Stochastic differential equation models in biology

Susanne Ditlevsen

Introduction This chapter is concerned with continuous time processes, which are often modeled as a system of ordinary differential equations. These models assume that the observed dynamics are driven exclusively by internal, deterministic mechanisms. However, real biological systems will always be exposed to influences that are not completely understood or not feasible to model explicitly, and therefore there is an increasing need to extend the deterministic models to models that embrace more complex variations in the dynamics. A way of modeling these elements is by including stochastic influences or noise. A natural extension of a deterministic differential equations model is a system of stochastic differential equations, where relevant parameters are modeled as suitable stochastic processes, or stochastic processes are added to the driving system equations. This approach assumes that the dynamics are partly driven by noise.

All biological dynamical systems evolve under stochastic forces, if we define stochasticity as the parts of the dynamics that we either cannot predict or understand or that we choose not to include in the explicit modeling. To be realistic, models of biological systems should include random influences, since they are concerned with subsystems of the real world that cannot be sufficiently isolated from effects external to the model. The physiological justification to include erratic behaviors in a model can be found in the many factors that cannot be controlled, like hormonal oscillations, blood pressure variations, respiration, variable neural control of muscle activity, enzymatic processes, energy requirements, the cellular metabolism, sympathetic nerve activity, or individual characteristics like body mass index, genes, smoking, stress impacts, etc. Also external influences, like small differences in the experimental procedure, temperature, differences in preparation and administration of drugs, if this is included in the experiment, or maybe the experiments are conducted by different experimentalists that inevitably will exhibit small differences in procedures within the protocols. Different sources of errors will require different modeling of the noise, and these factors should be considered as carefully as the modeling of the deterministic part, in order to make the model predictions and parameter values possible to interpret.

It is therefore essential to understand and investigate the influence of noise in the dynamics. In many cases the noise simply blurs the underlying dynamics without qualitatively affecting it, as is the case with measurement noise or in many linear systems. However, in nonlinear dynamical systems with system noise, the noise will often drastically change the corresponding deterministic dynamics. In general, stochastic effects influence the dynamics, and may enhance, diminish or even completely change the dynamic behavior of the system.

The Wiener process (or Brownian Motion) The most important stochastic process in continuous time is the Wiener process, also called Brownian Motion. It is used as a building block in more elaborate models. In 1828 the Scottish botanist Robert Brown observed that pollen grains suspended in water moved in an apparently random way, changing direction continuously. This was later explained by the pollen grains being bombarded by water molecules, and Brown only contributed to the theory with his name. The precise mathematical formulation to explain this phenomenon was given by Norbert Wiener in 1923.

The Wiener process can be seen as the limit of a random walk when the time steps and the jump sizes go to 0 in a suitable way, and can formally be defined as follows.

Definition 1 (Wiener process) *A stochastic process* $\{W(t)\}_{t\geq 0}$ *is called a* **Wiener process** *or a* **Brownian motion** *if*

i) $W(0) = 0$

ii) $\{W(t)\}_{t\geq 0}$ *has independent increments, i.e.*

$$W_{t_1}, W_{t_2} - W_{t_1}, \cdots, W_{t_k} - W_{t_{k-1}}$$

are independent random variables for all $0 \leq t_1, < t_2 < \cdots < t_k$.

iii) $W(t+s) - W(s) \sim N(0,t)$ *for all* $t > 0$.

Here, $N(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2. Thus, the Wiener process is a Gaussian process: a stochastic process X is called a *Gaussian process* if for any finite set of indices $t_1, \ldots, t_k$ the vector of random variables $(X(t_1), \ldots, X(t_k))$ is following a k-dimensional normal distribution. In fact, it can be shown that any continuous time stochastic process with independent increments and finite second moments: $E(X^2(t)) < \infty$ for all t, is a Gaussian process provided that $X(t_0)$ is Gaussian for some t_0.

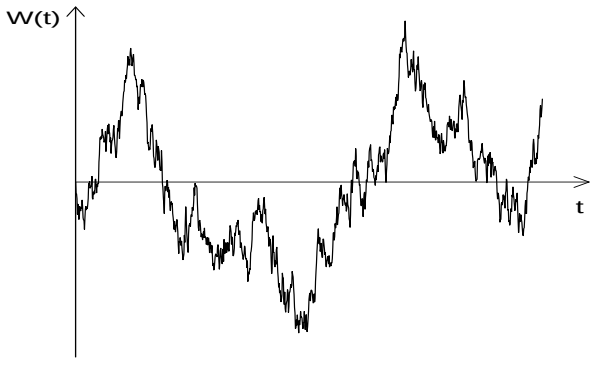

Figure 1: A Wiener sample path

The Wiener process is continuous with mean zero and variance proportional to the elapsed time: $\mathrm{E}(W(t)) = 0$ and $\mathrm{Var}(W(t)) = t$. If $X(t)$ is a *stationary* stochastic process, then $X(t)$ has the same distribution as $X(t+h)$ for all $h > 0$. Thus, the Wiener process cannot be stationary since the variance increases with t. The autocovariance function is given by $\mathrm{Cov}(W_t, W_s) = \min(s,t)$.

The sample paths of a Wiener process behave "wildly" in that they are *nowhere differentiable*. To see what that means define the *total variation* of a real-valued function f on an interval $[a,b] \subset \mathbb{R}$ by the quantity

$$V_a^b(f) \;=\; \sup \sum_{k=1}^{n} |f(t_k) - f(t_{k-1})|$$

where the supremum is taken over all finite partitions $a \le t_0 < \cdots < t_n \le b$ of $[a,b]$. When $V_a^b(f) < \infty$ we say that f is of *bounded variation* on $[a,b]$. Functions that behave sufficiently "nice" are of bounded variation, if for example f is differentiable it is of bounded variation. It turns out that the Wiener process is everywhere of unbounded variation. This happens because the increments $W(t+\Delta t) - W(t)$ is on the order of $\sqrt{\Delta t}$ instead of Δt since the variance is Δt. Heuristically we write

$$\begin{aligned} V_a^b(W) &= \sup \sum_{k=1}^{n} |W(t_k) - W(t_{k-1})| \\ &\geq \lim_{n\to\infty} \sum_{k=1}^{n} \left| W\left(a + \frac{k}{n}(b-a)\right) - W\left(a + \frac{(k-1)}{n}(b-a)\right) \right| \\ &\approx \lim_{n\to\infty} \sum_{k=1}^{n} \sqrt{\frac{1}{n}(b-a)} \;=\; \lim_{n\to\infty} \sqrt{n(b-a)} \;=\; \infty \end{aligned}$$

for *any* interval $[a,b]$. Trying to differentiate we see how this affects the limit

$$\lim_{\Delta t \to 0} \frac{|W(t+\Delta t) - W(t)|}{\Delta t} \;\approx\; \lim_{\Delta t \to 0} \frac{|\sqrt{\Delta t}|}{\Delta t} \;=\; \infty.$$

Now define the *quadratic variation* of a real-valued function f on $[a,b] \subset \mathbb{R}$ by

$$[f]_a^b \;=\; \sup \sum_{k=0}^{n} (f(t_k) - f(t_{k-1}))^2$$

where the supremum is taken as before. For functions of bounded variation the quadratic variation is always 0, and thus, if $[f]_a^b > 0$ then $V_a^b(f) = \infty$. The quadratic variation of a Wiener process over an interval $[s,t]$ equals $t-s$, and in the limit we therefore expect

$$\lim_{\Delta t \to 0} (W(t+\Delta t) - W(t))^2 \;\approx\; \Delta t. \tag{1}$$

Stochastic differential equations Assume the ordinary differential equation

$$\frac{dx}{dt} = a(x,t) \tag{2}$$

describes a one-dimensional dynamical system. Assume that $a(\cdot)$ fulfills conditions such that a unique solution exists, thus $x(t) = x(t; x_0, t_0)$ is a solution satisfying the initial condition $x(t_0) = x_0$. Given the initial condition, we know how the system behaves at all times t, even if we cannot find a solution analytically. We can always solve it numerically up to any desired precision. In many biological systems this is not realistic, and a more realistic model can be obtained if we allow for some randomness in the description.

A natural extension of a deterministic ordinary differential equations model is given by a stochastic differential equations model, where relevant parameters are randomized or modeled as random processes of some suitable form, or simply by adding a noise term to the driving equations of the system. This approach assumes that some degree of noise is present in the dynamics of the process. Here we will use the Wiener process. It leads to a mixed system with both a deterministic and a stochastic part in the following way:

$$dX_t = \mu(X_t, t)\, dt + \sigma(X_t, t)\, dW_t \tag{3}$$

where $X_t = X(t)$ is a stochastic process, not a deterministic function like in (2). This is indicated by the capital letter. Here $W_t = W(t)$ is a Wiener process and since it is nowhere differentiable, we need to define what the differential means. It turns out that it is very useful to write $dW_t = \xi_t dt$, where ξ_t is a white noise process, defined as being normally distributed for any fixed t and uncorrelated: $E(\xi_t \xi_s) = 0$ if $s \neq t$. Strictly speaking, the white noise process ξ_t does not exist as a conventional function of t, but could be interpreted as the generalized derivative of a Wiener process.

We call a process given by an equation of the form (3) for an *Ito process*. The functions $\mu(\cdot)$ and $\sigma(\cdot)$ can be nonlinear, where $\mu(\cdot)$ is the drift part or the deterministic component, and $\sigma(\cdot)$ is the diffusion part or the stochastic component (system noise), that may depend on the state of the system, X_t. If $\mu(\cdot)$ and $\sigma(\cdot)$ do not depend on t the process is called time-homogeneous. Eq. (3) should be interpreted in the following way:

$$X_t = X_{t_0} + \int_{t_0}^{t} \mu(X_s, s)\, ds + \int_{t_0}^{t} \sigma(X_s, s)\, dW_s \tag{4}$$

where X_{t_0} is a random variable independent of the Wiener process. It could simply be a constant. The first integral on the right hand side can be interpreted as an ordinary integral, but what is the second integral? The Wiener process is nowhere differentiable, so how do we give meaning to this differential?

Let us try the usual tricks from ordinary calculus, where we define the integral for a simple class of functions, and then extend by some approximation procedure to a larger class of functions. We want to define

$$\int_{t_0}^{t} f(s)\, dW_s. \tag{5}$$

If $f(t) \equiv \sigma$ is constant we would expect the integral (5) to equal $\sigma(W(t) - W(t_0))$. Note that this is a random variable with expectation 0 since the increments of a Wiener process has expectation 0. Now assume that $f(t)$ is a non-random step function of the form $f(s) = \sigma_j$ on $t_j \leq s < t_{j+1}$ for $j = 1, 2, \ldots, n$ where $t_0 = t_1 < t_2 < \cdots < t_{n+1} = t$. Then we define

$$\int_{t_0}^{t} f(s)\, dW_s \;=\; \sum_{j=1}^{n} \sigma_j \left(W(t_{j+1}) - W(t_j)\right).$$

It is natural to approximate a given function $f(t)$ by a step function. Define a partition Π_n of the interval $[t_0, t]$ by $t_0 = t_1 < t_2 < \cdots < t_{n+1} = t$ where $|\Pi_n| = \max\{|t_{j+1} - t_j| : j = 1, \ldots, n\}$ is the norm of the partition, and approximate

$$f(t) \;\approx\; f(t_j^*) \quad \text{for} \quad t_j \leq t < t_{j+1}$$

where the point t_j^* belongs to the interval $[t_j, t_{j+1}]$. Then we define

$$\int_{t_0}^{t} f(s)\, dW_s \;=\; \lim_{|\Pi_n| \to 0} \sum_{j=1}^{n} f(t_j^*) \left(W(t_{j+1}) - W(t_j)\right).$$

When $f(t)$ is stochastic it turns out that - unlike ordinary integrals - it makes a difference how t_j^* is chosen! To see this consider $f(t) = W_t$ and define two approximations: $t_j^* = t_j$, the left end point, and $t_j^* = t_{j+1}$, the right end point. Taking expectations we see that the two choices yield different results:

$$\begin{aligned} \mathrm{E}\left[\sum_{j=1}^{n} W(t_j)\left(W(t_{j+1}) - W(t_j)\right)\right] &= \sum_{j=1}^{n} \mathrm{E}\left[W(t_j)\left(W(t_{j+1}) - W(t_j)\right)\right] \\ &= \sum_{j=1}^{n} \mathrm{E}\left[W(t_j)\right] \mathrm{E}\left[W(t_{j+1}) - W(t_j)\right] \;=\; 0 \end{aligned}$$

because the Wiener process has independent increments with mean 0. On the other hand,

$$\begin{aligned} \mathrm{E}\left[\sum_{j=1}^{n} W(t_{j+1})\left(W(t_{j+1}) - W(t_j)\right)\right] &= \sum_{j=1}^{n} \mathrm{E}\left[\left(W(t_{j+1}) - W(t_j)\right)^2\right] \\ &= \sum_{j=1}^{n} (t_{j+1} - t_j) \;=\; t - t_0 \end{aligned}$$

where we have subtracted $\mathrm{E}\left[\sum_{j=1}^{n} W(t_j)\left(W(t_{j+1}) - W(t_j)\right)\right] = 0$ and rearranged in the first equality sign, and the second equality sign is the variance of the Wiener process. Two useful and common choices are the following:

- *The Ito integral:* $t_j^* = t_j$, the left end point.
- *The Stratonovich integral:* $t_j^* = (t_j + t_{j+1})/2$, the mid point.

There are arguments for using either one or the other, most of them rather technical and we will not enter in this discussion here. Fortunately, though, the difference between the two is a deterministic quantity and it is possible to calculate one integral from the other. Here we only use the Ito integral.

Some important examples of Ito processes are the following.

Wiener process with drift Imaging a particle suspended in water which is subject to the bombardment of water molecules. The temperature of the water will influence the force of the bombardment, and thus we need a parameter σ to characterize this. Moreover, there is a water current which drives the particle in a certain direction, and we will assume a parameter μ to characterize the drift. To describe this the Wiener process can be generalized to the process

$$dX_t \;=\; \mu\, dt + \sigma\, dW_t$$

which has solution

$$X_t \;=\; \mu\, t + \sigma\, W_t$$

and is thus normally distributed with mean μt and variance $\sigma^2 t$, as follows from the properties of the standard Wiener process.

Geometric Brownian motion Imaging a drug is supplied as a bolus to the blood stream and that the average metabolic process of the drug can be described by an exponential decay through the deterministic equation $x' = -ax$, where x is the concentration of the drug in plasma and a is the decay rate. Assume now that the decay rate fluctuates randomly due to the complex working of the enzymatic machinery involved in the breakdown of the drug. That could be described by letting a vary randomly as $a = \mu + \sigma\xi_t$, where ξ_t is a Gaussian white noise process. Then $\xi_t dt$ can be written as the differential of a Wiener process, dW_t. This leads to the model

$$dX_t \;=\; \mu X_t\, dt + \sigma X_t\, dW_t.$$

It is shown below that the explicit solution is

$$X_t \;=\; X_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right).$$

The process only takes positive values and X_t follows a log-normal distribution with parameters $(\mu - \sigma^2/2)t$ and $\sigma^2 t$.

Ornstein-Uhlenbeck process Imaging a process subject to a restoring force, i.e. the process is attracted to some constant level but is continuously perturbed by noise. An example is given by the membrane potential of a neuron that is constantly being perturbed by electrical impulses from the surrounding network, and at the same time is attracted to an equilibrium value depending on the resting potentials for different ions present at different concentrations inside the cell and in the interstitium. This leads to the model

$$dX_t = -\left(\frac{X_t - \alpha}{\tau}\right) dt + \sigma\, dW_t. \tag{6}$$

Here τ has units time, and is the typical time constant of the system. The autocorrelation is given by $\text{corr}(X_t, X_{t+s}) = e^{-s/\tau}$, and thus the autocorrelation has decreased with a factor of $1/e$ after τ units of time. It has the explicit solution

$$X_t = X_0 e^{-t/\tau} + \alpha(1 - e^{-t/\tau}) + e^{-t/\tau} \int_0^t e^{s/\tau} \sigma dW_s \tag{7}$$

and X_t given X_0 is normally distributed with mean $E(X_t) = X_0 e^{-t/\tau} + \alpha(1 - e^{-t/\tau})$ and variance $V(X_t) = \sigma^2\tau(1 - e^{-2t/\tau})/2$. If X_0 is normally distributed with mean α and variance $\sigma^2\tau/2$, then so is X_t for all t, i.e. X_t is stationary.

When the diffusion term does not depend on the state variable X_t as in the Wiener process with drift and the Ornstein-Uhlenbeck process, we say that it has *additive noise*. In this case the Ito and the Stratonovich integrals yield the same process, so it does not matter which calculus we choose. In the case of Geometric Brownian motion we say that it has *multiplicative noise*.

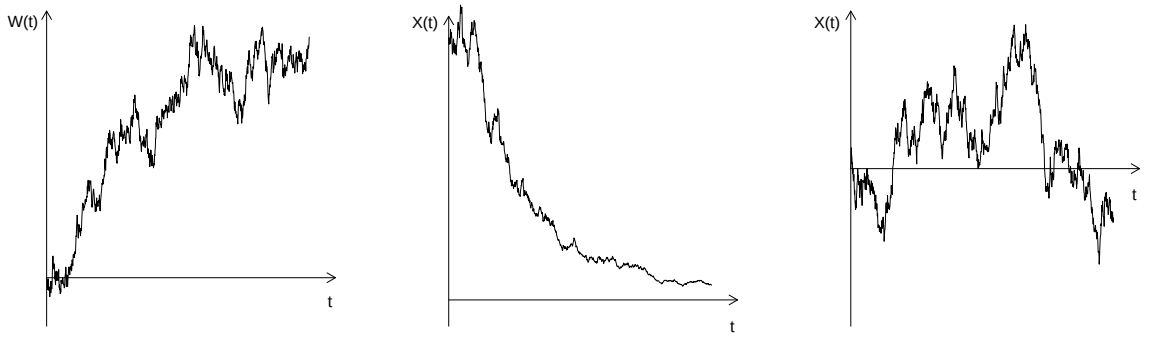

Figure 2: Sample paths from a Wiener process with drift (left), a Geometric Brownian motion (middle) and an Ornstein-Uhlenbeck process (right). Note how the amplitude of the noise does not change over time for the Wiener and the Ornstein-Uhlenbeck process, whereas for Geometric Brownian motion the amplitude of the noise is proportional to the state variable.

Existence and uniqueness To ensure the existence of a solution to (3) for $0 \leq t \leq T$ where T is fixed, the following is sufficient:

$$|\mu(t,x)| + |\sigma(t,x)| \quad \leq \quad C(1+|x|)$$

for some constant C. This ensures that X_t does not explode, i.e that $|X_t|$ does not tend to ∞ in finite time. To ensure uniqueness of a solution the Lipschitz condition is sufficient:

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \quad \leq \quad D|x-y|$$

for some constant D. Note that only *sufficient* conditions are stated, and in many biological applications these are too strict, and weaker conditions can be found. We will not treat these here, though. Note also that these conditions are fulfilled for the three processes described above.

Itos formula Stochastic differentials do not obey the ordinary chain rule as we know it from classical calculus. An additional term appears because $(dW_t)^2$ behaves like dt, see (1). We have

Theorem 2 (Itos formula) *Let X_t be an Ito process given by*

$$dX_t \quad = \quad \mu(t,X_t)dt + \sigma(t,X_t)\,dW_t$$

and let $f(t,x)$ be a twice continuously differentiable function in x and once continuously differentiable function in t. Then

$$dY_t \quad = \quad f(t,X_t)$$

is also an Ito process, and

$$dY_t \quad = \quad \frac{\partial f}{\partial t}(t,X_t)dt + \frac{\partial f}{\partial x}(t,X_t)dX_t + \frac{1}{2}\sigma^2(t,X_t)\frac{\partial^2 f}{\partial x^2}(t,X_t)dt.$$

Note that the first two terms on the right hand side correspond to the chain rule we know from classical calculus, but an extra term appears in stochastic calculus because the Wiener process is of unbounded variation, and thus the quadratic variation comes into play.

Example Let us calculate the integral $\int_0^t W_s dW_s$. From classical calculus we expect a term like $\frac{1}{2}W_t^2$ in the solution. Thus, we choose $f(t,x) = \frac{1}{2}x^2$ and $X_t = W_t$ and apply Itos formula to

$$Y_t \quad = \quad f(t,W_t) \;=\; \frac{1}{2}W_t^2.$$

We obtain

$$dY_t \;=\; \frac{\partial f}{\partial t}(t,W_t)dt + \frac{\partial f}{\partial x}(t,W_t)dW_t + \frac{1}{2}\sigma^2(t,W_t)\frac{\partial^2 f}{\partial x^2}(t,W_t)dt \;=\; 0 + W_t dW_t + \frac{1}{2}dt$$

because $\sigma^2(t,W_t) = 1$. Hence

$$Y_t \;=\; \frac{1}{2}W_t^2 \;=\; \int_0^t W_s dW_s + \frac{1}{2}\int_0^t ds \;=\; \int_0^t W_s dW_s + \frac{1}{2}t$$

and finally

$$\int_0^t W_s dW_s \;=\; \frac{1}{2}W_t^2 - \frac{1}{2}t.$$

Example Let us find the solution X_t to the Geometric Brownian motion

$$dX_t \;=\; \mu X_t\, dt + \sigma X_t\, dW_t.$$

Rewrite the equation as

$$\frac{dX_t}{X_t} \;=\; \mu\, dt + \sigma\, dW_t.$$

Thus, we have

$$\int_0^t \frac{dX_s}{X_s} \;=\; \mu\, t + \sigma\, W_t \tag{8}$$

which suggests to apply Itos formula on $f(t,x) = \log x$. We obtain

$$\begin{aligned} dY_t \;=\; d(\log X_t) \;&=\; \frac{\partial f}{\partial t}(t,X_t)dt + \frac{\partial f}{\partial x}(t,X_t)dX_t + \frac{1}{2}\sigma^2(t,X_t)\frac{\partial^2 f}{\partial x^2}(t,X_t)dt \\ &=\; 0 + \frac{1}{X_t}dX_t + \frac{1}{2}\sigma^2 X_t^2\left(-\frac{1}{X_t^2}\right)dt \;=\; \frac{dX_t}{X_t} - \frac{1}{2}\sigma^2 dt \end{aligned}$$

and thus

$$\frac{dX_t}{X_t} \;=\; d(\log X_t) + \frac{1}{2}\sigma^2 dt. \tag{9}$$

Integrating (9) and using (8) we finally obtain

$$\log\frac{X_t}{X_0} \;=\; \int_0^t \frac{dX_s}{X_s} - \frac{1}{2}\sigma^2 t \;=\; \mu\, t + \sigma\, W_t - \frac{1}{2}\sigma^2 t$$

and so

$$X_t \;=\; X_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma\, W_t\right\}.$$

Note that it is simply the exponential of a Wiener process with drift.

The solution (7) of the Ornstein-Uhlenbeck process can be found by multiplying both sides of (6) with $e^{-t/\tau}$ and then apply Itos formula to $e^{-t/\tau}X_t$. We will not do that here.

Monte Carlo simulations When no explicit solution is available we can approximate different characteristics of the process by simulation, such as sample paths, moments, qualitative behavior etc. Usually such simulation methods are based on discrete approximations of the continuous solution to a stochastic differential equation. Different schemes are available depending on how good we want the approximation to be, which comes at a price of computer time. Assume we want to approximate a solution to (3) in the time interval $[0, T]$. Consider the time discretization

$$0 = t_0 < t_1 < \cdots < t_j < \cdots < t_N = T$$

and denote the time steps by $\Delta_j = t_{j+1} - t_j$ and the increments of the Wiener process by $\Delta W_j = W_{t_{j+1}} - W_{t_j}$. Then $\Delta W_j \sim N(0, \Delta_j)$, which we can use to construct approximations by drawing normally distributed numbers from a random number generator. For simplicity assume that the process is time-homogenous.

The Euler-Maruyama scheme The simplest scheme is the stochastic analogue of the deterministic Euler scheme. Approximate the process X_t at the discrete time-points $t_j, 1 \leq j \leq N$ by the recursion

$$Y_{t_{j+1}} \quad = \quad Y_{t_j} + \mu(Y_{t_j})\Delta_j + \sigma(Y_{t_j})\Delta W_j \;\; ; \;\; Y_{t_0} = x_0$$

where $\Delta W_j = \sqrt{\Delta_j} \cdot Z_j$, with Z_j being standard normal variables with mean 0 and variance 1 for all j. This approximating procedure assumes that the drift and diffusion functions are constant between time steps, so obviously the approximation improves for smaller time steps. To evaluate the convergence things are more complicated for stochastic processes, and we operate with two criteria of optimality: the *strong* and the *weak* orders of convergence.

Consider the expectation of the absolute error at the final time instant T of the Euler-Maruyama scheme. It can be shown that there exist constants $K > 0$ and $\delta_0 > 0$ such that

$$E(|X_T - Y_{t_N}|) \quad \leq \quad K\delta^{0.5}$$

for any time discretization with maximum step size $\delta \in (0, \delta_0)$. We say that the approximating process Y *converges in the strong sense* with order 0.5. This is similar to how approximations are evaluated in deterministic systems, only here we take expectations, since X_T and Y_{t_N} are random variables. Compare with the Euler scheme for an ordinary differential equation which has order of convergence 1. Sometimes we do not need a close *pathwise* approximation, but only some function of the value at a given final time T (e.g. $E(X_T)$, $E(X_T^2)$ or generally $E(g(X_T))$). In this case we have that there exist constants $K > 0$ and $\delta_0 > 0$ such that for any polynomial g

$$|E\left(g(X_T) - E(g(Y_{t_N}))\right)| \quad \leq \quad K\delta$$

for any time discretization with maximum step size $\delta \in (0, \delta_0)$. We say that the approximating process Y *converges in the weak sense* with order 1.

The Milstein scheme To improve the accuracy of the approximation we add a second-order term that appears from Itos formula. Approximate X_t by

$$Y_{t_{j+1}} = \underbrace{\underbrace{Y_{t_j} + \mu(Y_{t_j})\Delta_j + \sigma(Y_{t_j})\Delta W_j}_{\text{Euler-Maruyama}} + \frac{1}{2}\sigma(Y_{t_j})\sigma'(Y_{t_j})\{(\Delta W_j)^2 - \Delta_j\}}_{\text{Milstein}}$$

where the prime $'$ denotes derivative. It is not obvious exactly how this term appears, but can be derived through *stochastic Taylor expansions*. The Milstein scheme converges in the strong sense with order 1, and could thus be regarded as the proper generalization of the deterministic Euler-scheme.

If $\sigma(X_t)$ does not depend on X_t the Euler-Maruyama and the Milstein scheme coincide.

Further reading

W. Horsthemke and R. Lefever. *Noise-induced Transitions: Theory and Applications in Physics, Chemistry, and Biology.* 2nd ed. Springer Series in Synergetics, Springer, 2006.

S.M. Iacus. *Simulation and Inference for Stochastic Differential Equations. With R Examples.* Springer Verlag, 2008.

P.E. Kloeden and E. Platen. *Numerical Solution of SDE through Computer Experiments*. Springer Verlag, 1997.

A. Longtin. *Effects of Noise on Nonlinear Dynamics*. Chapter in Nonlinear Dynamics in Physiology and Medicine, Eds: A. Beuter, L. Glass, M.C. Mackey and M.S. Titcombe. Springer Verlag, 2003.

B. Øksendal. *Stochastic Differential Equations. An introduction with applications*. Springer Verlag, 2007.