Probabilistic Numerical Methods 2024–2025

Lecture 7: SDEs and PDEs

julien.reygner@enpc.fr

We still work with $I = [0, T]$ or $I = [0, +\infty)$ and $b : I \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : I \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$ be measurable functions which are bounded on bounded subsets of $I \times \mathbb{R}^n$. We consider the system

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t, \tag{SDE}$$

complemented with the initial condition

$$X_0 = \xi. \tag{IC}$$

The associated differential operator is

$$L_t\phi(x) = \sum_{i=1}^n b_i(t, x)\frac{\partial \phi}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^n a_{ij}(t, x)\frac{\partial^2 \phi}{\partial x_i \partial x_j}(x),$$

with $a(t, x) := \sigma(t, x)\sigma^\top(t, x)$. It sometimes rewrites

$$L_t\phi(x) = b(t, x) \cdot \nabla\phi(x) + \frac{1}{2}a(t, x) : \nabla^2\phi(x),$$

where $A : B := \mathrm{tr}(A^\top B)$.

1 The Fokker–Planck equation

Let $(X_t)_{t \in I}$ be a solution to (SDE)–(IC). Let $\mu(t, \mathrm{d}y)$ denote the law of X_t, and let $\mu_0(\mathrm{d}y)$ be the law of ξ. Since the trajectories of $(X_t)_{t \geq 0}$ are continuous almost surely, $t \mapsto \mu(t, \mathrm{d}y)$ is continuous for the weak convergence topology on the space of probability measures on $\mathbb{R}^n$.

Proposition 1.1 (Fokker–Planck equation)**.** *For any $C_c^{1,2}(\mathbb{R}^n)$ function $\Phi : I \times \mathbb{R}^n \to \mathbb{R}$, for any $t \in I$,*

$$\int_{y \in \mathbb{R}^n} \Phi(t, y)\mu(t, \mathrm{d}y) = \int_{y \in \mathbb{R}^n} \Phi(0, y)\mu_0(\mathrm{d}y) + \int_{s=0}^t \int_{y \in \mathbb{R}^n} (\partial_t\Phi(s, y) + L_s\Phi(y))\, \mu(s, \mathrm{d}y)\mathrm{d}s.$$

Proof. Apply Ito's formula to $\Phi(s, X_s)$. □

This statement implies that $\mu(t, \mathrm{d}y)$ is a distributional solution to

$$\begin{cases} \partial_t\mu = L_t^*\mu, \\ \mu(0, \mathrm{d}y) = \mu_0(\mathrm{d}y), \end{cases}$$

which is usually called the Fokker–Planck equation. When b and a are smooth enough and $\mu(t, \mathrm{d}y)$ has a smooth enough density $p(t, y)$ with respect to the Lebesgue measure on $\mathbb{R}^n$, the latter satisfies

$$\partial_t p(t, y) = -\sum_{i=1}^n \partial_{x_i}(b_i(t, y)p(t, y)) + \frac{1}{2}\sum_{i,j=1}^n \partial_{x_i x_j}(a_{ij}(t, y)p(t, y))$$

in the classical sense.

Example: Brownian motion and heat equation. In particular in can be checked directly that $p(t, y) = (2\pi t)^{-d/2}\exp(-y^2/2t)$ satisfies

$$\begin{cases} \partial_t p = \frac{1}{2}\Delta p, \\ \lim_{t \downarrow 0} p = \delta_0, \quad \text{weakly.} \end{cases}$$

2 Feynman–Kac formulæ

2.1 Backward Cauchy problems

Here we take $I = [0, T]$ and assume that, for all $x \in \mathbb{R}^n$ and $t \in [0, T)$, there exists an Itô process $(X_s^{t,x})_{s\in[t,T]}$ such that, for all $s \in [0, T]$,

$$X_s^{t,x} = x + \int_{r=t}^{s} b(r, X_r^{t,x})\mathrm{d}r + \int_{r=t}^{s} \sigma(r, X_r^{t,x})\mathrm{d}B_r,$$

that is to say, a solution to (SDE) on $[t, T]$ which takes the value x at time t. This is in particular the case if the coefficients b and σ satisfy the assumptions of Ito's Theorem.

Theorem 2.1 (Feynman–Kac formula for children)**.** *Let $T > 0$ and $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. Assume that there exists a $C^{1,2}$ function $u : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ such that:*

(i) for any $t \in [0, T)$ and $x \in \mathbb{R}^n$, $(|\sigma^\top(s, X_s^{t,x})\nabla_x u(s, X_s^{t,x})|)_{s\in[t,T]} \in \mathbf{\Lambda}^2([t, T])$;

(ii) u solves the parabolic problem

$$\begin{cases} -\dfrac{\partial u}{\partial t}(t, x) = L_t u(t, x), & t \in [0, T], \quad x \in \mathbb{R}^n, \\ u(T, x) = f(x). \end{cases} \tag{1}$$

Then, for all $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$u(t, x) = \mathbb{E}\left[f(X_T^{t,x})\right].$$

Proof. For simplicity we write $X_s^{t,x} = X_s = (X_s^1, \dots, X_s^n)$. Let us fix $t \in [0, T]$ and apply Itô's formula to $u(s, X_s)$ for $s \in [t, T]$. We get

$$\begin{aligned} \mathrm{d}u(s, X_s) &= \left(\frac{\partial u}{\partial t}(s, X_s) + L_s u(s, X_s)\right)\mathrm{d}s + \sigma^\top(s, X_s)\nabla_x u(s, X_s) \cdot \mathrm{d}B_s \\ &= \sigma^\top(s, X_s)\nabla_x u(s, X_s) \cdot \mathrm{d}B_s, \end{aligned}$$

thanks to (ii). As a consequence,

$$u(T, X_T) = u(t, X_t) + \int_{s=t}^{T} \sigma^\top(s, X_s)\nabla_x u(s, X_s) \cdot \mathrm{d}B_s,$$

which rewrites

$$f(X_T) = u(t, x) + \int_{s=t}^{T} \sigma^\top(s, X_s)\nabla_x u(s, X_s) \cdot \mathrm{d}B_s.$$

The assumption (i) now ensures that

$$\mathbb{E}\left[\int_{s=t}^{T} \sigma^\top(s, X_s)\nabla_x u(s, X_s) \cdot \mathrm{d}B_s\right] = 0,$$

therefore

$$\mathbb{E}\left[f(X_T)\right] = u(t, x).$$

□

The Feynman–Kac formula shows that if one is interested in solving the PDE (1) in one point (t, x), a possible approach may be to simulate the trajectory of $(X_s^{t,x})_{s\in[t,T]}$ and then to compute the expectation $\mathbb{E}[f(X_T^{t,x})]$ by the Monte Carlo method, using discretisation seen last week.

Example 2.2 (The Black–Scholes model in mathematical finance)**.** *In mathematical finance, the* Black–Scholes model*assumes that the price of some asset (for instance, an action) is the solution* $(S_t)_{t\geq 0}$ *of the SDE*

$$\mathrm{d}S_t = \sigma S_t \mathrm{d}B_t,$$

whose solution writes $S_t = S_0 \exp(\sigma B_t - \sigma^2 t/2)$*. An* option *with* payoff *function* f *and* maturity T *is a contract between the bank and the client, where at time* T *the bank has to give the client the quantity* $f(S_T)$*. The price that the client has to pay to the bank at time* $t \leq T$ *in order to buy the option is given by* $u(t,s) = \mathbb{E}[f(S_T^{t,s})]$*, where* s *is the value of* S_t*. This quantity can be computed either by the Monte Carlo method, or by solving the parabolic problem*

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,s) + \dfrac{\sigma^2 s^2}{2}\dfrac{\partial^2 u}{\partial s^2}(t,s) = 0, & t \in [0,T), \quad s \geq 0, \\ u(T,s) = f(s). \end{cases}$$

🏠 **Exercise 2.3** (Feynman–Kac formula for grown-ups)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *and* $k, g : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ *be continuous functions, with* k *bounded from below. Assume that there exists a* $C^{1,2}$ *function* $u : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ *such that:*

(i) for any $t \in [0,T)$ *and* $x \in \mathbb{R}^n$*,* $(|\sigma^\top(s, X_s^{t,x}) \nabla_x u(s, X_s^{t,x})|)_{s\in[t,T]} \in \mathbf{\Lambda}^2([t,T])$*;*

(ii) u *solves the parabolic problem*

$$\begin{cases} -\dfrac{\partial u}{\partial t}(t,x) = L_t u(t,x) - k(t,x)u(t,x) + g(t,x), & t \in [0,T], \quad x \in \mathbb{R}^n, \\ \quad u(T,x) = f(x). \end{cases} \tag{2}$$

Show that for all $(t,x) \in [0,T] \times \mathbb{R}^n$*,*

$$u(t,x) = \mathbb{E}\left[f(X_T^{t,x}) \mathrm{e}^{-\int_{r=t}^T k(r, X_r^{t,x})\mathrm{d}r} + \int_{s=t}^T g(s, X_s^{t,x}) \mathrm{e}^{-\int_{r=t}^s k(r, X_r^{t,x})\mathrm{d}r} \mathrm{d}s \right].$$

Hint: start by applying Itô's formula to $u(s, X_s^{t,x})\mathrm{e}^{-\int_{r=t}^s k(r, X_r^{t,x})\mathrm{d}r}$.

The Feynman–Kac formula (in the form of Exercise 2.3) provides a probabilistic representation of a solution to (2) which satisfies the integrability condition that

$$\forall (t,x) \in [0,T] \times \mathbb{R}^n, \qquad \mathbb{E}\left[\int_{s=t}^T |\sigma^\top(s, X_s^{t,x}) \nabla_x u(s, X_s^{t,x})|^2 \mathrm{d}s \right] < +\infty. \tag{3}$$

It is therefore a *uniqueness* result for the PDE (2) in the class of solutions which satisfy (3). Thus, its application usually requires to find an *existence* result for a smooth solution to this Cauchy problem, in a PDE textbook. Se for example Section 6.5 in Friedman 1975.

2.2 Problems with boundaries

In this subsection, we assume that the coefficients b and σ do not depend on t and we denote by L the associated differential operator. We let D be an open and regular subset of $\mathbb{R}^n$ and, for any solution $(X_t^x)_{t\geq 0}$ of (SDE) with initial condition $x \in D$, we define the stopping time

$$\tau^x := \inf\{t \geq 0 : X_t^x \notin D\}.$$

Proposition 2.4 (Probabilistic interpretation of Dirichlet problem)**.** *Let* $f : \partial D \to \mathbb{R}$ *and* $k, g : \overline{D} \to \mathbb{R}^n$ *be continuous functions, with* $k \geq 0$*. Assume that there exists a* C^2 *function* $v : [0,T] \times \overline{D} \to \mathbb{R}$ *such that:*

(i) v *is bounded;*
(ii) *for any* $t > 0$ *and* $x \in D$, $(|\sigma^\top(X_s^x)\nabla v(X_s^x)|)_{s\in[0,t]} \in \mathbf{\Lambda}^2([0,t])$*;*
(iii) v *solves the elliptic problem*

$$\begin{cases} Lv(x) - k(x)v(x) = -g(x), & x \in D, \\ v(x) = f(x), & x \in \partial D. \end{cases} \tag{4}$$

Assume moreover that for any $x \in D$*:*
(iv) *the associated stopping time* τ^x *is finite, almost surely;*
(v) τ^x *and* g *satisfy*

$$\int_{t=0}^{+\infty} \mathbb{E}\left[\mathbb{1}_{\{t<\tau^x\}}|g(X_t)|\right]\mathrm{d}t < +\infty.$$

Then for all $x \in D$*,*

$$v(x) = \mathbb{E}\left[f(X_{\tau^x}^x)\mathrm{e}^{-\int_{r=0}^{\tau} k(X_r^x)\mathrm{d}r} + \int_{s=0}^{\tau^x} g(X_s^x)\mathrm{e}^{-\int_{r=0}^{s} k(X_r^x)\mathrm{d}r}\mathrm{d}s\right].$$

Proof. Let us fix $x \in D$ and write $X_t = X_t^x$, $\tau = \tau^x$. Itô's formula applied to $v(X_t)\mathrm{e}^{-\int_{r=0}^{t} k(X_r)\mathrm{d}r}$ yields

$$\begin{aligned} v(X_{t\wedge\tau})\mathrm{e}^{-\int_{r=0}^{t\wedge\tau} k(X_r)\mathrm{d}r} &= v(x) + \int_{s=0}^{t\wedge\tau} \mathrm{e}^{-\int_{r=0}^{s} k(X_r)\mathrm{d}r}\left(Lv(X_s) - k(X_s)v(X_s)\right)\mathrm{d}s \\ &\quad + \int_{s=0}^{t\wedge\tau} \mathrm{e}^{-\int_{r=0}^{s} k(X_r)\mathrm{d}r}\sigma^\top(X_s)\nabla v(X_s)\cdot\mathrm{d}B_s, \end{aligned}$$

so that

$$v(x) = \mathbb{E}\left[v(X_{t\wedge\tau})\mathrm{e}^{-\int_{r=0}^{t\wedge\tau} k(X_r)\mathrm{d}r} + \int_{s=0}^{t\wedge\tau} g(X_s)\mathrm{e}^{-\int_{r=0}^{s} k(X_r)\mathrm{d}r}\mathrm{d}s\right].$$

First, since $\tau < +\infty$ almost surely, v is bounded and $k \geq 0$, by the Dominated Convergence Theorem one has

$$\lim_{t\to+\infty}\mathbb{E}\left[v(X_{t\wedge\tau})\mathrm{e}^{-\int_{r=0}^{t\wedge\tau} k(X_r)\mathrm{d}r}\right] = \mathbb{E}\left[v(X_\tau)\mathrm{e}^{-\int_{r=0}^{\tau} k(X_r)\mathrm{d}r}\right] = \mathbb{E}\left[f(X_\tau)\mathrm{e}^{-\int_{r=0}^{\tau} k(X_r)\mathrm{d}r}\right].$$

Second, the integrability condition on g and τ allows to use the Dominated Convergence Theorem again to get

$$\lim_{t\to+\infty}\mathbb{E}\left[\int_{s=0}^{t\wedge\tau} g(X_s)\mathrm{e}^{-\int_{r=0}^{s} k(X_r)\mathrm{d}r}\mathrm{d}s\right] = \mathbb{E}\left[\int_{s=0}^{\tau} g(X_s)\mathrm{e}^{-\int_{r=0}^{s} k(X_r)\mathrm{d}r}\mathrm{d}s\right],$$

which completes the proof. □

Examples: find the PDE to solve in order to compte $\mathbb{E}[\tau^x]$ or $\mathbb{E}[\mathrm{e}^{-\lambda\tau^x}]$ for some $\lambda > 0$.

Example 2.5 (The committor function)**.** *Let* $(X_t^x)_{t\geq 0}$ *be the solution to the SDE* (SDE) *with coefficients* b *and* σ *which do not depend on* t*, and with deterministic initial condition* $x \in \mathbb{R}^n$*. Given two disjoint closed subsets* $A, B \subset \mathbb{R}^n$*, set*

$$\tau_A^x := \inf\{t \geq 0 : X_t^x \in A\}, \qquad \tau_B^x := \inf\{t \geq 0 : X_t^x \in B\},$$

and define

$$v(x) = \mathbb{P}(\tau_A^x < \tau_B^x).$$

In molecular dynamics, this function is called the committor *function*[1]*: in this context, X_t^x must be thought of as describing the* microscopic *state of a molecular system, and A and B describe particular* macroscopic *configurations. For example, in a protein-ligand system, X_t^x encodes the complete geometry of the protein-ligand, while A and B contain the states which correspond to the system being bound or unbound, respectively. Computing the committor function then allows to determine whether, given an initial state x, it is more likely that the system evolves toward one or the other configuration. Proposition 2.4 shows that under regularity assumptions and if $\tau_A^x \wedge \tau_B^x < \infty$, almost surely, then u solves the PDE*

$$\begin{cases} Lv(x) = 0, & x \in \mathbb{R}^n \setminus (A \cup B), \\ v(x) = 1, & x \in A, \\ v(x) = 0, & x \in B. \end{cases}$$

Remark 2.6. *The assumption that $k \geq 0$ is crucial in the statement of Proposition 2.4. Indeed, consider the case where $n = 1$, $D = (0,1)$ and $\mathrm{d}X_t = \mathrm{d}B_t$ so that $L = \frac{1}{2}\frac{\partial^2}{\partial x^2}$. It can be directly checked that for any $m \geq 1$, $v_m(x) = \sin(\pi m x)$ satisfies (4) with $k = -\frac{1}{2}(\pi m)^2 < 0$ and $f = g = 0$, so that applying the result of Proposition 2.4 would yield $v_m = 0$ on D.*

3 Solutions to SDEs as Markov processes

For simplicity we assume here that b and σ do not depend on time. We set $I = [0, +\infty)$ and assume that for any choice of $(\Omega, \mathcal{A}, \mathbb{P})$, $(\mathcal{F}_t)_{t\geq 0}$, $(B_t)_{t\geq 0}$, ξ, the system (SDE)–(IC) has a unique solution. Then there is a deterministic and measurable function $\mathcal{S} : \mathbb{R}^n \times C([0,+\infty), \mathbb{R}^d) \to C([0,+\infty), \mathbb{R}^n)$ such that $(X_t)_{t\geq 0} = \mathcal{S}(\xi, (B_t)_{t\geq 0})$, almost surely, and this function does not depend on the choice of $(\Omega, \mathcal{A}, \mathbb{P})$, $(\mathcal{F}_t)_{t\geq 0}$, $(B_t)_{t\geq 0}$, ξ: in particular, the law of $(X_t)_{t\geq 0}$ only depends on b, σ and the law of ξ.

In this section we use the notation $\mathbb{E}_x$ when we consider $X_0 = x$.

3.1 Markov property and semigroup

For any measurable and bounded function $f : \mathbb{R}^n \to \mathbb{R}$, we define $P_t f(x) = \mathbb{E}_x[X_t]$.

Proposition 3.1 (Markov property)**.** *For any $s, t \geq 0$, $\mathbb{E}[f(X_{t+s})|(X_r)_{r\leq t}] = \mathbb{E}[f(X_{t+s})|X_t] = P_s f(X_t)$, almost surely.*

Proof. The process $\overline{X}_s = X_{t+s}$ is a solution to the SDE with coefficients b and σ with initial condition $\overline{\xi} = X_t$, driven by the Brownian motion $\overline{B}_s = B_{t+s} - B_t$. So it writes $(\overline{X}_s)_{s\geq 0} = \mathcal{S}(\overline{\xi}, \overline{B})$, that is to say

$$(X_{t+s})_{s\geq 0} = \mathcal{S}(X_t, \overline{B}).$$

This shows that[2]

$$\mathbb{E}[f(X_{t+s})|(X_r)_{r\leq t}] = \mathbb{E}[f(X_{t+s})|X_t] = \mathbb{E}[f(\mathcal{S}_s(x, \overline{B}))]_{|x=X_t}.$$

But since $\overline{B}$ and B have the same law, we have

$$\mathbb{E}[f(\mathcal{S}_s(x, \overline{B}))] = \mathbb{E}[f(\mathcal{S}_s(x, B))] = \mathbb{E}_x[f(X_s)] = P_s f(x),$$

which concludes. □

[1] In potential theory, it is the *equilibrium potential*.

[2] You can check that if X is $\mathcal{F}$-measurable and Z is independent from $\mathcal{F}$, $\mathbb{E}[F(X,Z)|\mathcal{F}] = \mathbb{E}[F(X,Z)|X] = \mathbb{E}[F(x,Z)]_{|x=X}$, almost surely.

Corollary 3.2 (Semigroup property)**.** $P_0 = \mathrm{Id}$ *and for any* $s, t \geq 0$*,* $P_{t+s} = P_t \circ P_s$*.*

Let us define the *transition kernel* $p(t, x, \mathrm{d}y)$ of the SDE (SDE) by the identity

$$P_t f(x) = \int_{y \in \mathbb{R}^n} f(y) p(t, x, \mathrm{d}y);$$

in other words, it is the law of X_t when $X_0 = x$. The semigroup property translates, at the level of the transition kernel, as the identity

$$\forall s, t \geq 0, \quad \forall x \in \mathbb{R}^n, \qquad p(t+s, x, \mathrm{d}y) = \int_{z \in \mathbb{R}^n} p(s, x, \mathrm{d}z) p(t, z, \mathrm{d}y).$$

which is called the *Chapman–Kolmogorov equation*.

A last consequence of the Markov property is that if $X_0 \sim \mu_0$, then

$$\mathbb{E}[f(X_t)] = \mathbb{E}\left[\mathbb{E}[f(X_t)|X_0]\right] = \mathbb{E}[P_t f(X_0)] = \int_{x \in \mathbb{R}^n} P_t f(x) \mu_0(\mathrm{d}x) =: \mu_0 P_t f.$$

3.2 Infinitesimal generator and Kolmogorov equations

We follow here the Comets–Meyre book, Section 6.1.5.

We assume that b and σ are globally Lipschitz continuous. Then it turns out that, denoting by $\mathcal{C}_0$ the (Banach) space of continuous functions $f : \mathbb{R}^n \to \mathbb{R}$ converging to 0 when $|x| \to +\infty$, the semigroup $(P_t)_{t \geq 0}$ satisfies that $(t, f) \in [0, +\infty) \times \mathcal{C}_0 \mapsto P_t f \in \mathcal{C}_0$ is continuous. Let us define

$$\mathcal{D} := \{f \in \mathcal{C}_0, \lim_{t \downarrow 0} (P_t f - f)/t \text{ exists in } \mathcal{C}_0\},$$

and for any $f \in \mathcal{D}$, denote by $\mathcal{L}f$ the associated limit. We get the following statement from the Hille–Yosida Theorem.

Theorem 3.3 (Hille–Yosida)**.**

1. *$\mathcal{D}$ is dense in $\mathcal{C}_0$.*
2. *For any $t \geq 0$, $P_t \mathcal{D} \subset \mathcal{D}$.*
3. *For any $f \in \mathcal{D}$,*
$$\frac{\mathrm{d}}{\mathrm{d}t} P_t f = P_t \mathcal{L} f = \mathcal{L} P_t f. \tag{5}$$
$\mathcal{L}$ is called the infinitesimal generator *of the semigroup $(P_t)_{t \geq 0}$, and $\mathcal{D}$ is its domain.*
4. *The set of C^2_c functions $\mathbb{R}^n \to \mathbb{R}$ is contained in $\mathcal{D}$.*
5. *For any $f \in C^2_c$, $\mathcal{L}f = Lf$.*

The first identity in (5) is the Fokker–Planck equation. The second identity shows that if $u(t, x) = P_t f(x)$ is smooth enough, it satisfies the evolution equation

$$\begin{cases} \partial_t u = Lu, \\ u(0, x) = f(x), \end{cases}$$

which can be seen as a 'time-reversal' of the Feynman–Kac formula. Both equations also translate at the level of the transition kernel and write

$$\partial_t p(t, x, \mathrm{d}y) = L^*_{(y)} p(t, x, \mathrm{d}y) = L_{(x)} p(t, x, \mathrm{d}y),$$

where the notation $L_{(\cdot)}$ indicate that the operator L acts on the $\cdot$ variable. These equations are respectively called the *forward* and *backward* Kolmogorov equations.