Reinhard Mahnke, Jevgenijs Kaupužs,
and Ihor Lubashevsky

Physics of Stochastic Processes

How Randomness Acts in Time

WILEY-VCH Verlag GmbH & Co. KGaA

Reinhard Mahnke, Jevgenijs Kaupužs, and Ihor Lubashevsky

Physics of Stochastic Processes

Reinhard Mahnke, Jevgenijs Kaupužs,
and Ihor Lubashevsky

Physics of Stochastic Processes

How Randomness Acts in Time

WILEY-VCH Verlag GmbH & Co. KGaA

The Authors

Dr. Reinhard Mahnke
University of Rostock
Institute of Physics
Rostock, Germany
reinhard.mahnke@uni-rostock.de

Dr. Jevgenijs Kaupužs
University of Latvia
Mathematical and Computer Science
Riga, Latvia

Prof. Ihor Lubashevsky
Russian Academy of Sciences
Prokhorov General Physics Institute
Moscow, Russia

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche National-bibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Composition Laserwords Private Ltd., Chennai, India
Printing betz-druck GmbH, Darmstadt
Bookbinding Litges & Dopf GmbH, Heppenheim

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN: 978-3-527-40840-5

Contents

Physics of Stochastic Processes: How Randomness Acts in Time
Reinhard Mahnke, Jevgenijs Kaupužs and Ihor Lubashevsky

ISBN: 978-3-527-40840-5

Preface

A wide variety of systems in nature can be regarded as many-particle ensembles with extremely intricate dynamics of their elements. Numerous examples are known in physics, e.g. gases, fluids, superfluids, electrons and ions in conductors, semiconductors, plasma, nuclear matter in neutron stars, etc. Such macroscopic systems are typically formed of 10^{23}–10^{28} particles, with essentially erratic motion, so a description of the individual elements is really hopeless. However, it is not actually necessary for practical tasks because on the macroscopic level we are dealing only with cumulative effects expressed in macroscopic variables. At this level, details of the individual particle motion are averaged – only the mean characteristics are essential for a description of the system dynamics. The deviation of an individual particle from the mean behavior can then be taken into account, if necessary, in terms of random fluctuations characterized again by some mean parameters. It should be noted that many systems of a nonphysical nature, e.g. fish swarms and bird flocks, vehicle ensembles, pedestrians or stock markets can be regarded (leaving aside social aspects of their behavior) as ensembles of interacting particles.

There are several approaches to tackling many-particle systems. Dealing with a physical object whose dynamics is based on the Newtonian or Schrödinger equation, it is possible to start from the microscopic description and directly write down the corresponding governing equations. Then a rather small part of the system comprising, e.g. one, two, or three particles should be singled out and considered individually. The effect of the other elements on this selected part is taken into account on the average. Roughly speaking, it is in just this way that the notion of a thermal heat bath is introduced – a small part of the system under consideration is singled out and its interaction with the neighboring particles is simulated in terms of stochastic energy exchange with a certain reservoir characterized by some temperature. This approach is the most rigorous and, as a result, the most difficult way of constructing a bridge between the microscopic description dealing with individual particles (atoms, molecules, etc.) and the mesoscopic continuum fields, e.g. density, temperature, and pressure. Typically this bridge is implemented in the form of a partial differential equation or a system of such equations governing the distribution function of the particle or the collection of particles. We

Physics of Stochastic Processes: How Randomness Acts in Time
Reinhard Mahnke, Jevgenijs Kaupužs and Ihor Lubashevsky

ISBN: 978-3-527-40840-5

should point out that the many-particle ensembles governed by the laws of classical Newtonian mechanics exhibit chaotic dynamics rather than stochastic dynamics. The term chaotic refers to systems whose evolution from the initial conditions is rigorously determined by nonrandom dynamics. Given the initial conditions, the dynamics of such a deterministic system is formally predictable, so it is not stochastic in a rigorous sense. However, if the system trajectories are located inside a given bounded region and are nonperiodic, then their temporal and spatial structure is highly intricate and in fact looks like that of stochastic random paths. Moreover, these trajectories pass all the standard tests for randomness so, for practical purposes, they can be regarded as stochastic. This observation is actually one of the ways to justify introducing a thermal heat bath characterized by *stochastic* energy exchange (between the small part of the system under consideration and the surrounding particles treated as a random reservoir at a particular temperature). The same comments concerning the relationship between chaos and stochasticity should be addressed in time series analysis. Without knowledge of the origin it is practically impossible to distinguish between chaotic and stochastic behavior.

Another way to treat many-particle systems is to construct a collection of microscopic equations governing, e.g. the dynamics of individual particles where, for a given particle, the influence of the other particles is described in terms of both systematic and random forces. The notion of random forces again enables one to derive the corresponding partial differential equations for the distribution function of particles. Indeed the random forces should be introduced in such a way that these governing equations for the distribution function coincide with those obtained via the approach of the previous paragraph. As far as social, ecological, and economic systems are concerned, postulating the appropriate form of the random forces seems to be the only way to construct a mathematical description. This is due to such systems being open. Moreover the behavior of their elements is so intricate and multifactorial that a closed mathematical description is likely to be impossible.

The latter approach is precisely the main topic of this book. It is based on probability theory or, more specifically, on the notion of stochastic processes and the relevant mathematical constructions, which are the subject matter of Chapters 1 and 2 (see the layout of the book shown at the end of this preface, page XVII). On the microscopic level stochastic trajectories of the system motion are the basic elements of the probabilistic description. It is assumed that different stochastic realizations of the random force are independent and also that the motion of particles does not have long-time memory. The notion of stochastic trajectories has a long history, possibly going back to the scientific poem De Rerum Nature (On the Nature of Things, circa 60 BC) by Titus Lucretius Carus. Although very little is known about the Roman philosopher, it seems he described the random motion of dust particles in air. In 1785 Jan Ingenhousz observed the irregular motion of coal dust particles on the surface of alcohol. Then, in 1827, the British botanist Robert Brown also discovered random highly erratic motion of pollen particles floating in

water under the microscope. Since that time this phenomenon has been called *Brownian motion*. The generalization of the observed phenomena gave rise to the notion of random walks where the walker dynamics is governed by both regular and stochastic forces.

The first person who proposed a mathematical model for Brownian motion appears to be Thorvald N. Thiele in 1880. This was followed independently by Louis Bachelier in 1900 in his PhD thesis Théorie de la Spéculation devoted to a stochastic analysis of the stock and option markets. He worked out mathematically the idea that the stock market prices are essentially sums of independent, bounded random changes. The results put forward by Bechelier led to a flash of interest in stochastic processes and corresponding probabilistic approaches. However, it was Albert Einstein's independent research into the problem in his 1905 paper that brought the solution to the attention of physicists (see, e.g. *Brownian motion* – Wikipedia, The Free Encyclopedia, 23 October 2007).

The qualitative explanation of Brownian motion as a kinetic phenomenon was put forward by several authors. As mentioned above, it is possible to add random forces to the dynamical laws which were proposed for the first time by the French physicist Paul Langevin. (This resulted in a new mathematical field now known as stochastic differential equations.) The appropriate partial differential equations for the distribution function could then be derived based on the Langevin equation.

It is possible to develop the probabilistic description of a stochastic process in the opposite way – the equations governing the distribution function are postulated and the appropriate Langevin equation is constructed in order to give these equations. This idea was implemented for the first time by Albert Einstein deriving the diffusion equation for Brownian particles in his famous paper Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen published in *Annalen der Physik* (1905). The equation for diffusive motion was then developed by Adriaan Fokker (1914) and later more completely and generally by Max Planck (1918), leading to the transport equation now known as the Fokker–Planck equation. There are also approaches to describing random processes in discrete phase spaces based on ordinary differential equations (e.g. the probability balance law known as the master equation). If a stochastic process develops in discrete space and time the cellular automata models can be used, which form a distinct branch of the theory of stochastic processes. These problems and their mutual interrelationship are considered in Chapters 3–5 which adopt one of the main assumptions in the theory of stochastic processes, the Markovian approximation. According to this approximation, the displacement of a wandering particle on mesoscopic scales can be considered as the result of many small independent identically distributed steps. This reasoning is very close to what is now called a Kramers–Moyal expansion and has been used to derive the Fokker–Planck equation.

To elucidate the main notions of stochastic processes, Chapters 6 to 8 consider in detail some rather simple examples of discrete random walks and

continuous Brownian motion. In particular, they touch on the problem of reaching a boundary for the first time. This problem plays an essential role in many physical phenomena such as escaping from a potential well, anomalous diffusion in fractal media, heat diffusion in living tissue, etc.

As mentioned above, the notion of stochastic processes can form the initial mathematical description for objects of a nonphysical nature, e.g. social, ecological, and economic systems. This is a novel branch of science where only the first steps have been taken. It turns out that, in spite of their nonphysical nature, the cooperative phenomena in such systems (for example, self-organization processes in congested traffic or motion of pedestrians and social animals) exhibit a wide variety of properties commonly met in physical systems (for example in gas–liquid phase transitions, spinodal decomposition in solid solutions, ferromagnetic transitions, etc.). So, in some sense, the stochastic description of many-particle ensembles with strong interaction between their elements is of a more general nature than the basic laws of the corresponding mechanical systems.

These questions are considered in Chapters 9 and 10 dealing with the aggregation of particles out of an initially homogeneous situation. This phenomenon is well known in physics, as well as in other branches of the natural sciences and engineering. The formation of bound states as an aggregation process is due to self-organization. The formation of car clusters (jams) at overcritical densities in traffic flow is an analogous phenomenon in the sense that cars can be considered as (strong asymmetrically) interacting particles. The development of traffic jams in vehicular flow is an everyday example of the occurrence of nucleation and aggregation in a system of many point-like cars. Traffic jams are a typical signature of the complex behavior of the many-car system. The master equation approach to stochastic processes can be applied to describe the car-cluster formation on a road in partial analogy to droplet formation in a supersaturated vapor.

This jamming transition is very similar to conventional phase transitions appearing in the study of critical phenomena. Traffic-like collective movements are observed at almost all levels of biological systems. We study the energy balance of motorized particles in a many-car system. New dynamical features, such as steady state motion with energy flux, also appear. This phenomenon is also observed in a system of active Brownian particles with energy take-up and energy dissipation.

The last two Chapters 11 and 12 are devoted to some modern applications in the physics of stochastic processes. First, we consider nonequilibrium phase transition induced by noise or caused by dynamical traps. Probably, the former type of transition can only be described using the Langevin equation with multiplicative noise, that is, stochastic equations for which the intensity of the random forces depends on the system state. During the last few decades it has been demonstrated that the behavior of such systems can be rather complex; in particular, the appearance of new states can be induced by noise as its intensity increases and attains certain critical values. The second type of phase

transition seems to be a commonly encountered phenomenon in systems, for example, congested traffic flow, where the human factor is essential. Such transitions are due to the existence of some regions in the corresponding phase space where the system dynamics is stagnated. Following the notions introduced in the theory of Hamiltonian dynamics with complex behavior, these regions are called dynamical traps.

Finally, we turn to the kinetics of many-particle systems. The zero-range process, introduced in 1970 by Frank Spitzer as a system of interacting random walks, serves as a generic model in which rigorous large-scale description of the dynamics for arbitrary initial densities is possible in terms of a hydrodynamic equation for the coarse-grained particle density. It allows one to derive a criterion for phase separation in one-dimensional driven systems of interacting particles, e.g. in traffic flow, as well as to describe nontrivial features of stochastic dynamics like metastability.

Nowadays another aspect which should be taken into account is non-Gaussian behavior; that is, long-tail distributions which are observed in stock market data as well as in transportation theory. In this sense, applied sciences such as sociology and econophysics, biophysics and engineering, consider extreme events in nature and society and deal with effects (like material rupture) which can be investigated only by the probabilistic approach.

In concluding this preface, we would like to underline the spirit in which this book is intended. Here we are in agreement with other authors of books on random processes; in particular, A. J. Chorin and O. H. Hald in *Stochastic Tools in Mathematics and Science* state: 'When you asked alumni graduates from universities in Europe and US moving into nonacademic jobs in society and industry what they actually need in their business, you found that most of them did stochastic things like time series analysis, data processing etc., but that had never appeared in detail in university courses'. So the general aim of the present book is to provide stochastic tools for the multidisciplinary understanding of random events and to illustrate them with many beautiful applications in different disciplines ranging from econophysics to sociology. The central problem under consideration in this book is thus the theoretical modeling of complex systems, that is, many-particle systems with nondeterministic behavior. In contrast to the established classical deterministic approach based on trajectories, we develop and investigate probabilistic dynamics using stochastic tools, such as stochastic differential equations, Fokker–Planck and master equations, to obtain the probability density distribution. The stochastic technique provides an exact and more understandable background to describe complex systems.

The authors have been working for years on the problems to which this monograph is devoted. Nevertheless, the book is also the result of long-standing scientific cooperation with a number of colleagues from all over the world. The authors thank Werner Ebeling, Rudolf Friedrich, Vilnis Frishfelds, Namik Gusein-Zade, Peter Hänggi, Rosemary Harris, Andreas Heuer, Dirk Helbing, Alexander Ignatov, Andris Jakovičs, Holger Kantz,

Boris Kerner, Reinhart Kühne, Kai Nagel, Holger Nobach, Gerd Röpke, Yuri and Michael Romanovsky, Anri Rukhadze, Andreas Schadschneider, Michael Schreckenberg, Gunter M. Schütz, Lutz Schimansky-Geier, Yuki Sugiyama, Steffen Trimper, Peter Wagner and Hans Weber, for fruitful discussions.

Special thanks are due to Friedrich Liese from the Institute of Mathematics at Rostock University for delivering a joint lecture series on *Stochastic Processes* from the mathematical (F. Liese) as well as physical (R. Mahnke) points of view and for preparing Chapter 1 of this book – *Fundamental Concepts*.

The contents of this book took shape over several years, based on research and lectures performed at different locations. One of the recent lecture presentations took place in the summer term of 2007 at Rostock University. The authors have benefited from the contributions of a number of students. We would like to express our gratitude to the active participants, Michael Brüdgam, Matthias Florian, Peter Grünwald, Hannes Hartmann, Julia Hinkel, Bastian Holst, Thomas Kiesel, Susanne Killiches, Knut Klingbeil, Christof Liebe, Daniel Münzner, Ralf Remer, Elisabeth Schöne, Philipp Sperling, Marten Tolk, Andris Voitkans, Norman Wilken, and Mathias Winkel, together with many other students, PhD students and co-workers.

Finally, we would like to acknowledge Andrey Ushakov, a student from Moscow Technical University of Radiophysics, Engineering and Automation, who has contributed to Section 3.7 – *Three-Level System*.

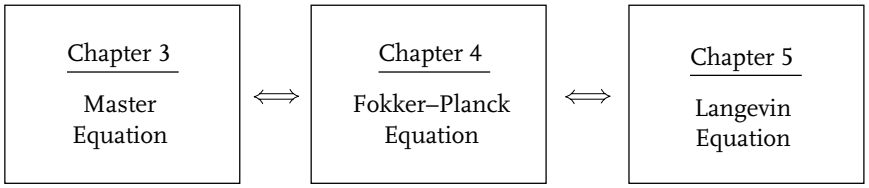
The authors acknowledge support from the Deutsche Forschungsgemeinschaft via grant MA 1508/8.

Rostock, Riga, Moscow
October 2008

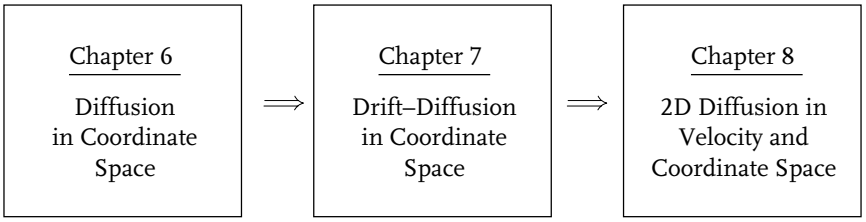
Reinhard Mahnke
Jevgenijs Kaupužs
Ihor Lubashevsky

Mathematics

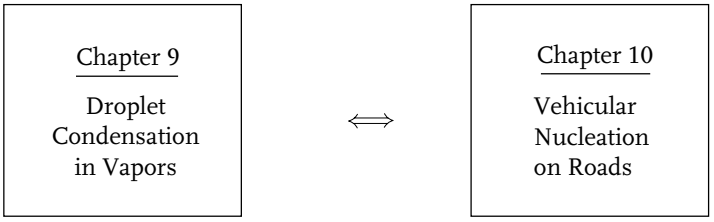
Physics



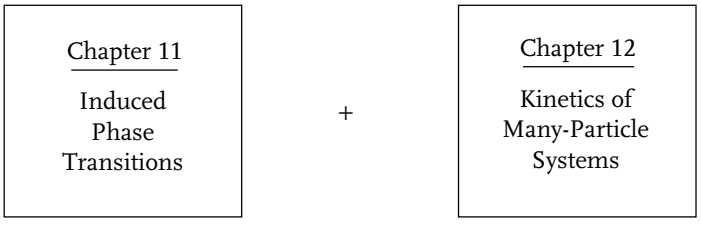
Applications I



Applications II



Applications III



Part I
Basic Mathematical Description

1
Fundamental Concepts

1.1
Wiener Process, Adapted Processes and Quadratic Variation

Stochastic processes represent a fundamental concept used to model the development of a physical or nonphysical system in time. It has turned out that the apparatus of stochastic processes is powerful enough to be applied to many other fields, such as economy, finance, engineering, transportation, biology and medicine.

To start with, we recall that a random variable X is a mapping $X : \Omega \to \mathbb{R}$ that assigns a real value to each elementary event $\omega \in \Omega$. The concrete value $X(\omega)$ is called a realization. It is the value we observe after the experiment has been done. To create a mathematical machine we suppose that a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ is given. Ω is the set of all elementary events and $\mathfrak{F}$ is the family of events we are interested in. It contains the set of all elementary events Ω and is assumed to be closed with respect to forming the complement and countable intersections and unions of events from this collection of events. Such families of sets or events are called σ-algebras. The character σ indicates that even the union or intersection of countably many sets belongs to $\mathfrak{F}$ as well. For mathematical reasons we have to assume that 'events generated by X', i.e. sets of the type $\{\omega : X(\omega) \in I\}$, where I is an open or closed or semi-open interval, are really events; i.e. such sets are assumed also to belong to $\mathfrak{F}$. Unfortunately the collection of all intervals of the real line is not closed with respect to the operation of union. The smallest collection of subsets of the real line that is a σ-algebra and contains all intervals is called the σ-algebra of Borel sets and will be denoted by $\mathfrak{B}$. It turns out that we have not only $\{\omega : X(\omega) \in I\} \in \mathfrak{F}$ for any interval but even $\{\omega : X(\omega) \in B\} \in \mathfrak{F}$ for every Borel set B. This fact is referred to as the $\mathfrak{F}$-measurability of X.

It turns out that for any random variable X and any continuous or monotone function g the function $Y(\omega) = g(X(\omega))$ is again a random variable. This statement remains true even if we replace g by a function from a larger class of functions, called the family of all measurable functions, to which not only the continuous functions but also the pointwise limit of continuous functions belong. This class of functions is closed with respect to 'almost all' standard manipulations with

Physics of Stochastic Processes: How Randomness Acts in Time
Reinhard Mahnke, Jevgenijs Kaupužs and Ihor Lubashevsky

ISBN: 978-3-527-40840-5

functions, such as linear combinations and products and finally forming new functions by plugging one function into another function.

The probability measure $\mathbb{P}$ is defined on $\mathfrak{F}$ and it assigns to each event $A \in \mathfrak{F}$ a number $P(A)$ called the probability of A. The mappings $A \mapsto P(A)$ satisfy the axioms of probability theory, i.e. P is a non-negative σ-additive set function on $\mathfrak{F}$ with $P(\Omega) = 1$.

We assume that the reader is familiar with probability theory at an introductory course level and in the following we use basic concepts and results without giving additional motivation or explanation.

Random variables or random vectors are useful concepts to model the random outcome of an experiment. But we have to include the additional variable 'time' when we are going to study random effects which change over time.

Definition 1.1 *By stochastic process we mean a family of random variables* $(X_t)_{t\geq 0}$ *which are defined on the probability space* $(\Omega, \mathfrak{F}, \mathbb{P})$.

By definition X_t is in fact a function of two variables $X_t(\omega)$. For fixed t this function of ω is a random variable. Otherwise, if we fix ω then we call the function of t defined by $t \mapsto X_t(\omega)$ a realization or a path. This means that the realization of a stochastic process is a function. Therefore stochastic processes are sometimes referred to as random functions. We call a stochastic process continuous if all realizations are continuous functions.

For the construction of a stochastic process, that is, of a suitable probability space, one needs the so-called finite dimensional distributions which are the distributions of random vectors $(X_{t_1}, \ldots, X_{t_n})$, where $t_1 < t_2 < \cdots < t_n$ is any fixed selection. For details of the construction we refer to Øksendal [175].

A fundamental idea of modeling experiments with several random outcomes in both probability theory and mathematical statistics is to start with independent random variables and to create a model by choosing suitable functions of these independent random variables. This fact explains why, in the area of stochastic processes, the particular processes with independent increments play an exceptional role. This, in combination with the fundamental meaning of the normal distribution in probability theory, makes clear the importance of the so-called Wiener process, which will now be defined.

Definition 1.2 *A stochastic process* $(W_t)_{t\geq 0}$ *is called a standard Wiener process or (briefly) Wiener process if:*

1) $W_0 = 0$,
2) $(W_t)_{t\geq 0}$ *has independent increments, i.e.* $W_{t_n} - W_{t_{n-1}}, \ldots, W_{t_2} - W_{t_1}, W_{t_1}$ *are independent for* $t_1 < t_2 < \cdots < t_n$,
3) *For all* $0 \leq s < t$, $W_t - W_s$ *has a normal distribution with expectation* $\mathbb{E}(W_t - W_s) = 0$ *and variance* $\mathbb{V}(W_t - W_s) = t - s$,
4) *All paths of* $(W_t)_{t\geq 0}$ *are continuous.*

The Wiener process is also called Brownian motion. This process is named after the biologist Robert Brown whose research dates back to the 1820s. The

mathematical theory began with Louis Bachelier (Théorie de la Spéculation, 1900) and later by Albert Einstein (Eine neue Bestimmung der Moleküldimensionen, 1905). Norbert Wiener (1923) was the first to create a firm mathematical basis for Brownian motion.

To study properties of the paths of the Wiener process we use the quadratic variation as a measure of the smoothness of a function.

Definition 1.3 *Let $f : [0, T] \to \mathbb{R}$ be a real function and $\mathfrak{z}_n : a = t_{0,n} < t_{1,n} < \cdots < t_{n,n} = b$, a sequence of partitions with*

$$\delta(\mathfrak{z}_n) := \max_{0 \le i \le n-1} (t_{i+1,n} - t_{i,n}) \to 0, \quad as \quad n \to \infty.$$

If $\lim_{n\to\infty} \sum_{i=0}^{n-1} (f(t_{i+1,n}) - f(t_{i,n}))^2$ exists and is independent of the concrete sequence of partitions then this limit is called the quadratic variation of f and will be denoted by $[f]_T$.

We show that the quadratic variation of a continuously differentiable function is zero.

Lemma 1.1 *If f is differentiable in $[0, T]$ and the derivative $f'(t)$ is continuous then $[f]_T = 0$.*

Proof. Put $C = \sup_{0\le t\le T} |f'(t)|$. Then $|f(t) - f(s)| \le C|t - s|$ and

$$\begin{aligned}\sum_{i=0}^{n-1} (f(t_{i+1,n}) - f(t_{i,n}))^2 &\le C^2 \sum_{i=0}^{n-1} (t_{i+1,n} - t_{i,n})^2 \\ &\le C^2 \delta(\mathfrak{z}_n) T \to_{n\to\infty} 0.\end{aligned}$$

If $(X_t)_{0\le t\le T}$ is a stochastic process then the quadratic variation $[X]_T$ is a random variable such that for any sequence of partitions $\mathfrak{z}_n$ with $\delta(\mathfrak{z}_n) \to 0$ it holds for $n \to \infty$

$$\sum_{i=0}^{n-1} (X_{t_{i+1,n}} - X_{t_{i,n}})^2 \to^{\mathbb{P}} [X]_T,$$

where $\to^{\mathbb{P}}$ is the symbol for stochastic convergence. Whether the quadratic variation of a stochastic process does or does not exist depends on the concrete structure of this process and has to be checked in a concrete situation and it is often more useful to deal with the convergence in mean square instead of the stochastic convergence. The relation between the two concepts provides the well known Chebyshev inequality which states that, for any random variables Z_n, Z

$$\mathbb{P}(|Z_n - Z| > \varepsilon) \le \frac{1}{\varepsilon^2} \mathbb{E}(Z_n - Z)^2.$$

Hence the mean square convergence $\mathbb{E}(Z_n - Z)^2 \to 0$ of Z_n to Z implies the stochastic convergence $\mathbb{P}(|Z_n - Z| > \varepsilon) \to 0$ of Z_n to Z.

Now we are going to calculate the quadratic variation of a Wiener process. To this end we need a well known fact. If V has a normal distribution with expectation μ and variance σ^2 then

$$\begin{aligned} \mathbb{E}V = \mu, \quad \mathbb{V}(V) = \mathbb{E}(V-\mu)^2 = \sigma^2 \\ \mathbb{E}(V-\mu)^3 = 0, \quad \mathbb{E}(V-\mu)^4 = 3\sigma^4. \end{aligned}$$

If $\mu = 0$ then

$$\begin{aligned} \mathbb{E}(V^2 - \sigma^2)^2 &= \mathbb{E}(V^4 - 2\sigma^2 V^2 + \sigma^4) \\ &= 3\sigma^4 - \sigma^4 = 2\sigma^4. \end{aligned} \tag{1.1}$$

Theorem 1.1 *If $(W_t)_{0\le t\le T}$ is a Wiener process then the quadratic variation*

$$[W]_T = T.$$

Proof. Let $\mathfrak{z}_n$ be a sequence of partitions of $[0, T]$ with $\delta(\mathfrak{z}_n) \to 0$ and put

$$Z_n = \sum_{i=0}^{n-1} (W_{t_{i+1,n}} - W_{t_{i,n}})^2.$$

From the definition of the Wiener process we get that $\mathbb{E}(W_{t_{i+1,n}} - W_{t_{i,n}})^2 = t_{i+1,n} - t_{i,n}$. As the variance of a sum of independent random variables is just the sum of the variances we get from the independent increments

$$\begin{aligned} \mathbb{E}(Z_n - t)^2 &= \mathbb{E}\left(\sum_{i=0}^{n-1} (W_{t_{i+1,n}} - W_{t_{i,n}})^2 - (t_{i+1,n} - t_{i,n})\right)^2 \\ &= \mathbb{V}(Z_n) = \sum_{i=0}^{n-1} \mathbb{V}((W_{t_{i+1,n}} - W_{t_{i,n}})^2) \\ &= \sum_{i=0}^{n-1} \mathbb{E}((W_{t_{i+1,n}} - W_{t_{i,n}})^2 - (t_{i+1,n} - t_{i,n}))^2 \\ &= 2\sum_{i=0}^{n-1} (t_{i+1,n} - t_{i,n})^2 \le 2\delta(\mathfrak{z}_n)T \to 0, \end{aligned}$$

where for the last equality we have used (1.1).

The statement $[W]_T = T$ is remarkable from different points of view. The exceptional fact is that the quadratic variation of this special stochastic process $(W_t)_{0\le t\le T}$ is a degenerate random variable, it is the deterministic value T. This value is non-zero. Therefore we may conclude from Lemma 1.1 that the paths of

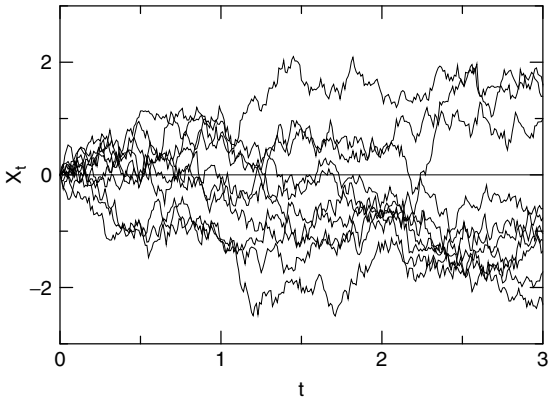


Figure 1.1 Collection of realizations X_t of the special stochastic process $(W_t)_{0\le t\le T}$ named after Norbert Wiener.

a Wiener process cannot be continuously differentiable as otherwise the quadratic variation must be zero. The fact that the quadratic variation is non-zero implies that the absolute value of an increment $W_t - W_s$ cannot be proportional to $t - s$. From here we may conclude that the paths of a Wiener process are continuous but not differentiable and therefore strongly fluctuating. The illustrative picture (see Figure 1.1) of simulated realizations of a Wiener process underlines this statement.

One of the main problems in the theory of stochastic processes is to find mathematical models that describe the evolution of a system in time and can especially be used to predict, of course not without error, the values in the future with the help of information about the process collected from the past. Here and in the sequel by 'the collected information' we mean the family of all events observable up to time t. This collection of events will be denoted by $\mathfrak{F}_t$, where we suppose that $\mathfrak{F}_t$ is a σ-algebra. It is clear that $\mathfrak{F}_s \subseteq \mathfrak{F}_t \subseteq \mathfrak{F}$. Such families of σ-algebras are referred to as a *filtration* and will be denoted by $(\mathfrak{F}_t)_{\ge 0}$. Each stochastic process $(X_t)_{t\ge 0}$ generates a filtration by the requirement that $\mathfrak{F}_t$ is the smallest σ-algebra that contains all events $\{X_s \in I\}$ where I is any interval and $0 \le s \le t$. This filtration will be denoted $\sigma((X_s)_{0\le s\le t})$. We call any stochastic process $(Y_t)_{t\ge 0}$ *adapted* to the filtration $(\mathfrak{F}_t)_{\ge 0}$ (short $\mathfrak{F}_t$-adapted) if all events that may be constructed by the process up to time t belong to the class of observable events, i.e. already belong to $\mathfrak{F}_t$. The formal mathematical condition is $\sigma((Y_s)_{0\le s\le t}) \subseteq \mathfrak{F}_t$ for every $t \ge 0$. If for any fixed t and any random variable Z all events $\{Z \in I\}$, $I \subseteq \mathbb{R}$, belong to $\mathfrak{F}_t$ and it holds that $\mathbb{E}Z^2 < \infty$ then there are $X_{t_{1,n}}, \dots, X_{t_{m_n,n}}$, $t_{i,j} \le t$ and (measurable) functions $f_n(X_{t_{1,n}}, \dots, X_{t_{m_n,n}})$ such that

$$\mathbb{E}(Z - f_n(X_{t_{1,n}}, \dots, X_{t_{m_n,n}}))^2 \to 0.$$

We omit the proof which would require additional results from measure theory. We denote by $\mathfrak{P}_t(X)$ the class of all such random variables. $\mathfrak{P}_t(X)$ may be considered as the *past of the process* $(X_t)_{t\ge 0}$.

Example 1.1 *Let $(W_t)_{t\geq 0}$ be a Wiener process and $\mathfrak{F}_t = \sigma((Y_s)_{0\leq s\leq t})$. The following processes are $\mathfrak{F}_t$-adapted $X_t = W_t^2$, $X_t = W_{0,5\cdot t}^2 + W_t^4$, $X_t = (W_t^4 / 1 + W_{0,1\cdot t}^2)$. The process W_{t+1} is not $\mathfrak{F}_t$-adapted.*

We fix the interval $[0, T]$, set $\mathfrak{F}_t = \sigma((W_s)_{0\leq s\leq t})$ and denote by $\mathfrak{E}_t(W) \subseteq \mathfrak{P}_t(W)$ the collection of all elementary $\mathfrak{F}_t$-adapted processes, that is of all processes that may be written as

$$Y_t = \sum_{i=0}^{n-1} X_{t_i} I_{[t_i, t_{i+1})}(t), \quad X_{t_i} \in \mathfrak{P}_{t_i}(W), \tag{1.2}$$

where $0 = t_0 < t_1 < \cdots < t_n$ and

$$I_{[a,b)}(t) = \begin{cases} 1 & \text{if} \quad a \leq t < b, \\ 0 & \text{if} \quad \text{else.} \end{cases}$$

The $\mathfrak{F}_t$-adeptness of the process Y_t follows from the fact that exclusively random variables X_{t_i} with $t_i \leq t$ appear in the sum. The process Y_t is piecewise constant, it has the value X_{t_i} in $[t_i, t_{i+1})$ and jumps at t_i with a height

$$\Delta Y_{t_i} = X_{t_i} - X_{t_{i-1}}.$$

1.2 The Space of Square Integrable Random Variables

By $\mathcal{H}_2$ we denote the space of all random variables X with $\mathbb{E}X^2 < \infty$. Here and in the sequel we identify random variables X and Y that take on different values only with probability zero, i.e. $\mathbb{P}(X \neq Y) = 0$. Set

$$\langle X, Y\rangle := \mathbb{E}(XY).$$

It is not hard to see that $\langle X, Y\rangle$ satisfies all conditions that are imposed on a scalar product, i.e. $\langle X, Y\rangle$ is symmetric in X and Y, it is linear in both X and Y, and it holds that

$$\langle X, X\rangle \geq 0,$$

where the equality is satisfied if and only if $X = 0$.

The norm of a random variable X is given by

$$\|X\| = \sqrt{\mathbb{E}X^2},$$

and the distance of X and Y is the norm of $X - Y$. Recall that a sequence of random variables X_n is said to be convergent in mean square to X if $\mathbb{E}(X_n - X)^2 = 0$. Hence this type of convergence is nothing other than the norm convergence $\lim_{n\to\infty} \|X_n - X\| = 0$. A sequence of random variables $\{X_n\}$ is said to be a Cauchy sequence if

$$\lim_{n,m\to\infty} \|X_n - X_m\| = 0.$$

For a proof of the following theorem we refer to Øksendal [175].

Theorem 1.2 *To each Cauchy sequence $X_n \in \mathcal{H}_2$ there is some $X \in \mathcal{Z}_2$ with*

$$\lim_{n\to\infty} \|X_n - X\| = 0,$$

i.e. the space is complete.

It is clear that $\mathcal{H}_2$ is a linear space. As we have already equipped $\mathcal{H}_2$ with a scalar product we get, together with the completeness, that $\mathcal{H}_2$ is a Hilbert space. This fact allows us to apply methods from the Hilbert space theory to problems of probability theory.

A subset $\mathcal{T} \subseteq \mathcal{H}_2$ is called closed, if every limit X of a sequence $X_n \in \mathcal{T}$ belongs to $\mathcal{T}$ again. If $\mathcal{L} \subseteq \mathcal{H}_2$ is a closed linear subspace of $\mathcal{H}_2$ then there is some element in $\mathcal{L}$ that best approximates X.

Theorem 1.3 *If $\mathcal{L} \subseteq \mathcal{H}_2$ is a closed linear subspace of $\mathcal{H}_2$, then to each $X \in \mathcal{H}_2$ there is a random variable in $\mathcal{L}$, denoted by $\Pi_{\mathcal{L}} X \in \mathcal{L}$ and called the projection of X on $\mathcal{L}$, such that*

$$\inf_{Y\in\mathcal{L}} \|X - Y\| = \|X - \Pi_{\mathcal{L}} X\|.$$

Proof. Let $Y_n \in \mathcal{L}$ be a minimum sequence, i.e.

$$\lim_{n\to\infty} \|X - Y_n\| = \inf_{Y\in\mathcal{L}} \|X - Y\|.$$

Then Y_{m_n} is a minimum sequence again. Because

$$\left\| X - \frac{1}{2}(Y_n + Y_{m_n}) \right\| \le \frac{1}{2}\|X - Y_n\| + \frac{1}{2}\|X - Y_{m_n}\|$$

$\frac{1}{2}(Y_n + Y_{m_n})$ is also a minimum sequence. Then

$$\lim_{n\to\infty} \left[\frac{1}{2}\|X - Y_n\|^2 + \frac{1}{2}\|X - Y_{m_n}\|^2 - \left\| X - \frac{1}{2}(Y_n + Y_{m_n}) \right\|^2 \right] = 0.$$

For any random variables U, V it holds that

$$\begin{aligned} \frac{1}{2}\|U\|^2 + \frac{1}{2}\|V\|^2 - \left\| \frac{1}{2}(U+V) \right\|^2 &= \mathbb{E}\left(\frac{1}{2}U^2 + \frac{1}{2}V^2 - \left(\frac{1}{2}(U+V) \right)^2 \right) \\ &= \frac{1}{4}\mathbb{E}(U-V)^2 = \frac{1}{4}\|U - V\|^2. \end{aligned}$$

Putting $U = X - Y_n$, $V = X - Y_{m_n}$ we arrive at

$$\begin{aligned} &\frac{1}{2}\|X - Y_n\|^2 + \frac{1}{2}\|X - Y_{m_n}\|^2 - \left\| X - \frac{1}{2}(Y_n + Y_{m_n}) \right\|^2 \\ &\quad = \frac{1}{4}\|Y_n - Y_{m_n}\|^2 \to 0. \end{aligned}$$

As m_n was an arbitrary sequence we see that Y_n is a Cauchy sequence and converges, by the completeness of $\mathcal{H}_2$, to some random variable $\Pi_{\mathcal{L}}X$ that belongs to $\mathcal{L}$ since $\mathcal{L}$ is closed by assumption.

Without going into detail we note that the projection $\Pi_{\mathcal{L}}X$ is uniquely determined in the sense that, for every $Z \in \mathcal{L}$ which also provides a best approximation, it holds that

$$\mathbb{P}(\Pi_{\mathcal{L}}X \neq Z) = 0. \tag{1.3}$$

The projection $\Pi_{\mathcal{L}}X$ can be also characterized with the help of conditions imposed on the error $X - \Pi_{\mathcal{L}}X$.

Corollary 1.1 *It holds that $Y = \Pi_{\mathcal{L}}X$ if and only if $Y \in \mathcal{L}$ and $Y - X \perp \mathcal{L}$, i.e.*

$$\langle Y - X, Z\rangle = 0 \quad \textit{for every } Z \in \mathcal{L}. \tag{1.4}$$

Proof. 1. Assume $Y = \Pi_{\mathcal{L}}X$. Then $Y \in \mathcal{L}$ by the definition of the projection.

We consider

$$g(t) = \left\|(X - Y) - tZ\right\|^2 = \|X - Y\|^2 + t^2\,\|Z\|^2 - 2t\,\langle Y - X, Z\rangle\,.$$

By the definition of $\Pi_{\mathcal{L}}X$ the function $g(t)$ attains its minimum at $t = 0$. Hence

$$g'(0) = -2\,\langle Y - X, Z\rangle = 0$$

which implies $\langle Y - X, Z\rangle = 0$.

2. If $Y \in \mathcal{L}$ satisfies (1.4) then for every $U \in \mathcal{L}$

$$\|X - U\|^2 = \|X - Y\|^2 + 2\,\langle X - Y, Y - U\rangle + \|Y - U\|^2\,.$$

As $Z = Y - U \in \mathcal{L}$ we see that the middle term vanishes. Hence the right-hand term is minimal if and only if $U = Y$.

The simplest prediction of a random variable X is a constant value. Which value a is the best one ? It is easy to see that the function

$$\varphi(a) = \mathbb{E}(X - a)^2$$

attains the minimum at $a_0 = \mathbb{E}X$. Consequently, if $\mathcal{L}$ consists of constant random variables only, then $\Pi_{\mathcal{L}}X = \mathbb{E}X$. This is the reason why, for any closed linear subspace, we call the projection $\Pi_{\mathcal{L}}X$ the *conditional expectation* given $\mathcal{L}$. In this case we tacitly assume that all constant random variables are contained in $\mathcal{L}$. As $\mathcal{L}$ is a linear space this is equivalent to the fact that $Z_0 \equiv 1 \in \mathcal{L}$. If this condition is satisfied then we write

$$\mathbb{E}(X|\mathcal{L}) := \Pi_{\mathcal{L}}X.$$

Choosing $Z = 1$ in (1.4) we get the following.