BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 34, Number 1, January 1997
S 0273-0979(97)00697-6

Introduction to the theory of diffusion processes, by N.V. Krylov, Transl. Math. Monographs, vol. 142, Amer. Math. Soc., Providence, RI, 1995, xi + 271 pp., \$95.00, ISBN 0-8218-4600-0

In its preface the author writes:

> *The aim of this book is to present the basic elements of the theory of diffusion processes, at the contemporary level of mathematical rigor, in a way understandable to a reader completely ignorant even of probability theory, to say nothing of the theory of stochastic processes. ...*

The probabilistic idea of a diffusion can be loosely described by the following model: imagine a pollen grain floating down a river. Its movement should, on average, follow the current of the river and be characterized by an ordinary differential equation of the form

$$dx(t) = b(t, x(t))dt, \tag{1}$$

where "$b(t, x)$" represents the "sensitivity to the current" at time t and location x. This can account, for example, for the difference between flat river bottoms (small b) and steep ones (large b, river rapids). Incorporation of t allows time to come into play, as currents change for example during the spring melting of snow. There is of course another force: that of the bombardment of the pollen grain by water molecules that causes rigorous but tiny oscillations of the particle as it floats downstream. This is the so-called Brownian motion component, named after the English botanist Robert Brown, who first discovered and described it in 1828 [2]. A model incorporating the Brownian motion element might look like

$$dx(t) = b(t, x(t))dt + \sigma(t, x(t))dB(t), \tag{2}$$

where B is a Brownian motion process.

Unaware of Brown's work, A. Einstein predicted Brownian motion in his 1905 paper [4]; he was concerned with demonstrating the molecular nature of matter and estimating Avogadro's number. He was not able to show his model existed as a rigorous stochastic process; this was done later by N. Wiener [9], who used the ideas of Borel and Lebesgue (measure theory) that were not available to Einstein. Nevertheless Wiener was unable to construct a satisfactory "stochastic integral" to make sense of the intuitive formulation equation (2). Thus the intuitive model (2), now ubiquitous, did not play a role in the early development of probabilistic diffusion theory.

As in all things, it was not clear at first exactly what a diffusion is. A very early approach was that of A. Fick (1855), who derived the mathematical equation for heat conduction and thus first put diffusions on a quantitative basis. Fick's first and second laws of diffusions in a one-dimensional setting are:

$$F = -D\frac{\partial C}{\partial x}, \tag{3}$$

1991 *Mathematics Subject Classification.* Primary 58G32, 60J60, 60J65; Secondary 35R60, 60G44, 60H15.

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2}, \tag{4}$$

where F is the rate of transfer per unit area of section, C the concentration of the diffusing substance, x the space coordinate, and D the diffusion coefficient (cf. Crank [3]). It would not be exaggerating to say that it was the second law of Fick (4) that laid the foundation of the works of Bachelier (1900) and later A. Einstein (1905) in which the probability density of the Brownian motion was proved to satisfy the heat equation (4).

After Wiener put Brownian motion (also now called the Wiener process) into rigorous mathematics, the development of diffusions was first carried out by A.N. Kolmogorov (1931) and W. Feller (1936). In a sense their methods were purely analytic and "*had little if anything to do with probability theory*" (cf. Strook and Varadhan [7, Introduction]). Roughly speaking, they proved the existence of a continuous (strong) Markov process by proving the existence of its transition density function as the solution of the forward and backward Kolmogorov equations—a pair of parabolic partial differential equations. The analytical approach, as well as the theory of parabolic PDEs, later benefited from the theory of semi-groups of linear operators developed by E. Hille and K. Yoshida in the 1940's. This now constitutes one of the fundamental approaches for the study of Markov processes.

Another method of studying a diffusion process, pioneered by P. Lévy (1939), is to study its sample paths. The milestone in this direction is undoubtedly Itô's invention of a *stochastic integral* and consequently *stochastic differential equations*. From 1942 to 1946, K. Itô published a sequence of papers developing a profound idea: most diffusion processes $x(t)$ can be representable by a stochastic differential equation of the type (2). Here $b(t,x)$ is the instantaneous mean, and $a(t,x) = \sigma(t,x)\sigma^{\mathrm{T}}(t,x)$ is the instantaneous variance of the diffusion at time t and position x, and B is a Brownian motion (Wiener process). Knowing that the paths of a Brownian motion are nowhere differentiable (proved by Wiener and Lévy in earlier years), Itô successfully established a new meaning of the second integration in the right side of (2), followed by a theory of stochastic calculus for the new "stochastic integral", including a change of variables formula known today as *Itô's formula* and a theory of stochastic differential equations. It should be mentioned here that the term "stochastic differential equation" was introduced almost a decade before Itô's first paper on the subject, in 1934 (and then 1938) by S. Bernstein. Along the lines of Bernstein, Gikhman (1951) proved the solvability of the Cauchy problem for second-order *degenerate* parabolic PDE, which has been viewed as one of the significant events that show the "advantage" of the probabilistic methods in analysis.

After almost a century since the first mathematical model of Brownian motion was established, especially after A. Kolmogorov's *Grundbegriffe der Wahrscheinlichkeitsrechnung* (1933) (from then on measure theory became the basic mathematical tool of probability theory, which made the latter more accessible to those mathematicians who may not have a strong skill in, if they do not hate, counting different ways to distribute balls in urns), the theory of diffusion processes has now reached the plateau where both analysts and probabilists can equally enjoy their joint venture on the subject. The main bridge that links the two fields is still the relationship between parabolic (or elliptic) partial differential equations and diffusion processes. Such a relationship has been developed extensively in the spirit of Kolmogorov's equations and Itô's intuition, as was explained in (2). One of the problems can be stated as follows. Consider the parabolic partial differential equation:

$$
\begin{aligned}
&\frac{\partial u}{\partial t} + \mathcal{L}u - cu + f = 0, \qquad \text{on } [0,T) \times \mathbb{R}^d, \\
&u(T,x) = g(x); \qquad x \in \mathbb{R}^d,
\end{aligned} \tag{5}
$$

where $\mathcal{L}$ is a second order differential operator:

$$
\mathcal{L}(t,x)\varphi(x) = \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(t,x)\varphi_{x^i,x^j}(x) + \sum_{i=1}^{d} b^i(t,x)\varphi_{x^i}(x) \tag{6}
$$

with $(a^{ij}) \geq 0$. Now suppose that $a = \sigma\sigma^{\mathrm{T}}$ for some matrix-valued function $\sigma = \sigma(t,x)$. Consider a diffusion process represented by the stochastic differential equation:

$$
\xi(t) = x + \int_s^t b(\tau, \xi(\tau))d\tau + \int_s^t \sigma(\tau, \xi(\tau))dw(\tau), \qquad 0 \leq s \leq t \leq T, \tag{7}
$$

where w is a standard Brownian motion and the second term on the right side of (7) is an Itô stochastic integral. Obviously the solution of (7) will depend on the initial state x and initial time s. To make such a dependence more explicit, we denote the solution of (7) by $\xi = \xi^{s,x}$. Now we ask the following question: how does $\xi^{s,x}$ relate to the differential equation (5)? It turns out that if we define a function

$$
v(s,x) = E\Big\{g(\xi^{s,x}(T))e^{-\int_s^T c(\tau,\xi^{s,x}(\tau))d\tau} + \int_s^T f(t,\xi^{s,x}(t))e^{-\int_s^t c(\tau,\xi^{s,x}(\tau))d\tau}dt\Big\}, \tag{8}
$$

then under suitable conditions on the functions b, σ, c, f and g, v is a solution to (7). Such a solution is called the *probabilistic solution* to the partial differential equation, and formula (8) is customarily called the *Feynman-Kač formula*, due to the fundamental work by both men on this subject during 1948–1950.

There are different ways to look at the above problem. The "traditional" way is to derive probabilistic theorems by using results in partial differential equations. For example, assuming *a priori* that the differential equation (5) has a (unique) classical solution $v \in C^{1,2}$, then a simple application of Itô's formula will show that v has to be of the form (8). This approach is, however, not satisfactory from the probabilists' point of view. In his remarkable paper in 1950, M. Kac [5] stated that "... once this (traditional approach) has been accomplished we find ourselves outside the field of probability and in a domain where methods of long standing are immediately available." For half a century the probabilists have been trying to study the aforementioned problem in another direction: prove analytic results using probabilistic methods. Early research in this direction includes Kac's 1950 paper on the connections between probability theory and differential and integral equations, in which, among other things, he derived the distribution of eigenvalues of certain integral equations, using probabilistic methods, and Gikhman's paper in 1951 on *degenerate* partial differential equations, which gave further evidence that a probabilistic method might work better than pure analytic ones.

Let us now come back to the problem regarding (5) and (8). The natural question to the reversal of the "traditional approach" then becomes: can one show directly that the function defined by (8) is really a solution to equation (5), without knowing a priori that the latter has a smooth solution? This question becomes more realistic when one does not have all the "nice" conditions on the coefficients that guarantee the existence and uniqueness of the classical solution to (5) by the theory of partial differential equations. (Here one should note that, compared to the methods of partial differential equations, the advantage of studying probabilistic solutions lies in that one has *a priori* a prospective candidate for the solution and the only thing remaining is to verify that it is indeed a solution.) To answer the question, however, one is required to investigate the differentiability of the function defined by (8) in both variables, which would necessarily lead to an interchange of limits (more precisely, interchange the derivatives and integrals). The problem will become more complicated when the differential equation (5) is defined on a domain (e.g., (5) holds for $(t, x) \in [0, T) \times D$, where $D \neq \mathbb{R}^d$). In that case the upper limits in the integrals in expression (8) have to be replaced by the "first exit time"

$$\tau(t, x) = \inf\{s > t : (s, X_s^x) \notin [t, T) \times D\}.$$

What makes the problem more formidable is that the stopping time $\tau(t, x)$, as a function of t and x, may not even be continuous, complicating any attempts to differentiate (8).

Here, finally, is where this book makes its contribution to the literature. As a main theme of the book, the diffusion processes are studied in a "probabilistic" way; that is, they are studied via Itô's stochastic differential equations. We note here that although there are few diffusion processes, such as reflecting Brownian motion (i.e., $X_t = |B_t|$, where B is a Brownian motion), that cannot be directly represented in the form (2), it is indeed a more convenient way for a beginner to access the notion of a diffusion process "probablistically" without being discouraged by too many probablistic subtleties, like those involved in the definition of a *Strong Markov Process*. The book contains ideas of the author that have not been systematically presented in any other standard texts. Tremendous efforts are made to explore the probabilistic solutions of partial differential equations, reflecting the interest of the author. As an "introduction" to the theory, the book is *elementary* enough, even for those who have not had serious training in probability theory (by which I mean those who, for example, do not necessarily have ready knowledge of why the number of car accidents occurring to a randomly selected driver in this country during a one-year period may have a negative binomial distribution). But on the other hand, the book is *rich* enough even for specialists in the field, as it contains many ideas which are different from the classical books on the subject. As a book with only 255 pages for its main body, it is condensed enough to furnish detailed proofs for all the theorems (once they are so named), ranging from those as elementary as Lebesgue's Dominated Convergence Theorem, the Monotone Class Theorem, and Fubini's theorem, to those as advanced as the regularity of probabilistic solutions of partial differential equations and quasiderivatives of the solutions to stochastic differential equations. Also found in passing are those "classical" results like the construction of the Wiener process, the Burkholder-Davis-Gundy inequalities (including a proof for multidimensional cases with dimension-independent constants, which is often over looked by standard texts), and existence and uniqueness theorem of stochastic differential equations under monotonicity conditions, ... (do not forget there are only 255 pages!). I would like to indicate here that even for the well-understood materials in modern stochastic analysis such as the notions of martingales, Itô's stochastic integrals, stochastic differentials, etc., one can still find many special flavors offered by the author, which not only make the book self-contained, but provide different angles of looking at the traditional results: among them, say, Itô's formula. Although this is a translated version of the original book (published in Russian in 1989), these ideas are still noteworthy today.

Finally, I would like to quote a sentence from the preface of the book of N. Ikeda and S. Watanabe [4] (well-known to most people who have been exposed to stochastic differential equations): "... *We especially regret that we could not include the important works by N.V. Krylov on the estimates for the distributions of stochastic integrals and their applications.* ..." Although in fact some estimates of this kind can be found in the author's earlier book (1980) [6], which is well-known to researchers in stochastic control and is more specialized, the book under review should be helpful to those who are interested in the subject but new to classical probability theory.

A last point: it might be surprisingly helpful to read the preface before you flip into any sections of the book for the first time! Not only does the preface contain the motivation of the author, which of course is normal, it contains a "user's guide". Among other things, remembering the numbering system of the book should probably be the first thing to do. A list of notations is also given at the end of the book.

References

1. L. Bachelier, *Théorie de la spéculation*, Ann. Sci. École Norm. Sup. **17** (1900), 21–86.
2. R. Brown, *A Brief Account of Microscopical Observations Made in the Months of June, July, and August, 1827, on the Particles Contained in the Pollen of Plants; and on the General Existence of Active Molecules in Organic and Inorganic Bodies*, Philosophical Magazine N. S. **4** (1928), 161–173.
3. J. Crank, *The Mathematics of Diffusion*, Oxford at the Clarendon Press (1956). MR **18:**616c
4. A. Einstein, *On the Movement of Small Particles Suspended in a Stationary Liquid Demanded by the Molecular-Kinetic Theory of Heat*, Ann. Physik **17** (1905) [See also A. Einstein: *Investigations on the Theory of the Brownian Movement* (R. Fürth, ed.) Dover, New York (1959)]. MR **17:**1035g
5. N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Publishing Company, 1989. MR **90m:**60069
6. M. Kac, *On Some Connections between Probability Theory and Differential and Integral Equations*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, University of California Press, 1951, pp. 189–215. MR **13:**568b
7. N.V. Krylov, *Controlled Diffusion Processes*, Springer-Verlag, 1980. MR **82a:**60062
8. D.W. Stroock and S.R.S. Varadhan, *Kiyosi Itô Selected Papers* (Stroock and Varadhan, eds.), Springer-Verlag, 1987. MR **88d:**60001
9. N. Wiener, *Differential Space*, J. Math. Phys. **2** (1923), 131–174.

Jin Ma
Purdue University
E-mail address: majin@math.purdue.edu