Manifolds and Geometric Integration Colloquia (MaGIC) March 1–4, 2011, Finse, Norway[1].

Numerical Methods for Stochastic Differential Equations

Stig Larsson
Department of Mathematical Sciences
Chalmers University of Technology
and University of Gothenburg
http://www.math.chalmers.se/~stig

1. Lecture 1

The main textbook is Klebaner [3]. For the computer exercises we refer to Higham [2]. It is very important to be able to do computer simulations with stochastic ODEs, so I do recommend that you do some of these computer exercises. In these lectures I will present the theoretical background.

1.1. **Introduction.** We study equations of the form

$$\begin{aligned} &\mathrm{d}X(t) = \mu(X(t),t)\,\mathrm{d}t + \sigma(X(t),t)\,\mathrm{d}B(t), \quad 0 \le t \le T, \\ &X(0) = X_0. \end{aligned}$$

This means that X is a stochastic process which solves the integral equation

$$X(t) = X_0 + \int_0^t \mu(X(s),s)\,\mathrm{d}s + \int_0^t \sigma(X(s),s)\,\mathrm{d}B(s), \quad t \in [0,T].$$

Here B is Brownian motion and the last term is the Itô integral. We next explain what this means.

1.2. **A very brief introduction to stochastic analysis.** Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. Here Ω is the sample space, $\mathcal{F}$ a σ-algebra on Ω and $\mathbf{P}$ a probability measure on Ω. A random variable is a measurable function $f \colon \Omega \to \mathbf{R}$. If $f \in L_1(\Omega)$, that is,

$$\int_\Omega |f(\omega)|\,\mathrm{d}\mathbf{P}(\omega) < \infty,$$

then we can compute the expected value

$$\mathbf{E}\big(f\big) = \int_\Omega f(\omega)\,\mathrm{d}\mathbf{P}(\omega) = \int_\Omega f\,\mathrm{d}\mathbf{P}.$$

A stochastic process is a mapping $X \colon [0,T] \times \Omega \to \mathbf{R}$ such that $X(t,\cdot)$ is a random variable for each t. It usual in stochastic mathematics to suppress the variable ω. Thus, we write $X = \{X(t)\}_{t\in[0,T]}$ or $X = \{X(t)\}_{t\in[0,\infty)}$.

A filtration $\{\mathcal{F}_t\}_{[0,T]}$ in $(\Omega, \mathcal{F}, \mathbf{P})$ is an increasing family of σ-algebras such that

$$\mathcal{F}_0 \subset \mathcal{F}_t \subset \cdots \mathcal{F}_T = \mathcal{F}.$$

Then $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{[0,T]}, \mathbf{P})$ is called a filtered probability space. A stochastic process $X = \{X(t)\}_{t\in[0,T]}$ on this space is adapted if $X(t)$ is $\mathcal{F}_t$-measurable for each t.

A stochastic process $X = \{X(t)\}_{t\in[0,T]}$ on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{[0,T]}, \mathbf{P})$ is called a martingale if it is adapted, if $\mathbf{E}\big(X(t)\big) < \infty$ for each t, and if

$$\mathbf{E}\big(X(t) \mid \mathcal{F}_s\big) = X(s), \quad s < t.$$

[1]Revised: February 27, 2011

In particular, by the law of double expectation,

$$\mathbf{E}\big(X(s)\big) = \mathbf{E}\big(\mathbf{E}\big(X(t) \mid \mathcal{F}_s\big)\big) = \mathbf{E}\big(X(t)\big), \quad s < t,$$

since $\mathcal{F}_s \subset \mathcal{F}_t$. Hence, a martingale has constant mean.

Brownian motion is a stochastic process $\{B(t)\}_{t\in[0,T]}$ on $(\Omega, \mathcal{F}, \mathbf{P})$ with the following properties:

- (continuous paths) $B(t)$ is a continuous function of t;
- (independent increments) $B(t) - B(s)$ is independent of $B(u)$ for $t > s \geq u$;
- (Gaussian increments) $B(t) - B(s)$ is normally distributed with mean 0 and variance $t - s$, $B(t) - B(s) \sim N(0, t - s)$, for $t > s$.

It is non-trivial to show that Brownian motion exists. It can be shown that it generates a filtration such that it becomes a martingale. We always work in this filtered probability space. In particular, if $B(0) = x$, then $\mathbf{E}\big(B(t)\big) = \mathbf{E}\big(x\big)$. We often use $B(0) = 0$. Brownian motion is not differentiable with respect to t at any point.

Let $X = \{X(t)\}_{t\in[0,T]}$ be a stochastic process which is adapted and square integrable

$$\int_0^T \mathbf{E}\big(X(t)^2\big)\,\mathrm{d}t < \infty.$$

Then the Itô integral can be defined:

$$\int_0^t X(s)\,\mathrm{d}B(s).$$

In a very simplified way we can say that it is constructed as a limit of

$$\sum_k X(t_k)(B(t_{k+1}) - B(t_k)) = \sum_k X(t_k)\Delta B_k$$

as the partition is refined. The limit is taken in the mean square sense, or $L_2(\Omega)$ sense:

$$\Big\| \sum_k X(t_k)\Delta B_k - \int_0^t X(s)\,\mathrm{d}B(s)\Big\|_{L_2}^2 = \mathbf{E}\Big(\Big| \sum_k X(t_k)\Delta B_k - \int_0^t X(s)\,\mathrm{d}B(s)\Big|^2\Big) \to 0.$$

Note that it is similar to the left endpoint rectangle rule. Using another interpolation point results in another stochastic integral. For example, with the midpoint rule we obtain the Stratonovich integral, which different from, but related to, the Itô integral.

The Itô integral is a stochastic process with the following properties:

- it is linear with respect to X;
- (zero mean) $\mathbf{E}\big(\int_0^t X(s)\,\mathrm{d}s\big) = 0$;
- (isometry)
$$\mathbf{E}\Big(\Big|\int_0^t X(s)\,\mathrm{d}B(s)\Big|^2\Big) = \int_0^t \mathbf{E}\big(|X(s)|^2\big)\,\mathrm{d}s;$$
- it is a martingale.

1.3. **Explicit solutions.** We will derive explicit formulas for solutions of some special cases of the stochastic differential equation. We need Itô's formula, which is a chain rule for Itô processes.

Assume that the process X satifies

$$\begin{aligned} \mathrm{d}X(t) &= \mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t), \quad 0 \leq t \leq T, \\ X(0) &= X_0, \end{aligned}$$

that is

$$X(t) = X_0 + \int_0^t \mu(s)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}B(s), \quad t \in [0,T].$$

Here μ, σ are processes such that the integrals exist. Such a process X is called an Itô process. If g is a smooth function then Itô's formula is

$$\begin{aligned}
\mathrm{d}g(X) &= g'(X)\,\mathrm{d}X + \tfrac{1}{2}g''(X)\,\mathrm{d}[X,X] \\
&= g'(X)(\mu\,\mathrm{d}t + \sigma\,\mathrm{d}B) + \tfrac{1}{2}g''(X)\,\sigma^2\,\mathrm{d}t \\
&= (g'(X)\mu + \tfrac{1}{2}g''(X)\,\sigma^2)\,\mathrm{d}t + g'(X)\sigma\,\mathrm{d}B.
\end{aligned}$$

Here we used the fact that the quadratic variation of X satisfies (formally)

$$\begin{aligned}
\mathrm{d}[X,X] &= \mathrm{d}[\mu(t)t, \mu(t)t] + 2\,\mathrm{d}[\mu(t)t, \sigma(t)B(t)] + \sigma^2\,\mathrm{d}[B,B] \\
&= \sigma^2\,\mathrm{d}t.
\end{aligned}$$

In integrated form (rigorously):

$$g(X(t)) = g(X(t_0)) + \int_{t_0}^t \Big(g'(X(s))\mu(s) + \tfrac{1}{2}g''(X(s))\,\sigma(s)^2\Big)\,\mathrm{d}s + \int_{t_0}^t g'(X(s))\sigma(s)\,\mathrm{d}B(s).$$

1.3.1. *Example 1.* Compute $\int_0^t B(s)\,\mathrm{d}B(s)$. We use $X = B$, that is, $\mu = 0$, $\sigma = 1$, and $g(x) = \frac{1}{2}x^2$, $g'(x) = x$, $g''(x) = 1$. Then

$$\Big[\tfrac{1}{2}B(s)^2\Big]_0^t = \int_0^t \tfrac{1}{2}\mathrm{d}s + \int_0^t B(s)\,\mathrm{d}B(s),$$

so that, with $B(0) = 0$,

$$\int_0^t B(s)\,\mathrm{d}B(s) = \tfrac{1}{2}B(t)^2 - \tfrac{1}{2}t.$$

This also means that $Y(t) = \int_0^t B(s)\,\mathrm{d}B(s) = \frac{1}{2}B(t)^2 - \frac{1}{2}t$ is a solution of

$$\mathrm{d}Y = B\,\mathrm{d}B,\ t > 0; \quad Y(0) = 0.$$

1.3.2. *Example 2.* The stochastic exponential. Assume that X satisfies

$$\mathrm{d}X = \beta X\,\mathrm{d}B,\ t > 0; \quad X(0) = X_0, \quad (\beta = \text{const}).$$

Here $\mu = 0$, $\sigma(t) = \beta X(t)$ and we take $g(x) = \log(x)$, $g'(x) = 1/x$, $g''(x) = -1/x^2$. We get

$$\Big[\log(X(s))\Big]_0^t = -\int_0^t \tfrac{1}{2}X(s)^{-2}(\beta X(s))^2\,\mathrm{d}s + \int_0^t X(s)^{-1}\beta X(s)\,\mathrm{d}B(s) = -\tfrac{1}{2}\beta^2 t + B(t),$$

so that

$$X(t) = X_0 \exp(\beta B(t) - \tfrac{1}{2}\beta^2 t).$$

1.3.3. *Example 3.* The Black-Scholes process. Assume that X satisfies

$$\mathrm{d}X = \alpha X\,Dt + \beta X\,\mathrm{d}B,\ t > 0; \quad X(0) = X_0, \quad (\alpha, \beta = \text{const}).$$

A similar calculation leads to

$$X(t) = X_0 \exp(\beta B(t) + (\alpha - \tfrac{1}{2}\beta^2)t).$$

1.3.4. *Example 4.* The Ornstein-Uhlenbeck process. Assume that X satisfies

$$\mathrm{d}X = \alpha X\,\mathrm{d}t + \beta\,\mathrm{d}B,\ t > 0; \quad X(0) = X_0, \quad (\alpha, \beta = \text{const}).$$

We want to multiply by the integrating factor $\mathrm{e}^{-\alpha t}$. We use Itô's formula for t-dependent function:

$$\mathrm{d}g(t, X(t)) = (g_t'(t, X) + g_x'(X)\mu + \tfrac{1}{2}g_{xx}''(X)\,\sigma^2)\,\mathrm{d}t + g_x'(X)\sigma\,\mathrm{d}B$$

and $g(t, x) = \mathrm{e}^{-\alpha t}x$ to get

$$\mathrm{d}(\mathrm{e}^{-\alpha t}X(t)) = \beta\,\mathrm{e}^{-\alpha t}\,\mathrm{d}B(t),$$

so that

$$\Big[\mathrm{e}^{-\alpha s}X(s)\Big]_0^t = \beta\int_0^t \mathrm{e}^{-\alpha s}\,\mathrm{d}B(s),$$

and

$$X(t) = \mathrm{e}^{\alpha t}X_0 + \beta\int_0^t \mathrm{e}^{\alpha(t-s)}\,\mathrm{d}B(s).$$

1.4. **Preparations.** Let X be a strong solution of the scalar Itô SDE:

$$\begin{aligned} &\mathrm{d}X(t) = \mu(X(t), t)\,\mathrm{d}t + \sigma(X(t), t)\,\mathrm{d}B(t), \quad 0 \le t \le T, \\ &X(0) = X_0. \end{aligned} \tag{1}$$

This means, see Klebaner Definition 5.1, that X solves the integral equation

$$X(t) = X_0 + \int_0^t \mu(X(s), s)\,\mathrm{d}s + \int_0^t \sigma(X(s), s)\,\mathrm{d}B(s), \quad t \in [0, T]. \tag{2}$$

Here B is Brownian motion and the last term is the Itô integral.

We assume that X_0 is a random variable and the coefficients are deterministic functions,

$$\begin{aligned} \mu &: \mathbf{R} \times [0, T] \to \mathbf{R}, \\ \sigma &: \mathbf{R} \times [0, T] \to \mathbf{R}, \end{aligned} \tag{3}$$

that are continuous and satisfy a global Lipschitz condition with respect to x,

$$\begin{aligned} |\mu(x, t) - \mu(y, t)| &\le L|x - y|, \\ |\sigma(x, t) - \sigma(y, t)| &\le L|x - y|, \end{aligned} \quad \forall x, y \in \mathbf{R},\ t \in [0, T], \tag{4}$$

and a linear growth bound,

$$\begin{aligned} |\mu(x, t)| &\le L\big(1 + |x|\big), \\ |\sigma(x, t)| &\le L\big(1 + |x|\big), \end{aligned} \quad \forall x \in \mathbf{R},\ t \in [0, T]. \tag{5}$$

Remark 1. The Lipschitz condition (4) is called *global* because it holds for all $x, y \in \mathbf{R}$ with the same constant L. Klebaner Theorem 5.4 assumes only a *local* Lipschitz condition, where the Lipschitz constant may depend on the size of x, y. We use a global condition in order to make the presentation simpler. Later on we will also assume a Lipschitz condition with respect to t, see (21). (There is a mistake in Theorem 5.4: the constants in (5.37) and (5.38) cannot be the same because the first one depends on N, $K = K_N$, while the second one is a global constant, independent of N.)

Remark 2. Note that (5) follows from (4) because

$$|\mu(x, t)| \le |\mu(0, t)| + |\mu(x, t) - \mu(0, t)| \le \max_{t\in[0,T]} |\mu(0, t)| + L|x - 0| \le \tilde{L}\big(1 + |x|\big) \tag{6}$$

with a new constant $\tilde{L}$. This is because (4) is global.

Remark 3. Systems of SDE of the form

$$\mathrm{d}X_i = \mu_i(X_1,\dots,X_n,t)\,\mathrm{d}t + \sum_{j=1}^{m} \sigma_{ij}(X_1,\dots,X_n,t)\,\mathrm{d}B_j(t), \quad i=1,\dots,n, \tag{7}$$

can be written in the form (1), and analyzed in the same way, with matrix notation

$$X=(X_1,\dots,X_n)^T \in \mathbf{R}^n, \quad \mu:\mathbf{R}^n\times[0,T]\to\mathbf{R}^n, \quad \sigma:\mathbf{R}^n\times[0,T]\to\mathbf{R}^{n\times m},$$

and $B=(B_1,\dots,B_m)^T$ an m-dimensional Brownian motion.

We quote Theorem 5.4 from Klebaner.

Theorem 1 (Existence and uniqueness)**.** *If X_0 is independent of $(B(t),\, 0\le t\le T)$ and $\mathbf{E}\big(|X_0|^2\big)<\infty$, then* (1) *has a unique strong solution X and*

$$\mathbf{E}\Big(\sup_{0\le t\le T}|X(t)|^2\Big) \le C\big(1+\mathbf{E}\big(|X_0|^2\big)\big), \tag{8}$$

where $C=C(L,T)$.

As a warm-up for the error analysis of numerical metods, we begin by proving a stability result. We will use this proof technique several times later. The bound (8) is proved in a similar way.

We need Gronwall's lemma.

Lemma 1 (Gronwall)**.** *Let A,B be constants with $B\ge 0$. If*

$$\phi(t)\le A + B\int_0^t \phi(s)\,\mathrm{d}s, \quad t\in[0,T],$$

then

$$\phi(t)\le A\mathrm{e}^{Bt}, \quad t\in[0,T].$$

Proof. $u(t):=A+B\int_0^t\phi(s)\,\mathrm{d}s,\ \phi(t)\le u(t),\ u'=B\phi\le Bu,\ u(t)\le u(0)\mathrm{e}^{Bt}=A\mathrm{e}^{Bt}$. □

We also need Doob's L_p martingale inequality, see Klebaner (7.37) p. 201.

Theorem 2 (Doob's inequality)**.** *If Y is a martingale, then*

$$\mathbf{E}\Big(\max_{0\le t\le T}|Y(t)|^p\Big) \le \Big(\frac{p}{p-1}\Big)^p \mathbf{E}\big(|Y(T)|^p\big), \quad 1<p<\infty. \tag{9}$$

With the L_p-norm, $\|X\|_{L_p} = \big(\mathbf{E}\big(|X|^p\big)\big)^{1/p} = \big(\int_\Omega |X(\omega)|^p\,\mathrm{d}\mathbf{P}(\omega)\big)^{1/p}$, this can also be written

$$\Big\|\max_{0\le t\le T}|Y(t)|\Big\|_{L_p} \le \frac{p}{p-1}\|Y(T)\|_{L_p}.$$

This gives a bound for the norm of the whole path in terms of the norm of its final value.

Finally we recall the isometry of the Itô integral, Klebaner (4.12),

$$\mathbf{E}\Big(\Big|\int_0^t f(s)\,\mathrm{d}B(s)\Big|^2\Big) = \int_0^t \mathbf{E}\big(|f(s)|^2\big)\,\mathrm{d}s. \tag{10}$$

We are now ready to prove stability with respect to perturbation of the initial value.

Theorem 3 (Stability)**.** *Let $\hat{X}$ be another solution of* (2) *with the same μ,σ,B but a different initial value $\hat{X}_0$. Then*

$$\mathbf{E}\Big(\max_{0\le t\le T}|\hat{X}(t)-X(t)|^2\Big) \le C\mathbf{E}\big(|\hat{X}_0-X_0|^2\big), \tag{11}$$

where $C=C(L,T)$.

Note that this implies uniqueness: two solutions with the same data μ, σ, B, X_0 must be equal.

Proof. We will use Gronwall's lemma with $\phi(t) = \mathbf{E}\big(\max_{0\le s\le t} |\hat{X}(s) - X(s)|^2\big)$. By (2), writing $\hat{\mu} = \mu(\hat{X}(z), z)$, $\mu = \mu(X(z), z)$, etc, and using the inequality $(a+b+c)^2 \le 3(a^2+b^2+c^2)$, we get

$$\begin{aligned}
\phi(t) &= \mathbf{E}\Big(\max_{0\le s\le t} |\hat{X}(s) - X(s)|^2\Big) \\
&= \mathbf{E}\Big(\max_{0\le s\le t} \Big|\hat{X}_0 - X_0 + \int_0^s \big(\mu(\hat{X}(z), z) - \mu(X(z), z)\big)\,\mathrm{d}z \\
&\quad + \int_0^s \big(\sigma(\hat{X}(z), z) - \sigma(X(z), z)\big)\,\mathrm{d}B(z)\Big|^2\Big) \\
&\le 3\Big\{\mathbf{E}\Big(|\hat{X}_0 - X_0|^2\Big) + \mathbf{E}\Big(\max_{0\le s\le t}\Big|\int_0^s (\hat{\mu} - \mu)\,\mathrm{d}z\Big|^2\Big) \\
&\quad + \mathbf{E}\Big(\max_{0\le s\le t}\Big|\int_0^s (\hat{\sigma} - \sigma)\,\mathrm{d}B\Big|^2\Big)\Big\}.
\end{aligned}$$

Here, by the Cauchy-Schwarz inequality and (4),

$$\begin{aligned}
\mathbf{E}\Big(\max_{0\le s\le t}\Big|\int_0^s (\hat{\mu} - \mu)\,\mathrm{d}z\Big|^2\Big) &\le \mathbf{E}\Big(\max_{0\le s\le t}\Big\{\int_0^s 1^2\,\mathrm{d}z \int_0^s |\hat{\mu} - \mu|^2\,\mathrm{d}z\Big\}\Big) \\
&= \mathbf{E}\Big(t\int_0^t |\hat{\mu} - \mu|^2\,\mathrm{d}z\Big) = t\int_0^t \mathbf{E}\big(|\hat{\mu} - \mu|^2\big)\,\mathrm{d}s \\
&\le L^2 T \int_0^t \mathbf{E}\big(|\hat{X}(s) - X(s)|^2\big)\,\mathrm{d}s \\
&\le L^2 T \int_0^t \mathbf{E}\big(\max_{0\le z\le s} |\hat{X}(z) - X(z)|^2\big)\,\mathrm{d}s.
\end{aligned}$$

For the other term we use Doob's inequality (9) with $p = 2$ and

$$Y(t) = \int_0^t \big(\sigma(\hat{X}(z), z) - \sigma(X(z), z)\big)\,\mathrm{d}B(z),$$

which is a martingale by Theorem 4.7 in Klebaner. Using also (10) and (4) we get

$$\begin{aligned}
\mathbf{E}\Big(\max_{0\le s\le t}\Big|\int_0^s (\hat{\sigma} - \sigma)\,\mathrm{d}B\Big|^2\Big) &\le 4\mathbf{E}\Big(\Big|\int_0^t (\hat{\sigma} - \sigma)\,\mathrm{d}B\Big|^2\Big) \\
&= 4\int_0^t \mathbf{E}\big(|\hat{\sigma} - \sigma|^2\big)\,\mathrm{d}s \\
&\le 4L^2 \int_0^t \mathbf{E}\big(|\hat{X}(s) - X(s)|^2\big)\,\mathrm{d}s \\
&\le 4L^2 \int_0^t \mathbf{E}\big(\max_{0\le z\le s} |\hat{X}(z) - X(z)|^2\big)\,\mathrm{d}s.
\end{aligned}$$

We now have

$$\mathbf{E}\big(\max_{0\le s\le t} |\hat{X}(s) - X(s)|^2\big) \le 3\mathbf{E}\big(|\hat{X}_0 - X_0|^2\big) + C\int_0^t \mathbf{E}\big(\max_{0\le z\le s} |\hat{X}(z) - X(z)|^2\big)\,\mathrm{d}s$$

with $C = 3L^2(T+4)$, or

$$\phi(t) \le 3\mathbf{E}\big(|\hat{X}_0 - X_0|^2\big) + C\int_0^t \phi(s)\,\mathrm{d}s, \quad t \in [0, T],$$

where $\phi(t) = \mathbf{E}\big(\max_{0\le s\le t} |\hat{X}(s) - X(s)|^2\big)$. Gronwall's lemma completes the proof. $\square$

We now prove that X is Hölder continuous in a mean square sense.

Theorem 4. *Under the assumptions of Theorem 1 we have*

$$\mathbf{E}\big(|X(t) - X(s)|^2\big) \le C\big(1 + \mathbf{E}\big(|X_0|^2\big)\big)|t-s|, \quad \forall t, s \in [0,T], \tag{12}$$

where $C = C(L,T)$.

Proof. We may assume that $s \le t$. Working as in the previous proof, but without using Doob's inequality, we then get

$$\begin{aligned}
\mathbf{E}\big(|X(t) - X(s)|^2\big) &= \mathbf{E}\Big(\Big| \int_s^t \mu(X(z),z)\,\mathrm{d}z + \int_s^t \sigma(X(z),z)\,\mathrm{d}B(z)\Big|^2\Big) \\
&\le 2\underbrace{(t-s)}_{\le T} \int_s^t \mathbf{E}\big(|\mu|^2\big)\,\mathrm{d}z + 2\int_s^t \mathbf{E}\big(|\sigma|^2\big)\,\mathrm{d}z \quad \Big\{\text{by (5)}\Big\} \\
&\le 2L^2(T+1)\int_s^t \mathbf{E}\big((1+|X(z)|)^2\big)\,\mathrm{d}z \\
&\le 4L^2(T+1)\int_s^t \big\{1 + \mathbf{E}\big(|X(z)|^2\big)\big\}\,\mathrm{d}z \quad \Big\{\text{by (8)}\Big\} \\
&\le C(L,T)\int_s^t \big\{1 + C(L,T)\big(1 + \mathbf{E}\big(|X_0|^2\big)\big)\big\}\,\mathrm{d}z \\
&\le C(L,T)(t-s)\big(1 + \mathbf{E}\big(|X_0|^2\big)\big).
\end{aligned}$$

$\square$

Remark 4. This means, with $\|X_0\|_{L_2} = \mathbf{E}\big(|X_0|^2\big)^{1/2} \le M$ and $C = C(M,L,T)$,

$$\|X(t) - X(s)\|_{L_2} = \Big(\mathbf{E}\big(|X(t) - X(s)|^2\big)\Big)^{1/2} \le C|t-s|^{1/2}. \tag{13}$$

So X satisfies a Hölder condition with exponent $\frac{1}{2}$. By inspection of the previous proof, we see that in the deterministic case, $\sigma = 0$, we have a stronger result, namely that X is Lipschitz continuous,

$$\|X(t) - X(s)\|_{L_2} = \Big(\mathbf{E}\big(|X(t) - X(s)|^2\big)\Big)^{1/2} \le C|t-s|. \tag{14}$$

1.5. **Strong convergence of Euler's method.** We introduce a mesh

$$0 = t_0 < t_1 < \dots < t_n < t_{n+1} < \dots < t_N = T, \quad h_n = t_{n+1} - t_n, \quad h = \max h_n \le 1, \tag{15}$$

and note that X satisfies

$$X(t_{n+1}) = X(t_n) + \int_{t_n}^{t_{n+1}} \mu(X(s),s)\,\mathrm{d}s + \int_{t_n}^{t_{n+1}} \sigma(X(s),s)\,\mathrm{d}B(s). \tag{16}$$

This motivates the Euler (or Euler-Maruyama) method, which defines $\{Y_n\}_{n=0}^N$ by

$$\begin{aligned}
&Y_0 \approx X_0, \\
&Y_{n+1} = Y_n + \mu(Y_n,t_n)\int_{t_n}^{t_{n+1}} \mathrm{d}s + \sigma(Y_n,t_n)\int_{t_n}^{t_{n+1}} \mathrm{d}B(s).
\end{aligned} \tag{17}$$

Since $\int_{t_n}^{t_{n+1}} \mathrm{d}s = \Delta t_n = h_n$, $\int_{t_n}^{t_{n+1}} \mathrm{d}B(s) = B(t_{n+1}) - B(t_n) = \Delta B_n$, this can be written

$$\begin{aligned}
&Y_0 \approx X_0, \\
&Y_{n+1} = Y_n + \mu(Y_n,t_n)h_n + \sigma(Y_n,t_n)\Delta B_n.
\end{aligned} \tag{18}$$

For the practical aspects concerning the computation in a Matlab environment we refer to [2]. Here we only note that $\Delta B_n = \sqrt{h_n}\xi_n$, where the $\xi_n \sim N(0,1)$ are independent Gaussian random variables with mean 0 and variance 1, which can be simulated on the computer by a random number generator.

Although we only compute the node values Y_n it is convenient for our proofs to define $Y(t)$ for $t \in [t_n, t_{n+1})$ by

$$Y(t) = Y_n + \mu(Y_n, t_n) \int_{t_n}^{t} \mathrm{d}s + \sigma(Y_n, t_n) \int_{t_n}^{t} \mathrm{d}B(s), \quad t \in [t_n, t_{n+1}). \tag{19}$$

Then Y is continuous on $[0,T], Y(t_n) = Y_n$ for all n, and

$$\begin{aligned} &Y(t) = Y_0 + \int_0^t \bar{\mu}(s)\,\mathrm{d}s + \int_0^t \bar{\sigma}(s)\,\mathrm{d}B(s), \quad 0 \le t \le T, \\ &\bar{\mu}(s) = \mu(Y_n, t_n),\ \bar{\sigma}(s) = \sigma(Y_n, t_n), \quad s \in (t_n, t_{n+1}). \end{aligned} \tag{20}$$

We now assume that the coefficients are globally Lipschitz continuous also with respect to t:

$$\begin{aligned} |\mu(x,t) - \mu(x,s)| \le L|t-s|, \\ |\sigma(x,t) - \sigma(x,s)| \le L|t-s|, \end{aligned} \quad \forall x \in \mathbf{R},\ t, s \in [0,T]. \tag{21}$$

Theorem 5 (Strong convergence)**.** *If*

$$\mathbf{E}\big(|Y_0 - X_0|^2\big) \le Kh, \tag{22}$$

and

$$\mathbf{E}\big(|X_0|^2\big) \le M, \tag{23}$$

then

$$\mathbf{E}\Big(\max_{0\le t\le T} |Y(t) - X(t)|^2\Big) \le Ch \tag{24}$$

with $C = C(L,K,M,T)$. In particular,

$$\mathbf{E}\big(|Y_n - X(t_n)|\big) \le Ch^{1/2}, \quad n = 0, \dots, N. \tag{25}$$

Proof. We first prove that (24) implies (25). We use the Cauchy-Schwarz inequality,

$$|\mathbf{E}(fg)| \le \sqrt{\mathbf{E}(f^2)}\sqrt{\mathbf{E}(g^2)},$$

with $f = 1$ and $g = |Y(t_n) - X(t_n)|$:

$$\mathbf{E}\big(|Y_n - X(t_n)|\big) \le \sqrt{\mathbf{E}\big(|Y(t_n) - X(t_n)|^2\big)} \le \sqrt{\mathbf{E}\Big(\max_{0\le t\le T} |Y(t) - X(t)|^2\Big)} \le Ch^{1/2}.$$

We now prove (24). Let $0 \le s \le t \le T$. As in the proof of Theorem 3, using (2), (20):

$$\begin{aligned} \mathbf{E}\Big(\max_{0\le s\le t} |Y(s) - X(s)|^2\Big) &= \Big\{\text{with } \mu(z) = \mu(X(z), z),\ \sigma(z) = \sigma(X(z), z)\Big\} \\ &= \mathbf{E}\Big(\max_{0\le s\le t} \Big\{\Big|Y_0 - X_0 + \int_0^s (\bar{\mu}(z) - \mu(z))\,\mathrm{d}z + \int_0^s (\bar{\sigma}(z) - \sigma(z))\,\mathrm{d}B(z)\Big|^2\Big\}\Big) \\ &\le 3\Big\{\mathbf{E}\big(|Y_0 - X_0|^2\big) + T\int_0^t \mathbf{E}\big(|\bar{\mu} - \mu|^2\big)\,\mathrm{d}s + 4\int_0^t \mathbf{E}\big(|\bar{\sigma} - \sigma|^2\big)\,\mathrm{d}s\Big\}. \end{aligned}$$

We split $\bar\mu-\mu$ into two parts and use the Lipschitz conditions (4) and (21),

$$\begin{aligned}|\bar\mu(s)-\mu(s)| &= |\mu(Y(t_n),t_n)-\mu(X(s),s)|\\
&\le |\mu(Y(t_n),t_n)-\mu(X(s),t_n)|+|\mu(X(s),t_n)-\mu(X(s),s)|\\
&\le L\big(|Y(t_n)-X(s)|+s-t_n\big)\\
&\le L\big(|Y(t_n)-X(t_n)|+|X(t_n)-X(s)|+h_n\big)\\
&\le L\big(\max_{0\le z\le s}|Y(z)-X(z)|+|X(t_n)-X(s)|+h_n\big),\quad \text{for } s\in(t_n,t_{n+1}).\end{aligned}$$

Using also (13) and $h_n\le 1$ we get

$$\begin{aligned}\mathbf{E}\big(|\bar\mu(s)-\mu(s)|^2\big) &\le 3L^2\Big(\mathbf{E}\big(\max_{0\le z\le s}|Y(z)-X(z)|^2\big)+\mathbf{E}\big(|X(t_n)-X(s)|^2\big)+h_n^2\Big)\\
&\le 3L^2\Big(\mathbf{E}\big(\max_{0\le z\le s}|Y(z)-X(z)|^2\big)+C(s-t_n)+h_n^2\Big)\\
&\le C\Big(\mathbf{E}\big(\max_{0\le z\le s}|Y(z)-X(z)|^2\big)+h_n\Big),\quad \text{for } s\in(t_n,t_{n+1}),\end{aligned}$$

with $C=C(M,L,T)$. The difference $\bar\sigma-\sigma$ is estimated in the same way.

Suppose that t belongs to the m-th mesh interval, $t\in(t_m,t_{m+1}]$, and denote $\tilde t_{n+1}=t_{n+1}\wedge t$ (the minimum of t and t_{n+1}). Then, since $h_n\le h$,

$$\begin{aligned}\int_0^t \mathbf{E}\big(|\bar\mu(s)-\mu(s)|^2\big)\,\mathrm{d}s &= \sum_{n=0}^m\int_{t_n}^{\tilde t_{n+1}}\mathbf{E}\big(|\bar\mu(s)-\mu(s)|^2\big)\,\mathrm{d}s\\
&\le C\sum_{n=0}^m\int_{t_n}^{\tilde t_{n+1}}\Big(\mathbf{E}\big(\max_{0\le z\le s}|Y(z)-X(z)|^2\big)+h_n\Big)\,\mathrm{d}s\\
&\le C\int_0^t\mathbf{E}\big(\max_{0\le z\le s}|Y(z)-X(z)|^2\big)\,\mathrm{d}s+Ch.\end{aligned}$$

We get the same bound for the σ-term. Therefore

$$\phi(t)\le Ch+C\int_0^t\phi(s)\,\mathrm{d}s,\quad t\in[0,T],$$

with $C=C(L,K,M,T)$, $\phi(t)=\mathbf{E}\big(\max_{0\le s\le t}|Y(s)-X(s)|^2\big)$. Gronwall's lemma completes the proof. □

We have now proved *strong convergence*, or *pathwise convergence*, of order $h^{1/2}$:

$$\Big\|\max_{0\le t\le T}|Y(t)-X(t)|\Big\|_{L_2}=\sqrt{\mathbf{E}\Big(\max_{0\le t\le T}|Y(t)-X(t)|^2\Big)}\le Ch^{1/2}, \tag{26}$$

in other words,

$$\Big\|\max_{0\le t\le T}|Y(t)-X(t)|\Big\|_{L_2}=O(h^{1/2})\quad\text{as } h\to 0.$$

Note that, due to the use of Gronwall's inequality, the constant in (26) grows exponentially with T, $C=\exp(C(L,K,M,T)T)$. By inspection of the previous proof we see that in the deterministic case, $\sigma=0$, we have convergence of order h:

$$\max_{0\le t\le T}|Y(t)-X(t)|\le Ch.$$

This is the classical result for Euler's method for deterministic ODEs.

In the next lecture we shall study *weak convergence* and prove that

$$\Big|\mathbf{E}\big(g(Y_n)-g(X(t_n))\big)\Big|\le Ch$$

for all smooth functions g. Thus the weak convergence order is h.

We shall also derive a higher order method: Milstein's method.

The presentation of some of the material in this lecture was inspired by [1]. For more details see [4].

2. Lecture 13

Recall the strong solution

$$X(t) = X_0 + \int_0^t \mu(X(s), s)\,\mathrm{d}s + \int_0^t \sigma(X(s), s)\,\mathrm{d}B(s), \quad t \in [0, T], \tag{27}$$

and Euler's method

$$Y(t) = Y_0 + \int_0^t \bar{\mu}(s)\,\mathrm{d}s + \int_0^t \bar{\sigma}(s)\,\mathrm{d}B(s), \quad t \in [0, T], \tag{28}$$

where

$$\bar{\mu}(s) = \mu(Y(t_n), t_n), \quad \bar{\sigma}(s) = \sigma(Y(t_n), t_n), \quad s \in (t_n, t_{n+1}), \tag{29}$$

are piecewise constant, "frozen", functions.

We have proved *strong convergence*, or *pathwise convergence*:

$$\Big\| \max_{0 \le t \le T} |Y(t) - X(t)| \Big\|_{L_2} = \sqrt{\mathbf{E}\Big(\max_{0 \le t \le T} |Y(t) - X(t)|^2 \Big)} \le C h^{1/2}. \tag{30}$$

We are often not interested in individual paths but we would like to compute the expected value of some quantity that the depends on $X(t)$, i.e., we want to compute $\mathbf{E}\big(g(X(t))\big)$ for some function g.

We shall prove that

$$\Big| \mathbf{E}\big(g(Y(t)) - g(X(t)) \big) \Big| \le Ch$$

for all smooth functions g. This is called *weak convergence*. Thus the weak convergence order is h. We will use Kolmogorov's backward equation in the proof.

We shall also derive the Itô-Taylor expansion and use it to obtain a numerical method of higher order: Milstein's method.

All of this rests on Itô's formula.

2.1. Kolmogorov's backward equation.

We recall Itô's formula for functions of x, t (see Klebaner, Theorem 4.18 or Theorem 6.1). If X satisfies

$$\mathrm{d}X(t) = \mu(X(t), t)\,\mathrm{d}t + \sigma(X(t), t)\,\mathrm{d}B(t)$$

and $u = u(x, t)$ is twice differentiable in x and once in t, then (denoting the partial derivatives $u_x = \frac{\partial u}{\partial x},\ u_{xx} = \frac{\partial^2 u}{\partial x^2},\ u_t = \frac{\partial u}{\partial t}$)

$$\begin{aligned} \mathrm{d}u(X(t), t) = \Big(u_t(X(t), t) + \mu(X(t), t)\, u_x(X(t), t) + \tfrac{1}{2}\sigma^2(X(t), t)\, u_{xx}(X(t), t) \Big)\,\mathrm{d}t \\ + \sigma(X(t), t)\, u_x(X(t), t)\,\mathrm{d}B(t). \end{aligned} \tag{31}$$

This really means

$$u(X(t), t) - u(X(s), s) = \int_s^t \big(u_t + \mu u_x + \tfrac{1}{2}\sigma^2 u_{xx} \big)\,\mathrm{d}z + \int_s^t \sigma u_x\,\mathrm{d}B(z), \quad 0 \le s \le t, \tag{32}$$

where the integrands are evaluated at $(X(z), z)$.

Theorem 6 (Kolmogorov's backward equation, Klebaner Theorem 6.6)**.** *Let*

$$\mathrm{d}X(t) = \mu(X(t),t)\,\mathrm{d}t + \sigma(X(t),t)\,\mathrm{d}B(t), \quad t \in [0,T], \tag{33}$$

and let $u(x,t)$ be a sufficiently smooth solution of

$$\begin{aligned} &u_t(x,t) + \mu(x,t)u_x(x,t) + \tfrac{1}{2}\sigma^2(x,t)u_{xx}(x,t) = 0, \quad x \in \mathbf{R},\ t < T, \\ &u(x,T) = g(x). \end{aligned} \tag{34}$$

Then

$$u(x,t) = \mathbf{E}\big(g(X(T)) \mid X(t) = x\big). \tag{35}$$

Note: Under appropriate assumptions we can prove that (34) has a unique solution. Thus, the only solution is given by (35).

The notation in (35) is somewhat confusing but it is the traditional one. To explain what it means, we let $Z(s) = Z(s;x,t)$ denote the unique strong solution of

$$\begin{aligned} &\mathrm{d}Z(s) = \mu(Z(s),s)\,\mathrm{d}s + \sigma(Z(s),s)\,\mathrm{d}B(s), \quad t \le s \le T, \\ &Z(t) = x. \end{aligned}$$

In particular, the solution of (1) is $X(t) = Z(t; X_0, 0)$. Now (35) can be written

$$u(x,t) = \mathbf{E}\big(g(Z(T;x,t))\big). \tag{36}$$

Note that the variable x in (35), (36) is deterministic, while X_0 in (1) may be a random variable.

Proof. Itô's formula applied to $u(Z(s;x,t),s)$ and (34) give

$$\begin{aligned} u(Z(T;x,t),T) - u(Z(t;x,t),t) &= \int_t^T \big(u_t + \mu u_x + \tfrac{1}{2}\sigma^2 u_{xx}\big)\,\mathrm{d}s + \int_t^T \sigma u_x\,\mathrm{d}B \\ &= \int_t^T \sigma u_x\,\mathrm{d}B. \end{aligned}$$

Here the Itô integral is a martingale with respect to T and hence its expected value is zero, so that

$$\begin{aligned} &\mathbf{E}\big(u(Z(T;x,t),T)\big) - \mathbf{E}\big(u(Z(t;x,t),t)\big) \\ &\quad = \mathbf{E}\Big(\int_t^T \sigma(Z(s;x,t),s)u_x(Z(s;x,t),s)\,\mathrm{d}B(s)\Big) = 0. \end{aligned}$$

Since $u(Z(T;x,t),T) = g(Z(T;x,t))$ by (34) and $\mathbf{E}\big(u(Z(t;x,t),t)\big) = \mathbf{E}\big(u(x,t)\big) = u(x,t)$, we get

$$\mathbf{E}\big(g(Z(T;x,t))\big) = u(x,t),$$

which is (36). $\square$

The proof requires that u is sufficiently smooth and bounded so that the Itô formula holds. In particular, we need to guarantee that the Itô integral $\int_t^T \sigma u_x\,\mathrm{d}B$ is defined. This holds if we assume that u_x is continuous and bounded,

$$|u_x(x,t)| \le C, \quad \forall x \in \mathbf{R},\ t \in [0,T].$$

Then the process $s \mapsto \sigma(Z(s;x,t),s)u_x(Z(s;x,t),s)$ is adapted and square integrable,

$$\mathbf{E}\big(|\sigma(Z(s;x,t),s)u_x(Z(s;x,t),s)|^2\big) \le CL^2\big(1 + \mathbf{E}\big(|Z(s;x,t)|^2\big)\big) < \infty,$$

where we also used (5) and (8). Hence, $\int_t^T \sigma u_x\,\mathrm{d}B$ is defined.

Note that the proof actually shows that the expected value of $u(y,s)$ is constant along the path $y = Z(s;x,t)$, $s \in [0,T]$, in the (y,s)-plane. More precisely,

$$\mathbf{E}\Big(u(Z(s;x,t),s)\Big) = u(x,t), \quad s \in [0,T]. \tag{37}$$

According to Definition 5.14 and Theorem 5.15 in Klebaner, there is a unique function $p(y,s,x,t)$, $t < s \le T$, $x, y \in \mathbf{R}$, (a fundamental solution) such that the unique solution of (34) is given by

$$u(x,t) = \int_{-\infty}^{\infty} p(y,T,x,t) g(y) \, \mathrm{d}y, \quad x \in \mathbf{R}, \ t < T. \tag{38}$$

The function $(y,s) \mapsto p(y,s,x,t)$ (with x,t fixed) satisfies the forward equation, see Klebaner (5.5a),

$$\frac{\partial p}{\partial s} + \frac{\partial}{\partial y}\Big(\mu(y,s)p\Big) - \frac{1}{2}\frac{\partial^2}{\partial y^2}\Big(\sigma^2(y,s)p\Big) = 0, \quad y \in \mathbf{R}, \ t < s \le T. \tag{39}$$

The initial value at $s = t$ is the Dirac measure at x:

$$p(y,t,x,t) = \delta_x(y), \quad y \in \mathbf{R}. \tag{40}$$

In the simplified case when $\mu = 0$, $\sigma^2 = 1$, (39), (40) become

$$\begin{aligned} &p_s - \tfrac{1}{2} p_{yy} = 0, && y \in \mathbf{R}, \ s > t, \\ &p(y,t,x,t) = \delta_x(y), && y \in \mathbf{R}, \end{aligned} \tag{41}$$

which is the forward heat equation with solution

$$p(y,s,x,t) = \frac{1}{\sqrt{2\pi(s-t)}} \, \mathrm{e}^{-\frac{(x-y)^2}{2(s-t)}}, \quad s > t, \ x, y \in \mathbf{R},$$

the Gauss kernel. This is proved by taking the Fourier transform of (41) with respect to y. Then (38) becomes

$$u(x,t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(T-t)}} \, \mathrm{e}^{-\frac{(x-y)^2}{2(T-t)}} g(y) \, \mathrm{d}y.$$

2.2. **Weak convergence of Euler's method.** In the following theorem we make rather strong assumptions on the coefficients and the initial value in order to make the proof simpler. These can be relaxed.

Theorem 7 (Weak convergence)**.** *Assume that μ, σ, and g have sufficiently many bounded derivatives. Let X and Y be solutions of* (27) *and* (28) *with $Y_0 = X_0$ a deterministic variable. Then*

$$\Big|\mathbf{E}\big(g(Y(T))\big) - \mathbf{E}\big(g(X(T))\big)\Big| \le Ch. \tag{42}$$

Proof. Let $u = u(x,t)$ be the solution of the Kolmogorov backward equation (34). Then by (35), (36) we have

$$u(x,t) = \mathbf{E}\big(g(Z(T;x,t))\big), \tag{43}$$

and, in particular, since X_0 is deterministic,

$$u(X_0, 0) = \mathbf{E}\big(g(Z(T;X_0,0))\big) = \mathbf{E}\big(g(X(T))\big), \tag{44}$$

Since Y satisfies (27), Itô's formula (32) gives

$$
\begin{aligned}
&u(Y(T),T) - u(Y(0),0) \\
&\quad = \int_0^T \Big(u_t(Y(t),t) + \bar{\mu}(t) u_x(Y(t),t) + \tfrac{1}{2}\bar{\sigma}^2(t) u_{xx}(Y(t),t) \Big)\,\mathrm{d}t \\
&\qquad + \int_0^T \bar{\sigma}(t) u_x(Y(t),t)\,\mathrm{d}B(t).
\end{aligned}
$$

From (34) we get

$$u_t(Y(t),t) = -\mu(Y(t),t) u_x(Y(t),t) - \tfrac{1}{2}\sigma^2(Y(t),t) u_{xx}(Y(t),t),$$

so that

$$u(Y(T),T) - u(Y(0),0) = \int_0^T \Big((\bar{\mu} - \mu) u_x + \tfrac{1}{2}(\bar{\sigma}^2 - \sigma^2) u_{xx} \Big)\,\mathrm{d}t + \int_0^T \bar{\sigma} u_x \,\mathrm{d}B(t),$$

where the integrands are evaluated at $(Y(t),t)$. Take expectation here, use $u(Y(T),T) = g(Y(T))$ from (34), use $Y_0 = X_0$ and (44) to get

$$\mathbf{E}\big(u(Y(0),0)\big) = \mathbf{E}\big(u(Y_0,0)\big) = \mathbf{E}\big(u(X_0,0)\big) = \mathbf{E}\big(g(X(T))\big),$$

and use $\mathbf{E}\big(\int_0^T \bar{\sigma} u_x \,\mathrm{d}B\big) = 0$. We get

$$
\begin{aligned}
\mathbf{E}\big(g(Y(T))\big) - \mathbf{E}\big(g(X(T))\big) &= \int_0^T \Big\{ \mathbf{E}\big((\bar{\mu} - \mu) u_x\big) + \tfrac{1}{2}\mathbf{E}\big((\bar{\sigma}^2 - \sigma^2) u_{xx}\big) \Big\}\,\mathrm{d}t \\
&= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \Big\{ \mathbf{E}\big((\bar{\mu} - \mu) u_x\big) + \tfrac{1}{2}\mathbf{E}\big((\bar{\sigma}^2 - \sigma^2) u_{xx}\big) \Big\}\,\mathrm{d}t.
\end{aligned}
$$

We shall show

$$\sup_{t \in (t_n, t_{n+1})} \Big| \mathbf{E}\big((\bar{\mu}(t) - \mu(Y(t),t)) u_x(Y(t),t) \big) \Big| \le C h_n, \tag{45}$$

$$\sup_{t \in (t_n, t_{n+1})} \tfrac{1}{2} \Big| \mathbf{E}\big((\bar{\sigma}^2(t) - \sigma^2(Y(t),t)) u_{xx}(Y(t),t) \big) \Big| \le C h_n, \tag{46}$$

which implies the desired result:

$$\mathbf{E}\big(g(Y(T))\big) - \mathbf{E}\big(g(X(T))\big) \le C \sum_{n=0}^{N-1} h_n \int_{t_n}^{t_{n+1}} \,\mathrm{d}t \le CT \max h_n = CTh.$$

To prove (45) we define

$$v(x,t) = \big(\mu(Y_n,t_n) - \mu(x,t)\big) u_x(x,t),$$

and note that (45) means (recall that $\bar{\mu}(t) = \mu(Y(t_n),t_n)$ for $t \in (t_n,t_{n+1})$)

$$\Big| \mathbf{E}\big(v(Y(t),t)\big) \Big| \le C h_n, \quad t \in (t_n, t_{n+1}).$$

Since Y satisfies (27), Itô's formula (32) gives

$$
\begin{aligned}
&v(Y(t),t) - v(Y(t_n),t_n) \\
&\quad = \int_{t_n}^t \Big(v_t(Y(s),s) + \bar{\mu}(s) v_x(Y(s),s) + \tfrac{1}{2}\bar{\sigma}^2(s) v_{xx}(Y(s),s) \Big)\,\mathrm{d}s \\
&\qquad + \int_{t_n}^t \bar{\sigma}(s) v_x(Y(s),s)\,\mathrm{d}B(s).
\end{aligned}
$$

Here $v(Y(t_n),t_n)=0$ and after taking expectation we get

$$\begin{aligned}\big|\mathbf{E}\big(v(Y(t),t)\big)\big| &= \Big|\int_{t_n}^{t}\mathbf{E}\Big(v_t(Y(s),s)+\bar{\mu}(s)v_x(Y(s),s)+\tfrac{1}{2}\bar{\sigma}^2(s)v_{xx}(Y(s),s)\Big)\,\mathrm{d}s\Big|\\ &\le h_n \max_{\mathbf{R}\times[0,T]}\Big(|v_t|+|\mu|\,|v_x|+\tfrac12|\sigma^2|\,|v_{xx}|\Big),\end{aligned}$$

where we used that

$$\begin{aligned}\Big|\int_{t_n}^{t}\mathbf{E}\big(v_t(Y(s),s)\big)\,\mathrm{d}s\Big| &\le \int_{t_n}^{t}\mathbf{E}\big(|v_t(Y(s),s)|\big)\,\mathrm{d}s\\ &\le \int_{t_n}^{t}\max_{y\in\mathbf{R},s\in[0,T]}|v_t(y,s)|\,\mathrm{d}t \le h_n \max_{\mathbf{R}\times[0,T]}|v_t|,\end{aligned}$$

and similarly for the other terms. This proves (45) if we have the global bounds

$$\begin{aligned}&|\mu(x,t)|\le C,\quad x\in\mathbf{R},\ t\in[0,T],\\ &|\sigma(x,t)|\le C,\quad x\in\mathbf{R},\ t\in[0,T],\\ &|v_t(x,t)|+|v_x(x,t)|+|v_{xx}(x,t)|\le C,\quad x\in\mathbf{R},\ t\in[0,T],\end{aligned}$$

where the last one in its turn follows from the bounds

$$\begin{aligned}&|\mu_t(x,t)|+|\mu_x(x,t)|+|\mu_{xx}(x,t)|\le C,\quad x\in\mathbf{R},\ t\in[0,T],\\ &|u_{xt}(x,t)|+|u_x(x,t)|+|u_{xx}(x,t)|+|u_{xxx}(x,t)|\le C,\quad x\in\mathbf{R},\ t\in[0,T].\end{aligned}$$

Such bounds can be proved if μ,σ,g are sufficiently smooth and bounded. This proves (45), and (46) is proved in the same way. □

This proof was inspired by [5].

If $Y_0=X_0$ is a random variable, then we must first show

$$u(X_0,0)=\mathbf{E}\big(g(X(T))\mid X_0\big),$$

so that (by the law of double expectation)

$$\mathbf{E}\big(u(X_0,0)\big)=\mathbf{E}\big(\mathbf{E}\big(g(X(T))\mid X_0\big)\big)=\mathbf{E}\big(g(X(T))\big)$$

instead of (44). The proof can then be continued in the same way. We skip the details.

2.3. **Monte-Carlo method, sampling error.** In practice we compute several sample paths $\{Y(t,\omega_j)\}_{j=1}^M$, $\omega_j\in\Omega$, and approximate the expected value $\mathbf{E}\big(g(X(T))\big)$ by the average $\frac{1}{M}\sum_{j=1}^M g(Y(T,\omega_j))$. This is called the Monte-Carlo method. The total error is then the sum of the *discretization error* and the *statistical error (sampling error)*:

$$\begin{aligned}&\Big|\frac{1}{M}\sum_{j=1}^M g(Y(T,\omega_j))-\mathbf{E}\big(g(X(T))\big)\Big|\\ &\quad\le \Big|\mathbf{E}\big(g(Y(T))\big)-\mathbf{E}\big(g(X(T))\big)\Big|+\Big|\frac{1}{M}\sum_{j=1}^M\Big(g(Y(T,\omega_j))-\mathbf{E}\big(g(Y(T))\big)\Big)\Big|.\end{aligned}$$

We have shown that the first part is $\le Ch$, and by using the *central limit theorem* it can be shown that the sampling error is $\le \frac{C}{\sqrt{M}}$, where the constant depends on the variance of $g(Y(T))$. See [5].

2.4. **A numerical method of higher order: Milstein's method.** We derive the first steps of the *Itô-Taylor expansion*. We consider for simplicity the autonomous equation

$$\mathrm{d}X(t) = \mu(X(t))\,\mathrm{d}t + \sigma(X(t))\,\mathrm{d}B(t),$$

where $\mu = \mu(x), \sigma = \sigma(x)$ do not depend on t. Then

$$X(t) = X(t_0) + \int_{t_0}^{t} \mu(X(s))\,\mathrm{d}s + \int_{t_0}^{t} \sigma(X(s))\,\mathrm{d}B(s). \tag{47}$$

Itô's formula is, see Klebaner (4.53),

$$\begin{aligned} f(X(t)) &= f(X(t_0)) + \int_{t_0}^{t} \Big(\mu(X(s))f'(X(s)) + \tfrac{1}{2}\sigma^2(X(s))f''(X(s))\Big)\,\mathrm{d}s \\ &\quad + \int_{t_0}^{t} \sigma(X(s))f'(X(s))\,\mathrm{d}B(s) \\ &= f(X(t_0)) + \int_{t_0}^{t} L_0 f(X(s))\,\mathrm{d}s + \int_{t_0}^{t} L_1 f(X(s))\,\mathrm{d}B(s), \end{aligned} \tag{48}$$

where the differential operators L_0 and L_1 are defined by

$$\begin{aligned} L_0 f &= \mu f' + \tfrac{1}{2}\sigma^2 f'', \\ L_1 f &= \sigma f'. \end{aligned}$$

In particular, with $f(x) = x$ we get $L_0 f = \mu,\ L_1 f = \sigma$ and we retrieve (47). Next we take $f = \mu$ and $f = \sigma$ in (48) and insert the result into (47). We get

$$\begin{aligned} X(t) &= X(t_0) + \int_{t_0}^{t} \Big(\mu(X(t_0)) + \int_{t_0}^{s} L_0\mu(X(z))\,\mathrm{d}z + \int_{t_0}^{s} L_1\mu(X(z))\,\mathrm{d}B(z)\Big)\,\mathrm{d}s \\ &\quad + \int_{t_0}^{t} \Big(\sigma(X(t_0)) + \int_{t_0}^{s} L_0\sigma(X(z))\,\mathrm{d}z + \int_{t_0}^{s} L_1\sigma(X(z))\,\mathrm{d}B(z)\Big)\,\mathrm{d}B(s) \\ &= X(t_0) + \mu(X(t_0))\int_{t_0}^{t} \mathrm{d}s + \sigma(X(t_0))\int_{t_0}^{t} \mathrm{d}B(s) + R_1(t,t_0), \end{aligned}$$

where the remainder is

$$\begin{aligned} R_1(t,t_0) &= \int_{t_0}^{t}\int_{t_0}^{s} L_0\mu(X(z))\,\mathrm{d}z\,\mathrm{d}s \\ &\quad + \int_{t_0}^{t}\int_{t_0}^{s} L_1\mu(X(z))\,\mathrm{d}B(z)\,\mathrm{d}s \\ &\quad + \int_{t_0}^{t}\int_{t_0}^{s} L_0\sigma(X(z))\,\mathrm{d}z\,\mathrm{d}B(s) \\ &\quad + \int_{t_0}^{t}\int_{t_0}^{s} L_1\sigma(X(z))\,\mathrm{d}B(z)\,\mathrm{d}B(s). \end{aligned}$$

This is the motivation for Euler's method.

Next use Itô's formula with $f = L_1\sigma = \sigma\sigma'$ and insert into the last term in the remainder to get

$$\begin{aligned} X(t) &= X(t_0) + \mu(X(t_0))\int_{t_0}^{t} \mathrm{d}s + \sigma(X(t_0))\int_{t_0}^{t} \mathrm{d}B(s) \\ &\quad + \sigma(X(t_0))\sigma'(X(t_0))\int_{t_0}^{t}\int_{t_0}^{s} \mathrm{d}B(z)\,\mathrm{d}B(s) + R_2(t,t_0), \end{aligned}$$

where the new remainder $R_2(t,t_0)$ consists of five complicated terms.

This is the motivation for Milstein's method. We have here

$$\int_{t_0}^{t} \mathrm{d}s = t - t_0, \quad \int_{t_0}^{t} \mathrm{d}B(s) = B(t) - B(t_0),$$

and the iterated Itô integral, see Klebaner Example 4.12 on p. 107,

$$\int_{t_0}^{t}\int_{t_0}^{s} \mathrm{d}B(z)\,\mathrm{d}B(s) = \int_{t_0}^{t} \Big(B(s) - B(t_0)\Big)\,\mathrm{d}B(s) = \frac{1}{2}\Big(\big(B(t) - B(t_0)\big)^2 - (t - t_0)\Big).$$

The Milstein method defines $\{Y_n\}_{n=0}^{N}$ by $Y_0 = X_0$ and

$$Y_{n+1} = Y_n + \mu(Y_n)h_n + \sigma(Y_n)\Delta B_n + \tfrac{1}{2}\sigma(Y_n)\sigma'(Y_n)\big((\Delta B_n)^2 - h_n\big),$$

where

$$h_n = t_{n+1} - t_n, \quad \Delta B_n = B(t_{n+1}) - B(t_n).$$

It can be shown that it converges with strong order h:

$$\mathbf{E}\big(|Y_n - X(t_n)|\big) \le Ch.$$

Milstein's method for a system of SODE (7) involves iterated integrals

$$\int_{t_0}^{t}\int_{t_0}^{s} \mathrm{d}B_i(z)\,\mathrm{d}B_j(s),$$

which are difficult to evaluate.

This idea can be extended to derive methods of higher order but these are not very useful in practice because the coefficients are many times iterated integrals, for example, $\int_{t_0}^{t} \cdots \int_{t_0}^{z_2} \mathrm{d}B(z_1) \cdots \mathrm{d}B(z_n)\,\mathrm{d}s_1 \cdots \mathrm{d}s_k$, which are difficult to evaluate.

More material on iterated Itô formulas and Itô-Taylor expansions can be found in [4].

3. Exercises

Exercise 1. Prove the estimate (8).

Exercise 2. **Existence.** Prove Theorem 1. That is, show that there is a unique strong solution on the interval $[0,T]$.

Hint: note that equation (2) is a fixed point equation $X = G(X)$ and show that the operator

$$G(Y)(t) = X_0 + \int_0^t \mu(Y(s),s)\,\mathrm{d}s + \int_0^t \sigma(Y(s),s)\,\mathrm{d}B(s),$$

is a contraction on the Banach space

$$W_{[0,\tau]} = \Big\{Y \in C([0,\tau], L_2(\Omega)) : Y \text{ is adapted to the filtration generated by } B\Big\}$$

with norm

$$\|Y\|_{W_{[0,\tau]}} = \max_{0\le t\le\tau} \|Y(t)\|_{L_2} = \max_{0\le t\le\tau} \sqrt{\mathbf{E}\big(|Y(t)|^2\big)}$$

provided that τ is small enough. Hence, we obtain a solution on a (short) interval $[0,\tau]$. Repeat and obtain solutions on $[\tau, 2\tau]$, $[2\tau, 3\tau]$, and so on until we cover the whole interval $[0,T]$.

Exercise 3. Repeat the proof of Theorem 3 but without using Doob's inequality and with $\phi(t) = \mathbf{E}\big(|\hat{X}(t) - X(t)|^2\big)$. Prove the weaker result

$$\mathbf{E}\big(|\hat{X}(t) - X(t)|^2\big) \le C\mathbf{E}\big(|\hat{X}_0 - X_0|^2\big), \quad t \in [0,T],$$

and hence

$$\max_{0 \le t \le T} \mathbf{E}\big(|\hat{X}(t) - X(t)|^2\big) \le C\mathbf{E}\big(|\hat{X}_0 - X_0|^2\big). \tag{49}$$

Note that the maximum is now outside the expected value and that (11) implies (49).

Exercise 4. Prove (37).

Exercise 5. Note that the partial differential equations in (39) and (34) may be written as $Lp = 0$ and $L^*u = 0$, respectively, where $L = \frac{\partial}{\partial t} + \frac{\partial}{\partial x}(\mu(x,t)\cdot) - \frac{1}{2}\frac{\partial^2}{\partial x^2}(\sigma^2(x,t)\cdot)$ and $L^* = -\frac{\partial}{\partial t} - \mu(x,t)\frac{\partial}{\partial x} - \frac{1}{2}\sigma^2(x,t)\frac{\partial^2}{\partial x^2}$. Show that the operators L and L^* are formally adjoint in the sense that

$$\int_0^T \int_{\mathbf{R}} (L\phi)\psi \,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_{\mathbf{R}} \phi(L^*\psi)\,\mathrm{d}x\,\mathrm{d}t \quad \forall \phi, \psi \in C_0^\infty(\mathbf{R} \times (0,T)). \tag{50}$$

Exercise 6. Show that (38) follows from (34) and (39), (40). Hint: multiply (34) in the form $0 = u_s + \mu u_y + \frac{1}{2}\sigma^2 u_{yy}$ by $p(y,s,x,t)$, integrate $\int_t^T \int_{\mathbf{R}} \cdots \,\mathrm{d}y\,\mathrm{d}s$, integrate by parts and use (39), (40).

Exercise 7. Derive Milstein's method for a system of SODE (7).

Exercise 8. **Matlab computations.** Read Higham and do (at least) the programs `bpath1.m`, `bpath2.m`, `bpath3.m`, `stint.m`, `em.m`, `emstrong.m`.

Exercise 9. **The Black-Scholes process.**

$$\begin{aligned} \mathrm{d}X(t) &= rX(t)\,\mathrm{d}t + \sigma X(t)\,\mathrm{d}B(t) \\ X(0) &= X_0 \end{aligned}$$

with constants $r > 0$, $\sigma > 0$. (a) Show that the unique strong solution is given by the formula

$$X(t) = X_0 \exp((r - \tfrac{1}{2}\sigma^2)t + \sigma B(t))$$

(b) Write a Matlab program and solve the equation by the Euler method. Plot several sample paths and compare with the solution formula. Examine strong convergence. (This is `emstrong.m`.)

Exercise 10. **The Ornstein-Uhlenbeck process.** The Langevin equation is

$$\begin{aligned} \mathrm{d}X(t) &= -\alpha X(t)\,\mathrm{d}t + \sigma\,\mathrm{d}B(t) \\ X(0) &= X_0 \end{aligned}$$

with constants $\alpha > 0$, $\sigma > 0$. (a) Show that the unique strong solution is given by the formula

$$X(t) = \mathrm{e}^{-\alpha t}X_0 + \sigma \int_0^t \mathrm{e}^{-\alpha(t-s)}\,\mathrm{d}B(s)$$

(b) Write a Matlab program and solve the equation by the Euler method. Plot several sample paths and compare with the solution formula. Examine strong convergence.

References

[1] L. C. Evans, "An introduction to stochastic differential equations", lecture notes, Department of Mathematics, University of California, Berkeley. http://www.math.berkeley.edu/~evans/SDE.course.pdf

[2] D. J. Higham "An Algorithmic Introduction to Numerical Simulation of SDE", SIAM Review 43 (2001), 525–522. link, Matlab files and corrections

[3] F. C. Klebaner, "Introduction to Stochastic Calculus with Applications", second edition, Imperial College Press, 2005.

[4] P. E. Kloeden and E. Platen "Numerical Solution of Stochastic Differential Equations", Springer 1995.

[5] K. S. Moon, J. Goodman, A. Szepessy, R. Tempone, and G. Zouraris, "Stochastic and Partial Differential Equations with Adapted Numerics", Chapter 5, lecture notes, NADA, Royal Institute of Technology, Stockholm.
http://www.math.kth.se/~szepessy/sdepde.pdf
http://www.math.kth.se/~szepessy/sdepde.ps