1054-4 (alk. paper)
1. Stochastic differential equations. I. Title.

QA274.23.E93 2013
519.2—dc23

2013024818

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at `http://www.ams.org/`

10 9 8 7 6 5 4 3 2 1 18 17 16 15 14 13

An Introduction to Stochastic Differential Equations

PREFACE

This book is a revision of lecture notes for a course on stochastic differential equations (SDE) that I have taught several times over the past decades at the University of Maryland, the University of California, Berkeley, and the Mathematical Sciences Research Institute.

My intention has been to survey, honestly but with some omission of precise detail, the basics of the Itô stochastic calculus and the foundations of stochastic differential equations, with particular emphasis upon applications to partial differential equations (PDE).

I assume my readers to be fairly conversant with measure-theoretic mathematical analysis but do not assume any particular knowledge of probability theory (which I develop very rapidly in Chapter 2). I downplay most measure theory issues but do emphasize the probabilistic interpretations. I "prove" many formulas by confirming them in easy cases (for simple random variables or for step functions) and then just stating that by approximation these rules hold in general. This whirlwind introduction is of course no substitute for a solid graduate level course in probability; but this book should provide enough background and motivation for students who lack the preparation to tackle the standard SDE text Øksendal [**O**].

Thanks to my colleague Fraydoun Rezakhanlou, who has taught from these notes and added several improvements, and to Lisa Goldberg, who several years ago gave my class with several lectures on financial applications. Jonathan Weare provided several computer simulations illustrating the text. Thanks also to many readers of the online version who have found errors, especially Robert Piche, who provided me with an extensive list of typos and suggestions.

For this printing as a book, the notes have been retyped and reformatted; I have also updated the references and made many improvements in the presentation. I have, as usual, received great help from everyone at the American Mathematical Society, especially Sergei Gelfand, Stephen Moye, Arlene O'Sean, Tom Costa and Chris Thivierge.

I will post a list of errors on my homepage, accessible through the math.berkeley.edu website.

I have been supported by the NSF during the writing of this book, most recently by the grants DMS-1001724 and DMS-1301661.

LCE
July 2013
Berkeley

CONTENTS

Chapter 1

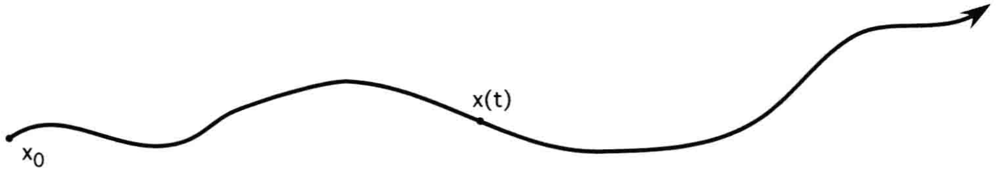
INTRODUCTION

1.1. DETERMINISTIC AND RANDOM DIFFERENTIAL EQUATIONS

Fix a point $x_0 \in \mathbb{R}^n$ and consider then the ordinary differential equation

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{b}(\mathbf{x}(t)) & (t > 0) \\ \mathbf{x}(0) = x_0, \end{cases}$$

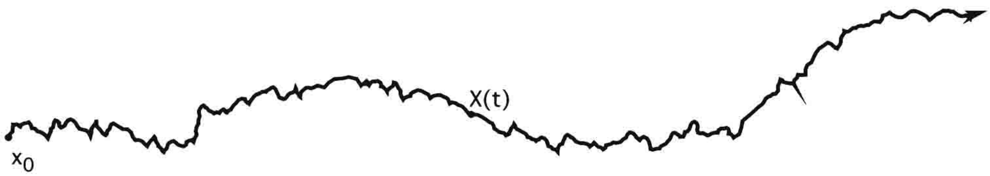
where $\mathbf{b} : \mathbb{R}^n \to \mathbb{R}^n$ is a given smooth vector field and the solution is the trajectory $\mathbf{x} : [0, \infty) \to \mathbb{R}^n$, where $\mathbf{x} = \mathbf{x}(t)$ is a function of time t. The dot means differentiation: $\dot{} = \frac{d}{dt}$.



Trajectory of the differential equation ODE

We call $\mathbf{x}(t)$ the *state of the system* at time $t \geq 0$. Under reasonable assumptions on the vector field $\mathbf{b}$, the ordinary differential equation (ODE) has a solution, uniquely determined by the initial condition x_0.

In many applications, however, the experimentally measured trajectories of systems modeled by (ODE) do not in fact behave as predicted: the observed state seems to more or less follow the trajectory predicted by (ODE), but is apparently subject also to random perturbations.



Trajectory of a stochastic differential equation (SDE)

Hence it seems reasonable to modify (ODE), somehow to include the possibility of random effects disturbing the system. A *formal* way to do so is to write

$$(1) \qquad \begin{cases} \dot{\mathbf{X}}(t) = \mathbf{b}(\mathbf{X}(t)) + \mathbf{B}(\mathbf{X}(t))\boldsymbol{\xi}(t) \quad (t > 0) \\ \mathbf{X}(0) = x_0, \end{cases}$$

where

$$\mathbf{B} : \mathbb{R}^n \to \mathbb{M}^{n\times m} \text{ (= space of } n \times m \text{ matrices)}$$

and

$$\boldsymbol{\xi}(\cdot) := m\text{-dimensional “white noise”}.$$

We then have these **mathematical problems:**

- Define “white noise” $\boldsymbol{\xi}(\cdot)$.
- Define what it means for $\mathbf{X}(\cdot)$ to solve (1).
- Show that (1) has a solution and discuss uniqueness, asymptotic behavior, dependence upon x_0, $\mathbf{b}$, $\mathbf{B}$, etc.

This book develops the rigorous mathematical theory to address these and many related questions.

1.2. STOCHASTIC DIFFERENTIALS

Let us first study (1) in the case $m = n$, $x_0 = 0$, $\mathbf{b} \equiv 0$, and $\mathbf{B} \equiv I$. The solution in this setting turns out to be n-dimensional *Brownian motion* (or *Wiener process*), denoted $\mathbf{W}(\cdot)$. Thus we may symbolically write

$$(2) \qquad \dot{\mathbf{W}}(\cdot) = \boldsymbol{\xi}(\cdot),$$

thereby asserting that *“white noise” is the time derivative of Brownian motion.*

Now return to the general form of equation (1), write $\frac{d}{dt}$ instead of the dot:

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{b}(\mathbf{X}(t)) + \mathbf{B}(\mathbf{X}(t))\frac{d\mathbf{W}(t)}{dt},$$

and formally multiply by "dt":

$$\text{(SDE)} \qquad \begin{cases} d\mathbf{X}(t) = \mathbf{b}(\mathbf{X}(t))dt + \mathbf{B}(\mathbf{X}(t))d\mathbf{W}(t) \\ \mathbf{X}(0) = x_0. \end{cases}$$

The terms "$d\mathbf{X}$" and "$\mathbf{B}d\mathbf{W}$" are called *stochastic differentials*, and the expression (SDE), properly interpreted, is a *stochastic differential equation.*

We say that $\mathbf{X}(\cdot)$ *solves* (SDE) provided

$$\text{(3)} \quad \mathbf{X}(t) = x_0 + \int_0^t \mathbf{b}(\mathbf{X}(s))\, ds + \int_0^t \mathbf{B}(\mathbf{X}(s))\, d\mathbf{W} \quad \text{for all times } t > 0.$$

To make sense of all this, we must:

- Construct Brownian motion $\mathbf{W}(\cdot)$: see Chapter 3.
- Define the *stochastic integral* $\int_0^t \cdots d\mathbf{W}$: see Chapter 4.
- Show that (3) has a solution, etc.: see Chapter 5.

And once all this is accomplished, there will still remain these **modeling problems:**

- Does (SDE) truly model the physical situation?
- Is the term $\boldsymbol{\xi}(\cdot)$ in "really" white noise or is it rather some ensemble of smooth but highly oscillatory functions? See Chapter 6.

As we will see later, these questions are subtle, and different answers can yield completely different solutions of (SDE).

1.3. ITÔ'S CHAIN RULE

Part of the trouble is the strange form of the chain rule in the stochastic calculus. To illustrate this, let us assume $n = m = 1$ and $X(\cdot)$ solves the SDE

$$\text{(4)} \qquad dX = b(X)dt + dW.$$

Suppose next that $u : \mathbb{R} \to \mathbb{R}$ is a given smooth function, $u = u(x)$. We ask: what stochastic differential equation does

$$Y(t) := u(X(t)) \quad (t \geq 0)$$

solve? Offhand, we would guess from (4) that

$$dY = u'dX = u'bdt + u'dW,$$

according to the usual chain rule, where $' = \frac{d}{dx}$.

This is wrong, however! In fact, as we will later see, Brownian motion is so irregular that

$$\text{(5)} \qquad dW \approx (dt)^{1/2}$$

in some heuristic sense. Consequently if we compute dY and keep all terms of order dt or $(dt)^{\frac{1}{2}}$, we obtain from (4) that

$$\begin{aligned} dY &= u'dX + \frac{1}{2}u''(dX)^2 + \cdots \\ &= u'(bdt + dW) + \frac{1}{2}u''(bdt + dW)^2 + \cdots \\ &= \left(u'b + \frac{1}{2}u''\right) dt + u'dW + \{\text{terms of order } (dt)^{3/2} \text{ and higher}\}. \end{aligned}$$

Here we used the "fact" that $(dW)^2 = dt$, which follows from (5). Hence

$$du(X) = \left(u'b + \frac{1}{2}u''\right) dt + u'dW, \tag{6}$$

with the extra term "$\frac{1}{2}u''dt$" not present in ordinary calculus. ▷

The strange looking expression (6) is an instance of *Itô's chain rule*, also known as *Itô's formula*. A major goal of this book is to provide a rigorous interpretation for calculations like these, involving stochastic differentials.

EXAMPLE 1. According to Itô's chain rule (6), the solution of the stochastic differential equation

$$\begin{cases} dY = YdW \\ Y(0) = 1 \end{cases}$$

is

$$Y(t) := e^{W(t)-\frac{t}{2}}$$

and *not* what might seem the obvious guess, namely $\hat{Y}(t) := e^{W(t)}$. □

EXAMPLE 2. Let $S(t)$ denote the (random) price of a stock at time $t \geq 0$. A standard model assumes that $\frac{dS}{S}$, the relative change of price, evolves according to the SDE

$$\frac{dS}{S} = \mu dt + \sigma dW$$

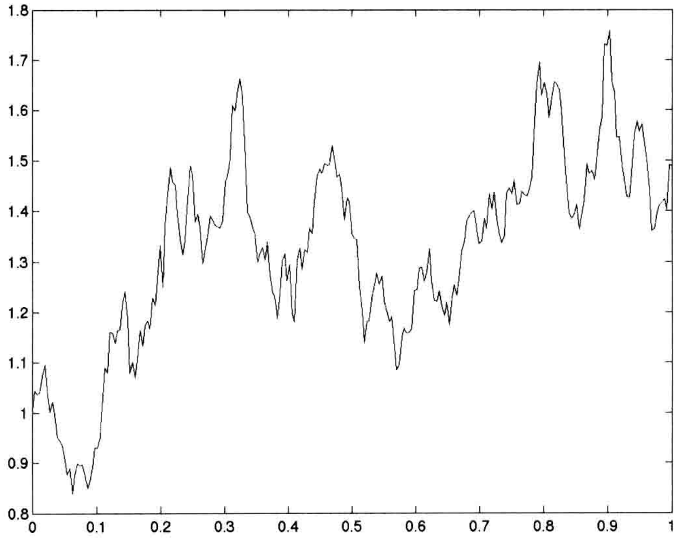
for certain constants $\mu > 0$ and σ, called respectively the *drift* and the *volatility* of the stock. In other words,

$$\begin{cases} dS = \mu S dt + \sigma S dW \\ S(0) = s_0, \end{cases}$$

where s_0 is the starting price. Using Itô's chain rule (6) once again, we can check that the solution is

$$S(t) = s_0 e^{\sigma W(t) + \left(\mu - \frac{\sigma^2}{2}\right)t}.$$

We will return to this example several times later. □



A trajectory for stock prices

Chapter 2

A CRASH COURSE IN PROBABILITY THEORY

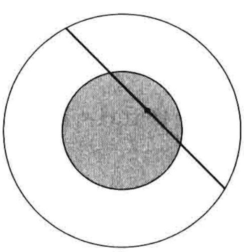
This chapter is an *extremely* rapid introduction to the measure-theoretic foundations of probability theory. See the Notes and Suggested Reading at the back of the book for recommendations of good textbooks that can provide full details of the proofs that we will sometimes omit.

2.1. BASIC DEFINITIONS

Let us begin with a puzzle:

2.1.1. Bertrand's paradox. Take a circle of radius 2 inches in the plane and choose a chord of this circle at random. What is the probability this chord intersects the concentric circle of radius 1 inch?

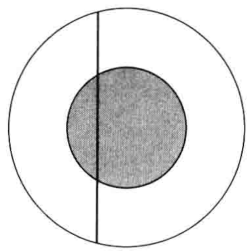
Solution #1. Any such chord (provided it does not hit the center) is uniquely determined by the location of its midpoint:



Thus

$$\text{probability of hitting the inner circle} = \frac{\text{area of inner circle}}{\text{area of larger circle}} = \frac{1}{4}.$$

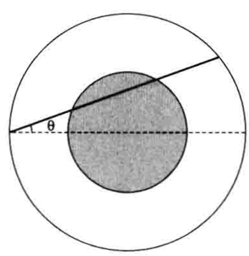
Solution #2. By symmetry under rotation we may assume the chord is vertical. The diameter of the large circle is 4 inches and the chord will hit the small circle if it falls within its 2-inch diameter:



Hence

$$\text{probability of hitting the inner circle} = \frac{2 \text{ inches}}{4 \text{ inches}} = \frac{1}{2}.$$

Solution #3. By symmetry we may assume one end of the chord is at the far left point of the larger circle. The angle θ that the chord makes with the horizontal lies between $\pm\frac{\pi}{2}$, and the chord hits the inner circle if θ lies between $\pm\frac{\pi}{6}$:



Therefore

$$\text{probability of hitting the inner circle} = \frac{\frac{2\pi}{6}}{\frac{2\pi}{2}} = \frac{1}{3}. \qquad \square$$

2.1.2. Probability spaces. Bertrand's paradox shows that we must carefully define what we mean by the term "random". The correct way to do so is by introducing as follows the precise mathematical structure of a *probability space*.

We start with a nonempty set, denoted Ω, certain subsets of which we will in a moment interpret as being "events".