Theorem (Itô's formula):
Let X_t be given by
$$dX_t = b(t,\omega)\, dt + \sigma(t,\omega)\, dW_t$$
and let $f(t,x)$ be a $C^{1,2}$ function.
then the process $Y_t = f(X_t)$ satisfies

$$\boxed{\begin{aligned} df(X_t) = \Big(&\frac{\partial f}{\partial t}(t,X_t) + \frac{\partial f}{\partial x}(t,X_t)\, b(t,\omega) \\ &+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,X_t)\, \sigma^2(t,\omega) \Big) dt + \\ &+ \frac{\partial f}{\partial x}(t,X_t)\, \sigma(t,\omega)\, dW_t . \end{aligned}}$$

this can be generalised to higher dimensions.
Consider
$$dX_t = b(t,\omega)\, dt + \sigma(t,\omega)\, dW_t$$
with
$$b : [0,T] \times \Omega \to \mathbb{R}^d$$
$$\sigma : [0,T] \times \Omega \to \mathbb{R}^{d\times m}$$
$W_t \to m$-dimensional BM
$X_t \to d$-dimensional process

$\Rightarrow \int_0^t \sigma(s,\omega)\, dW_s \to d$-vector of Itô integrals:
$$\left[\int_0^t \sigma(s,\omega)\, dW_s\right]_i = \int_0^t \sum_{j=1}^m (\sigma)^{i,j}\, dW_s^j .$$

Itô's formula can be generalised as follows

$$dY_t = \partial_t f(t,X_t)\, dt + \sum_{i=1}^d \partial_{x_i} f(t,X_t)\, dX_t^i$$
$$+ \frac{1}{2} \sum_{i,j=1}^d \partial_{x_i}\partial_{x_j} f(t,X_t) \sum_{l=1}^m \sigma^{i,l}\sigma^{j,l}\, dt .$$

4.3. Stochastic Differential Equations

Let W_t, $t \geq 0$ be a Brownian motion. An equation of the form

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dW_t \qquad (*)$$
$$X_0 = \eta$$

where $b(t,x)$ (drift coefficient) and $\sigma(t, X_t)$ (diffusion coefficient) are given functions, η is a given random variable and X_t is the unknown is called a (Itô) stochastic differential equation (SDE) driven by Brownian motion.

The solution of this equation (if it exists) is called an Itô diffusion.

Definition: A process X_t is called a strong solution of the SDE $(*)$ if it is $\mathcal{G}_t$-adapted and $\forall t > 0$, the integrals

$$\int_0^t b(s, X_s)\,ds \quad \text{and} \quad \int_0^t \sigma(s, X_s)\,dW_s$$

exist and

$$X_t = \eta + \int_0^t b(s, X_s)\,ds + \int_0^t \sigma(s, X_s)\,dW_s .$$

This solution is unique if whenever $\hat{X}_t$ is another strong solution, $\mathbb{P}(X_t \neq \hat{X}_t) = 0$, $\forall t \geq 0$.

Theorem: If $b(t,x)$ and $\sigma(t,x)$ satisfy the conditions:

(i) $\forall T, N$, $\exists\, K = K(N,T)$ such that $\forall\, |x|, |y| \leq N$, $0 \leq t \leq T$
$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq K|x-y|$$

(ii) $|b(t,x)| + |\sigma(t,x)| \leq K(1 + |x|)$

(iii) η is independent of $(W_t,\ 0 \leq t \leq T)$, $\mathbb{E}\eta^2 < \infty$

then there exists a strong solution X_t of the SDE $(*)$. X_t has continuous paths and

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} X_t^2\right) \leq C(1 + \mathbb{E}\eta^2), \quad C = C(K,T).$$

Note that:

- Condition (i) means that $b(t,\cdot)$ and $\sigma(t,\cdot)$ are locally Lipschitz continuous uniformly in time. If b and σ have continuous first derivatives, then this condition holds.
- the strong solution X_t of (*) is a Markov process (it is a diffusion process)
- the proof of this theorem is quite similar to the proof of the analogous Picard existence theorem for ODEs.
- We can convert the Itô SDE (*) to a Stratonovich SDE by performing a drift correction

$$dX_t = \left(b(t,X_t) - \frac{1}{2}\sigma(t,X_t)\frac{\partial\sigma}{\partial x}(t,X_t)\right)dt + \sigma(t,X_t)\circ dW_t.$$

examples of SDEs

1) Geometric Brownian Motion (gBM)

$$dX_t = \lambda X_t\, dt + \sigma X_t\, dW_t, \qquad X_0 = x.$$

We can rewrite this as

$$\underbrace{\frac{dX_t}{X_t}}_{\text{reminds of the derivative of } Y_t = \ln X_t!} = \lambda\, dt + \sigma\, dW_t.$$

Consider $Y_t = f(X_t)$, $f(x) = \ln x$.

We have $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = \frac{1}{x}$ and $\frac{\partial^2 f}{\partial x^2} = -\frac{1}{2x^2}$.

We can use Itô's formula:

$$d(\ln X_t) = \left(0 + \frac{1}{X_t}\lambda X_t - \frac{1}{2X_t^2}\sigma^2 X_t^2\right)dt$$

$$+ \frac{1}{X_t}\sigma X_t\, dW_t$$

$$= \left(\lambda - \frac{\sigma^2}{2}\right)dt + \sigma\, dW_t.$$

Which we can integrate!

$$\ln\left(\frac{X_t}{X_0}\right) = \left(\lambda - \frac{\sigma^2}{2}\right) t + \sigma W_t$$

and this gives

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

Note that we saw this before.
the law of X_t is a log-normal, with mean $\mathbb{E}(X_t) = X_0 e^{\lambda t}$ and variance $V(X_t) = X_0^2 e^{2\lambda t}(e^{\sigma^2 t} - 1)$.
the gBM is the most widely used model of stock price behaviour in mathematical finance.

2) Ornstein Uhlenbeck process (OU).

$$dX_t = -\alpha X_t \, dt + \sigma \, dW_t, \quad X_0 = x, \quad \alpha, \sigma > 0.$$

We can solve this analytically too. To see how, consider the ODE analogue

$$\frac{dx}{dt} = -\alpha x + f(t).$$

the solution is

$$x(t) = e^{-\alpha t} x(0) + \int_0^t e^{-\alpha(t-s)} f(s) \, ds.$$

Similarly, the solution to the OU equation is

$$X_t = e^{-\alpha t} X_0 + \sigma \int_0^t e^{-\alpha(t-s)} \, dW_s.$$

We can check this using Itô's formula.

If $X_0 = x$ is deterministic, we have

$$\mathbb{E}(X_t) = e^{-\alpha t} x$$

$$\mathrm{Var}(X_t) = \mathbb{E}(X_t - \mathbb{E} X_t)^2 = \mathbb{E}\left(\sigma^2 \left(\int_0^t e^{-\alpha(t-s)} \, dW_s\right)^2\right)$$

$$= \sigma^2 e^{-2\alpha t} \int_0^t e^{2\alpha s} \, ds = \sigma^2 e^{-2\alpha t} \frac{e^{2\alpha t} - 1}{2\alpha} = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha t}\right)$$

↑
Itô isometry!

So, if x is deterministic,

$$X_t \sim \mathcal{N}\left(e^{-\alpha t} x, \ \sigma^2/2\alpha \left(1 - e^{-2\alpha t}\right)\right).$$

Note that: the Ornstein-Uhlenbeck process describes the motion of a Brownian particle moving within a fluid, with random "kicks" due to friction with neighboring particles.

- It is a mean-reverting process (the mean acts as an equilibrium for the process)
- In mathematical finance, it is used to model interest rates and currency exchange rates.

3) Other interesting examples:

→ Cox-Ingersoll-Ross model for interest rates

$$dX_t = \alpha(\beta - X_t)\,dt + \sigma\sqrt{X_t}\,dW_t$$

→ Stochastic Verhulst equation for population dynamics

$$dX_t = (\lambda X_t - X_t^2)\,dt + \sigma X_t\,dW_t$$

→ Langevin equation (OU with particle having potential energy V)

$$dQ_t = P_t\,dt$$

$$dP_t = (-\lambda P_t - V'(Q_t))\,dt + \sigma\,dW_t$$

Can be written as

$$dX_t = b(X_t)\,dt + \Sigma\,dW_t, \quad X_t = (Q_t, P_t),$$

$$b(q,p) = \begin{pmatrix} p \\ -\lambda p - V'(q) \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & \sigma \end{pmatrix}.$$

Note that noise only acts on P_t.

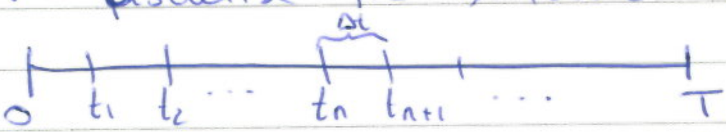
We can also obtain evolution equations for the expectation of functionals of the solution of the SDE (*) and for its probability density. We will not do it in this module, but if interested ask me for references!

4.4. Numerical Solution of SDEs

Similarly to the ODEs case, even though there are some examples of SDEs that can be solved analytically, most of the time this is not possible (for example, for nonlinear drift or diffusion coefficients) and we must resort to numerical techniques.

To this end, we will construct numerical approximations X_n of the solution X_{t_n} of our SDE.

We discretise the interval $[0,T]$:



$t_n = n\Delta t, \quad n = 0, \dots, N \qquad T = N\Delta t.$

and will see how to define X_n.

4.4.1. The Euler-Maruyama Scheme

This is the simplest approach for numerical solution of SDEs. It is the analogue of the explicit Euler method for ODEs.

Since X_t is Markovian, we have, in each time interval $[t_n, t_{n+1}]$:

$$X(t_{n+1}) = X(t_n) + \int_{t_n}^{t_{n+1}} b(X(s))\,ds + \int_{t_n}^{t_{n+1}} \sigma(X(s))\,dW_s$$

The E-M method consists in assuming that b and σ are sufficiently smooth, so that they do not vary too much over $[t_n, t_{n+1}]$ (for sufficiently small Δt). Then we can do the following approximation:

$$X(t_{n+1}) = X(t_n) + b(X_n)\int_{t_n}^{t_{n+1}} dt + \sigma(X_n)\int_{t_n}^{t_{n+1}} dW_s$$

or

$$X_{n+1} = X_n + b(X_n)\Delta t + \sigma(X_n)\Delta W_n,$$

X_n (left), consistent with Itô integral

$\Delta W_n = W(t_{n+1}) - W(t_n) \sim \mathcal{N}(0, \Delta t)$

Note that ΔW_n are Brownian increments. They are i.i.d. and easy to sample. A typical algorithm is the following.

Algorithm: Euler-Maruyama Method

Let X_0 be the initial state and $\Delta t > 0$.

For $n = 0$ to N:

1) Sample $\xi \sim \mathcal{N}(0,1)$
2) Set $X_{n+1} = b(X_n)\Delta t + \sqrt{\Delta t}\, \sigma(X_n)\, \xi$.

We will discuss discretisation error and stability later.

4.4.2. The Milstein Scheme:

The Milstein scheme improves on the approximation

$$\int_{t_n}^{t_{n+1}} \sigma(X_u)\, dW_u \approx \sigma(X_n) \int_{t_n}^{t_{n+1}} dW_u .$$

In order to do that, we ~~use~~ the Itô formula and the following Lemma:

Lemma: Let $\Delta t > 0$, $n \in \mathbb{N}$ and $\Delta W_n = W_{(n+1)\Delta t} - W_{n\Delta t}$. Then

$$\int_{n\Delta t}^{(n+1)\Delta t} \int_{n\Delta t}^{s} dW_u\, dW_s = \frac{1}{2}\left(\Delta W_n^2 - \Delta t\right).$$

Proof: Let $f(x) = x^2$ and apply Itô's formula to $f(W_t)$:

$$dW_t^2 = 2 W_t\, dW_t + dt .$$

Therefore

$$\int_{n\Delta t}^{(n+1)\Delta t} W_t\, dW_t = \frac{1}{2}\left(\int_{n\Delta t}^{(n+1)\Delta t} dW_t^2 - \int_{n\Delta t}^{(n+1)\Delta t} dt\right)$$

$$= \frac{1}{2}\left(W^2_{(n+1)\Delta t} - W^2_{n\Delta t} + \Delta t\right)$$

Furthermore,

$$\int_{n\Delta t}^{(n+1)\Delta t}\int_{n\Delta t}^{s} dW_u\, dW_s = \int_{n\Delta t}^{(n+1)\Delta t} W_s\, dW_s - W_{n\Delta t}(W_{(n+1)\Delta t} - W_{n\Delta t})$$

$$= \frac{1}{2} W_{(n+1)\Delta t}^2 + \frac{1}{2} W_{n\Delta t} - W_{n\Delta t} W_{n+1\Delta t} - \frac{1}{2}\Delta t$$

$$= \frac{1}{2}\left(\Delta W_n^2 - \Delta t\right). \qquad \Delta W_n^2 = (W_{n+1} - W_n)^2.$$

So now we can apply Itô's lemma.
Define

$$\boxed{\mathcal{L} f(x) = b(x)\frac{\partial f}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}}$$ so that

$$df(X_t) = \left(\frac{\partial f}{\partial t}(t, X_t) + \mathcal{L} f(x)\right)dt + \frac{\partial f}{\partial x}\sigma\, dW_t.$$

<u>Note that</u>: $\mathcal{L}$ is the generator of the SDE.
Its adjoint is given by $\mathcal{L}^* f = -\frac{\partial}{\partial x}\left(-b(x) f + \frac{1}{2}\frac{\partial}{\partial x}(\sigma^2 f)\right)$

As before, we write

$$X_{n+1} = X_n + \int_{t_n}^{t_{n+1}} b(X_s)\,ds + \int_{t_n}^{t_{n+1}} \sigma(X_s)\,dW_s \qquad (\star)$$

where $X_n = X_{t_n} = X_{n\Delta t}$. But now, we apply Itô's formula:

$$b(X_s) = b(X_n) + \int_{t_n}^{s} \mathcal{L} b(X_u)\,du + \int_{t_n}^{s} \frac{\partial b}{\partial x}(X_u)\sigma(X_u)\,dW_u.$$

$$\sigma(X_s) = \sigma(X_n) + \int_{t_n}^{s} \mathcal{L}\sigma(X_u)\,du + \int_{t_n}^{s} \sigma(X_u)\frac{\partial \sigma}{\partial x}(X_u)\,dW_u.$$

then we can substitute these in $(\star)$

$$X_{n+1} = X_n + b(X_n)\Delta t + \sigma(X_n)\Delta W_n +$$

$$\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} \mathcal{L} b(X_u)\,du\,ds + \int_{t_n}^{t_{n+1}}\int_{t_n}^{s} b'(X_u)\sigma(X_u)\,dW_u\,ds +$$

$$+\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} \mathcal{L}\sigma(X_u)\,du\,dW_s + \int_{t_n}^{t_{n+1}}\int_{t_n}^{s} \sigma'(X_u)\sigma(X_u)\,dW_u\,dW_s.$$

Note that: this is exact: we have not made any approximations!

Now, we will use the fact that for $\alpha, \beta \geq 0$, we have $(\Delta t)^{\alpha}(\Delta W_n)^{\beta} = O((\Delta t)^{\alpha+\beta/2})$ to discard terms which are higher than Δt.
This means that terms like $du\,ds$, $dW_u\,ds$, $du\,dW_s$ are discarded and all that remains is the last integral:

$$X_{n+1} = X_n + b(X_n)\Delta t + \sigma(X_n)\Delta W_n + \int_{t_n}^{t_{n+1}}\int_{t_n}^{s} \sigma'\sigma\,dW_u\,dW_s + \boxed{\theta(\Delta t)}$$

and we can do the same approximation as before:

$$X_{n+1} = X_n + b(X_n)\Delta t + \sigma(X_n)\Delta W_n + (\sigma'\sigma)(X_n)\iint dW_u\,dW_s + \theta(\Delta t)$$

$$\Rightarrow \boxed{X_{n+1} \approx X_n + b(X_n)\Delta t + \sigma(X_n)\Delta W_n + (\sigma'\sigma)(X_n)(\Delta W_n^2 - \Delta t)}$$

Noting that $W_{(n+1)\Delta t} - W_{n\Delta t} \sim \mathcal{N}(0,\Delta t)$, we have that $\Delta W_n^2 \sim \Delta t\,\xi_n^2$, where $\xi_n \sim \mathcal{N}(0,1)$. We can now describe the Milstein Scheme:

Algorithm: Milstein Scheme:
Let X_0 be the initial state and $\Delta t > 0$
For $n = 0$ to N
1) Sample $\xi \sim \mathcal{N}(0,1)$
2) Set

$$X_{n+1} = X_n + b(X_n)\Delta t + \sqrt{\Delta t}\,\sigma(X_n)\xi + \frac{\Delta t}{2}(\sigma'\sigma)(X_n)(\xi^2-1)$$

Note that:

- If the SDE has additive noise, i.e., σ is constant (independent of X_t), then $\sigma' = 0$ and the Milstein scheme reduces to the Euler-Maruyama scheme
- We can derive an analogous Milstein approximation for multivariate processes, but in general we will have to deal with terms of the form

$$A_{ij} = \int_s^t \int_s^r dW_i(p)\, dW_j(r) - \int_s^t \int_s^r dW_j(p)\, dW_i(r)$$

(Lévy area terms) which come from non-diagonal terms in σ, and cannot be handled in the same manner as before.

In the next section, we will define two notions of discretisation error and see how these two methods compare.
Then we will study the stability of these schemes.

Additional information on stochastic Taylor schemes (not examinable)

The Euler–Maruyama and Milstein schemes both belong to the class of strong Taylor schemes for stochastic differential equations. In particular, the associated update formulae can be viewed as truncated stochastic Taylor expansions, as we explain below.

Taylor methods for *ordinary* differential equations. We first recall the proof for obtaining the remainder of a truncated Taylor series of a smooth function, because a very similar reasoning can be applied to obtain the remainder of an Itô-Taylor expansion.

Theorem 1 (Taylor's formula with integral remainder)**.** *Let $x(t)$ be a smooth function on $[0,T]$ and let $t^* \in [0,T]$. It holds that*

$$x(t) = \sum_{i=0}^{n} \frac{(t-t^*)^i}{i!}\, x^{(i)}(t^*) + \int_{t^*}^{t} \frac{(t-t^*)^n}{n!}\, x^{(n+1)}(s)\, \mathrm{d}s.$$

Proof. By the fundamental theorem of analysis, it holds that

$$x(t) = x(t^*) + \int_{t^*}^{t} x'(s_1)\, \mathrm{d}s_1. \tag{1}$$

Since x' is also a smooth function, we can apply the fundamental theorem of analysis again to obtain

$$x'(s_1) = x'(t^*) + \int_{t^*}^{s_1} x''(s_2)\, \mathrm{d}s_2.$$

Continuing in this fashion and substituting in (1) leads to the equation

$$x(t) = \sum_{i=0}^{n} c_i(t)\, x^{(i)}(t^*) + \int_{t^*}^{t}\int_{t^*}^{s_1} \cdots \int_{t^*}^{s_n} x^{(n+1)}(s_{n+1})\, \mathrm{d}s_{n+1} \dots \mathrm{d}s_1\, \mathrm{d}t, \qquad n = 1, 2, \dots$$

where the coefficients $c_i(t)$ can be expressed as multiple integrals

$$c_i(t) = \int_{t^*}^{t}\int_{t^*}^{s_1} \cdots \int_{t^*}^{s_{i-1}} \mathrm{d}s_i \dots \mathrm{d}s_2\, \mathrm{d}s_1 = \frac{(t-t^*)^i}{i!}.$$

To conclude the proof, it suffices to rewrite the remainder in a simpler form, which is left as an exercise (but is not very important for our purposes):

$$\int_{t^*}^{t}\int_{t^*}^{s_1} \cdots \int_{t^*}^{s_n} x^{(n+1)}(s_{n+1})\, \mathrm{d}s_{n+1} \dots \mathrm{d}s_1\, \mathrm{d}t = \int_{t^*}^{t} \frac{(t-t^*)^n}{n!}\, x^{(n+1)}(s)\, \mathrm{d}s.$$

□

Assume now that $b(\cdot)$ is a smooth function and consider the equation:

$$x'(t) = b(x(t)).$$

Let us emphasize again that the deterministic version of Itô's formula is simply the chain rule, which in integral form leads to

$$f(x(t)) = f(x(t^*)) + \int_{t^*}^{t} b(x(s))\, f'(x(s))\, \mathrm{d}s.$$

Introducing the operator $\mathcal{L} = b(x)\,\partial_x$, the Taylor expansion of $x(t)$ can be written as follows:

$$x(t) = x(t^*) + b(x(t^*))\,(t - t^*) + \mathcal{L}b(x(t^*))\,\frac{(t-t^*)^2}{2} + \mathcal{L}\mathcal{L}b(x(t^*))\frac{(t-t^*)^3}{6} + \dots,$$

where here and later expressions such as $\mathcal{L}\mathcal{L}b(x(t^*))$ should be read as $(\mathcal{L}(\mathcal{L}b))(x(t^*))$. Remember from the proof of Theorem 1 that the factors $(t-t^*)^i/i!$ originate from multiple integrals. Numerical methods for ODEs can be defined by simply keeping more and more terms in the Taylor series. The simplest scheme of that form is the explicit Euler scheme, which is based on the update following update formula:

$$\hat{x}_{n+1} = \hat{x}_n + \Delta t\, b(\hat{x}_n).$$

Similarly, the second-order Taylor scheme reads

$$\hat{x}_{n+1} = \hat{x}_n + \Delta t\, b(\hat{x}_n) + \frac{1}{2}\Delta t^2\, b'(\hat{x}_n)\, b(\hat{x}_n).$$

Taylor methods for *stochastic* differential equations. Taylor methods for SDEs are very similar in spirit to Taylor methods for ODEs. Instead of relying on deterministic Taylor expansions, these methods are based on *Itô–Taylor expansions*, also known as *stochastic Taylor expansions*. We will refrain here from presenting a theorem as general as Theorem 1 for Itô Taylor expansions, not because it would be difficult to do so but because the notations necessary for stating a general result in a compact manner are quite cumbersome. In contrast with Taylor's formula presented above, which is valid for any smooth function $x(t)$, the Itô–Taylor expansion applies only to stochastic processes that solve an SDE. Let us assume that $b(\cdot)$ and $\sigma(\cdot)$ are smooth, globally Lipschitz functions and let X_t denote the unique strong solution of

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0,$$

where x_0 is independent of the Brownian motion. The construction of an Itô–Taylor expansion is very similar to that of a Taylor expansion: the only difference is that, instead of the fundamental theorem of analysis employed in (1), Itô's formula is used. For any $t^* \in [0,T]$ and $t \geq t^*$, write

$$X_t = X_{t^*} + \int_{t^*}^{t} b(X_{s_1})\,\mathrm{d}s_1 + \int_{t^*}^{t} \sigma(X_{s_1})\,\mathrm{d}W_{s_1}.$$

Applying Itô's formula, we obtain

$$\begin{aligned} b(X_{s_1}) &= b(X_{t^*}) + \int_{t^*}^{s_1} \mathcal{L}b(X_{s_2})\,\mathrm{d}s_2 + \int_{t^*}^{s_1} \mathcal{N}b(X_{s_2})\,\mathrm{d}W_{s_2} \\ \sigma(X_{s_1}) &= \sigma(X_{t^*}) + \int_{t^*}^{s_1} \mathcal{L}\sigma(X_{s_2})\,\mathrm{d}s_2 + \int_{t^*}^{s_1} \mathcal{N}\sigma(X_{s_2})\,\mathrm{d}W_{s_2}, \end{aligned}$$

where the operators are defined by $\mathcal{L} = b(x)\,\partial_x + \frac{1}{2}\sigma(x)^2\,\partial_x^2$ and $\mathcal{N} = \sigma(x)\,\partial_x$. In order to write the truncation error associated with the Milstein scheme, we apply Itô's formula again to the integrands in the previous equation:

$$
\begin{aligned}
\mathcal{L}b(X_{s_2}) &= \mathcal{L}b(X_{t^*}) + \int_{t^*}^{s_2} \mathcal{L}(\mathcal{L}b)(X_{s_3})\,\mathrm{d}s_3 + \int_{t^*}^{s_2} \mathcal{N}(\mathcal{L}b)(X_{s_3})\,\mathrm{d}W_{s_3},\\
\mathcal{L}\sigma(X_{s_2}) &= \mathcal{L}\sigma(X_{t^*}) + \int_{t^*}^{s_2} \mathcal{L}(\mathcal{L}\sigma)(X_{s_3})\,\mathrm{d}s_3 + \int_{t^*}^{s_2} \mathcal{N}(\mathcal{L}\sigma)(X_{s_3})\,\mathrm{d}W_{s_3},\\
\mathcal{N}b(X_{s_2}) &= \mathcal{N}b(X_{t^*}) + \int_{t^*}^{s_2} \mathcal{L}(\mathcal{N}b)(X_{s_3})\,\mathrm{d}s_3 + \int_{t^*}^{s_2} \mathcal{N}(\mathcal{N}b)(X_{s_3})\,\mathrm{d}W_{s_3},\\
\mathcal{N}\sigma(X_{s_2}) &= \mathcal{N}\sigma(X_{t^*}) + \int_{t^*}^{s_2} \mathcal{L}(\mathcal{N}\sigma)(X_{s_3})\,\mathrm{d}s_3 + \int_{t^*}^{s_2} \mathcal{N}(\mathcal{N}\sigma)(X_{s_3})\,\mathrm{d}W_{s_3}.
\end{aligned}
$$

Collecting all terms and writing all the integrals without simplifying, we obtain

$$
\begin{aligned}
X_t = {}& X_{t^*} + b(X_{t^*}) \int_{t^*}^{t} \mathrm{d}s_1 + \sigma(X_{t^*}) \int_{t^*}^{t} \mathrm{d}W_{s_1}\\
&+ \mathcal{L}b(X_{t^*}) \int_{t^*}^{t}\int_{t^*}^{s_1} \mathrm{d}s_2\,\mathrm{d}s_1 + \mathcal{N}b(X_{t^*}) \int_{t^*}^{t}\int_{t^*}^{s_1} \mathrm{d}W_{s_2}\,\mathrm{d}s_1\\
&+ \mathcal{L}\sigma(X_{t^*}) \int_{t^*}^{t}\int_{t^*}^{s_1} \mathrm{d}s_2\,\mathrm{d}W_{s_1} + \mathcal{N}\sigma(X_{t^*}) \int_{t^*}^{t}\int_{t^*}^{s_1} \mathrm{d}W_{s_2}\,\mathrm{d}W_{s_1}\\
&+ \int_{t^*}^{t}\int_{t^*}^{s_1}\int_{t^*}^{s_2} \mathcal{L}\mathcal{L}b(X_{s_3})\,\mathrm{d}s_3\,\mathrm{d}s_2\,\mathrm{d}s_1 + \int_{t^*}^{t}\int_{t^*}^{s_1}\int_{t^*}^{s_2} \mathcal{N}\mathcal{L}b(X_{s_3})\,\mathrm{d}W_{s_3}\,\mathrm{d}s_2\,\mathrm{d}s_1\\
&+ \int_{t^*}^{t}\int_{t^*}^{s_1}\int_{t^*}^{s_2} \mathcal{L}\mathcal{N}b(X_{s_3})\,\mathrm{d}s_3\,\mathrm{d}W_{s_2}\,\mathrm{d}s_1 + \int_{t^*}^{t}\int_{t^*}^{s_1}\int_{t^*}^{s_2} \mathcal{N}\mathcal{N}b(X_{s_3})\,\mathrm{d}W_{s_3}\,\mathrm{d}W_{s_2}\,\mathrm{d}s_1\\
&+ \int_{t^*}^{t}\int_{t^*}^{s_1}\int_{t^*}^{s_2} \mathcal{L}\mathcal{L}\sigma(X_{s_3})\,\mathrm{d}s_3\,\mathrm{d}s_2\,\mathrm{d}W_{s_1} + \int_{t^*}^{t}\int_{t^*}^{s_1}\int_{t^*}^{s_2} \mathcal{N}\mathcal{L}\sigma(X_{s_3})\,\mathrm{d}W_{s_3}\,\mathrm{d}s_2\,\mathrm{d}W_{s_1}\\
&+ \int_{t^*}^{t}\int_{t^*}^{s_1}\int_{t^*}^{s_2} \mathcal{L}\mathcal{N}\sigma(X_{s_3})\,\mathrm{d}s_3\mathrm{d}W_{s_2}\,\mathrm{d}W_{s_1} + \int_{t^*}^{t}\int_{t^*}^{s_1}\int_{t^*}^{s_2} \mathcal{N}\mathcal{N}\sigma(X_{s_3})\,\mathrm{d}W_{s_3}\,\mathrm{d}W_{s_2}\,\mathrm{d}W_{s_1}.
\end{aligned}
$$

The terms in green are the ones retained for the Milstein scheme, and the other terms constitute the truncation error. By now, it should be clear that this procedure can be iterated to construct more and more accurate numerical schemes for SDEs, but observe that the number of terms increases exponentially as we increase the multiplicity of the integrals! Note also that, in contrast with the deterministic Taylor expansion, it is not always possible to simplify the multiple integrals; in high-order schemes, these have to be approximated, for example by Karhunen–Loève expansion of the Brownian motion.

Scaling of the multiple integrals. To understand how a multiple integral scales with the time step of the numerical method, and thereby decide whether or not to keep it in a numerical scheme of a given strong order, we can use the scaling property of Brownian motion: for any $c > 0$

$$V_t := \frac{1}{\sqrt{c}} W_{ct}$$

is another Brownian motion. Suppose that we would like to find the scaling with respect to the time step Δt of the following integral:

$$\int_{t_i}^{t_{i+1}}\int_{t_i}^{s_1}\int_{t_i}^{s_2} \mathrm{d}W_{s_3}\,\mathrm{d}W_{s_2}\,\mathrm{d}W_{s_1}. \tag{2}$$

Applying successively the changes of variables $u_1 = s_1 - t_i$, $u_2 = s_2 - t_i$ and $u_3 = s_3 - t_i$, we obtain

$$\begin{aligned}\int_{t_i}^{t_{i+1}}\int_{t_i}^{s_1}\int_{t_i}^{s_2} \mathrm{d}V_{s_3}\,\mathrm{d}W_{s_2}\,\mathrm{d}W_{s_1} &\overset{\text{law}}{=} \int_0^{\Delta t}\int_{t_i}^{t_i+u_1}\int_{t_i}^{s_2} \mathrm{d}W_{s_3}\,\mathrm{d}W_{s_2}\,\mathrm{d}V_{u_1} \\ &\overset{\text{law}}{=} \int_0^{\Delta t}\int_0^{u_1}\int_{t_i}^{t_i+u_2} \mathrm{d}W_{s_3}\,\mathrm{d}V_{u_2}\,\mathrm{d}V_{u_1} \\ &\overset{\text{law}}{=} \int_0^{\Delta t}\int_0^{u_1}\int_0^{u_2} \mathrm{d}V_{u_3}\,\mathrm{d}V_{u_2}\,\mathrm{d}V_{u_1},\end{aligned}$$

for another Brownian motion V. Then, employing successively the changes of variables $z_1 = u_1/\Delta t$, $z_2 = u_2/\Delta t$ and $z_3 = u_3/\Delta t$

$$\begin{aligned}\int_0^{\Delta t}\int_0^{u_1}\int_0^{u_2} \mathrm{d}V_{u_3}\,\mathrm{d}V_{u_2}\,\mathrm{d}V_{u_1} &\overset{\text{law}}{=} \int_0^{1}\int_0^{z_1\Delta t}\int_0^{u_2} \mathrm{d}W_{z_3}\,\mathrm{d}V_{u_2}\,\mathrm{d}V_{z_1\Delta t} \\ &\overset{\text{law}}{=} \int_0^{1}\int_0^{z_1}\int_0^{z_2\Delta t} \mathrm{d}V_{z_3}\,\mathrm{d}V_{z_2\Delta t}\,\mathrm{d}V_{z_1\Delta t} \\ &\overset{\text{law}}{=} \int_0^{1}\int_0^{z_1}\int_0^{z_2} \mathrm{d}V_{z_3\Delta t}\,\mathrm{d}V_{z_2\Delta t}\,\mathrm{d}V_{z_1\Delta t},\end{aligned}$$

Finally, using the scaling property of Brownian motion,

$$\int_0^{1}\int_0^{z_1}\int_0^{z_2} \mathrm{d}V_{z_3\Delta t}\,\mathrm{d}V_{z_2\Delta t}\,\mathrm{d}V_{z_1\Delta t} \overset{\text{law}}{=} (\Delta t)^{3/2} \int_0^{1}\int_0^{z_1}\int_0^{z_2} \mathrm{d}B_{z_3}\,\mathrm{d}B_{z_2}\,\mathrm{d}B_{z_1},$$

where B_t is another Brownian motion. The integral multiplying $(\Delta t)^{3/2}$ no longer depends on the time step. From this calculation we can deduce, for example, that the mean and standard deviation of the triple integral (2) scale as $\mathcal{O}((\Delta t)^{3/2})$. In general, this approach can be employed to show that a multiple integral comprising n_t time integrals and n_w Itô integrals scales as $\Delta t^{n_t + \frac{n_w}{2}}$.

4.5. Discretisation error

The discretisation schemes we described are not exact. The distribution of the random vector $(X_0, X_{t_1}, \dots, X_{t_N})$ is not the same as that of $(X_0, X_{\Delta t}, \dots, X_{N\Delta t})$, although we expect that the difference vanishes (in some sense) as $\Delta t \to 0$. Since both the exact approximation $(X_0, X_{t_1}, \dots, X_{t_N})$ and the numerical approximation $(X_0, X_{\Delta t}, \dots, X_{N\Delta t})$ are random variables, we need to define what we mean by quantifying the error.

The two most common and useful concepts of error for approximations of SDEs are

- Strong convergence (mean of the error, related to the SDE paths)
- Weak convergence (error of the mean, related to statistics of the solution).

4.5.1. Strong convergence

Let $X_t = (X_0, X_{t_1}, \dots, X_{t_N})$ be the solution to the SDE (*) and $\hat{X}_n = (X_0, X_{\Delta t}, \dots, X_{N\Delta t})$ be a numerical approximation, using the same Brownian motion as X_t does.

The strong error of the approximation $\hat{X}_n$ at time $N\Delta t$ is given by

$$\boxed{e_{strong} = \mathbb{E}\,|X_{N\Delta t} - \hat{X}_{N\Delta t}|}$$

for N sufficiently large. For a fixed realisation of a BM, $|X_{N\Delta t} - \hat{X}_{N\Delta t}|$ measures the distance between the two solutions after N time steps. The strong error averages this distance over all realisations of BM. (Hence mean of the error!)

We say that a numerical approximation $\hat{X}_n$ has strong error of order r if, for all $N>0$ there exists $\delta = \delta(N)$ and $K = K(N,\delta)$ such that for $\Delta t \le \delta$

$$\boxed{\mathbb{E}\,|\hat{X}_{n\Delta t} - X_{n\Delta t}| \le K(\Delta t)^r, \quad \forall n \le N.}$$

Let $\hat{X}_n^{EM}$ be the Euler-Maruyama approximation of X_t and $\hat{X}_n^{Mil}$ be the Milstein approximation. It is possible to prove (using Itô's Lemma, Itô isometry and Gronwall's inequality) that given $N>0$, $\exists K>0$:

$$e_{strong}^{EM} = \mathbb{E}\,|\hat{X}^{EM}_{n\Delta t} - X_{n\Delta t}| \le K\sqrt{\Delta t}, \quad \forall n \le N,$$

for all sufficiently small Δt. Therefore, the EM approximation has strong error of order $1/2$. This is under appropriate conditions on the coefficients b and σ. Similarly, for the Milstein scheme

$$e_{strong}^{Mil} = \mathbb{E}\,|\hat{X}^{Mil}_{n\Delta t} - X_{n\Delta t}| \le K\Delta t \quad \forall n \le N$$

and the Milstein scheme has strong error of order 1.

Note that:

- If $\sigma(X_t) = \sigma$, constant, then the two methods coincide $\Rightarrow$ EM has strong order of convergence 1.
- Strong convergence requires the paths to be close at all times.
- the processes need to be driven by the same Brownian Motion to compute strong error.
- even though strong convergence involves an expected value, it has implications for individual simulations: From Markov's inequality, we have $\mathbb{P}(|X| > a) \le \mathbb{E}|X|/a$ $\Rightarrow$ by taking $a = \Delta t^{1/4}$ we see that since EM has strong order of convergence $1/2$, we have $\mathbb{P}(|X_{n\Delta t} - \hat{X}_{n\Delta t}| > \Delta t^{1/4}) \le C\Delta t^{1/4}$ $\Rightarrow$ error is small with probability close to 1.

4.5.2. Weak error

In certain applications such as filtering and statistical inference, strong convergence is needed. However, quite often we are only interested in the calculation of statistical quantities of the solutions, such as moments or their density, rather than the accuracy of the paths. This motivates the concept of weak error.

We say that an approximation $\hat{X}_n$ has weak order of convergence r if for all $N \in \mathbb{N}$ there exists $K > 0$ such that

$$\boxed{e_{weak} = |\mathbb{E} f(\hat{X}_{n\Delta t}) - \mathbb{E} f(X_{n\Delta t})| \leq K \Delta t^r, \quad \forall n \leq N}$$

for Δt sufficiently small and for all f in some class of functions, usually $C_p^l(\mathbb{R})$ — l times continuously differentiable functions which have at most polynomial growth.

Under appropriate assumptions in $b(t, X_t)$, $\sigma(t, X_t)$ and the class of functions, it can be shown that

$$e_{weak}^{EM} = \sup_{f \in \mathcal{F}} |\mathbb{E} f(\hat{X}^{EM}_{n\Delta t}) - \mathbb{E} f(X_{n\Delta t})| \leq K \Delta t$$

(the E-M method has weak order of convergence 1)

and

$$e_{weak}^{Mil} = \sup_{f \in \mathcal{F}} |\mathbb{E} f(\hat{X}^{Mil}_{n\Delta t}) - \mathbb{E} f(X_{n\Delta t})| \leq K \Delta t$$

(the Milstein scheme has weak order of convergence 1)

Note that:

- We can prove that EM and Milstein have weak order of convergence 1 using the Feynman-Kac formula, which relates SDEs to PDEs for the functions we need to compute
- Weak convergence only requires the probability distributions to converge. Therefore, one can use different Brownian Motions for each approximation, or even a random process which is not BM but has increments with the same mean and variance
- Strong convergence $\Rightarrow$ Weak convergence, but in general weak order of convergence $\geq$ strong order of convergence

We can define the weak Euler-Maruyama method by replacing the Brownian Motion increments by V_j, where $V_j = 1$ or -1 with probability $1/2$. $\sqrt{\Delta t}\, V_j$ has the same mean and variance as ΔW_j.
$\rightarrow$ This offers no strong convergence but is more efficient than EM, if we are only interested in weak convergence.

Note that: Clark and Cameron (1980) showed that the Milstein Scheme has the optimal convergence rate (in L^2) among all discretisation schemes that have the same computational cost, i.e., which only use r.v. equal to ΔW.
See also Rümelin 1982.

It is possible to derive higher order schemes using a Runge-Kutta approach, but this requires generating more r.v.s or evaluating many derivatives of b and σ.

! It is important to note that, when evaluating the error numerically, we are implicitly assuming that a number of other sources of error are negligible. For example, common sources of error are

- Sampling error → arises from approximating an expected value by a sample mean → it is the same error as we had in MC approx.
- Random number bias → comes from inherent errors in the RNG
- Rounding error → roundoff errors.

The most significant of the three is the sampling error. We can see that, for example, for weak error,

$$e_{weak} = \left| \mathbb{E} f(X_T) - \frac{1}{M} \sum_{m=1}^{M} f(\hat{X}^m_{N\Delta t}) \right| \leq$$

$$\leq \left| \mathbb{E} f(X_T) - \mathbb{E} f(\hat{X}_T) \right| + \left| \mathbb{E} f(\hat{X}_T) - \frac{1}{M} \sum f(\hat{X}^m_{N\Delta t}) \right|$$

$$\leq \underbrace{C_1 (\Delta t)^r}_{\text{weak error}} + \underbrace{C_2 M^{-1/2}}_{\text{MC error.}}$$

So we need to take this into account (optimise over N and M) when solving the SDEs.

4.6. Stability analysis of numerical schemes

The concepts of strong and weak convergence concern the accuracy of a numerical method over an interval. However, it is important that if the solution of an SDE remains bounded for all time, then so does the numerical approximation. This is the concept of stability of the numerical scheme.

Weak and strong convergence do not usually imply stability since the constant K depends on T (and usually increases with it).

A good illustration of when things can go wrong with stability is the geometric Brownian motion, or its deterministic counterpart:

$$\frac{dX}{dt} = \lambda X \longrightarrow \text{it is stable } \left(\lim_{t\to\infty} X(t) = 0\right) \text{ if and only if } \lambda < 0 \quad (\text{or } \mathrm{Re}(\lambda) < 0).$$

For random variables, we need to define what we mean by "$\lim_{t\to\infty} X(t) = 0$"

Definition: A stochastic process X_t is said to be mean-square stable if

$$\lim_{t\to\infty} \mathbb{E}\,|X_t|^2 = 0.$$

and asymptotically stable if $\forall\, X_0 \neq 0$,

$$\mathbb{P}\left(\lim_{t\to\infty} |X_t| = 0\right) = 1.$$

We will illustrate these concepts with the geometric Brownian motion.

Consider the gBM

$$dX_t = \lambda X_t\, dt + \sigma X_t\, dW_t, \qquad X_0 = x.$$

Its solution is given by $X_t = X_0\, e^{(\lambda - \sigma^2/2)t + \sigma W_t}$, from where we can conclude that

$$\mathbb{E}\left(|X_t|^2\right) = X_0^2 \exp\left[(2\lambda + \sigma^2)t\right]$$

Clearly, X_t is mean-square stable if and only if

$$\boxed{2\lambda + \sigma^2 < 0}$$

On the other hand, it is also easy to check that X_t is asymptotically stable if

$$\boxed{\lambda - \sigma^2/2 < 0.}$$

The main question now is whether these conditions are sufficient for the EM and Milstein schemes to be mean-square or asymptotically stable!

Consider the Euler-Maruyama approximation:

$$\begin{aligned} \hat{X}_{n+1} &= \hat{X}_n + \lambda \hat{X}_n \Delta t + \sigma \hat{X}_n \Delta W_n \\ &= (1 + \lambda \Delta t + \sigma \Delta W_n)\, \hat{X}_n \end{aligned}$$

Doing this iteratively, we can write

$$\boxed{\hat{X}_{n+1} = \prod_{j=0}^{n} (1 + \lambda \Delta t + \sigma \Delta W_j)\, X_0}$$

If X_0 is deterministic, we can compute the second moment of $\hat{X}_{n+1}$ as

$$\mathbb{E}\left(\hat{X}_{n+1}^2\right) = \prod_{j=0}^{n} \mathbb{E}(1 + \lambda \Delta t + \sigma \Delta W_j)^2\, X_0.$$

Now $\mathbb{E}\left((1 + \lambda\Delta t + \sigma \Delta W_j)^2\right) = 1 + 2\lambda\Delta t + \lambda^2 \Delta t^2 + \sigma^2 \Delta t$

$\mathbb{E}\Delta W_j = 0$, $\mathbb{E}\Delta W_j^2 = \Delta t$. $\qquad = 1 + \Delta t\,(2\lambda + \sigma^2 + \lambda^2 \Delta t)$

So, in order for $\mathbb{E}(X_{n+1}^2)$ to converge to zero, we need that $\mathbb{E}\left((1+2\lambda\Delta t+\sigma\Delta W_j)^2\right)<1$, or equivalently, $2\lambda+\sigma^2+\lambda^2\Delta t<0$. This implies that we need to choose Δt such that

$$\boxed{0<\Delta t<\frac{-2(\lambda+\sigma^2/2)}{\lambda^2}}$$

Note that this condition is more restrictive than $2\lambda+\sigma^2<0$ for the true solution! So EM is <u>not</u> always stable, even if the solution is!!
This is not completely surprising, as the same happens for the explicit Euler scheme for ODEs!

<u>Note that</u>: Using the strong LLN and the law of the iterated logarithm, one can prove that the EM scheme is asymptotically stable if and only if

$$\mathbb{E}\left(\log\left|1+\Delta t\lambda+\sqrt{\Delta t}\,\sigma\mathcal{N}(0,1)\right|\right)<0.$$

exercise: Repeat the above argument to identify the stability region of the Milstein scheme ⇒ identify conditions for which the scheme is mean-square stable.

<u>Note that</u>: The reason why EM (and Milstein!) are not always stable when the true solution is is because <u>they are explicit!</u> So as in the ODEs case, this motivates us to consider <u>implicit schemes</u>. These usually have better stability properties, but their implementation requires the solution of an additional algebraic equation at each time step ⟶ uses Newton-Raphson algorithm.

Treating the diffusion coefficient implicitly involves computing reciprocals of Gaussian r.v.s, which do not have moments ($\int e^{x^2}dx$ is not finite!). So we treat the drift implicitly and diffusion explicitly.

The θ-Euler-Maruyama Method (θEM)

Consider a time step $\Delta t > 0$ and $X_0 \in \mathbb{R}$. The θ-Euler-Maruyama approximation X_n of $X(\Delta t n)$ is given by

$$\boxed{X_{n+1} = X_n + \left((1-\theta) b(X_n) + \theta\, b(X_{n+1})\right) dt + \sigma(X_n) \Delta W_n}$$

where $\theta \in [0,1]$ is a parameter which controls the "degree of implicitness" in the drift term.

Note that:

- $\theta = 0$ reduces to the EM scheme. $\theta = 1$ corresponds to a fully implicit EM scheme
- We can define a θ-Milstein scheme similarly:

$$\boxed{X_{n+1} = X_n + \left((1-\theta) b(X_n) + \theta\, b(X_{n+1})\right) dt + \sigma(X_n) \Delta W_n + \frac{\sigma \sigma'}{2} (X_n)\left((\Delta W_n)^2 - \Delta t\right).}$$

We can repeat the steps we did before for EM and obtain (for gBM)

$$X_{n+1} = \prod_{j=0}^{n} \frac{((1-\theta)\lambda + 1)\Delta t + \sigma \Delta W_n}{1 - \theta \lambda \Delta t} X_0$$

from where we can conclude that the θ-EM scheme is mean-square stable for the gBM when

$$\boxed{2\lambda + \sigma^2 + \Delta t (1 - 2\theta) \lambda^2 < 0.}$$

In particular, if $\theta = 1/2$, this reduces to $2\lambda + \sigma^2 < 0 \Rightarrow$ it is mean-square stable whenever the gBM is! Independently of Δt!

Furthermore, if $\theta < 1/2$ the region of stability of θEM is contained in the region of stability of SDE. However, if $\theta > 1/2$, the region of stability of θ-EM contains that of the SDE!

Exercise 1. Remembering that

$$B_t := \begin{cases} 0 & \text{if } t = 0, \\ t\, W_{1/t} & \text{if } t > 0 \end{cases}$$

defines a Brownian motion and that, by definition, Brownian motions have almost surely continuous paths, show that the geometric Brownian motion solving

$$\mathrm{d}X_t = \lambda\, X_t\, \mathrm{d}t + \sigma\, X_t\, \mathrm{d}W_t, \qquad X_0 = x_0, \qquad \mathbb{P}(x_0 = 0) = 0,$$

is asymptotically stable if and only if

$$\lambda - \frac{\sigma^2}{2} < 0.$$

Proof. The exact solution is given by

$$X_t = x_0 \exp\left(\left(\lambda - \frac{\sigma^2}{2}\right) t + \sigma W_t\right).$$

Employing the change of variable $s = 1/t$, we observe that, almost surely,

$$\lim_{t\to\infty} \left(\lambda - \frac{\sigma^2}{2} + \sigma \frac{W_t}{t}\right) = \lim_{s\to 0} \left(\lambda - \frac{\sigma^2}{2} + \sigma\, s W_{1/s}\right) = \lim_{s\to 0} \left(\lambda - \frac{\sigma^2}{2} + \sigma\, B_s\right) = \lambda - \frac{\sigma^2}{2}.$$

We conclude that, almost surely,

$$\lim_{t\to\infty} |X_t| = \lim_{t\to\infty} |x_0| \exp\left(t\left(\lambda - \frac{\sigma^2}{2} + \sigma \frac{W_t}{t}\right)\right) = \begin{cases} 0 & \text{if } \lambda - \frac{\sigma^2}{2} < 0, \\ \infty & \text{if } \lambda - \frac{\sigma^2}{2} > 0. \end{cases}$$

On the other hand, when $\lambda - \sigma^2/2 = 0$,

$$\mathbb{P}(|X_t| > |x_0|) = \mathbb{P}\left(\mathrm{e}^{\sigma W_t} > 1\right) = \mathbb{P}(W_t > 0) = \frac{1}{2},$$

and so it is clear that $\mathbb{P}(\lim_{t\to\infty} X_t = 0) \leq 1/2$. □

Exercise 2. Show that the Euler–Maruyama approximation with time step Δt of the geometric Brownian motion is asymptotically stable if

$$E := \mathbb{E}\left[\log |1 + \lambda\, \Delta t + \sigma \sqrt{\Delta t}\, \xi|\right] < 0, \qquad \xi \sim \mathcal{N}(0,1).$$

Optionally (not examinable), use the law of iterated logarithm to show that this condition is also necessary.

Proof. The Euler–Maruyama approximation satisfies

$$\hat{X}_n = \hat{X}_0 \prod_{i=0}^{n-1} (1 + \lambda\, \Delta t + \sigma\, \Delta W_i),$$

Taking absolute values and the logarithm,

$$\log |\hat{X}_n| = \log |\hat{X}_0| + \sum_{i=0}^{n-1} \log |1 + \lambda\, \Delta t + \sigma\, \Delta W_i|,$$

Since $\log|1+\lambda\,\Delta t+\sigma\,\Delta W_i| \in L^1(\Omega)$ (because $x \mapsto \log x$ is in $L^1(\mathbb{R}_{>0})$), the strong law of large numbers implies that, almost surely,

$$\lim_{n\to\infty}\left[\frac{1}{n}\sum_{i=0}^{n-1}\log|1+\lambda\,\Delta t+\sigma\,\Delta W_i|\right] = \mathbb{E}[\log|1+\lambda\,\Delta t+\sigma\,\Delta W_i|].$$

We conclude that, if $E<0$, then

$$\lim_{n\to\infty}\left[\log|\hat{X}_n|\right] = \log|\hat{X}_0| + \lim_{n\to\infty} n\left(\frac{1}{n}\sum_{i=0}^{n-1}\log|1+\lambda\,\Delta t+\sigma\,\Delta W_i|\right) = -\infty$$

almost surely, and so $\lim_{n\to\infty}|\hat{X}_n| = 0$ almost surely. Similarly, $\lim_{n\to\infty}|\hat{X}_n| = \infty$ almost surely if $E>0$. To complete the optional part of the exercise, it remains to examine the case $E=0$. In this case, introducing the variance $s^2 = \operatorname{var}\left[\log|1+\lambda\,\Delta t+\sigma\sqrt{\Delta t}\,\xi|\right]$ and using the law of iterated logarithm, we obtain

$$\lim_{n\to\infty}\left[\log|\hat{X}_n|\right] = \log|\hat{X}_0| + \lim_{n\to\infty}\left[s\sqrt{2n\log\log n}\underbrace{\left(\frac{\frac{1}{s}\sum_{i=0}^{n-1}\log|1+\lambda\,\Delta t+\sigma\,\Delta W_i|}{\sqrt{2n\log\log n}}\right)}_{\to 1\text{ a.s.}}\right] = \infty,$$

almost surely. $\square$