Vincenzo Capasso
David Bakstein

An Introduction to Continuous-Time Stochastic Processes

Theory, Models, and Applications to Finance, Biology, and Medicine

Birkhäuser
Boston • Basel • Berlin

Vincenzo Capasso
MIRIAM (Milan Research Centre for
Industrial and Applied Mathematics)
University of Milan
Department of Mathematics
20133 Milan
Italy

David Bakstein
MIRIAM (Milan Research Centre for
Industrial and Applied Mathematics)
University of Milan
Department of Mathematics
20133 Milan
Italy

AMS Subject Classifications: 93-xx, 93-02, 93B05, 93B07, 93B10, 93B11, 93B17, 93B25, 93B27, 93B36, 93B40, 93B51, 93B52, 93B55, 93B60, 93C05, 93C15, 93C35, 93C55, 93C70

Library of Congress Cataloging-in-Publication Data
Capasso, V. (Vincenzo), 1945-
An introduction to continuous-time stochastic processes : theory, models, and applications to finance, biology, and medicine / Vincenzo Capasso, David Bakstein.
p. cm — (Modeling and simulation in science, engineering and technology)
Includes bibliographical references and index.
ISBN 0-8176-3234-4 (acid-free paper)
1. Stochastic processes. I. Bakstein, David, 1975- II. Title. III. Modeling and simulation in science, engineering & technology.

QA274.C36 2004
519.2'3–dc22 2003063634

ISBN 0-8176-3234-4 Printed on acid-free paper.

Printed in the United States of America. (SB)

9 8 7 6 5 4 3 2 1 SPIN 10925249

www.birkhauser.com

Contents

6

Applications to Biology and Medicine

6.1 Population Dynamics: Discrete-in-Space–Continuous-in-Time Models

In the chapter on stochastic processes the Poisson process was introduced as an example of an RCLL nonexplosive counting process. Furthermore, we reviewed a general theory of counting processes as point processes on the real line within the framework of martingale theory and dynamics. Indeed, for these processes, under the usual regularity assumptions, we can invoke the Doob–Meyer decomposition theorem (see (2.79) onwards) and claim that any nonexplosive RCLL process $(X_t)_{t\in\mathbb{R}_+}$ satisfies a generalized stochastic differential equation of the form

$$dX_t = dA_t + dM_t \tag{6.1}$$

subject to a suitable initial condition. Here A is the compensator of the process representing the model of "evolution" and M is a martingale representing the "noise."

As was mentioned in the sections on counting and marked point processes, a counting process $(N_t)_{t\in\mathbb{R}_+}$ is a random process that counts the occurrence of certain events over time, namely N_t being the number of such events having occurred during the time interval $]0,t]$. We have noticed that a nonexplosive counting process is RCLL with upward jumps of magnitude 1 and we impose the initial condition $N_0 = 0$, almost surely. Since we deal with those counting processes that satisfy the conditions of Theorem 2.87 (local Doob–Meyer decomposition theorem), a nondecreasing predictable process $(A_t)_{t\in\mathbb{R}_+}$ (the compensator) exists such that $(N_t - A_t)_{t\in\mathbb{R}_+}$ is a right-continuous local martingale. Further, we assume that the compensator is absolutely continuous with respect to the usual Lebesgue measure on $\mathbb{R}_+$. In this case we say that $(N_t)_{t\in\mathbb{R}_+}$ has a (predictable) intensity $(\lambda_t)_{t\in\mathbb{R}_+}$ such that

$$A_t = \int_0^t \lambda_s ds \qquad \text{for any } t \in \mathbb{R}_+,$$

and the stochastic differential equation (6.1) can be rewritten as

$$dX_t = \lambda_t dt + dM_t. \tag{6.2}$$

If the process is integrable and λ is left-continuous with right limits (LCRL), one can easily show that

$$\lambda_t = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[N_{t+\delta t} - N_t | \mathcal{F}_{t-}] \qquad \text{a.s.}$$

and, if we further assume the simplicity of the process, we also have

$$\lambda_t = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} P(N_{t+\delta t} - N_t = 1 | \mathcal{F}_{t-}) \qquad \text{a.s.;}$$

i.e., $\lambda_t dt$ is the conditional probability of a new *event* during $[t, t+dt]$ given the history of the process during $[0, t]$. It really represents the model of evolution of the counting process, similar to classical deterministic differential equations.

Example 6.1. Let X be a nonnegative real random variable with absolutely continuous probability law having density f, cumulative distribution function F, survival function $S = 1 - F$, and hazard rate function $\alpha(t) = \frac{f(t)}{S(t)}$, $t > 0$. Assume

$$\int_0^t \alpha(s) ds = -\ln(1 - F(t)) < +\infty \qquad \text{for any } t \in \mathbb{R}_+,$$

but

$$\int_0^{+\infty} \alpha(t) dt = +\infty.$$

Define the univariate process N_t by

$$N_t = I_{[X \le t]}(t)$$

and let $(\mathcal{N}_t)_{t \in \mathbb{R}_+}$ be the filtration the process generates; i.e.,

$$\mathcal{N}_t = \sigma(N_s, s \le t) = \sigma\left(X \wedge t, I_{[X \le t]}(t)\right).$$

Define the left-continuous adapted process Y_t by

$$Y_t = I_{[X \ge t]}(t) = 1 - N_{t-}.$$

It can be easily shown (see, e.g., Andersen et al. (1993)) that N_t admits

$$A_t = \int_0^t Y_s \alpha(s) ds$$

as a compensator and hence N_t has stochastic intensity λ_t defined by

$$\lambda_t = Y_t \alpha(t), \qquad t \in \mathbb{R}_+.$$

In other words,

$$N_t - \int_0^{X \wedge t} \alpha(s) ds$$

is a local martingale. Here $\alpha(t)$ is a deterministic function, while Y_t, clearly, is a predictable process. This is a first example of what is known as a multiplicative intensity model.

Example 6.2. Let X be a random time as in the previous example, and let U be another random time, i.e., a nonnegative real random variable. Consider the random variable $T = X \wedge U$ and define the processes

$$N_t = I_{[T \le t]} I_{[X \le U]}(t)$$

and

$$N_t^U = I_{[T \le t]} I_{[U < X]}(t)$$

and the filtration

$$\mathcal{N}_t = \sigma\left(N_s, N_s^U, s \le t\right).$$

The hazard rate function α of X is known as the *net hazard rate*; it is given by

$$\alpha(t) = \lim_{h \to 0+} \frac{1}{h} P[t \le X \le t + h | X > t].$$

On the other hand, the quantity

$$\alpha^+(t) = \lim_{h \to 0+} \frac{1}{h} P[t \le X \le t + h | X > t, U > t]$$

is known as the *crude hazard rate*, whenever the limit exists. In this case

$$N_t - \int_0^t I_{[T \ge t]} \alpha(s) ds$$

is a local martingale.

Birth-and-Death Processes

A Markov birth-and-death process provides an example of a bivariate counting process. Let $(X_t)_{t \in \mathbb{R}_+}$ be the size of a population subject to a birth rate λ and a death rate μ. Then the infinitesimal transition probabilities are

$$P\left(X_{t+\delta t} = j | X_t = h\right) = \begin{cases} \lambda h \delta t + o(\delta t) & \text{if } j = h+1, \\ \mu h \delta t + o(\delta t) & \text{if } j = h-1, \\ 1 - (\lambda h + \mu h)\delta t + o(\delta t) & \text{if } j = h, \\ o(\delta t) & \text{otherwise.} \end{cases}$$

Let $N_t^{(1)}$ and $N_t^{(2)}$ be the number of births and deaths, respectively, up to time $t \geq 0$, assuming $N_0^{(1)} = 0$ and $N_0^{(2)} = 0$. Then

$$(\mathbf{N}_t)_{t\in\mathbb{R}_+} = \left(N_t^{(1)}, N_t^{(2)}\right)$$

is a bivariate counting process with intensity process $(\lambda X_{t-}, \mu X_{t-})_{t\in\mathbb{R}_+}$ (see Figures 6.1 and 6.2). This is an example of a formulation of a Markov process with countable state space as a counting process. In particular, we may write a stochastic differential equation for X_t as follows:

$$dX_t = \lambda X_{t-}dt - \mu X_{t-}dt + dM_t,$$

where M_t is a suitable martingale noise.

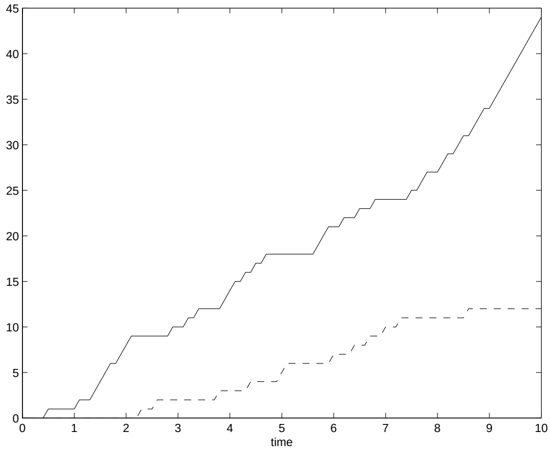

Fig. 6.1. Simulation of a birth-and-death process with birth rate $\lambda = 0.2$, death rate $\mu = 0.05$, initial population $X_0 = 10$, time step $dt = 0.1$, and interval of observation $[0, 10]$. The continuous line represents the number of births $N_t^{(1)}$; the dashed line represents the number of deaths $N_t^{(2)}$.

A Model for Software Reliability

Let N_t denote the number of software failures detected during the time interval $]0, t]$ and suppose that F is the true number of faults existing in the software at time $t = 0$. In the Jelinski–Moranda model (see Jelinski and Moranda (1972)) it is assumed that N_t is a counting process with intensity

$$\lambda_t = \rho(F - N_{t-}),$$

where ρ is the individual failure rate (see Figure 6.3). One may note that this model corresponds to a pure death process in which the total initial population F usually is unknown, as is the rate ρ.

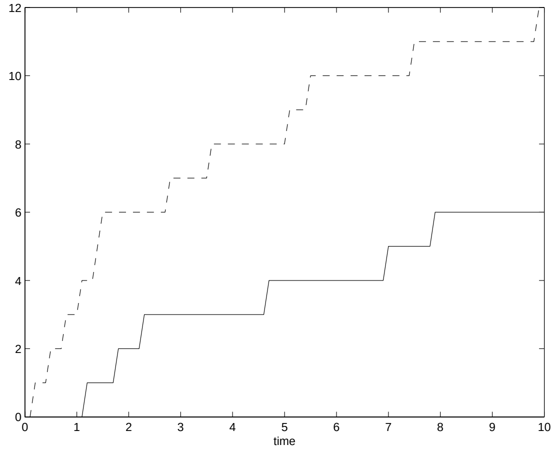

Fig. 6.2. Simulation of a birth-and-death process with birth rate $\lambda = 0.09$, death rate $\mu = 0.2$, initial population $X_0 = 10$, time step $dt = 0.1$, and interval of observation $[0, 10]$. The continuous line represents the number of births $N_t^{(1)}$; the dashed line represents the number of deaths $N_t^{(2)}$.

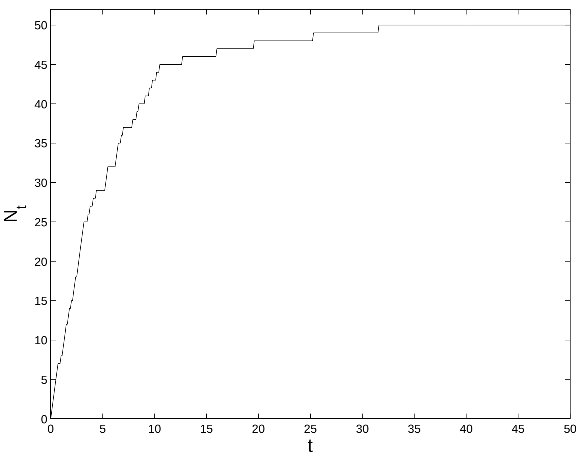

Fig. 6.3. Simulation of a model for software reliability: individual failure rate $\rho = 0.2$, true initial number of faults $F = 50$, time step $dt = 0.1$, and interval of observation $[0, 50]$.

Contagion: The Simple Epidemic Model

Epidemic systems provide models for the transmission of a contagious disease within a population. In the "simple epidemic model" (Bailey (1975) and Becker (1989)) the total population N is divided into two main classes:

(S) the class of susceptibles, including those individuals capable of contracting the disease and becoming infectives themselves;
(I) the class of infectives, including those individuals who, having contracted the disease, are capable of transmitting it to susceptibles.

Let I_t denote the number of individuals who have been infected during the time interval $]0, t]$. Assume that individuals become infectious themselves immediately upon infection and remain so for the entire duration of the epidemic. Suppose that at time $t = 0$ there are S_0 susceptible individuals and I_0 infectives in the community. The classical model based on the *law of mass action* (see, e.g., Bailey (1975) or Capasso (1993)) assumes that the counting process I_t has stochastic intensity

$$\lambda_t = \beta_t(I_0 + I_{t-})(S_0 - I_{t-}),$$

which is appropriate when the community is mixing uniformly. Here β_t is called the *infection rate*(see Figure6.4).

It can be noted that formally this corresponds to writing $N(t)$ with the stochastic differential equation

$$dI_t = \beta_t(I_0 + I_{t-})(S_0 - I_{t-})dt + dM_t,$$

where M_t is a suitable martingale noise. In this case we obtain

$$\langle M \rangle_t = \int_0^t \lambda_s ds$$

for the variation process $\langle M \rangle_t$, so that

$$M_t^2 - \int_0^t \lambda_s ds$$

is a zero mean martingale. As a consequence

$$Var[M_t] = E\left[\int_0^t \lambda_s ds\right] = E[I_t].$$

More general models can be found in Capasso (1990) and references therein.

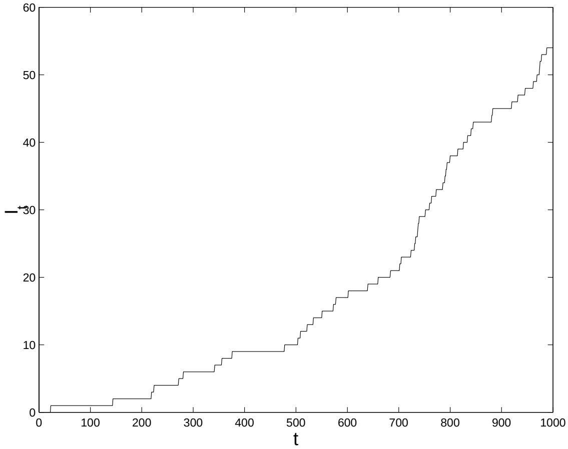

Fig. 6.4. Simulation of a simple epidemic (SI) model: initial number of susceptibles $S_0 = 500$, initial number of infectives $I_0 = 4$, infection rate (constant) $\beta = 5 \cdot 10^{-6}$, time step $dt = 1$, interval of observation $[0, 1000]$.

Contagion: The General Stochastic Epidemic

For a wide class of epidemic models the total population $(N_t)_{t\in\mathbb{R}_+}$ includes three subclasses. In addition to the classes of susceptibles $(S_t)_{t\in\mathbb{R}_+}$ and infectives $(I_t)_{t\in\mathbb{R}_+}$ already introduced in the simple model, a third class is considered, i.e.,

(R), the class of *removals*. This comprises those individuals who, having contracted the disease, and thus being already infectives, are no longer in the position of transmitting the disease to other susceptibles because of death, immunization, or isolation. Let us denote the number of removals as $(R_t)_{t\in\mathbb{R}_+}$.

The process $(S_t, I_t, R_t)_{t\in\mathbb{R}_+}$ is modelled as a multivariate jump Markov process valued in $E' = \mathbb{N}^3$. Actually, if we know the behavior of the total population process N_t, because

$$S_t + I_t + R_t = N_t \qquad \text{for any } t \in \mathbb{R}_+,$$

we need to provide a model only for the bivariate process $(S_t, I_t)_{t\in\mathbb{R}_+}$, which is now valued in $E = \mathbb{N}^2$. The only nontrivial elements of a resulting intensity matrix Q (see chapter on Markov processes) are given by

- $q_{(s,i),(s+1,i)} = \alpha$, birth of a susceptible;
- $q_{(s,i),(s-1,i)} = \gamma s$, death of a susceptible;
- $q_{(s,i),(s,i+1)} = \beta$, birth of an infective;
- $q_{(s,i),(s,i-1)} = \delta i$, removal of an infective;

- $q_{(s,i),(s-1,i+1)} = \kappa si$, infection of a susceptible.

For $\alpha = \beta = \gamma = 0$, we have the so-called *general stochastic epidemic* (see, e.g., Bailey (1975) and Becker(1989)). In this case the total population is constant (assume $R_0 = 0$; see Figure 6.5):

$$N_t \equiv N = S_0 + I_0 \qquad \text{for any } t \in \mathbb{R}_+.$$

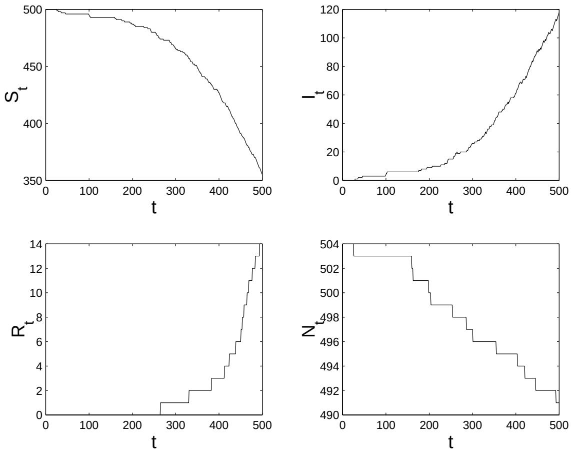

Fig. 6.5. Simulation of an SIR epidemic model with vital dynamics: initial number of susceptibles $S_0 = 500$, initial number of infectives $I_0 = 4$, initial number of removed $R_0 = 0$, birth rate of susceptibles $\alpha = 10^{-4}$, death rate of a susceptible $\gamma = 5 \cdot 10^{-5}$, birth rate of an infective $\beta = 10^{-5}$, rate of removal of an infective $\delta = 8.5 \cdot 10^{-4}$, infection rate of a susceptible $k = 1.9 \cdot 10^{-5}$, time step $dt = 1$, interval of observation $[0, 500]$.

Contagion: Diffusion of Innovations

When a new product is introduced in a market, its diffusion is due to a process of adoption by individuals who are aware of it. Classical models of innovation diffusion are very similar to epidemic systems, even though in this case rates of adoption (infection) depend upon specific marketing and advertising strategies (see Capasso, Di Liddo, and Maddalena (1994) and Mahajan and Wind (1986)). In this case the total population N of possible consumers is divided into the following main classes:

(S) the class of potential adopters, including those individuals capable of adopting the new product, thus themselves becoming adopters;

(A) the class of adopters, those individuals who, having adopted the new product, are capable of transmitting it to potential adopters.

Let A_t denote the number of individuals who, by time $t \geq 0$, have already adopted a new product that has been put on the market at time $t = 0$. Suppose that at time $t = 0$ there are S_0 potential adopters and A_0 adopters in the market. In the basic models it is assumed that all consumers are homogeneous with respect to their inclination to adopt the new product. Moreover, all adopters are homogeneous in their ability to persuade others to try new products, and adopters never lose interest, but continue to inform those consumers who are not aware of the new product. Under these assumptions the classical model for the adoption rate is again based on the law of mass action (see Bartholomew (1976)), apart from an additional parameter $\lambda_0(t)$ that describes adoption induced by external actions, independent of the number of adopters, such as advertising, price reduction policy, etc. Then the stochastic intensity for this process is given by

$$\lambda(t) = (\lambda_0(t) + \beta_t A_{t-})(S_0 - A_{t-}),$$

which is appropriate when the community is mixing uniformly. Here β_t is called the *adoption rate*(see Figure 6.6).

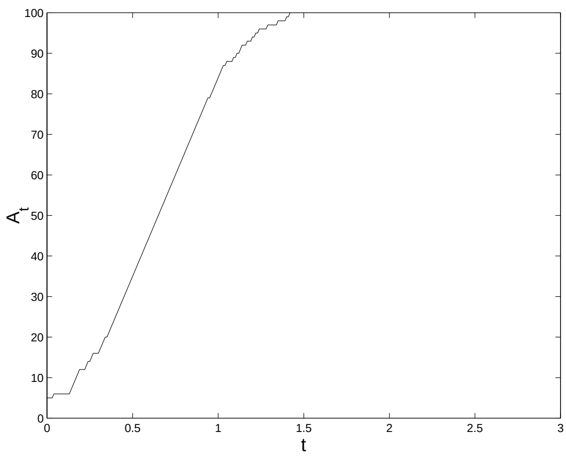

Fig. 6.6. Simulation of the contagion model for diffusion of innovations: external influence $\lambda_0(t) = 5 \cdot 10^{-4} t$, adoption rate (constant) $\beta = 0.05$, initial potential adopters $S_0 = 100$, initial adopters $A_0 = 5$, time step $dt = 0.01$, interval of observation $[0, 3]$.

Inference for Multiplicative Intensity Processes

Let

$$dN_t = \alpha_t Y_t dt + dM_t$$

be a stochastic equation for a counting process N_t, where the noise is a zero mean martingale. Furthermore, let

$$B_s = \frac{J_{s-}}{Y_s} \qquad \text{with} \qquad J_s = I_{[Y_s>0]}(s).$$

B_t is, like Y_t, a predictable process, so that by the integration theorem,

$$M_t^* = \int_0^t B_s dM_s$$

is itself a zero mean martingale. It can be noted that

$$M_t^* = \int_0^t B_s dM_s = \int_0^t B_s dN_s - \int_0^t \alpha_s J_{s-} ds,$$

so that

$$E\left[\int_0^t B_s dN_s\right] = E\left[\int_0^t \alpha_s J_{s-} ds\right];$$

i.e., $\int_0^t B_s dN_s$ is an unbiased estimator of $E[\int_0^t \alpha_s J_{s-} ds]$. If α is constant and we stop the process at a time T such that $Y_t > 0, t \in [0, T]$, then

$$\hat{\alpha} = \frac{1}{T}\int_0^T \frac{dN_s}{Y_s}$$

is an unbiased estimator of α. This method of inference is known as Aalen's method (Aalen (1978)) (the reader may also refer to Andersen et al. (1993) for an extensive application of this method to the statistics of counting processes).

Inference for the Simple Epidemic Model

We may apply the above procedure to the simple epidemic model as discussed in Becker (1989). Let

$$B_s = \frac{I_{[S_s>0]}(s)}{I_{s-}S_{s-}}$$

and suppose β is constant. Let T be such that $S_t > 0, t \in [0, T]$. Then an unbiased estimator for β would be

$$\hat{\beta} = \frac{1}{T}\int_0^T \frac{dI_s}{S_{s-}I_{s-}} = \frac{1}{T}\frac{1}{S_0 I_0} + \frac{1}{(S_0-1)(I_0-1)} + \cdots + \frac{1}{(S_T+1)(I_T+1)}.$$

The standard error (SE) of $\hat{\beta}$ is

$$\frac{1}{T}\left(\int_0^T B_s^2 dI_s\right)^2.$$

By the central limit theorem for martingales (see Rebolledo (1980)) we can also deduce that

$$\frac{\hat{\beta}-\beta}{SE(\hat{\beta})}$$

has an asymptotic $N(0,1)$ distribution, which leads to confidence intervals and hypothesis testing on the model in the usual way (see Becker (1989) and references therein).

Inference for a General Epidemic Model

In Yang (1985) a model was proposed as an extension of the *general epidemic model* presented above. The epidemic process is modelled in terms of a multivariate jump Markov process $(S_t, I_t, R_t)_{t\in\mathbb{R}_+}$, or simply $(S_t, I_t)_{t\in\mathbb{R}_+}$, when the total population is constant, i.e.,

$$N_t := S_t + I_t + R_t = N + 1.$$

In this case, if we further suppose that $S_0 = N$, $I_0 = 1$, $R_0 = 0$, instead of using (S_t, I_t), the epidemic may be described by the number of infected individuals (not including the initial case) $M_1(t)$ and the number of removals $M_2(t) = R_t$ during $]0,t]$, $t \in \mathbb{R}_+^*$. Since we are dealing with a finite total population, the number of infected individuals and the number of removals are bounded, so that

$$E[M_k(t)] \leq N + 1, \qquad k = 1, 2.$$

The processes $M_1(t)$ and $M_2(t)$ are submartingales with respect to the history $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ of the process, i.e., the filtration generated by all relevant processes. We assume that the two processes admit multiplicative stochastic intensities of the form

$$\begin{aligned} \Lambda_1(t) &= \kappa G_1(t-)(N - M_1(t-)), \\ \Lambda_2(t) &= \delta(1 + M_1(t-) - M_2(t-)), \end{aligned}$$

respectively, where $G_1(t)$ is a known function of infectives in circulation at time t. It models the release of pathogen material by infected individuals. Hence

$$Z_k(t) = M_k(t) - \int_0^t \Lambda_k(s)ds, \qquad k = 1, 2,$$

are orthogonal martingales with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. As a consequence, Aalen's unbiased estimators for the infection rate κ and the removal rate δ are given by

$$\hat{\kappa} = \frac{M_1(t)}{B_1(t-)}, \quad \hat{\delta} = \frac{M_1(t)}{B_1(t-)},$$

where

$$B_1(t) = \int_0^t G_1(s)(N - M_1(s))ds,$$

$$B_2(t) = \int_0^t (1 + M_1(s) - M_2(s))ds.$$

Theorem 1.3 in Jacobsen (1982), page 163, gives conditions for a multivariate martingale sequence to converge to a normal process. If such conditions are met, then as $N \to \infty$,

$$\begin{pmatrix} \sqrt{B_1(t)}(\hat{\kappa} - \kappa) \\ \sqrt{B_2(t)}(\hat{\delta} - \delta) \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Gamma \right),$$

where

$$\Gamma = \begin{pmatrix} \kappa & 0 \\ 0 & \delta \end{pmatrix}.$$

In general, it is not easy to verify the conditions of this theorem. They surely hold for the simple epidemic model presented above, where $\delta = 0$. Related results are given in Ethier and Kurtz (1986) and Wang (1977) for a scaled infection rate $\kappa \to \frac{\kappa}{N}$ (see the following section). See also Capasso (1990) for additional models and related inference problems.

6.2 Population Dynamics: Continuous Approximation of Jump Models

A more realistic model than the general stochastic epidemic of the preceding section, which takes into account a rescaling of the force of infection due to the size of the total population, is the following (see Capasso (1993)):

$$q_{(s,i),(s-1,i+1)} = \frac{\kappa}{N} si = N\kappa \frac{s}{N}\frac{i}{N}.$$

We may also rewrite

$$q_{(s,i),(s,i-1)} = \delta N \frac{i}{N} = \frac{i}{N},$$

so that both transition rates are of the form

$$q_{k,k+l}^{(N)} = N\beta_l \left(\frac{k}{N} \right)$$

for

$$k = (s, i)$$

and

$$k + l = \begin{cases} (s, i-1), \\ (s-1, i+1). \end{cases}$$

This model is a particular case of the following situation:

Let $E = \mathbb{Z}^d \cup \{\Delta\}$, where Δ is the point at infinity of $\mathbb{Z}^d$, $d \geq 1$. Further, let

$$\beta_l : \mathbb{Z}^d \to \mathbb{R}_+, \qquad l \in \mathbb{Z}^d,$$

$$\sum_{l \in \mathbb{Z}^d} \beta_l(k) < +\infty \qquad \text{for each} \qquad k \in \mathbb{Z}^d.$$

For f defined in $\mathbb{Z}^d$, and vanishing outside a finite subset of $\mathbb{Z}^d$, let

$$\mathcal{A}f(x) = \begin{cases} \sum_{l \in \mathbb{Z}^d} \beta_l(x)(f(x+l) - f(x)), & x \in \mathbb{Z}^d, \\ 0, \quad x = \Delta. \end{cases}$$

Let $(Y_l)_{l \in \mathbb{Z}^d}$ be a family of independent standard Poisson processes. Let $X(0) \in \mathbb{Z}^d$ be nonrandom and suppose

$$X(t) = X(0) + \sum_{l \in \mathbb{Z}^d} l Y_l \left(\int_0^t \beta_l(X(s)) ds \right), \qquad t < \tau_\infty, \tag{6.3}$$

$$X(t) = \Delta, \qquad t \geq \tau_\infty, \tag{6.4}$$

where

$$\tau_\infty = \inf\{t | X(t-) = \Delta\}.$$

The following theorem holds (see Ethier and Kurtz (1986), page 327).

Theorem 6.3. *1. Given $X(0)$, the solution of system (6.3)–(6.4) above is unique.*
2. If $\mathcal{A}$ is a bounded operator, then X is a solution of the martingale problem for $\mathcal{A}$.

As a consequence, for our class of models for which

$$q^{(N)}_{k,k+l} = N \beta_l \left(\frac{k}{N} \right), \qquad k \in \mathbb{Z}^d, l \in \mathbb{Z}^d,$$

we have that the corresponding Markov process, which we shall denote by $\hat{X}^{(N)}$, satisfies, for $t < \tau_\infty$:

$$\hat{X}^{(N)}(t) = \hat{X}^{(N)}(0) + \sum_{l \in \mathbb{Z}^d} l Y_l \left(N \int_0^t \beta_l \left(\frac{\hat{X}^{(N)}(s)}{N} \right) ds \right),$$

where the Y_l are independent standard Poisson processes. By setting

$$F(x) = \sum_{l \in \mathbb{Z}^d} l \beta_l(x), \qquad x \in \mathbb{R}^d,$$

and

$$X^{(N)} = \frac{1}{N}\hat{X}^{(N)},$$

we have

$$X^{(N)}(t) = X^{(N)}(0) + \sum_{l\in\mathbb{Z}^d} \frac{l}{N}\tilde{Y}_l\left(N\int_0^t \beta_l\left(X^{(N)}(s)\right)ds\right) + \int_0^t F(X^{(N)}(s))ds, \tag{6.5}$$

where

$$\tilde{Y}_l(u) = Y_l(u) - u$$

is the centered standard Poisson process. The state space for $X^{(N)}$ is

$$E_N = E \cap \left\{\frac{k}{N}, k \in \mathbb{Z}^d\right\}$$

for $E \subset \mathbb{R}^d$. We require that $x \in E_N$ and $\beta_(x) > 0$ imply $x + \frac{l}{N} \in E_N$. The generator for $X^{(N)}$ is

$$\begin{aligned}
&\mathcal{A}^{(N)}f(x)\\
&= \sum_{l\in\mathbb{Z}^d} N\beta_l(x)\left(f\left(x+\frac{l}{N}\right) - f(x)\right)\\
&= \sum_{l\in\mathbb{Z}^d} N\beta_l(x)\left(f\left(x+\frac{l}{N}\right) - f(x) - \frac{l}{N}\nabla f(x)\right) + F(x)\nabla f(x), \qquad x \in E_N.
\end{aligned}$$

By the strong law of large numbers, we know that

$$\lim_{N\to\infty}\sup_{u\le v}\left|\frac{1}{N}\tilde{Y}_l(Nu)\right| = 0, \qquad \text{a.s.}$$

for any $v \ge 0$. As a consequence, the following theorem holds (Ethier and Kurtz (1986), page 456).

Theorem 6.4. *Suppose that for each compact $K \subset E$,*

$$\sum_{l\in\mathbb{Z}^d} |l| \sup_{x\in K}\beta_l(x) < +\infty,$$

and there exists $M_K > 0$ such that

$$|F(x) - F(y)| \le M_K|x-y|, \qquad x, y \in K;$$

suppose $X^{(N)}$ satisfies equation (6.5) above, with

$$\lim_{N\to\infty} X^{(N)}(0) = x_0 \in \mathbb{R}^d.$$

Then, for every $t \geq 0$,

$$\lim_{N\to\infty} \sup_{s\leq t} \left| X^{(N)}(s) - x(s) \right| = 0 \qquad a.s.,$$

where $x(t)$, $t \in \mathbb{R}_+$ *is the unique solution of*

$$x(t) = x_0 + \int_0^t F(x(s))ds, \qquad t \geq 0,$$

wherever it exists.

For the application of the above theorem to the general stochastic epidemic introduced at the beginning of this section see problem 6.9. For a graphical illustration of the above see Figures 6.7 and 6.8. Further, and interesting examples may be found in section 6.4 of Tan (2002).

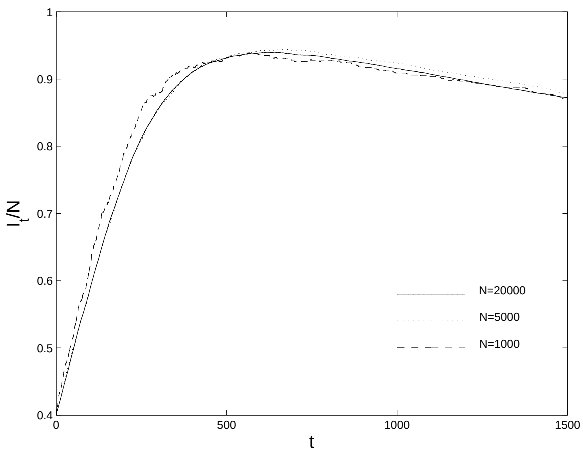

Fig. 6.7. Continuous approximation of a jump model: general stochastic epidemic model with $S_0 = 0.6N$, $I_0 = 0.4N$, $R_0 = 0$, rate of removal of an infective $\delta = 10^{-4}$; infection rate of a susceptible $k = 8 \cdot 10^{-3}N$; time step $dt = 10^{-2}$; interval of observation $[0, 1500]$. The three lines represent the simulated I_t/N as a function of time t for three different values of N.

6.3 Population Dynamics: Individual-Based Models

The scope of this chapter is to introduce the reader to the modeling of a system of a large but still finite population of individuals subject to mutual interaction and random dispersal. These systems may well describe the collective

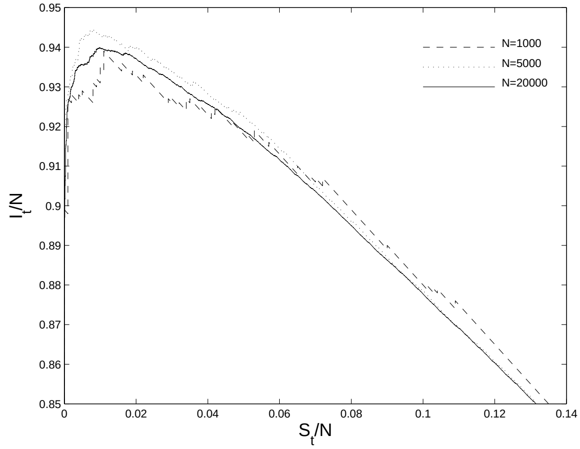

Fig. 6.8. Continuous approximation of a jump model: the same model as in Figure 6.7 of a general stochastic epidemic model with $S_0 = 0.6N$, $I_0 = 0.4N$, $R_0 = 0$, rate of removal of an infective $\delta = 10^{-4}$, infection rate of a susceptible $k = 8 \cdot 10^{-3}N$, time step $dt = 10^{-2}$, interval of observation $[0, 1500]$. The three lines represent the simulated trajectory $(S_t/N, I_t/N)$ for three different values of N.

behavior of individuals in herds, swarms, colonies, armies, etc. (examples can be found in Burger, Capasso, and Morale (2003), Durrett and Levin (1994), Flierl et al. (1999), Gueron, Levin, and Rubenstein (1996), Okubo (1986), Skellam (1951)). It is interesting to observe that under suitable conditions the behavior of such systems in the limit of the number of individuals tending to infinity may be described in terms of nonlinear reaction-diffusion systems. We may then claim that while stochastic differential equations may be utilized for modeling populations at the *microscopic* scale of individuals (Lagrangian approach), partial differential equations provide a *macroscopic* Eulerian description of population densities.

Up to now, Kolmogorov equations like that of Black–Scholes were linear partial differential equations; in this chapter we derive nonlinear partial differential equations for density-dependent diffusions. This field of research, already well established in the general theory of statistical physics (see, e.g., De Masi and Presutti (1991), Donsker and Varadhan (1989), Méléard (1996)), has gained increasing attention, since it also provides the framework for the modelling, analysis, and simulation of agent-based models in economics and finance (see, e.g., Epstein and Axtell (1996)).

The Empirical Distribution

We start from the Lagrangian description of a system of $N \in \mathbb{N} \setminus \{0, 1\}$ particles. Suppose the kth particle ($k \in \{1, \ldots, N\}$) is located at $X_N^k(t)$, at time

$t \geq 0$. Each $(X_N^k(t))_{t\in\mathbb{R}_+}$ is a stochastic process valued in the state space $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$, $d \in \mathbb{N} \setminus \{0\}$, on a common probability space $(\Omega, \mathcal{F}, P)$. An equivalent description of the above system may be given in terms of the (random) measures $\epsilon_{X_N^k(t)}$ $(k = 1, 2, \ldots, N)$ on $\mathcal{B}_{\mathbb{R}^d}$ such that, for any real function $f \in C_0(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} f(y)\epsilon_{X_N^k(t)}(dy) = f\left(X_N^k(t)\right).$$

As a consequence, information about the collective behavior of the N particles is provided by the so-called *empirical measure*, i.e., the random measure on $\mathbb{R}^d$:

$$X_N(t) := \frac{1}{N}\sum_{k=1}^{N} \epsilon_{X_N^k(t)}, \qquad t \in \mathbb{R}_+.$$

This measure may be considered as the empirical spatial distribution of the system. It is such that for any $f \in C_0(\mathbb{R}^d)$:

$$\int_{\mathbb{R}^d} f(y)[X_N(t)](dy) = \frac{1}{N}\sum_{k=1}^{N} f\left(X_N^k(t)\right).$$

In particular, given a region $B \in \mathcal{B}_{\mathbb{R}^d}$, the quantity

$$[X_N(t)](B) := \frac{1}{N}\left(\sharp\left\{X_N^k(t) \in B\right\}\right)$$

denotes the relative frequency of individuals, out of N, that at time t stay in B. This is why the measure-valued process

$$X_N : t \in \mathbb{R}_+ \to X_N(t) = \frac{1}{N}\sum_{k=1}^{N} \epsilon_{X_N^k(t)} \in \mathcal{M}_{\mathbb{R}^d} \tag{6.6}$$

is called the process of *empirical distributions* of the system of N particles.

The Evolution Equations

The Lagrangian description of the dynamics of the system of interacting particles is given via a system of stochastic differential equations. Suppose that for any $k \in \{1, \ldots, N\}$, the process $(X_N^k(t))_{t\in\mathbb{R}_+}$ satisfies the stochastic differential equation

$$dX_N^k(t) = F_N[X_N(t)](X_N^k(t))dt + \sigma_N dW^k(t), \tag{6.7}$$

subject to a suitable initial condition $X_N^k(0)$, which is an $\mathbb{R}^d$-valued random variable. Thus we are assuming that the kth particle is subject to random dispersal, modelled as a Brownian motion W^k. In fact, we suppose that W^k,

$k = 1, \ldots, N$, is a family of independent standard Wiener processes. Furthermore the common variance σ_N^2 may depend on the total number of particles.

The drift term is defined in terms of a given function

$$F_N : \mathcal{M}_{\mathbb{R}^d} \to C(\mathbb{R}^d)$$

and it describes the "interaction" of the kth particle located at $X_N^k(t)$ with the random field $X_N(t)$ generated by the whole system of particles at time t. An evolution equation for the empirical process $(X_N(t))_{t\in\mathbb{R}_+}$ can be obtained thanks to Itô's formula. For each individual particle $k \in \{1, \ldots, N\}$ subject to its stochastic differential equation, given $f \in C_b^2(\mathbb{R}^d \times \mathbb{R}_+)$, we have

$$\begin{aligned} f\left(X_N^k(t), t\right) &= f\left(X_N^k(0), 0\right) + \int_0^t F_N[X_N(s)]\left(X_N^k(s)\right) \nabla f\left(X_N^k(s), s\right) ds \\ &\quad + \int_0^t \left[\frac{\partial}{\partial s} f\left(X_N^k(s), s\right) + \frac{\sigma_N^2}{2} \triangle f\left(X_N^k(s), s\right)\right] ds \\ &\quad + \sigma_N \int_0^t \nabla f\left(X_N^k(s), s\right) dW^k(s). \end{aligned} \tag{6.8}$$

Correspondingly, for the empirical process $(X_N(t))_{t\in\mathbb{R}_+}$, we get the following weak formulation of its evolution equation. For any $f \in C_b^{2,1}(\mathbb{R}^d \times \mathbb{R}_+)$ we have

$$\begin{aligned} \langle X_N(t), f(\cdot, t)\rangle &= \langle X_N(0), f(\cdot, 0)\rangle + \int_0^t \langle X_N(s), F_N[X_N(s)](\cdot)\nabla f(\cdot, s)\rangle\, ds \\ &\quad + \int_0^t \left\langle X_N(s), \frac{\sigma_N^2}{2} \triangle f(\cdot, s) + \frac{\partial}{\partial s} f(\cdot, s)\right\rangle ds \\ &\quad + \frac{\sigma_N}{N} \int_0^t \sum_k \nabla f\left(X_N^k(s), s\right) dW^k(s). \end{aligned} \tag{6.9}$$

In the previous expressions, we have used the notation

$$\langle \mu, f\rangle = \int f(x)\mu(dx), \tag{6.10}$$

for any measure μ on $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$ and any (sufficiently smooth) function $f : \mathbb{R}^d \to \mathbb{R}$.

The last term of (6.9) is a martingale with respect to the process's $(X_N(t))_{t\in\mathbb{R}_+}$ natural filtration. Hence we may apply Doob's inequality (see Proposition 2.69) such that

$$E\left[\sup_{t\leq T} |M_N(f, t)|^2\right] \leq \frac{4\sigma_N^2 \|\nabla f\|_\infty^2 T}{N}.$$

This shows that, for N sufficiently large, the martingale term, which is the only source of stochasticity of the evolution equation for $(X_N(t))_{t\in\mathbb{R}_+}$, tends to zero, for N tending to infinity, since ∇f is bounded in $[0,T]$, and $\frac{\sigma_N^2}{N} \to 0$ for N tending to infinity. Under these conditions we may conjecture that a limiting measure-valued deterministic process $(X_\infty(t))_{t\in\mathbb{R}_+}$ exists, whose evolution equation (in weak form) is

$$\langle X_\infty(t), f(\cdot,t)\rangle = \langle X_\infty(0), f(\cdot,0)\rangle + \int_0^t \langle X_\infty(s), F[X_\infty(s)](\cdot)\nabla f(\cdot,s)\rangle\, ds$$
$$+ \int_0^t \left\langle X_\infty(s), \frac{\sigma_\infty^2}{2}\triangle f(\cdot,s) + \frac{\partial}{\partial s} f(\cdot,s)\right\rangle ds$$

for $\sigma_\infty^2 \geq 0$. Actually, various nontrivial mathematical problems arise in connection with the existence of a limiting measure-valued process $(X_\infty(t))_{t\in\mathbb{R}_+}$. A typical resolution includes the following:

1. Prove the existence of a deterministic limiting measure-valued process $(X_\infty(t))_{t\in\mathbb{R}_+}$.
2. Prove the absolute continuity of the limiting measure with respect to the usual Lebesgue measure on $\mathbb{R}^d$.
3. Provide an evolution equation for the density $p(x,t)$.

In the following subsections we will show how the above procedure has been carried out in particular cases.

A "Moderate" Repulsion Model

As an example we consider the system (due to Oelschläger (1990))

$$dX_N^k(t) = -\frac{1}{N}\sum_{m=1, m\neq k}^{N} \nabla V_N\left(X_N^k(t) - X_N^m(t)\right) dt + dW^k(t), \tag{6.11}$$

where W^k, $k = 1,\ldots,N$, represent N independent standard Brownian motions valued in $\mathbb{R}^d$ (here all variances are set equal to 1). The kernel V_N is chosen of the form

$$V_N(\mathbf{x}) = \chi_N^d V_1(\chi_N \mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^d, \tag{6.12}$$

where V_1 is a symmetric probability density with compact support in $\mathbb{R}^d$ and

$$\chi_N = N^{\frac{\beta}{d}}, \qquad \beta \in]0,1[. \tag{6.13}$$

With respect to the general structure introduced in the preceding subsection on evolution equations, we have assumed that the drift term is given by

$$F_N[X_N(t)]\left(X_N^k(t)\right) = [\nabla V_N * X_N(t)]\left(X_N^k(t)\right)$$
$$= -\frac{1}{N}\sum_{m=1, m\neq k}^{N} \nabla V_N\left(X_N^k(t) - X_N^m(t)\right).$$

System (6.11) describes a population of N individuals, subject to random dispersal (Brownian motion) and to repulsion within the range of the kernel V_N. The choice of the scaling (6.12) in terms of the parameter β means that the range of interaction of each individual with the rest of the population is a decreasing function of N (correspondingly, the strength is an increasing function of N). On the other hand, the fact that β is chosen to belong to $]0,1[$ is relevant for the limiting procedure. It is known as *moderate interaction* and allows one to apply suitable convergence results (*laws of large numbers*) (see Oelschläger (1985)).

For the sake of useful regularity conditions, we assume that

$$V_1 = W_1 * W_1,$$

where W_1 is a symmetric probability density with compact support in $\mathbb{R}^d$, satisfying the condition

$$\int_{\mathbb{R}^d} (1 + |\lambda|^\alpha)|\widetilde{W_1}(\lambda)|^2 d\lambda < \infty \tag{6.14}$$

for some $\alpha > 0$ (here $\widetilde{W_1}$ denotes the Fourier transform of W_1). Henceforth we also make use of the following notations:

$$W_N(x) = \chi_N^d W_1(\chi_N x), \tag{6.15}$$
$$h_N(x,t) = (X_N(t) * W_N)(x), \tag{6.16}$$
$$V_N(x) = \chi_N^d V_1(\chi_N x) = (W_N * W_N)(x), \tag{6.17}$$
$$g_N(x,t) = (X_N(t) * V_N)(x) = (h_N(\cdot,t) * W_N)(x), \tag{6.18}$$

so that system (6.11) can be rewritten as

$$dX_N^k(t) = -\nabla g_N(X_N^k(t),t)dt + dW^k(t), \qquad k = 1,\dots,N. \tag{6.19}$$

The following theorem holds.

Theorem 6.5. *Let*

$$X_N(t) = \frac{1}{N}\sum_{k=1}^N \epsilon_{X_N^k(t)}$$

be the empirical process associated with system (6.11). Assume that

1. *condition (6.14) holds;*
2. $\beta \in]0, \frac{d}{d+2}[$*;*
3. $$\sup_{N\in\mathbb{N}} E\left[\langle X_N(0), \varphi_1\rangle\right] < \infty, \qquad \varphi_1(x) = (1+x^2)^{1/2}; \tag{6.20}$$
4. $$\sup_{N\in\mathbb{N}} E\left[||h_N(\cdot,0)||_2^2\right] < \infty; \tag{6.21}$$

5.
$$\lim_{N\to\infty} \mathcal{L}\left(X_N(0)\right) = \epsilon_{\Xi_0} \text{ in } \mathcal{M}(\mathcal{M}(\mathbb{R}^d)), \tag{6.22}$$

where Ξ_0 *is a probability measure having a density* $p_0 \in C_b^{2+\alpha}(\mathbb{R}^d)$ *with respect to the usual Lebesgue measure on* $\mathbb{R}^d$.

Then the empirical process X_N *converges to* X_∞*, which admits a density satisfying the evolution equation*

$$\begin{aligned} \frac{\partial}{\partial t} p(x,t) &= \frac{1}{2}\left(\Delta p(x,t)\right)^2 + \frac{1}{2}\Delta p(x,t), \\ &= \nabla(p(x,t)\nabla p(x,t)) + \frac{1}{2}\Delta p(x,t), \\ p(\cdot,0) &= p_0. \end{aligned} \tag{6.23}$$

More precisely,

$$\lim_{N\to\infty} \mathcal{L}(X_N) = \epsilon_{\Xi} \text{ in } \mathcal{M}(C([0,T],\mathcal{M}(\mathbb{R}^d))), \tag{6.24}$$

where

$$\Xi = (\Xi(t))_{0\le t\le T} \in C([0,T],\mathcal{M}(\mathbb{R}^d))$$

admits a density

$$p \in C_b^{2+\alpha,1+\frac{\alpha}{2}}(\mathbb{R}^d \times [0,T]),$$

which satisfies

$$\begin{aligned} \frac{\partial}{\partial t} p(x,t) &= \frac{1}{2}\nabla(1+2p(x,t))\nabla p(x,t), \\ p(x,0) &= p_0(x). \end{aligned} \tag{6.25}$$

It can be observed that equation (6.25) includes nonlinear terms, as in the porous media equation (see Oelschläger (1990)). This is due to the repulsive interaction between particles, which in the limit produces a density-dependent diffusion. A linear diffusion persists because the variance of the Brownian motions in the individual equations was kept constant. We will see in a second example how it may vanish when the individual variances tend to zero for N tending to infinity. We will not provide a detailed proof of Theorem 6.5, even though we are going to provide a significant outline of it, leaving further details to the referred literature.

By proceeding as in the previous subsection, a straightforward application of Doob's inequality for martingales (Proposition 2.69) justifies the vanishing of the noise term in the following evolution equation for the empirical measure $(X_N(t))_{t\in\mathbb{R}_+}$:

$$\langle X_N(t), f(\cdot,t)\rangle = \langle X_N(0), f(\cdot,0)\rangle + \int_0^t \langle X_N(s), \nabla g_N(\cdot,s)\nabla f(\cdot,s)\rangle\, ds$$

$$+\int_0^t \left\langle X_N(s), \frac{\sigma^2}{2}\triangle f(\cdot, s) + \frac{\partial}{\partial s} f(\cdot, s) \right\rangle ds$$
$$+\frac{\sigma}{N}\int_0^t \sum_k \nabla f\left(X_N^k(s), s\right) dW^k(s) \qquad (6.26)$$

for a given $T > 0$ and any $f \in C_b^{2,1}(\mathbb{R}^d \times [0,T])$. The major difficulty in a rigorous proof of Theorem 6.5 comes from the nonlinear term

$$\Xi_{N,f}(t) = \int_0^t \langle X_N(s), \nabla g_N(\cdot, s)\nabla f(\cdot, s)\rangle\, ds. \qquad (6.27)$$

If we rewrite (6.27) in an explicit form we get

$$\Xi_{N,f}(t) = \int_0^t \frac{1}{N^2} \sum_{k,m=1}^{N} \nabla V_N\left(X_N^k(s) - X_N^m(s)\right) \nabla f\left(X_N^k(s), s\right) ds. \qquad (6.28)$$

Since for $\beta > 0$ the kernel $V_N \to \delta_0$, namely the Dirac delta function, this shows that, in the limit, even small changes of the relative position of neighbouring particles may have a considerable effect on $\Xi_{N,f}(t)$. But in any case, the regularity assumptions made on the kernel V_N let us state the following lemma which provides sufficient estimates about g_N and h_N as defined above.

Lemma 6.6. *Under the assumptions 2 and 4 of Theorem 6.5, the following holds:*

$$E\left[\sup_{t\leq T} \|h_N(\cdot, t)\|_2^2 + \int_0^t \left\langle X_N(s), |\nabla g_N(\cdot, s)|^2\right\rangle ds + \int_0^t \|\nabla h_N(\cdot, t)\|_2^2 ds\right] < \infty. \qquad (6.29)$$

As a consequence the sequence $\{h_N(\cdot, t) : N \in \mathbb{N}\}$ *is relatively compact in* $L^2(\mathbb{R}^d)$.

A significant consequence of the above lemma is the following one.

Lemma 6.7. *With* X_N *as above, the sequence* $\mathcal{L}(X_N)$ *is relatively compact in the space* $\mathcal{M}(C([0,T], \mathcal{M}(\mathbb{R}^d)))$.

By Lemma 6.7 we may claim that a subsequence of $(\mathcal{L}(X_N))_{N\in\mathbb{N}}$ exists, which converges to a probability measure on the space $\mathcal{M}(C([0,T], \mathcal{M}(\mathbb{R}^d)))$ (see the appendix on convergence of probability measures). The Skorohod representation Theorem 1.158 then assures that a process X_∞^k exists in $C([0,T], \mathcal{M}(\mathbb{R}^d))$ such that

$$\lim_{l\to\infty} X_{N_l} = X_\infty^k, \qquad \text{almost surely with respect to } P.$$

If we can assure the uniqueness of the limit, then all X_∞^k will coincide with some X_∞.

Remark 6.8. We need to notice that a priori the limiting process X_∞ may still be a random process in $C([0,T],\mathcal{M}(\mathbb{R}^d))$.

The proof of Theorem 6.5 is now based on the proof of the two following lemmas. Uniqueness of X_∞ is a consequence of Lemma 6.10.

Lemma 6.9. *Under the assumptions of Theorem 6.5 the random variable $X_\infty(t)$ admits almost surely with respect to P a density $h_\infty(\cdot,t)$ with respect to the usual Lebesgue measure on $\mathbb{R}^d$ for any $t \in [0,T]$. Moreover,*

$$\langle X_\infty(t), f\rangle = \langle X_\infty(0), f\rangle - \frac{1}{2}\int_0^t \langle \nabla h_\infty(\cdot,s), (1+2h_\infty(\cdot,s))\nabla f\rangle\, ds,$$

with $0 \leq t \leq T$, $f \in C_b^1(\mathbb{R}^d)$, almost surely with respect to P.

This shows that if we assume that $X_\infty(0)$ admits a deterministic density p_0 at time $t=0$, then $(X_\infty(t))_{t\in[0,T]}$ satisfies a deterministic evolution equation and is thus itself a deterministic process on $C([0,T],\mathcal{M}(\mathbb{R}^d))$. From the general theory we know that equation (6.23) admits a unique solution $p \in C_b^{2+\alpha,1+\alpha/2}(\mathbb{R}^d \times [0,T])$. We can now state the following lemma.

Lemma 6.10.

$$\|h_\infty(\cdot,t) - p(\cdot,t)\|_2^2 \leq C\int_0^t \|h_\infty(\cdot,s) - p(\cdot,s)\|_2^2 ds.$$

Due to Gronwall's Lemma 4.3 we may then state that

$$\sup_{t\leq T} \|h_\infty(\cdot,t) - p(\cdot,t)\|_2^2 = 0,$$

which concludes the proof of Theorem 6.5.

Ant Colonies

As another example, we consider a model for ant colonies. The latter provide an interesting concept of *aggregation* of individuals. According to a model proposed in Morale, Capasso, and Oelschläger (2004), (1998) (see also Burger, Capasso, and Morale (2003)) (based on an earlier model by Grünbaum and Okubo (1994)), in a colony or in an army (in which case the model may be applied to any cross section) ants are assumed to be subject to two conflicting *social forces*: long-range attraction and short-range repulsion. Hence we consider the following basic assumptions:

(i) Particles tend to aggregate subject to their interaction within a range of size $R_a > 0$ (finite or not). This corresponds to the assumption that each particle is capable of perceiving the others only within a suitable sensory range; in other words, each particle has a limited knowledge of the spatial distribution of its neighbors.

(ii) Particles are subject to repulsion when they come "too close" to each other.

We may express assumptions (i) and (ii) by introducing in the drift term F_N in (6.7) two additive components (see Warburton and Lazarus (1991)): F_1, responsible for aggregation, and F_2, for repulsion, such that

$$F_N = F_1 + F_2.$$

The Aggregation Term F_1

We introduce a convolution kernel $G_a : \mathbb{R}^d \to \mathbb{R}_+$, having a support confined to the ball centered at $0 \in \mathbb{R}^d$ and radius $R_a \in \bar{\mathbb{R}}_+$ as the range of sensitivity for aggregation, independent of N. A *generalized gradient* operator is obtained as follows. Given a measure μ on $\mathbb{R}^d$, we define the function

$$[\nabla G_a * \mu](x) = \int_{\mathbb{R}^d} \nabla G_a(x-y)\mu(dy), \qquad x \in \mathbb{R}^d,$$

as the classical convolution of the gradient of the kernel G_a with the measure μ. Furthermore, G_a is such that

$$G_a(x) = \widehat{G}_a(|x|), \tag{6.30}$$

with $\widehat{G}_a$ a decreasing function in $\mathbb{R}_+$. We assume that the aggregation term F_1 depends on such a generalized gradient of $X_N(t)$ at $X_N^k(t)$:

$$F_1[X_N(t)]\left(X_N^k(t)\right) = [\nabla G_a * X_N(t)]\left(X_N^k(t)\right). \tag{6.31}$$

This means that each individual feels this generalized gradient of the measure $X_N(t)$ with respect to the kernel G_a. The positive sign for F_1 and (6.30) expresses a force of attraction of the particle in the direction of increasing concentration of individuals.

We emphasize the great generality provided by this definition of a generalized gradient of a measure μ on $\mathbb{R}^d$. By using particular shapes of G_a, one may include angular ranges of sensitivity, asymmetries, etc. at a finite distance (see Gueron et al (1996)).

The Repulsion Term F_2

As far as repulsion is concerned we proceed in a similar way by introducing a convolution kernel $V_N : \mathbb{R}^d \to \mathbb{R}_+$, which determines the range and the strength of influence of neighbouring particles. We assume (by anticipating a limiting procedure) that V_N depends on the total number N of interacting particles. Let V_1 be a continuous probability density on $\mathbb{R}^d$ and consider the scaled kernel $V_N(x)$ as defined in (6.12), again with $\beta \in]0,1[$. It is clear that

$$\lim_{N\to+\infty} V_N = \delta_0, \tag{6.32}$$

where δ_0 is Dirac's delta function. We define

$$\begin{aligned} F_2[X_N(t)]\left(X_N^k(t)\right) &= -\left(\nabla V_N * X_N(t)\right)\left(X_N^k(t)\right) \\ &= -\frac{1}{N}\sum_{m=1}^{N} \nabla V_N\left(X_N^k(t) - X_N^m(t)\right). \end{aligned} \tag{6.33}$$

This means that each individual feels the gradient of the population in a small neighborhood. The negative sign for F_2 expresses a drift towards decreasing concentration of individuals. In this case the range of the repulsion kernel decreases to zero as the size N of the population increases to infinity.

The Diffusion Term

In this model randomness may be due to both external sources and "social" reasons. The external sources could, for instance, be unpredictable irregularities of the environment (like obstacles, changeable soils, varying visibility). On the other hand, the innate need of interaction with peers is a social reason. As a consequence, randomness can be modelled by a multidimensional Brownian motion $\mathbf{W}_t$.

The coefficient of $d\mathbf{W}_t$ is a matrix function depending upon the distribution of particles or some environmental parameters. Here, we take into account only the intrinsic stochasticity due to the need of each particle to interact with others. In fact, experiments carried out on ants have shown this need. Hence, simplifying the model, we consider only one Brownian motion dW_t with the variance of each particle σ_N depending on the total number of particles, not on their distribution. We could interpret this as an approximation of the model by considering all the stochasticities (also the ones due to the environment) modeled by $\sigma_N dW_t$.

Since σ_N expresses the intrinsic randomness of each individual due to its need for social interaction, it should be decreasing as N increases. Indeed, if the number of particles is large, the mean free path of each particle may reduce down to a limiting value that may eventually be zero:

$$\lim_{N\to\infty} \sigma_N = \sigma_\infty \geq 0. \tag{6.34}$$

Scaling Limits

Let us discuss the two choices for the interaction kernel in the aggregation and repulsion terms, respectively. They anticipate the limiting procedure for N tending to infinity. Here we are focusing on two types of scaling limits, the *McKean–Vlasov limit*, which applies to the long-range aggregation, and the

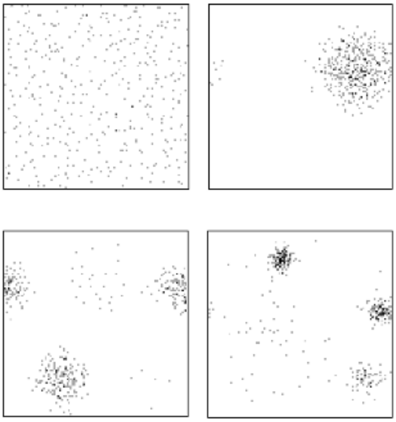

Fig. 6.9. A simulation of the long-range aggregation (6.31) and short-range repulsion (6.33) model for the ant colony with diffusion.

moderate limit, which applies to the short-range repulsion. In the previous subsection we have already considered the moderate limit case.

Mathematically the two cases correspond to the choice made on the interaction kernel. In the moderate limit case (see, e.g., Oelschläger (1985)) the kernel is scaled with respect to the total size of the population N via a parameter $\beta \in]0,1[$. In this case the range of interaction among particles is reduced to zero for N tending to infinity. Thus any particle interacts with many (of order $\frac{N}{\alpha(N)}$) other particles in a small volume (of order $\frac{1}{\alpha(N)}$), where both $\alpha(N)$ and $\frac{N}{\alpha(N)}$ tend to infinity. In the McKean–Vlasov case (see, e.g., Méléard (1996)) $\beta = 0$, so that the range of interaction is independent of N, and as a consequence any particle interacts with order N other particles.

This is why in the moderate limit we may speak of *mesoscale*, which lies between the *microscale* for the typical volume occupied by each individual and the *macroscale* applicable to the typical volume occupied by the total population. Obviously, it would be possible also to consider interacting particle systems rescaled by $\beta = 1$. This case is known as the hydrodynamic case, for which we refer to the literature (De Masi and Presutti (1991), Donsker and Varadhan (1989)).

The case $\beta > 1$ is less significant in population dynamics. It would mean that the range of interaction decreases much faster than the typical distance between neighboring particles. So most of the time particles do not approach sufficiently close to feel the interaction.

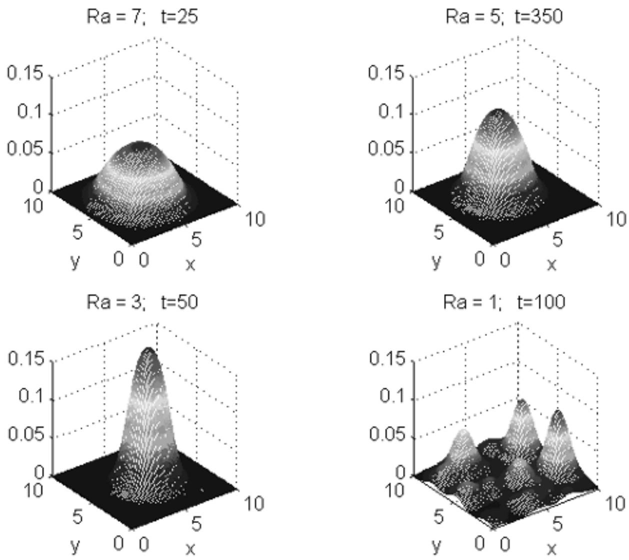

Fig. 6.10. A simulation of the long-range aggregation (6.31) and short-range repulsion (6.33) model for the ant colony with diffusion (smoothed empirical distribution).

Evolution Equations

Again, the fundamental tool for deriving an evolution equation for the empirical measure process is Itô's formula. As in the previous case, the time evolution of any function $f\left(X_N^k(t), t\right)$, $f \in C_b^2(\mathbb{R}^d \times \mathbb{R}_+)$, of the trajectory $\left(X_N^k(t)\right)_{t\in\mathbb{R}_+}$ of the individual particle, subject to the stochastic differential equation (6.7), is given by (6.8). By taking into account expressions (6.31) and (6.33) for F_1 and F_2 and (6.10), then from (6.8), we get the following weak formulation of the time evolution of $X_N(t)$ for any $f \in C_b^{2,1}(\mathbb{R}^d \times [0, \infty[)$:

$$\begin{aligned}
\langle X_N(t), f(\cdot,t)\rangle = \langle X_N(0), f(\cdot,0)\rangle &+ \int_0^t \langle X_N(t), (X_N(s) * \nabla G_a) \cdot \nabla f(\cdot, s)\rangle \, ds \\
&- \int_0^t \langle X_N(t), \nabla g_N(\cdot, s) \cdot \nabla f(\cdot, s)\rangle \, ds \\
&+ \int_0^t \left\langle X_N(t), \frac{\sigma_N^2}{2} \triangle f(\cdot, s) + \frac{\partial}{\partial s} f(\cdot, s) \right\rangle ds \\
&+ \frac{\sigma_N}{N} \int_0^t \sum_k \nabla f\left(X_N^k(s), s\right) dW^k(s), \qquad (6.35)
\end{aligned}$$

$$g_N(x,t) = (X_N(t) * V_N)(x). \qquad (6.36)$$

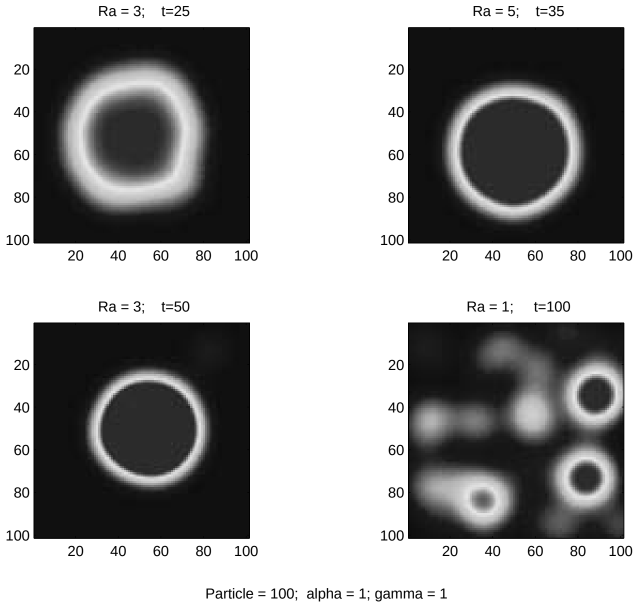

Fig. 6.11. A simulation of the long-range aggregation (6.31) and short-range repulsion (6.33) model for the ant colony with diffusion (two-dimensional projection of the smoothed empirical distribution).

Also for this case we may proceed as in the previous subsection on evolution equations with the analysis of the last term in (6.35). The process

$$M_N(f,t) = \frac{\sigma_N}{N} \int_0^t \sum_k \nabla f \left(X_N^k(s), s \right) dW^k(s), \qquad t \in [0,T],$$

is a martingale with respect to the process's $(X_N(t))_{t \in \mathbb{R}_+}$ natural filtration. By applying Doob's inequality (Proposition 2.69), we obtain

$$E \left[\sup_{t \le T} |M_N(f,t)| \right]^2 \le \frac{4\sigma_N^2 \|\nabla f\|_\infty^2 T}{N}.$$

Hence, by assuming that σ_N remains bounded as in (6.34), $M_N(f,\cdot)$ vanishes in the limit $N \to \infty$. This is again the essential reason of the deterministic limiting behavior of the process, since then its evolution equation will no longer be perturbed by Brownian noise.

We will not go into more details at this point. The procedure is the same as for the previous model. But here we confine ourselves to a formal convergence procedure. Indeed, let us suppose that the empirical process $(X_N(t))_{t \in \mathbb{R}_+}$ tends, as $N \to \infty$, to a deterministic process $(X(t))_{t \in \mathbb{R}_+}$, which for any t is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$, with density $\rho(x,t)$:

$$\begin{aligned}\lim_{N\to\infty}\langle X_N(t), f(\cdot,t)\rangle &= \langle X(t), f(\cdot,t)\rangle \\ &= \int f(x,t)\rho(x,t)dx, \qquad t\geq 0.\end{aligned}$$

As a formal consequence we get

$$\begin{aligned}\lim_{N\to\infty} g_N(x,t) &= \lim_{N\to\infty}(X_N(t)*V_N)(x) = \rho(x,t),\\ \lim_{N\to\infty}\nabla g_N(x,t) &= \nabla\rho(x,t),\\ \lim_{N\to\infty}(X_N(t)*\nabla G_a)(x) &= (X(t)*\nabla G_a(x))\\ &= \int \nabla G_a(x-y)\rho(y,t)dy.\end{aligned}$$

Hence, by applying the above limits, from (6.35) we obtain

$$\begin{aligned}&\int_{\mathbb{R}^d} f(x,t)\rho(x,t)dx\\ &= \int_{\mathbb{R}^d} f(x,0)\rho(x,0)dx\\ &\quad +\int_0^t ds\int_{\mathbb{R}^d} dx\,[(\nabla G_a*\rho(\cdot,s))(x)-\nabla\rho(x,s)]\cdot\nabla f(x,s)\rho(x,s)\\ &\quad +\int_0^t ds\int_{\mathbb{R}^d} dx\left[\frac{\partial}{\partial s}f(x,s)\rho(x,s)+\frac{\sigma_\infty^2}{2}\triangle f(x,s)\rho(x,s)\right], \qquad (6.37)\end{aligned}$$

where σ_∞ is defined as in (6.34).

It can be observed that (6.37) is a weak version of the following equation for the spatial density $\rho(x,t)$:

$$\begin{aligned}\frac{\partial}{\partial t}\rho(x,t) &= \frac{\sigma_\infty^2}{2}\triangle\rho(x,t)+\nabla\cdot(\rho(x,t)\nabla\rho(x,t))\\ &\quad -\nabla\cdot[\rho(x,t)(\nabla G_a*\rho(\cdot,t))(x)], \qquad x\in\mathbb{R}^d, t\geq 0, \quad (6.38)\\ \rho(x,0) &= \rho_0(x). \quad (6.39)\end{aligned}$$

In the degenerate case, i.e., if (6.34) holds with equality, equation (6.38) becomes

$$\frac{\partial}{\partial t}\rho(x,t) = \nabla\cdot(\rho(x,t)\nabla\rho(x,t)) - \nabla\cdot[\rho(x,t)(\nabla G_a*\rho(\cdot,t))(x)]. \qquad (6.40)$$

As in the preceding subsection on moderate repulsion, we need to prove existence and uniqueness of a sufficiently regular solution to equation (6.40). We refer to Burger, Capasso, and Morale (2003) or Nagai and Mimura (1983) and also to Carrillo (1999) for a general discussion of this topic.

A Law of Large Numbers in Path Space

In this section we supplement our results on the asymptotics of the empirical processes by a law of large numbers in path space. This means that we study the *empirical measures in path space*

$$X_N = \frac{1}{N}\sum_{k=1}^{N} \epsilon_{X_N^k(\cdot)},$$

where $X_N^k(\cdot) = (X_N^k(t))_{0\le t\le T}$ denotes the entire path of the kth particle in the time interval $[0,T]$. The particles move continuously in $\mathbb{R}^d$. Moreover, X_N is a measure on the space $\mathcal{C}([0,T],\mathbb{R}^d)$ of continuous functions from $[0,T]$ to $\mathbb{R}^d$. As in the case of empirical processes, one can prove the convergence of X_N to some limit Y. The proof can be achieved with a few additional arguments from the limit theorem for the empirical processes.

By heuristic considerations in Morale, Capasso, and Oelschläger (2004) we get a convergence result for the empirical distribution of the drift $\nabla g_N(\cdot,t)$ of the individual particles

$$\lim_{N\to\infty}\int_0^T \langle X_N(t), |\nabla g_N(\cdot,t) - \nabla\rho(\cdot,t)|\rangle\, dt = 0, \tag{6.41}$$

$$\lim_{N\to\infty}\int_0^T \langle X_N(t), |X_N(t) * \nabla G_a - \nabla G_a * \rho(\cdot,t)|\rangle\, dt = 0.$$

So equation (6.41) allows us to replace the drift

$$\nabla g_N(\cdot,t) - X_N(t) * \nabla G_a$$

with the function

$$\nabla\rho(\cdot,t) - \nabla G_a * \rho(\cdot,t)$$

for large N. Hence, for most k, we have $X_k(t) \sim Y(t)$, uniformly in $t\in[0,T]$, where $Y = Y(t)$, $0\le t\le T$, is the solution of

$$dY(t) = [\nabla G_a * \rho(\cdot,t)(Y(t)) - \nabla\rho(Y(t))]\, dt + \sigma_\infty dW^k(t), \tag{6.42}$$

with the initial condition, for each $k = 1,\dots,N$,

$$Y(0) = X_N^k(0). \tag{6.43}$$

So, not only does the density follow the deterministic equation (6.38), which presents the memory of the fluctuations by means of the term $\frac{\sigma_\infty}{2}\triangle\rho$, but also the stochasticity of the movement of each particle is preserved.

For the degenerate case $\sigma_\infty = 0$, the Brownian motion vanishes as $N\to\infty$. From (6.42) the dynamics of a single particle depend on the density of the whole system. This density is the solution of (6.40), which does not contain any diffusion term. So, not only do the dynamics of a single particle become deterministic, but neither is there any memory of the fluctuations present, when the number of particles N is finite. The following result confirms these heuristic considerations (see Morale, Capasso, and Oelschläger (2004)).

Theorem 6.11. *For the stochastic system (6.7)–(6.33) make the same assumptions as in Theorem 6.5. Then we obtain*

$$\lim_{N\to\infty} E\left[\frac{1}{N}\sum_{k=1}^{N}\sup_{t\leq T}\left|X_N^k(t)-Y(t)\right|\right]=0, \tag{6.44}$$

where Y is the solution of (6.42) with the initial solution (6.43) for each $k=1,\ldots,N$ and ρ is the density of the limit of the empirical processes; i.e., it is the solution of (6.40).

Price Herding

As an example of herding in economics we present a model for price herding that has been applied to simulate the prices of cars; see Capasso, Morale, and Sioli (2003). The model is based on the assumption that prices of products of a similar nature and within the same market segment tend to aggregate within a given interaction kernel, which characterizes the segment itself. On the other hand, unpredictable behavior of individual prices may be modelled as a family of mutually independent Brownian motions. Hence we suppose that in a segment of N prices, for any $k\in\{1,\ldots,N\}$ the price $X_N^k(t)$, $t\in\mathbb{R}_+$, satisfies

$$\frac{dX_N^k(t)}{X_N^k(t)}=F_k[\mathbf{X}(t)]\left(X_N^k(t)\right)dt+\sigma_k(\mathbf{X}(t))dW^k(t).$$

As usual, for a population of prices it is more convenient to consider the evolution of rates. For the force of interaction F_k, which depends upon the vector of all individual prices

$$\mathbf{X}(t):=\left(X_N^1(t),\ldots,X_N^N(t)\right),$$

we assume the following model, similar to the ant colony of the previous subsection:

$$F_k[\mathbf{X}(t)]\left(X_N^k(t)\right)=\frac{1}{N}\sum_{j=1}^{N}\frac{1}{A_{jk}}\left(\frac{I_j(t)}{I_k(t)}\right)^{\beta_{jk}}\nabla K_a\left(X_N^k(t)-X_N^j(t)\right), \tag{6.45}$$

which includes the following ingredients:

(a) The aggregation kernel

$$K_a(x)=\frac{1}{\sqrt{2\pi a^2}}e^{-\frac{x^2}{2a^2}},$$
$$\nabla K_a(x)=-\frac{x}{a^2}\frac{1}{\sqrt{2\pi a^2}}e^{-\frac{x^2}{2a^2}}.$$

(b) The sensitivity coefficient for aggregation

$$\frac{1}{A_{jk}}\left(\frac{I_j(t)}{I_k(t)}\right)^{\beta_{jk}}$$

depending (via the parameters A_{jk} and β_{jk}) on the relative market share $I_j(t)$ of the product j with respect to the market share $I_k(t)$ of product k. Clearly, a stronger product will be less sensitive to the prices of competing weaker products.

(c) The coefficient $\frac{1}{N}$ takes into account possible crowding effects, which are also modulated by the coefficients A_{jk}.

As an additional feature a model for inflation may be included in F_k. Given a general rate of inflation $(\alpha_t)_{t\in\mathbb{R}_+}$, F_k may include a term $s_k\alpha_t$ to model via s_k the specific sensitivity of price k. We leave the analysis of the model to the reader, who may refer to Capasso, Morale, and Sioli (2003) for details.

Data are shown in Figure 6.12; parameter estimates are given in Tables 6.1, 6.2, and 6.3; Figure 6.13 shows the simulated car prices based on such estimates.

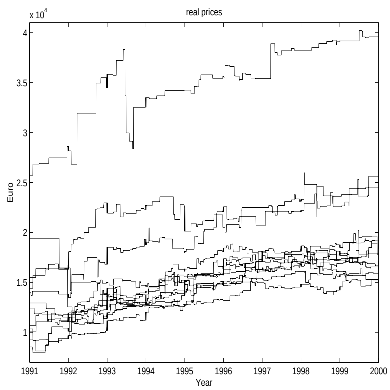

Fig. 6.12. Time series of prices of a segment of cars in Italy during the years 1991–2000 (source: *Quattroruote Magazine, Editoriale Domus, Milan, Italy).*

6.4 Neurosciences

Stein's Model of Neural Activity

The main component of Stein's model (Stein (1965), (1967)) is the depolarization V_t for $t \in \mathbb{R}_+$. A nerve cell is said to be *excited* (or *depolarized*), if $V_t > 0$, and *inhibited*, if $V_t < 0$. In the absence of other events V_t decays according to

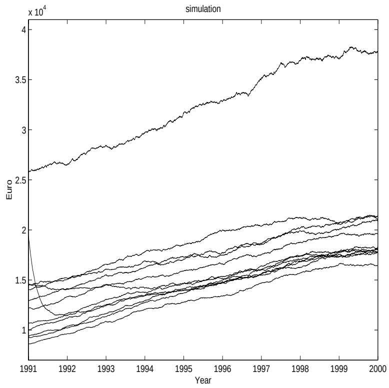

Fig. 6.13. Simulated car prices.

Parameter	Method of Estimation	Estimate	St. Dev.
$X_1(0)$	ML	1.6209E+00	5.8581E-02
$X_2(0)$	ML	8.4813E-01	6.0740E-03
$X_3(0)$	ML	7.4548E-01	2.3420E-02
$X_4(0)$	ML	1.0189E+00	1.2273E-01
$X_5(0)$	ML	1.4164E+00	1.4417E-01
$X_6(0)$	ML	2.4872E+00	6.2947E-02
$X_7(0)$	ML	1.2084E+00	4.7545E-02
$X_8(0)$	ML	1.0918E+00	4.7569E-02
a	ML	5.0767E+03	6.5267E+02

Table 6.1. Estimates for the price herding model (6.45) for the initial conditions $X_k(0)$ and the range of the kernel a.

$$\frac{dV}{dt} = -\alpha V,$$

where $\alpha = 1/\tau$ is the reciprocal of the nerve membrane time constant $\tau > 0$.

In the resting state (initial condition) $V_0 = 0$. Afterwards jumps may occur at random times according to independent Poisson processes $(N_t^E)_{t\in\mathbb{R}_+}$ and $(N_t^I)_{t\in\mathbb{R}_+}$ with intensities λ_E and λ_I, respectively, assumed to be strictly positive real constants. If an excitation (a jump) occurs for N^E, at some time $t_0 > 0$, then

$$V_{t_0} - V_{t_0-} = a_E,$$

whereas if an inhibition (again a jump) occurs for N^I, then

$$V_{t_0} - V_{t_0-} = -a_I,$$

where a_E and a_I are nonnegative real numbers. When V_t attains a given value $\theta > 0$ (the *threshold*), the cell fires. Upon firing V_t is reset to zero along with

Parameter	Method of Estimation	Estimate	St. Dev.
A_{12}	ML	1.0649E-03	3.0865E-02
A_{13}	ML	1.1489E-04	4.1737E-04
A_{14}	ML	1.5779E-03	5.4687E-02
A_{15}	ML	7.6460E-04	1.8381E-02
A_{16}	ML	1.2908E-03	4.0634E-02
A_{17}	ML	1.8114E-03	6.5617E-02
A_{18}	ML	1.5956E-03	5.5572E-02
A_{23}	ML	1.0473E-04	7.2687E-05
A_{24}	ML	1.7397E-04	6.0809E-04
A_{25}	ML	1.7550E-04	5.1100E-04
A_{26}	ML	1.2080E-03	3.7392E-02
A_{27}	ML	9.4809E-04	2.6037E-02
A_{28}	ML	2.7277E-04	2.0135E-03
A_{34}	ML	4.0404E-04	5.5468E-03
A_{35}	ML	1.8136E-04	8.6471E-04
A_{36}	ML	9.5558E-03	4.9764E-01
A_{37}	ML	1.0341E-04	4.4136E-05
A_{38}	ML	7.0953E-04	1.6428E-02
A_{45}	ML	1.0066E-03	2.8485E-02
A_{46}	ML	1.3354E-04	1.3632E-03
A_{47}	ML	2.5239E-04	1.6979E-03
A_{48}	ML	1.1232E-03	3.3652E-02
A_{56}	ML	2.3460E-03	9.2592E-02
A_{57}	ML	1.0143E-03	2.8898E-02
A_{58}	ML	1.1026E-03	3.2724E-02
A_{67}	ML	1.8560E-03	6.8275E-02
A_{68}	ML	2.2820E-03	8.9278E-02
A_{78}	ML	6.4630E-04	1.4003E-02

Table 6.2. Estimates for the price herding model (6.45) for the parameters A_{ij}.

N^E and N^I and the process restarts along the previous model. By collecting all of the above assumptions, the subthreshold evolution equation for V_t may be written in the following form:

$$dV_t = -\alpha V_t dt + a_E dN_t^E - a_I dN_t^I,$$

subject to the initial condition $V_0 = 0$. The model is a particular case of a more general (stochastic) evolution equation of the form

$$dX_t = \alpha(X_t)dt + \int_{\mathbb{R}} \gamma(X_t, u)N(dt, du), \tag{6.46}$$

where N is a marked Poisson process on $\mathbb{R}_+ \times \mathbb{R}$ (in (6.46) the integration is over u). In Stein's model $\alpha(x) = -\alpha x$, with $\alpha > 0$ (or simply $\alpha(x) = -x$, if we assume $\alpha = 1$); $\gamma(x, u) = u$, and the marked Poisson process N has intensity measure

Parameter	Method of Estimation	Estimate	St. Dev.
β_{12}	ML	6.8920E-01	5.8447E+00
β_{13}	ML	2.3463E+00	2.7375E+00
β_{14}	ML	7.2454E-01	6.6182E+00
β_{15}	ML	8.4049E-01	6.2349E+00
β_{16}	ML	7.7929E-01	5.6565E+00
β_{17}	ML	6.6793E-01	5.4208E+00
β_{18}	ML	7.6508E-01	5.8422E+00
β_{23}	ML	2.4531E+00	4.5883E-01
β_{24}	ML	1.6924E+00	6.8734E+00
β_{25}	ML	1.6262E+00	5.7128E+00
β_{26}	ML	1.2122E+00	2.1666E+00
β_{27}	ML	7.5140E-01	7.4760E+00
β_{28}	ML	1.3537E+00	6.0109E+00
β_{34}	ML	1.2444E+00	8.1509E+00
β_{35}	ML	1.7544E+00	8.4976E+00
β_{36}	ML	1.0572E+00	8.0208E+00
β_{37}	ML	2.4730E+00	1.9801E-01
β_{38}	ML	1.0674E+00	8.4626E+00
β_{45}	ML	7.5781E-01	6.7267E+00
β_{46}	ML	2.2121E+00	6.9754E+00
β_{47}	ML	1.7360E+00	6.4971E+00
β_{48}	ML	8.1043E-01	6.1451E+00
β_{56}	ML	7.1269E-01	4.5857E+00
β_{57}	ML	7.7251E-01	6.3947E+00
β_{58}	ML	7.0792E-01	6.5014E+00
β_{67}	ML	8.4060E-01	6.8871E+00
β_{68}	ML	8.1190E-01	6.0759E+00
β_{78}	ML	1.0794E+00	8.4994E+00

Table 6.3. Estimates for the price herding model (6.45) for the parameters β_{ij}.

$$\Lambda((s,t)\times B) = (t-s)\int_B \phi(u)du \qquad \text{for any } s,t\in\mathbb{R}_+, s<t, B\subset\mathcal{B}_{\mathbb{R}}.$$

Here

$$\phi(u) = \lambda_E\delta_0(u-a_E) + \lambda_I\delta_0(u+a_I),$$

with δ_0 the standard Dirac delta distribution. The infinitesimal generator $\mathcal{A}$ of the Markov process $(X_t)_{t\in\mathbb{R}_+}$ given by (6.46) is given by

$$\mathcal{A}f(x) = \alpha(x)\frac{\partial f}{\partial x}(x) + \int_{\mathbb{R}}(f(x+\gamma(x,u)) - f(x))\phi(u)du$$

for any test function f in the domain of $\mathcal{A}$.

The firing problem may be seen as a first passage time through the threshold $\theta > 0$. Let $A =]-\infty, \theta[$. Then the random variable of interest is

$$T_A(x) = \inf\{t\in\mathbb{R}_+ | X_t \in A, X_0 = x\in A\},$$

Parameter	Method of Estimation	Estimate	St. Dev.
s_1	ML	2.0267E-03	2.1858E-04
s_2	ML	5.1134E-03	1.6853E-03
s_3	ML	3.6238E-03	2.5305E-03
s_4	ML	3.6777E-03	2.3698E-03
s_5	ML	1.0644E-04	1.1132E-04
s_6	ML	5.4133E-03	1.2452E-03
s_7	ML	1.0769E-04	1.4414E-04
s_8	ML	2.1597E-03	2.8686E-03
σ_1	MAP	7.0000E-03	2.9073E-06
σ_2	MAP	7.0000E-03	2.9766E-06
σ_3	MAP	7.0000E-03	3.0128E-06
σ_4	MAP	7.0000E-03	2.9799E-06
σ_5	MAP	7.0000E-03	3.0025E-06
σ_6	MAP	7.0000E-03	2.9897E-06
σ_7	MAP	7.0000E-03	2.8795E-06
σ_8	MAP	7.0000E-03	2.9656E-06

Table 6.4. Estimates for the price herding model (6.45) of s_k and σ_k.

which is the first exit time from A. If the indicated set is empty, then we set $T_A(x) = +\infty$. The following result holds

Theorem 6.12. (Tuckwell (1976), Darling and Siegert (1953)). *Let $(X_t)_{t\in\mathbb{R}_+}$ be a Markov process satisfying (6.46) and assume that the existence and uniqueness conditions are fulfilled. Then the distribution function*

$$F_A(x,t) = P(T_A(x) \le t)$$

satisfies

$$\frac{\partial F_A}{\partial t}(x,t) = \mathcal{A}F_A(\cdot,t)(x), \qquad x \in A, t > 0,$$

subject to the initial condition

$$F_A(x,0) = \begin{cases} 0 & \textit{for } x \in A, \\ 1 & \textit{for } x \notin A, \end{cases}$$

and boundary condition

$$F_A(x,t) = 1, \qquad x \notin A, x \ge 0.$$

Corollary 6.13. If the moments

$$\mu_n(x) = E\left[(T_A(x))^n\right], \qquad n \in \mathbb{N}^*,$$

exist, they satisfy the recursive system of equations

$$\mathcal{A}\mu_n(x) = -n\mu_{n-1}(x), \qquad x \in A, \tag{6.47}$$

subject to the boundary conditions

$$\mu_n(x) = 0, \qquad x \notin A.$$

The quantity $\mu_0(x)$, $x \in A$, is the probability of X_t exiting from A in a finite time. It satisfies the equation

$$\mathcal{A}\mu_0(x) = 0, \qquad x \in A, \tag{6.48}$$

subject to

$$\mu_0(x) = 1, \qquad x \notin A.$$

The following lemma is due to Gihman and Skorohod (1972).

Lemma 6.14. *If there exists a bounded function g on $\mathbb{R}$ such that*

$$\mathcal{A}g(x) \leq -1, \qquad x \in A, \tag{6.49}$$

then $\mu_1 < \infty$ and $P(T_A(x) < +\infty) = 1$.

As a consequence of Lemma 6.14 a neuron in Stein's model fires in a finite time with probability 1 and with finite mean interspike interval. This is due to the fact that the solution of (6.48) is $\mu_0(x) = 1$, $x \in \mathbb{R}$, and this satisfies (6.49). The mean first passage time through θ for an initial value x satisfies, by (6.47):

$$-x\frac{d\mu_1}{dx}(x) + \lambda_E\mu_1(x + a_E) + \lambda_I\mu_1(x - a_I) - (\lambda_E + \lambda_I)\mu_1(x) = -1, \tag{6.50}$$

with $x < \theta$ and boundary condition

$$\mu_1(x) = 0, \qquad \text{for } x \geq \theta.$$

The solution of (6.50) is discussed in Tuckwell (1989), where a diffusion approximation of the original Stein's model of neuron firing is also analyzed.

6.5 Exercises and Additions

6.1. Consider a birth-and-death process $(X(t))_{t\in\mathbb{R}_+}$ valued in $\mathbb{N}$, as in section 6.1. In integral form the evolution equation for X will be

$$X(t) = X(0) + \alpha\int X(s-)ds + M(t),$$

where $\alpha = \lambda - \mu$ is the survival rate and $M(t)$ is a martingale. Show that

1.
$$\langle M\rangle(t) = \langle M, M\rangle(t) = (\lambda + \mu)\int_0^t X(s-)ds.$$

2.
$$E[X(t)] = X(0)e^{\alpha t}.$$

3. $X(t)e^{-\alpha t}$ is a square-integrable martingale.
4. $Var[X(t)e^{-\alpha t}] = X(0)\dfrac{\lambda+\mu}{\lambda-\mu}(1-e^{-\alpha t})$.

6.2. (*Age-dependent birth-and-death process*). An age-dependent population can be divided into two subpopulations, described by two marked counting processes. Given $t > 0$, $U^{(1)}(A_0, t)$ describes those individuals who already existed at time $t = 0$ with ages in $A_0 \in \mathcal{B}_{\mathbb{R}_+}$ and are still alive at time t; and $U^{(2)}(T_0, t)$ describes those individuals who are born during $T_0 \in \mathcal{B}_{\mathbb{R}+}$, $T_0 \subset [0, t]$ and are still alive at time t. Assume that the age-specific death rate is $\mu(a),\ a \in \mathbb{R}_+$, and that the birth process $B(T_0), T_0 \in \mathcal{B}_{\mathbb{R}+}$ admits stochastic intensity

$$\alpha(t_0) = \int_0^{+\infty} \beta(a_0 + t_0)U^{(1)}(da_0, t_0-) + \int_0^{t_0-} \beta(t_0 - \tau)U^{(2)}(d\tau, t_0-),$$

where $\beta(a),\ a \in \mathbb{R}_+$ is the age-specific fertility rate. Assume now that suitable densities u_0 and b exist on $\mathbb{R}_+$ such that

$$E[U^{(1)}(A_0, 0)] = \int_{A_0} u_0(a)da$$

and
$$E[B(T_0)] = \int_{T_0} b(\tau)d\tau.$$

Show that the following *renewal equation* holds for any $s \in \mathbb{R}_+$:

$$b(s) = \int_0^{+\infty} da\, u_0(a)\, n(s+a)\, \beta(a+s) + \int_0^s d\tau\, \beta(s-\tau)\, n(s-\tau)\, b(\tau),$$

where $n(t) = \exp\{-\int_0^t \mu(\tau)d\tau\}$, $t \in \mathbb{R}_+$. The reader may refer to Capasso (1988).

6.3. Let $\bar{E}$ be the closure of an open set $E \subset \mathbb{R}^d$ for $d \geq 1$. Consider a spatially structured birth-and-death process associated with the marked point process defined by the random measure on $\mathbb{R}^d$:

$$\nu(t) = \sum_{i=1}^{I(t)} \varepsilon_{X^i(t)},$$

where $I(t),\ t \in \mathbb{R}_+$, denotes the number of individuals in the total population at time t; and $X^i(t)$ denotes the random location of the ith individual in $\bar{E}$. Consider the process defined by the following parameters:

1. $\mu : \bar{E} \to \mathbb{R}_+$ is the spatially structured death rate;
2. $\gamma : \bar{E} \to \mathbb{R}_+$ is the spatially structured birth rate;
3. for any $x \in \bar{E}$, $D(x, \cdot) : \mathcal{B}_{\mathbb{R}^d} \to [0, 1]$ is a probability measure such that $\int_{\bar{E}\setminus\{x\}} D(x, dz) = 1$; $D(x, A)$ for $x \in \bar{E}$ and $A \in \mathcal{B}_{\mathbb{R}^d}$ represents the probability that an individual born in x will be dispersed in A.

Show that the infinitesimal generator of the process is the operator L defined as follows: for any sufficiently regular test function ϕ

$$L\phi(\nu) = \int_{\bar{E}} \nu(dx) \int_{\mathbb{R}^d} \gamma(x) D(x, dz)[-\phi(\nu) + \phi(\nu + \varepsilon_{x+z})] \\ +\mu(x)[-\phi(\nu) + \phi(\nu - \varepsilon_x)].$$

(The reader may refer to Fournier and Méléard (2003) for further analysis.)

6.4. Let X be an integer-valued random variable, with probability distribution $p_k = P(X = k),\ k \in \mathbb{N}$. The probability generating function of X is defined as

$$g_X(s) = E[s^X] = \sum_{k=0}^{\infty} s^k p_k,\ |s| \le 1.$$

Consider a homogeneous birth-and-death process $X(t), t \in \mathbb{R}_+$, with birth rate λ and death rate μ, and initial value $X(0) = k_0 > 0$. Show that the probability generating function $G_X(s; t)$ of $X(t)$ satisfies the partial differential equation

$$\frac{\partial}{\partial t} G_X(s; t) + (1 - s)(\lambda s - \mu) \frac{\partial}{\partial s} G_X(s; t) = 0,$$

subject to the initial condition

$$G_X(s; 0) = s^{k_0}.$$

6.5. Consider now a nonhomogeneous birth-and-death process $X(t), t \in \mathbb{R}_+$, with time-dependent birth rate $\lambda(t)$ and death rate $\mu(t)$, and initial value $X(0) = k_0 > 0$. Show that the probability generating function $G_X(s; t)$ of $X(t)$ satisfies the partial differential equation

$$\frac{\partial}{\partial t} G_X(s; t) + (1 - s)(\lambda(t)s - \mu(t)) \frac{\partial}{\partial s} G_X(s; t) = 0,$$

subject to the initial condition

$$G_X(s; 0) = s^{k_0}.$$

Evaluate the probability of extinction of the population. (The reader may refer to Chiang (1968).)

6.6. Consider the general epidemic process as defined in section 6.1 with infection rate $\kappa = 1$ and removal rate δ. Let $G_Z(x, y; t)$ denote the probability generating function of the random vector $\mathbf{Z}(t) = (S(t), I(t))$, where $S(t)$ denotes the number of susceptibles at time $t \geq 0$ and $I(t)$ denotes the number of infectives at time $t \geq 0$. Assume that $S(0) = s_0$ and $I(0) = i_0$, and let $p(m, n; t) = P(S(t) = m, I(t) = n)$. The joint probability generating function G will be defined as

$$G_{\mathbf{Z}}(x, y; t) = E[x^{S(t)} y^{I(t)}] = \sum_{m=0}^{s_0} \sum_{n=0}^{s_0+i_0-m} p(m, n; t)\, x^m\, y^n.$$

Show that it satisfies the partial differential equation

$$\frac{\partial}{\partial t} G_{\mathbf{Z}}(x, y; t) = y(y - x) \frac{\partial^2}{\partial x \partial y} G_{\mathbf{Z}}(x, y; t) + \delta(1 - y) \frac{\partial}{\partial y} G_{\mathbf{Z}}(x, y; t),$$

subject to the initial condition

$$G_{\mathbf{Z}}(x, y; 0) = x^{s_0} y^{i_0}.$$

6.7. Consider a discrete birth-and-death chain $(Y_n^{(\Delta)})_{n \in \mathbb{N}}$ valued in $S = \{0, \pm\Delta, \pm 2\Delta, \ldots\}$, with step size $\Delta > 0$, and denote by $p_{i,j}$ the one-step transition probabilities

$$p_{ij} = P\left(Y_{n+1}^{(\Delta)} = j\Delta \,\middle|\, Y_n^{(\Delta)} = i\Delta \right) \quad \text{for } i, j \in \mathbb{Z}.$$

Assume that the only nontrivial transition probabilities are

1. $p_{i,i-1} = \gamma_i := \frac{1}{2}\sigma^2 - \frac{1}{2}\mu\Delta$,

2. $p_{i,i+1} = \beta_i := \frac{1}{2}\sigma^2 + \frac{1}{2}\mu\Delta$,

3. $p_{i,i} = 1 - \beta_i - \gamma_i = 1 - \sigma^2$;

where σ^2 and μ are strictly positive real numbers. Note that for Δ sufficiently small, all rates are nonnegative. Consider now the rescaled (in time) process $(Y_{n/\varepsilon}^{(\Delta)})_{n \in \mathbb{N}}$, with $\varepsilon = \Delta^2$; show (formally and possibly rigorously) that the rescaled process weakly converges to a diffusion on $\mathbb{R}$ with drift μ and diffusion coefficient σ^2.

6.8. With reference to the previous problem, show that the same result may be obtained (with suitable modifications) also in the case in which the drift and the diffusion coefficient depend upon the state of the process. For this case show that the probability $\psi(x)$ that the diffusion process reaches c before d, when starting from a point $x \in (c, d) \subset \mathbb{R}$, is given by

$$\psi(x) = \frac{\int_x^d \exp\left\{-\int_c^z \left(2\frac{\mu(y)}{\sigma^2(y)}\right) dy\right\} dz}{\int_c^d \exp\left\{-\int_c^z \left(2\frac{\mu(y)}{\sigma^2(y)}\right) dy\right\} dz}.$$

The reader may refer, e.g., to Bhattacharya and Waymire (1990).

6.9. Consider the general stochastic epidemic with the rescaling proposed at the beginning of section 6.2. Derive the asymptotic ordinary differential system corresponding to Theorem 6.4.

Nomenclature

	"increasing" is used with the same meaning as "nondecreasing"; "decreasing" is used with the same meaning as "nonincreasing." In the strict cases "strictly increasing/strictly decreasing" is used.
$(\Omega, \mathcal{F}, P)$	probability space with Ω a set, $\mathcal{F}$ a σ-algebra of parts of Ω, and P a probability measure on $\mathcal{F}$
$(E, \mathcal{B}_E)$	measurable space with E a set and $\mathcal{B}_E$ a σ-algebra of parts of E
$:=$	equal by definition
$< f, g >$	scalar product of two elements f and g in an Hilbert space
$< M, N >$	predictable covariation of the martingales M and N
$< M >, < M, M >$	predictable variation of the martingale M
$[a, b[$	semiopen interval closed at extreme a and open at extreme b
$[a, b]$	closed interval of extremes a and b
$\bar{\mathbb{R}}$	extended set of real numbers; i.e., $\mathbb{R} \cup \{-\infty, +\infty\}$
$\bar{A}$	closure of a set A depending upon the context
$\bar{C}$	the complement of the set C depending upon the context
Δ	Laplace operator
δ_x	Dirac delta-function localized at x
δ_{ij}	Kronecker delta; i.e., $= 1$ for $i = j$, $= 0$ for $i \neq j$
$\emptyset$	the empty set
ϵ_x	Dirac delta-measure localized at x
$\equiv$	coincide
$\exp\{x\}$	exponential function e^x
$\int^*$	integral of a nonnegative measurable function, finite or not
$\lim_{s \downarrow t}$	limit for s decreasing while tending to t
$\lim_{s \uparrow t}$	limit for s increasing while tending to t
$\mathbb{C}$	the complex plane
$\mathbb{N}$	the set of natural nonnegative integers
$\mathbb{N}^*$	the set of natural (strictly) positive integers

$\mathbb{Q}$	the set of rational numbers
$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{R}_+$	the set of positive (nonnegative) real numbers
$\mathbb{R}_+^*$	the set of (strictly) positive real numbers
$\mathbb{Z}$	the set of all integers
$\mathcal{A}$	infinitesimal generator of a semigroup
$\mathcal{B}_{\mathbb{R}^n}$	σ-algebra of Borel sets on $\mathbb{R}^n$
$\mathcal{B}_E$	σ-algebra of Borel sets generated by the topology of E
$\mathcal{D}_A$	domain of definition of an operator A
$\mathcal{F}_t$ or $\mathcal{F}_t^X$	history of a process $(X_t)_{t\in\mathbb{R}_+}$ up to time t; i.e., the σ-algebra generated by $\{X_s, s \leq t\}$
$\mathcal{F}_{t+}$	$\bigcap_{s>t}\mathcal{F}_t$
$\mathcal{F}_{t-}$	σ-algebra generated by $\sigma(X_s, s < t)$
$\mathcal{F}_X$	σ-algebra generated by the random variable X
$\mathcal{L}(X)$	probability law of X
$\mathcal{L}^p(P)$	set of integrable functions with respect to the measure P
$\mathcal{M}(\mathcal{F}, \bar{\mathbb{R}}_+)$	set of all $\mathcal{F}$-measurable functions with values in $\bar{\mathbb{R}}_+$
$\mathcal{M}(E)$	set of all measures on E
$\mathfrak{P}(\Omega)$	the set of all parts of a set Ω
$\xrightarrow[n]{P}$ or P-lim	convergence in probability
$\xrightarrow[n]{\mathcal{W}}$	weak convergence
$\xrightarrow[n]{a.s.}$	almost sure convergence
$\xrightarrow[n]{d}$	convergence in distribution
$\xrightarrow[n]{P}$	convergence in probability
∇	gradient
Ω	the underlying sample space
ω	an element of the underlying sample space
$\otimes$	product of σ-algebras or product of measures
∂A	boundary of a set A
Φ	cumulative distribution function of a standard normal probability law
$\mathrm{sgn}\{x\}$	sign function; 1, if $x > 0$; 0, if $x = 0$; -1, if $x < 0$
$\sigma(\mathcal{R})$	σ-algebra generated by the family of events $\mathcal{R}$
$\square$	end of a proof
$\|a\|$	absolute value of a number a; or modulus of a complex number a
$\|A\|$ or $\sharp(A)$	cardinal number (number of elements) of a finite set A
$\|\|x\|\|$	the norm of a point x
$]a, b[$	open interval of extremes a, b
$]a, b]$	semiopen interval open at extreme a and closed at extreme b
$a \vee b$	maximum of two numbers

A'	transpose of a matrix A
$A \setminus B$	the set of elements of A that do not belong to B
$a \wedge b$	minimum of two numbers
$B(x, r)$ or $B_r(x)$	the open ball centered at x and having radius r
$C(A)$	set of continuous functions from A to $\mathbb{R}$
$C(A, B)$	set of continuous functions from A to B
$C^k(A)$	set of functions from A to $\mathbb{R}$ with continuous derivatives up to order k
$C^{k+\alpha}(A)$	set of functions from A to $\mathbb{R}$ whose k-th derivatives are Lipschitz continuous with exponent α
$C_0(A)$	set continuous functions on A with compact support
$C_b(A)$ or $BC(A)$	set of bounded continuous functions on A
$Cov[X, Y]$	the covariance of two random variables X and Y
$E[\cdot]$	expected value with respect to an underlying probability law clearly identifiable from the context
$E[Y\|\mathcal{F}]$	conditional expectation of a random variable Y with respect to the σ-algebra $\mathcal{F}$
$E_P[\cdot]$	expected value with respect to the probability law P
$E_x[\cdot]$	expected value conditional upon a given initial state x in a stochastic process
$f * g$	convolution of functions f and g
$f \circ X$ or $f(X)$	a function f composed with a function X
$f\|_A$	the restriction of a function f to the set A
f^-, f^+	negative (positive) part of f; i.e., $f^- = \max\{-f, 0\}$ ($f^+ = \max\{f, 0\}$)
$f^{-1}(B)$	the preimage of the set B by the function f
F_X	cumulative distribution function of a random variable X
$H \bullet X$	stochastic Stieltjes integral of the process H with respect to the stochastic process X
I_A	indicator function associated with a set A; i.e., $I_A(x) = 1$, if $x \in A$ otherwise $I_A(x) = 0$
$L^p(P)$	set of equivalence classes of a.e. equal integrable functions with respect to the measure P
$N(\mu, \sigma^2)$	normal (Gaussian) random variable with mean μ and variance σ^2
$O(\Delta)$	of the same order as Δ
$o(\delta)$	of higher order with respect to δ
P-a.s.	almost surely with respect to the measure P
$P(A\|B)$	conditional probability of an event A with respect to an event B
$P * Q$	convolution of measures P and Q
$P \ll Q$	the measure P is absolutely continuous with respect to the measure Q
$P \sim Q$	the measure P is equivalent to the measure Q
P_X	probability law of a random variable X

P_x	probability law conditional upon a given initial state x in a stochastic process
$Var[X]$	the variance of a random variable X
W_t	standard Brownian motion, Wiener process
$X \sim P$	the random variable X has probability law P
a.e.	almost everywhere
a.s.	almost surely

Index