CLASSICS IN MATHEMATICS

Kiyosi Itô · Henry P. McKean, Jr.

Diffusion Processes and their Sample Paths

扩散过程及其样本轨道

Springer

世界图书出版公司
www.wpcbj.com.cn

CLASSICS IN MATHEMATICS

Kiyosi Itô · Henry P. McKean, Jr.

Diffusion Processes and their Sample Paths

Springer

Kiyosi Itô Henry P. McKean, Jr.

Diffusion Processes and their Sample Paths

Reprint of the 1974 Edition

Springer

图书在版编目（CIP）数据

扩散过程及其样本轨道 = Diffusion Processes and their Sample Paths：英文/（日）伊藤清著．—影印本．—北京：世界图书出版公司北京公司，2010.2

数学经典教材

ISBN 978-7-5100-0526-8

Ⅰ.①扩… Ⅱ.①伊… Ⅲ.①扩散过程—教材—英文 Ⅳ.0211.6

中国版本图书馆 CIP 数据核字（2010）第 011244 号

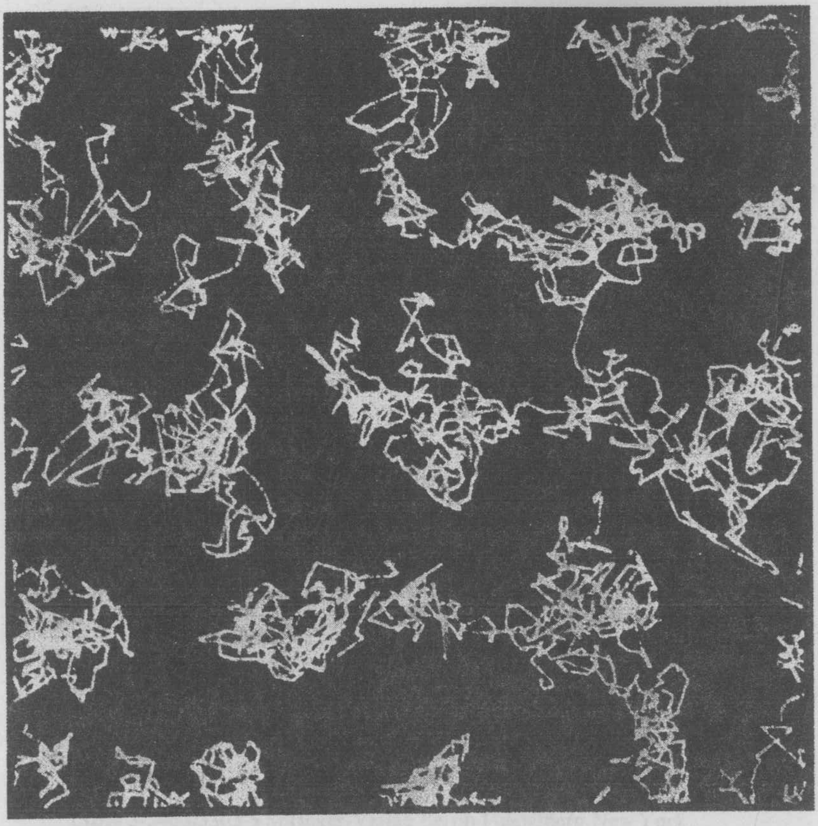
书　　名：	Diffusion Processes and their Sample Paths
作　　者：	Kiyosi Itô et al.
中 译 名：	扩散过程及其样本轨道
责任编辑：	高蓉　刘慧
出 版 者：	世界图书出版公司北京公司
印 刷 者：	三河国英印务有限公司
发　　行：	世界图书出版公司北京公司（北京朝内大街 137 号　100010）
联系电话：	010-64021602，010-64015659
电子信箱：	kjb@wpcbj.com.cn
开　　本：	24 开
印　　张：	14.5
版　　次：	2010 年 01 月
版权登记：	图字：01-2009-1075
书　　号：	978-7-5100-0526-8/O·742　　定　　价：38.00 元

DEDICATED TO

P. LÉVY

WHOSE WORK HAS BEEN

OUR SPUR AND ADMIRATION



Computer-simulated molecular motions, reminiscent of 2-dimensional Brownian motion. [From Alder, B. J., and T. E. Wainwright: Molecular motions. Scientific American **201**, no. 4, 113–126 (1959)].

Preface

ROBERT BROWN, an English botanist, observed (1828) that pollen grains suspended in water perform a continual swarming motion (see, for example, D'ARCY THOMPSON [*1*: 73—77]).

L. BACHELIER (1900) derived the law govering the position of a single grain performing a 1-dimensional BROWNian motion starting at $a \in R^1$ at time $t = 0$:

$$1) \qquad P_a[x(t) \in db] = g(t, a, b)\, db \qquad (t, a, b) \in (0, +\infty) \times R^2,$$

where g is the source (GREEN) function

$$2) \qquad g(t, a, b) = \frac{e^{-(b-a)^2/2t}}{\sqrt{2\pi t}}$$

of the problem of heat flow:

$$3) \qquad \frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial a^2} \qquad (t > 0).$$

BACHELIER also pointed out the MARKOVian nature of the BROWNian path expressed in

$$4) \qquad P_a[a_1 \leqq x(t_1) < b_1, a_2 \leqq x(t_2) < b_2, \ldots, a_n \leqq x(t_n) < b_n]$$
$$= \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} g(t_1, a, \xi_1)\, g(t_2 - t_1, \xi_1, \xi_2) \ldots$$
$$g(t_n - t_{n-1}, \xi_{n-1}, \xi_n)\, d\xi_1\, d\xi_2 \ldots d\xi_n \qquad 0 < t_1 < t_2 < \cdots < t_n$$

and used it to establish the law of maximum displacement

$$5) \qquad P_0\left[\max_{s \leqq t} x(s) \leqq b\right] = 2 \int_0^b \frac{e^{-a^2/2t}}{\sqrt{2\pi t}}\, da \qquad t > 0, b \geqq 0$$

(see BACHELIER [*1*]).

A. EINSTEIN (1905) also derived 1) from statistical mechanical considerations and applied it to the determination of molecular diameters (see, for example, A. EINSTEIN [*1*]).

BACHELIER was unable to obtain a clear picture of the BROWNian motion and his ideas were unappreciated at the time; nor is this surprising because the precise definition of the BROWNian motion involves a measure on the path space, and it was not until 1909 that E. BOREL published his classical memoir [1] on BERNOULLI trials. But as soon as the ideas of BOREL, LEBESGUE, and DANIELL appeared, it was possible to put the BROWNian motion on a firm mathematical foundation; this was achieved in 1923 by N. WIENER [1].

Consider the space of *continuous* paths $w: t \in [0, +\infty) \to R^1$ with coordinates $x(t) = w(t)$ and let $\mathbf{B}$ be the smallest BOREL algebra of subsets B of this path space which includes all the simple events $B = (w: a \leqq x(t) < b)$ $(t \geqq 0, a < b)$. WIENER established the existence of non-negative BOREL measures $P_a(B)$ $(a \in R^1, B \in \mathbf{B})$ for which 4) holds; among other things, this result attaches a precise meaning to BACHELIER's statement that *the Brownian path is continuous.*

P. LÉVY [2] found another construction of the BROWNian motion and, in his 1948 monograph [3], gave a profound description of the fine structure of the individual BROWNian path. LÉVY's results, with several complements due to D. B. RAY [4] and ourselves, will be explained in chapters 1 and 2, with special attention to the standard BROWNian *local time* (the *mesure du voisinage* of P. LÉVY):

6) $$\mathfrak{t}(t, a) = \lim_{b \downarrow a} \frac{\text{measure } (s: a \leqq x(s) < b, s \leqq t)}{2(b-a)}.$$

Given a STURM-LIOUVILLE operator $\mathfrak{G} = (c_2/2) D^2 + c_1 D$ $(c_2 > 0)$ on the line, the source (GREEN) function $p = p(t, a, b)$ of the problem

7) $$\frac{\partial u}{\partial t} = \mathfrak{G} u \qquad (t > 0)$$

shares with the GAUSS kernel g of 2) the properties

8a) $$0 \leqq p$$

8b) $$\int_{R^1} p(t, a, b)\, db = 1$$

8c) $$p(t, a, b) = \int_{R^1} p(t-s, a, c)\, p(s, c, b)\, dc \qquad t > s > 0.$$

Soon after the publication of WIENER's monograph [3] in 1930, the associated stochastic motions (diffusions) analogous to the BROWNian motion ($\mathfrak{G} = D^2/2$) made their debut; the names of W. FELLER and A. N. KOLMOGOROV stand out in this connection. At a later date (1946), K. ITÔ [2] proved that if

9) $$|c_1(b) - c_1(a)| + |\sqrt{c_2(b)} - \sqrt{c_2(a)}| < \text{constant} \times |b-a|,$$

then the motion associated with $\mathfrak{G} = (c_2/2)\, D^2 + c_1\, D$ is identical in law to the *continuous* solution of

10) $$a(t) = a(0) + \int_0^t c_1(a)\, ds + \int_0^t \sqrt{c_2(a)}\, db$$

where b is a standard BROWNian motion.

W. FELLER took the lead in the next development.

Given a MARKOVian motion with sample paths $w: t \to x(t)$ and probabilities $P_a(B)$ on a linear interval Q, the operators

11) $$H_t: f \to \int P_a[x(t) \in db]\, f(b)$$

constitute a *semi-group*:

12) $$H_t = H_{t-s}\, H_s \qquad (t \geqq s),$$

and as E. HILLE [*1*] and K. YOSIDA [*1*] proved,

13) $$H_t = e^{t\mathfrak{G}} \qquad (t > 0)$$

with a suitable interpretation of the exponential, where $\mathfrak{G}$ is the so-called *generator*.

D. RAY [*2*] proved that if the motion is *strict Markov* (*i.e.*, if it *starts afresh* at certain stochastic (MARKOV) times including the passage times $\mathfrak{m}_a = \min(t: x(t) = a)$, *etc.*), then $\mathfrak{G}$ is *local* if and only if the motion has *continuous* sample paths, substantiating a conjecture of W. FELLER; this combined with FELLER's papers [*4, 5, 7, 9*] implies that the generator of a strict MARKOVian motion with continuous paths (diffusion) can be expressed as a *differential operator*

14) $$(\mathfrak{G}\, u)\,(a) = \lim_{b \downarrow a} \frac{u^+(b) - u^+(a)}{m(a,\, b]},$$

where m is a non-negative BOREL measure positive on open intervals and, with a change of scale, $u^+(a) = \lim_{b \downarrow a} (b - a)^{-1}[u(b) - u(a)]$, except at certain singular points where $\mathfrak{G}$ degenerates to a differential operator of degree $\leqq 1$.

E. B. DYNKIN [*1*] also arrived at the idea of a strict MARKOV process, derived an elegant formula for $\mathfrak{G}$, and used it to make a simple (probabilistic) proof of FELLER's expression for $\mathfrak{G}$; the names of R. BLUMENTHAL [*1*] and G. HUNT [*1*] and the monographs of E. B. DYNKIN [*6, 8*] should also be mentioned in this connection.

Our plan is the following.

BROWNian motion is discussed in chapters 1 and 2 and then, in chapter 3, the general linear *diffusion* is introduced as a strict MARKOVian motion with continuous paths on a linear interval subject to possible annihilation of mass. $\mathfrak{G}$ is computed in great detail in chapter 4 using probabilistic methods similar to those of E. B. DYNKIN [5]; in the so-called non-singular case it is a differential operator

15) $$(\mathfrak{G}\, u)(a) = \lim_{b \downarrow a} \frac{u^+(b) - u^+(a) - \int_{(a,b]} u\, dk}{m(a,b]}$$

with u^+ and m as in 14), where now k is the (non-negative) BOREL measure that governs the annihilation of mass (such generators occur somewhat disguised in W. FELLER [9]).

Given $\mathfrak{G}$ as in 14) and a standard BROWNian motion with sample paths $w: t \to x(t)$, if $\mathfrak{t}$ is P. LÉVY's *Brownian local time* and if $\mathfrak{f}^{-1}$ is the inverse function of the local time integral

16) $$\mathfrak{f}(t) = \int \mathfrak{t}(t, \xi)\, m(d\xi),$$

then the motion $x(\mathfrak{f}^{-1})$ is identical in law to the diffusion attached to $\mathfrak{G}$, as will be proved in chapter 5, substantiating a suggestion of H. TROTTER; B. VOLKONSKIĬ [1] has obtained the same *time substitution* in a less explicit form.

Given $\mathfrak{G}$ as in 15), the associated motion can be obtained by *killing* the paths $x(\mathfrak{f}^{-1})$ described above at a (stochastic) time $\mathfrak{m}_\infty$ with conditional law

17) $$P_\bullet[\mathfrak{m}_\infty > t \mid x(\mathfrak{f}^{-1})] = e^{-\int \mathfrak{t}(\mathfrak{f}^{-1}(t), \xi)\, k(d\xi)}$$

as is also proved in chapter 5; in the special case of the *elastic Brownian motion* on $[0, +\infty)$ with generator $\mathfrak{G} = D^2/2$ subject to the condition $\gamma\, u(0) = (1-\gamma)\, u^+(0)$ $(0 < \gamma < 1)$,

$$\mathfrak{f} = \int_0^{+\infty} \mathfrak{t}(t, \xi)\, 2d\xi = \text{measure}\,(s : x(s) \geqq 0,\ s \leqq t),$$

$x(\mathfrak{f}^{-1})$ is identical in law to the classical *reflecting Brownian motion* $x^+ = |x|$, and 17) takes the simple form

18) $$P_\bullet[\mathfrak{m}_\infty > t \mid x^+] = e^{-\frac{\gamma}{1-\gamma}\mathfrak{t}^+(t,0)} \qquad (\mathfrak{t}^+ = 2\mathfrak{t}),$$

substantiating a conjecture of W. FELLER: that the elastic BROWNian motion ought to be the same as the reflecting BROWNian motion killed at the instant a certain increasing functional $e(\mathfrak{Z}_+ \cap [0, t))$ of the time t and the visiting set $\mathfrak{Z}^+ = (t : x^+(t) = 0)$ hits a certain level.

Details about the fine structure of the sample paths of the general linear diffusion with emphasis on local times, will be found in chapter 6. BROWNian motion in several dimensions is treated in chapter 7, and in chapter 8, the reader will find some glimpses of the general higher-dimensional diffusion.

Acknowledgements. W. FELLER has our best thanks, his ideas run through the whole book, and we shall think it a success if it pleases him.

We have also to thank R. BLUMENTHAL and G. HUNT who placed at our disposal their then unpublished results on MARKOV times, and H. TROTTER with whom we had many helpful conversations.

This book was begun at Princeton and the Institute for Advanced Study (1954/56) with the partial support of the Office of Ordnance Research, and continued at Kyôto (1957/58) with the aid of a Fulbright grant, at Hanover, N. H. and Cambridge, Mass. (1960) with the support of the Office of Naval Research, and at Stanford, Calif. with the support of the National Science Foundation (1962/63); to those institutions and agencies, also, we extend our warmest thanks. Finally, we must thank the staff of Springer-Verlag for their meticulous labors and cordial attitude to what must have been a difficult job.

Kyôto, Japan, and Cambridge, Mass., November 1964

K. Itô H. P. McKean, Jr.

The present edition is the same as the first except for the correction of numerous errors. Among those who helped us in this task, we would particularly like to thank F. B. KNIGHT.

September 1973 K. I., H. P. McK.

Contents

Prerequisites

The reader is expected to have about the same mathematical background as is needed to read COURANT-HILBERT [*1*, *2*]. Besides this, he should have mastered most of W. FELLER's book on probability [*3*] plus the topics listed below (see A. N. KOLMOGOROV [*2*] for a helpful outline and some of the proofs).

Algebras. Given a space W, a class A of its subsets is said to be an *algebra* if

$$1) \qquad W \in \mathsf{A}$$

$$2) \qquad A - B, A \cup B, A \cap B \in \mathsf{A} \quad \textit{in case} \quad A, B \in \mathsf{A}.$$

A is said to be a *Borel algebra* if, in addition,

$$3) \qquad \bigcup_{n \geqq 1} B_n, \bigcap_{n \geqq 1} B_n \in \mathsf{A} \quad \textit{in case} \quad B_n \in \mathsf{A}\,(n \geqq 1).$$

Borel extension. A class A of subsets of W is contained in a least *Borel algebra* B. B is the so-called *Borel extension* of A.

Probability measures. Consider a non-negative set function $P(C)$ defined on an algebra A. P is said to be a *probability measure* if

$$1) \qquad P(W) = 1$$

and

$$2) \qquad P(A \cup B) = P(A) + P(B) \qquad A, B \in \mathsf{A}, \quad A \cap B = \varnothing.$$

P is said to be a *Borel probability measure* if, in addition,

$$3\,\text{a}) \qquad P(B_n) \downarrow 0 \quad (n \uparrow +\infty) \quad B_n \downarrow \varnothing, B_n \in \mathsf{A}\ (n \geqq 1)$$

or, what is the same,

$$3\,\text{b}) \quad P\Big(\bigcup_{n \geqq 1} B_n\Big) = \sum_{n \geqq 1} P(B_n) \qquad B_n \cap B_m = \varnothing\,(n < m), \quad B_n \in \mathsf{A}(n \geqq 1), \bigcup_{n \geqq 1} B_n \in \mathsf{A}.$$

Kolmogorov extension. Given a BOREL probability measure P on an algebra A, there is a unique BOREL probability measure Q on the BOREL extension B of A that coincides with P on A: $Q(B) = \inf \sum_{n \geqq 1} P(A_n)$, where the infemum is taken over all coverings $\bigcup_{n \geqq 1} A_n$ of $B \in \mathsf{B}$ with $A_n \in \mathsf{A}\,(n \geqq 1)$; the triple (W, B, Q) is said to be a *probability space.*

Measurable functions. Given a space W and a BOREL algebra B of its subsets, a function $f: W \to [-\infty, +\infty]$ is said to be *measurable* B (or BOREL) if the ordinate set $f^{-1}[a, b) \in \mathsf{B}$ for each choice of $a < b$.

Integrals. Given a probability space (W, B, Q), the *integral* or *expectation* of a non-negative B measurable function f is defined to be

$$\begin{aligned} E(f) &= \int_W f\,dQ \\ &= \lim_{n\uparrow+\infty} \sum_{l\geqq 1} l\,2^{-n} Q[f^{-1}[(l-1)\,2^{-n}, l\,2^{-n})] \qquad Q[f^{-1}(+\infty)] = 0 \\ &= +\infty \qquad\qquad Q[f^{-1}(+\infty)] > 0. \end{aligned}$$

E, applied to such non-negative functions, satisfies

1) $$E(f) \geqq 0$$

2) $$E(1) = 1$$

3) $$E(f_1 + f_2) = E(f_1) + E(f_2)$$

4a) $$E(f_n) \uparrow E(f) \qquad f_n \uparrow f$$

4b) $$E(f_n) \downarrow E(f) \qquad f_n \downarrow f,\ E(f_1) < +\infty,$$

5) $$E(\underline{\lim}\, f_n) \leqq \underline{\lim}\, E(f_n).$$

$E(f, B) = E(B, f)$ is short for $\int_B f\,dQ$.

Products. Given probability spaces (W_1, B_1, Q_1) and (W_2, B_2, Q_2), the class A of finite sums A of disjoint rectangles $B_1 \times B_2$ ($B_1 \in \mathsf{B}_1$, $B_2 \in \mathsf{B}_2$) is an algebra and

$$Q(B_1 \times B_2) \equiv Q_1(B_1) \times Q_2(B_2)$$

can be extended to a BOREL probability measure on A; the *product* $Q_1 \times Q_2$ is the KOLMOGOROV extension of this Q to the BOREL extension $\mathsf{B} \equiv \mathsf{B}_1 \times \mathsf{B}_2$ of A.

Fubini's theorem. Given a $\mathsf{B}_1 \times \mathsf{B}_2$ measurable function f from $W_1 \times W_2$ to $[0, +\infty)$,

$$\int_{W_1\times W_2} f\,dQ_1 \times Q_2 = \int_{W_1} dQ_1 \int_{W_2} f\,dQ_2 = \int_{W_2} dQ_2 \int_{W_1} f\,dQ_1.$$

Infinite products. Given probability spaces (W_n, B_n, Q_n) $(n \geqq 1)$,

$$\mathsf{A}_n : A = B \times W_{n+1} \times W_{n+2} \times \textit{etc.}, \quad B \in \mathsf{B}_1 \times \mathsf{B}_2 \times \cdots \times \mathsf{B}_n$$

is a BOREL algebra and

$$Q(A) = Q_1 \times Q_2 \times \cdots \times Q_n(B)$$

is a BOREL probability measure on the algebra $\mathsf{A} = \bigcup_{n\geqq 1} \mathsf{A}_n$; the *infinite product* $\times_{n\geqq 1} Q_n$ is defined as the KOLMOGOROV extension of Q to the BOREL extension $\times_{n\geqq 1} \mathsf{B}_n$ of A.

Independence. Given a probability space (W, B, P), the BOREL algebras $\mathsf{B}_1, \mathsf{B}_2 \subset \mathsf{B}$ are said to be independent if

$$P(B_1 \cap B_2) = P(B_1)\, P(B_2) \qquad B_1 \in \mathsf{B}_1, B_2 \in \mathsf{B}_2;$$

the BOREL algebras B_n $(n \geqq 1)$ are independent, if, for each $n \geqq 1$, B_n is independent of the BOREL extension of $\bigcup_{l \neq n} \mathsf{B}_l$; the sets $B_n \in \mathsf{B}$ $(n \geqq 1)$ are independent if the algebras $\mathsf{B}_n = [\varnothing, B_n, W - B_n, W]$ $(n \geqq 1)$ are such; the B measurable functions f_n $(n \geqq 1)$ are independent if the BOREL extensions F_n of their ordinate sets $f_n^{-1}[a, b)$ $(a < b)$ are such; the B measurable function f is independent of the BOREL algebra $\mathsf{A} \subset \mathsf{B}$ if the BOREL extension of its ordinates sets is independent of A; *etc.* $E(f_1 f_2) = E(f_1)\, E(f_2)$ if f_1 is independent of f_2.

Kolmogorov 01 law. If A_n $(n \geqq 1)$ are independent subalgebras of B, if B_n is the BOREL extension of $\bigcup_{l \geqq n} \mathsf{A}_l$, and if $B \in \bigcap_{n \geqq 1} \mathsf{B}_n$, then $P(B) = 0$ or 1.

Strong law of large numbers. Given independent, non-negative, B measurable functions $f_n (n \geqq 1)$ with the same distribution,

$$P\left[\lim_{n \uparrow +\infty} n^{-1} \sum_{k \leqq n} f_k = E(f_1)\right] = 1.$$

Borel-Cantelli lemmas. Given events $B_n \in \mathsf{B}$ $(n \geqq 1)$, if

$$\sum_{l \geqq 1} P(B_l) < +\infty, \quad \text{then} \quad P\left[\bigcap_{n \geqq 1} \bigcup_{l \geqq n} B_l\right] = 0; \quad \text{if} \quad \sum_{l \geqq 1} P(B_l) = +\infty$$

and if the events are independent, then $P\left[\bigcap_{n \geqq 1} \bigcup_{l \geqq n} B_l\right] = 1$.

Conditional expectations. Given a BOREL subalgebra A of B and a non-negative B measurable function f, *the conditional expectation* $E(f \mid \mathsf{A})$ *of f relative to* A is defined to be the class of A measurable functions g such that $E(g, A) = E(f, A)$ $(A \in \mathsf{A})$; two such functions differ on a null set $\in \mathsf{A}$. $E(f \mid \mathsf{A})$ is the RADON-NIKODYM derivative of $Q(A) = E(f, A)$ with respect to the restriction of P to A. $E(f \mid \mathsf{A})$ (applied to non-negative f) satisfies

1) $$E(f \mid \mathsf{A}) \geqq 0$$

2) $$E(1 \mid \mathsf{A}) = 1$$

3) $$E(f_1 + f_2 \mid \mathsf{A}) = E(f_1 \mid \mathsf{A}) + E(f_2 \mid \mathsf{A})$$

4) $$E(f_n \mid \mathsf{A}) \uparrow E(f \mid \mathsf{A}) \qquad f_n \uparrow f$$

5) $$E(E(f \mid \mathsf{A}_2) \mid \mathsf{A}_1) = E(f \mid \mathsf{A}_1) \qquad \mathsf{A}_2 \supset \mathsf{A}_1$$

and

6) $$E(E(f \mid \mathsf{A})) = E(f);$$

in addition,

7) $$E(f \mid \mathsf{A}) = E(f)$$