

[SDE III] Existence, Uniqueness, and the SDE-PDE Bridge

Anqiao Ouyang

May 26, 2026

1 Introduction

In the previous two articles we built up the basic machinery: Brownian motion as a continuous, nowhere-differentiable driving process, the Itô integral as the right way to integrate against it, and the Itô formula as the chain rule for stochastic dynamics. We also discussed numerical schemes (Euler–Maruyama, Milstein), and along the way we used several results: existence of solutions, the Markov property, the parabolic operator $\partial_t + \mathcal{L}$ acting on conditional expectations, that were stated but not proved in full.

This article fills in those gaps and pushes one step further. We address three intertwined questions:

1. Under what conditions does an SDE

$$dX_t = a(t, X_t) dt + b(t, X_t) dB_t, \quad X_0 = x_0$$

have a solution, and is that solution unique? What does “solution” even mean here?

2. Once a solution exists, what is its probabilistic structure? In particular, why is it Markov, and what is the right object to read off the law of X_t from?
3. How does the SDE talk to deterministic PDEs? The conditional expectation $u(t, x) = \mathbb{E}[f(X_T) \mid X_t = x]$ already appeared as a magic ingredient in the weak-convergence proof of the EM scheme; here we explain why it solves a parabolic equation, derive the Fokker–Planck equation for the marginal density, and connect everything to the Feynman–Kac formula and Girsanov’s theorem.

The thread tying these together is the **infinitesimal generator** $\mathcal{L} = a\partial_x + \frac{1}{2}b^2\partial_{xx}$. It encodes the SDE; it determines the transition semigroup; it is the spatial operator in both Kolmogorov equations; and Feynman–Kac is essentially the statement that $\partial_t + \mathcal{L}$ has a stochastic representation.

2 Existence and Uniqueness

2.1 Strong vs. Weak Solutions

Two notions of solution coexist in SDE theory, and they answer different questions.

Definition 2.1 (Strong solution). Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a Brownian motion B , and the filtration $\{\mathcal{F}_t\}$ generated by B (augmented to satisfy the usual conditions). A **strong solution** to

$$dX_t = a(t, X_t) dt + b(t, X_t) dB_t, \quad X_0 = \xi, \quad (1)$$

is an $\{\mathcal{F}_t\}$ -adapted process X with continuous sample paths such that

$$X_t = \xi + \int_0^t a(s, X_s) ds + \int_0^t b(s, X_s) dB_s, \quad \forall t \in [0, T] \text{ a.s.}$$

In words: the driving Brownian motion is fixed up front, and the solution is a measurable functional of its path. This is the right notion when you want to simulate, since the EM scheme builds \bar{X}^δ pathwise from a single Brownian sample.

Definition 2.2 (Weak solution). A **weak solution** is a tuple $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}, B, X)$ such that B is an $\{\mathcal{F}_t\}$ -Brownian motion on this space and X satisfies the integral equation above. Here the Brownian motion and the probability space are part of the answer, not the question.

Weak solutions are sufficient for any question about the *law* of X (distributional properties, moments, transition probabilities), since one is free to choose a convenient space. Strong solutions are needed when the relationship between X and a specific B matters (simulation, filtering, pathwise control).

A strong solution is automatically a weak solution. The converse can fail: there are SDEs (Tanaka’s example $dX_t = \text{sgn}(X_t) dB_t$ is the canonical one) which admit a weak solution but no strong solution. Throughout this article we focus on strong solutions, since the standard Lipschitz framework delivers them directly.

2.2 The Picard–Lindelöf Argument, Stochastically

In the deterministic case, existence and uniqueness for $\dot{x} = f(t, x)$ with f Lipschitz in x is proved by Picard iteration: define $x_t^{(n+1)} = x_0 + \int_0^t f(s, x_s^{(n)}) ds$ and show $\{x^{(n)}\}$ is Cauchy in $C([0, T])$. The stochastic version uses the same skeleton, with $L^2(\Omega)$ replacing pointwise estimates and the Itô isometry doing the heavy lifting on the diffusion term.

Let us write down the assumptions in one place. These are the same Conditions 1–4 used implicitly in [II] for the convergence of the EM scheme.

- **(L) Lipschitz:** there exists $L > 0$ with $|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq L|x - y|$ for all $t \in [0, T]$, $x, y \in \mathbb{R}$.
- **(G) Linear growth:** there exists $C > 0$ with $|a(t, x)|^2 + |b(t, x)|^2 \leq C(1 + |x|^2)$.
- **(I) Initial data:** ξ is \mathcal{F}_0 -measurable with $\mathbb{E}|\xi|^2 < \infty$.

Theorem 2.3 (Existence and uniqueness of strong solutions). *Under (L), (G), (I), the SDE (1) admits a unique strong solution $X \in L_T^2(\Omega)$ with continuous sample paths. Moreover,*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t|^2 \right] \leq K (1 + \mathbb{E}|\xi|^2)$$

for a constant K depending only on L, C, T .

The proof splits into three estimates: an a priori moment bound, the Cauchy property of the Picard sequence, and pathwise uniqueness. We sketch the structure here and defer details to Appendix B.

Step 1: A priori bound. Define $\Phi : L_T^2(\Omega) \rightarrow L_T^2(\Omega)$ by

$$(\Phi X)_t = \xi + \int_0^t a(s, X_s) ds + \int_0^t b(s, X_s) dB_s.$$

We claim Φ maps L_T^2 into itself. By Cauchy–Schwarz on the drift integral and the Itô isometry on the diffusion integral,

$$\mathbb{E}|(\Phi X)_t|^2 \leq 3\mathbb{E}|\xi|^2 + 3T \mathbb{E} \int_0^t |a(s, X_s)|^2 ds + 3 \mathbb{E} \int_0^t |b(s, X_s)|^2 ds.$$

Plugging in (G),

$$\mathbb{E}|(\Phi X)_t|^2 \leq 3\mathbb{E}|\xi|^2 + 3(T+1)C \int_0^t (1 + \mathbb{E}|X_s|^2) ds,$$

which is finite for $X \in L_T^2$.

Step 2: Picard iteration. Set $X_t^{(0)} \equiv \xi$ and $X^{(n+1)} = \Phi X^{(n)}$. A standard computation using the Lipschitz bound (L) yields

$$\mathbb{E} \left[\sup_{s \leq t} |X_s^{(n+1)} - X_s^{(n)}|^2 \right] \leq K_1 \int_0^t \mathbb{E} \left[\sup_{r \leq s} |X_r^{(n)} - X_r^{(n-1)}|^2 \right] ds.$$

The Doob L^2 maximal inequality is what lets us pull the sup inside the expectation on the diffusion piece. Iterating n times,

$$\mathbb{E} \left[\sup_{s \leq T} |X_s^{(n+1)} - X_s^{(n)}|^2 \right] \leq \frac{(K_1 T)^n}{n!} M_0,$$

where $M_0 = \mathbb{E} \sup_{s \leq T} |X_s^{(1)} - X_s^{(0)}|^2 < \infty$. The series $\sum_n \sqrt{(K_1 T)^n / n!}$ converges, so $\{X^{(n)}\}$ is Cauchy in the Banach space \mathcal{S}_T^2 of adapted continuous processes with norm $\|X\|_{\mathcal{S}^2} = (\mathbb{E} \sup_{s \leq T} |X_s|^2)^{1/2}$. Its limit X is a strong solution.

Step 3: Uniqueness. Suppose X, Y both solve (1). Mimicking Step 2 on $X - Y$ and applying Gronwall (Appendix A) gives $\mathbb{E} \sup_{s \leq T} |X_s - Y_s|^2 = 0$, hence $X = Y$ indistinguishably.

A few remarks are in order.

- The Lipschitz assumption can be substantially weakened. **Yamada–Watanabe** shows pathwise uniqueness for 1D SDEs when b is only ρ -Hölder with $\rho \geq 1/2$ and a modulus condition $\int_0^\varepsilon \rho^{-2} = \infty$. The square root in the CIR model $dX_t = \kappa(\theta - X_t) dt + \sigma\sqrt{X_t} dB_t$ lives on this boundary. it is *not* Lipschitz, but Yamada–Watanabe still gives uniqueness.
- Without linear growth, solutions can blow up in finite time. The **explosion** time $\tau_\infty = \lim_{n \rightarrow \infty} \inf\{t : |X_t| \geq n\}$ may be finite a.s. (e.g. $dX_t = X_t^2 dt$ with $X_0 > 0$). Standard practice is to localize: solve up to τ_n , then check $\mathbb{P}(\tau_\infty > T) = 1$.
- The bound (G) is essentially sharp for moment control. Without it, even when a solution exists pathwise, $\mathbb{E}|X_t|^2$ may be infinite.

3 The Markov Property and the Generator

3.1 Markovianity

The deterministic ODE $\dot{x} = f(t, x)$ has a trivial flow property: the trajectory after time s depends only on x_s , not on how we got there. The same is true for solutions to time-homogeneous SDEs, but now “depends only on X_s ” must be promoted to a statement about *conditional distributions* rather than functional values.

Theorem 3.1 (Markov property). *Let X be the strong solution of (1) with coefficients $a(x), b(x)$ time-homogeneous and Lipschitz. Then for any bounded measurable $f : \mathbb{R} \rightarrow \mathbb{R}$ and any $0 \leq s \leq t \leq T$,*

$$\mathbb{E}[f(X_t) \mid \mathcal{F}_s] = \mathbb{E}[f(X_t) \mid X_s] \quad a.s.$$

Sketch. The key observation is that the SDE solution starting at time s from value X_s depends only on the Brownian increments $B_u - B_s$ for $u \geq s$, which are independent of \mathcal{F}_s . Writing $X_t = \Psi(s, t, X_s, (B_u - B_s)_{s \leq u \leq t})$ where Ψ is the deterministic functional that maps the initial condition and the driving path to the solution (this is well-defined by the existence/uniqueness theorem and the path-by-path nature of Picard iteration), we have $\mathbb{E}[f(X_t) \mid \mathcal{F}_s] = \Phi(X_s)$ where

$$\Phi(x) = \mathbb{E}[f(\Psi(s, t, x, (B_u - B_s)_{s \leq u \leq t}))].$$

The right side is a deterministic function of x , hence $\sigma(X_s)$ -measurable, and equals $\mathbb{E}[f(X_t) \mid X_s]$ by the tower property. \square

Time-inhomogeneous SDEs are still Markov, but the transition probability depends on both endpoints (s, t) rather than just the difference $t - s$.

3.2 The Semigroup of Transition Operators

The Markov property invites a functional-analytic reformulation. For time-homogeneous coefficients, define the operator $P_t : B_b(\mathbb{R}) \rightarrow B_b(\mathbb{R})$ (bounded measurable functions) by

$$P_t f(x) = \mathbb{E}_x[f(X_t)] := \mathbb{E}[f(X_t) \mid X_0 = x].$$

By the Markov property and time-homogeneity, $\{P_t\}_{t \geq 0}$ is a **semigroup**:

$$P_{t+s} f(x) = \mathbb{E}_x[f(X_{t+s})] = \mathbb{E}_x[\mathbb{E}[f(X_{t+s}) \mid \mathcal{F}_s]] = \mathbb{E}_x[P_t f(X_s)] = P_s(P_t f)(x).$$

The interpretation: P_t propagates the observable f backward in time by an amount t , in the sense that $P_t f(x)$ tells you the expected payoff of $f(X_t)$ from initial position x .

3.3 The Infinitesimal Generator

The semigroup is naturally analyzed via its infinitesimal generator. Heuristically,

$$\mathcal{L}f(x) = \lim_{t \downarrow 0} \frac{P_t f(x) - f(x)}{t}.$$

For an SDE with smooth coefficients, the generator coincides with the parabolic operator we have been writing all along.

Theorem 3.2 (Generator of an Itô diffusion). *Let X solve $dX_t = a(X_t) dt + b(X_t) dB_t$ with a, b Lipschitz, and let $f \in C_b^2(\mathbb{R})$. Then*

$$\mathcal{L}f(x) = a(x)f'(x) + \frac{1}{2}b^2(x)f''(x),$$

and the limit holds uniformly in x .

Proof. By Itô’s formula applied to $f(X_t)$ with $X_0 = x$,

$$f(X_t) - f(x) = \int_0^t (af' + \frac{1}{2}b^2f'')(X_s) ds + \int_0^t b(X_s)f'(X_s) dB_s.$$

Taking expectations the stochastic integral vanishes (its integrand is in L_T^2 since f', b are bounded on the range of X_s over $[0, t]$, by linear growth). Dividing by t ,

$$\frac{P_t f(x) - f(x)}{t} = \frac{1}{t} \mathbb{E}_x \int_0^t \mathcal{L}f(X_s) ds \xrightarrow{t \downarrow 0} \mathcal{L}f(x)$$

by continuity of $\mathcal{L}f(X_s)$ in s (since X is continuous and $\mathcal{L}f$ is continuous and bounded). \square

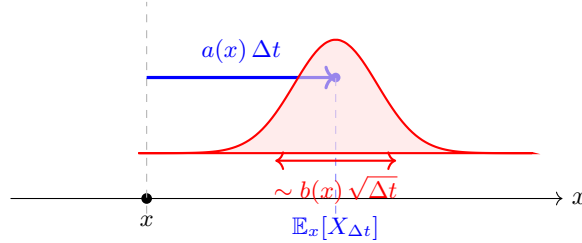


Figure 1: The generator $\mathcal{L} = a\partial_x + \frac{1}{2}b^2\partial_{xx}^2$ visualized locally. Starting at x , over a small time Δt the drift contributes a deterministic displacement of order $a(x)\Delta t$ (blue arrow), while the diffusion adds a Gaussian fluctuation of order $b(x)\sqrt{\Delta t}$ (red cloud). For small Δt , $\sqrt{\Delta t} \gg \Delta t$, so the diffusion dominates, which is why \mathcal{L} must contain a second-order term, with coefficient $\frac{1}{2}b^2$ to balance the Δt scaling of the variance.

Note the asymmetry: the drift contributes a first-order term, the diffusion a second-order term. This is a direct consequence of the Itô calculus identity $(dB_t)^2 = dt$, diffusion “moves twice as fast” as drift on the infinitesimal scale, which is what produces the $\frac{1}{2}b^2$ factor.

4 Kolmogorov’s Backward and Forward Equations

The generator \mathcal{L} governs two PDEs, dual to one another, that completely determine the law of X .

4.1 Backward Equation

Fix a terminal time T and a payoff $f \in C_b(\mathbb{R})$. Define

$$u(t, x) = \mathbb{E}[f(X_T) \mid X_t = x] = P_{T-t}f(x).$$

This is the conditional expectation that already appeared in [II] inside the weak-convergence argument for the EM scheme. The backward equation says it solves a deterministic PDE.

Theorem 4.1 (Kolmogorov backward equation). *Assume a, b are Lipschitz with all derivatives bounded, and $f \in C_P^2$ (polynomial growth). Then $u(t, x) = \mathbb{E}_{t,x}[f(X_T)]$ is in $C^{1,2}([0, T] \times \mathbb{R})$ and satisfies*

$$\begin{cases} \partial_t u(t, x) + \mathcal{L}u(t, x) = 0, & (t, x) \in [0, T] \times \mathbb{R}, \\ u(T, x) = f(x). \end{cases}$$

Sketch. Apply Itô’s formula to $u(s, X_s)$ between $s = t$ and $s = T$, conditional on $X_t = x$:

$$u(T, X_T) - u(t, x) = \int_t^T (\partial_s u + \mathcal{L}u)(s, X_s) ds + \int_t^T b(X_s) \partial_x u(s, X_s) dB_s.$$

Taking expectations conditional on $X_t = x$, the diffusion term has zero mean (by the moment estimate that follows from polynomial growth of $\partial_x u$). The left side is $\mathbb{E}_{t,x}[f(X_T)] - u(t, x) = 0$ by definition of u . Hence

$$\mathbb{E}_{t,x} \int_t^T (\partial_s u + \mathcal{L}u)(s, X_s) ds = 0$$

for all (t, x) . Differentiating in t at $s = t$ and using continuity yields $(\partial_t u + \mathcal{L}u)(t, x) = 0$. The boundary condition $u(T, x) = f(x)$ is immediate.

The regularity claim $u \in C^{1,2}$ is the hard part, and is why we assumed bounded derivatives of a, b . It can be proved by either differentiating the SDE in x (the **stochastic flow** approach) or by appealing to the standard theory of uniformly parabolic PDEs. \square

The name “backward” refers to the fact that the PDE evolves in the time variable that runs backward from the terminal data T : as t decreases from T , we accumulate more information about $f(X_T)$.

4.2 Forward Equation (Fokker–Planck)

If we instead fix the initial law $X_0 \sim \rho_0$ and ask for the density $\rho(t, x)$ of X_t , we get the dual equation.

Theorem 4.2 (Fokker–Planck equation). *Suppose X_t has a smooth density $\rho(t, x)$ for each $t > 0$ (which holds e.g. when b^2 is uniformly elliptic, $b^2 \geq \beta > 0$). Then ρ satisfies*

$$\partial_t \rho(t, x) = -\partial_x(a(x)\rho(t, x)) + \frac{1}{2} \partial_{xx}^2(b^2(x)\rho(t, x)) = \mathcal{L}^* \rho,$$

with initial condition $\rho(0, \cdot) = \rho_0$.

Sketch. Pick any test function $\varphi \in C_c^\infty(\mathbb{R})$. By Itô,

$$\mathbb{E}\varphi(X_t) - \mathbb{E}\varphi(X_0) = \mathbb{E} \int_0^t \mathcal{L}\varphi(X_s) ds.$$

Differentiating in t :

$$\frac{d}{dt} \int \varphi(x)\rho(t, x) dx = \int \mathcal{L}\varphi(x)\rho(t, x) dx = \int \varphi(x) \mathcal{L}^* \rho(t, x) dx,$$

where we integrated by parts to define the formal adjoint $\mathcal{L}^* \rho = -\partial_x(a\rho) + \frac{1}{2} \partial_{xx}^2(b^2 \rho)$. Since φ is arbitrary, $\partial_t \rho = \mathcal{L}^* \rho$ in the distributional sense; smoothness of ρ then upgrades this to a classical PDE. \square

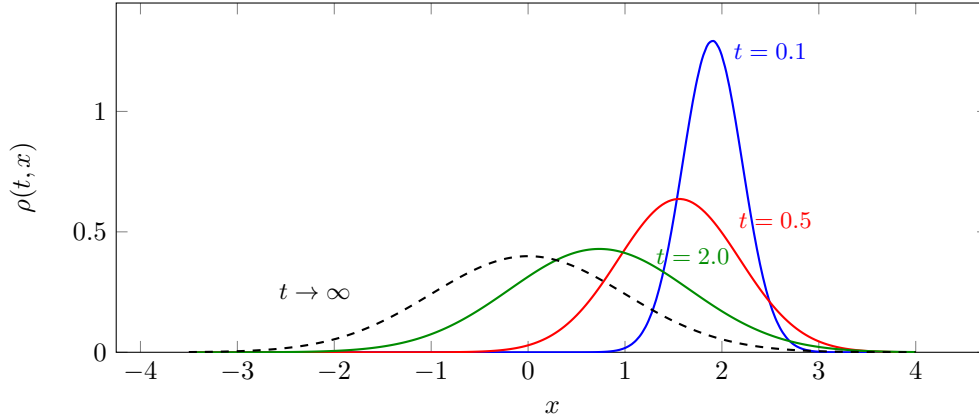


Figure 2: Solution of the Fokker–Planck equation for the OU process $dX_t = -\frac{1}{2}X_t dt + dB_t$ started from $X_0 = 2$ (Dirac mass). The density translates leftward toward the mean $\mu = 0$ at rate θ and broadens at rate σ until it relaxes to the stationary $\mathcal{N}(0, 1)$ (dashed). Both motions are encoded in \mathcal{L}^* : the $-\partial_x(a\rho)$ term shifts mass, the $\frac{1}{2}\partial_{xx}^2(b^2\rho)$ term spreads it.

4.3 Backward vs. Forward at a Glance

It is worth seeing the two equations side by side:

	Backward (Kolmogorov)	Forward (Fokker–Planck)
Unknown	$u(t, x) = \mathbb{E}_{t,x}[f(X_T)]$	$\rho(t, x) = \text{density of } X_t$
PDE	$\partial_t u + \mathcal{L}u = 0$	$\partial_t \rho = \mathcal{L}^* \rho$
Data	terminal: $u(T, \cdot) = f$	initial: $\rho(0, \cdot) = \rho_0$
Variable in \mathcal{L}	starting point x	current point x
Solves in time	backward from T	forward from 0

The dichotomy is the standard one between the Schrödinger and Heisenberg pictures, between covariant and contravariant evolution: u evolves observables, ρ evolves states, and the two operators $\mathcal{L}, \mathcal{L}^*$ are dual under the pairing $\langle u, \rho \rangle = \int u \rho dx$.

4.4 Stationary Distributions

A **stationary** (or **invariant**) distribution π satisfies $\mathcal{L}^*\pi = 0$. If $X_0 \sim \pi$, then $X_t \sim \pi$ for all t .

For 1D SDEs, the stationary density (when it exists and is integrable) can be written explicitly. From $\mathcal{L}^*\pi = -\partial_x(a\pi) + \frac{1}{2}\partial_{xx}^2(b^2\pi) = 0$, integrate once to get a first-order ODE

$$\frac{1}{2}(b^2\pi)' - a\pi = \text{const},$$

and for stationary densities decaying at infinity the constant is zero. This gives

$$\pi(x) \propto \frac{1}{b^2(x)} \exp\left(\int^x \frac{2a(y)}{b^2(y)} dy\right).$$

Example: OU process. For $dX_t = \theta(\mu - X_t) dt + \sigma dB_t$ the formula gives

$$\pi(x) \propto \exp\left(-\frac{\theta(x - \mu)^2}{\sigma^2}\right),$$

which is $\mathcal{N}(\mu, \sigma^2/(2\theta))$, exactly matching the long-time limit derived in [II] from the explicit solution.

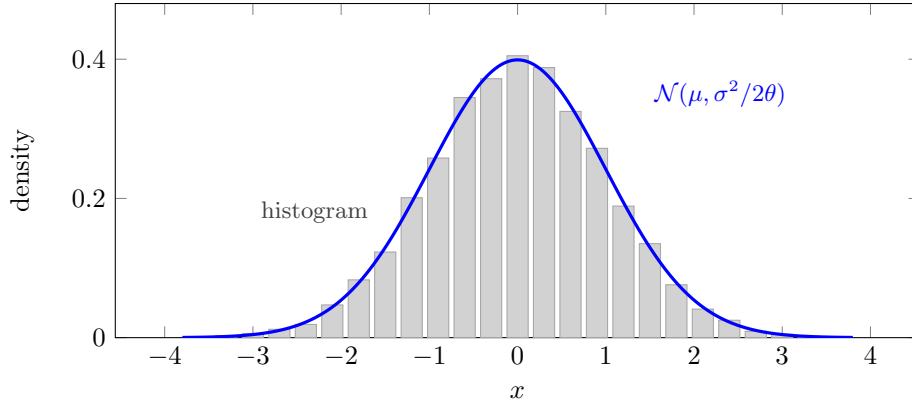


Figure 3: Long-run distribution of an OU process ($\mu = 0$, $\theta = 0.5$, $\sigma = 1$). The histogram (gray bars) is built from sampling a single long simulated trajectory; the blue curve is the analytic stationary density $\mathcal{N}(\mu, \sigma^2/2\theta)$. Ergodicity ensures the time-average over one path matches the space-average against π .

5 The Feynman–Kac Formula

The backward equation can be extended to handle discounting and source terms. The result is the **Feynman–Kac formula**, which gives a stochastic representation for solutions of linear parabolic PDEs.

Theorem 5.1 (Feynman–Kac). *Let a, b, r, g, f be smooth with appropriate boundedness, and let $u \in C^{1,2}([0, T] \times \mathbb{R})$ solve*

$$\begin{cases} \partial_t u(t, x) + \mathcal{L}u(t, x) - r(x)u(t, x) + g(x) = 0, & (t, x) \in [0, T] \times \mathbb{R}, \\ u(T, x) = f(x). \end{cases}$$

Then

$$u(t, x) = \mathbb{E}_{t,x} \left[f(X_T) e^{-\int_t^T r(X_s) ds} + \int_t^T g(X_s) e^{-\int_t^s r(X_\tau) d\tau} ds \right],$$

where X solves (1).

Proof. Consider the process $M_s = u(s, X_s) e^{-\int_t^s r(X_\tau) d\tau} + \int_t^s g(X_\tau) e^{-\int_t^\tau r(X_u) du} d\tau$ for $s \in [t, T]$. Applying Itô’s formula to $u(s, X_s) e^{-\int_t^s r d\tau}$ and combining with the differential of the second term, the drift collapses to $(\partial_t u + \mathcal{L}u - ru + g)(s, X_s) e^{-\int_t^s r} = 0$ by the PDE. Hence M is a (local) martingale.

Under the boundedness assumptions M is a true martingale, so $\mathbb{E}_{t,x} M_t = \mathbb{E}_{t,x} M_T$. The left side is $u(t, x)$; the right side is exactly the bracketed expression in the theorem (note $u(T, X_T) = f(X_T)$). \square

The lemma quoted (without proof) in [II] for the weak-convergence argument was the special case $r = g = 0$. Feynman–Kac generalizes that lemma in two directions:

- the multiplicative discount $e^{-\int r}$ allows position-dependent “killing” or “cost of carry”;
- the additive source g allows running cash flows.

Example: pricing under geometric Brownian motion. The Black–Scholes setup takes $dS_t = rS_t dt + \sigma S_t dB_t$ (under the risk-neutral measure, we will derive what that means in the next section). The price at time t of a European option paying $f(S_T)$ at time T is

$$V(t, S) = e^{-r(T-t)} \mathbb{E}[f(S_T) \mid S_t = S].$$

By Feynman–Kac with constant r and $g = 0$, V satisfies

$$\partial_t V + rS \partial_S V + \frac{1}{2} \sigma^2 S^2 \partial_{SS}^2 V - rV = 0,$$

which is the **Black–Scholes equation**. The closed-form Black–Scholes pricing formula is just the explicit evaluation of the conditional expectation for $f(S) = (S - K)^+$, using lognormality of S_T .

6 Girsanov’s Theorem: Changing the Drift by Reweighting Probability

We end with a result that has no PDE analogue but is indispensable in finance and statistics: under a suitable change of measure, an SDE can have its drift modified arbitrarily, while remaining driven by Brownian motion under the new measure.

6.1 The Doleans–Dade Exponential

Given an adapted process θ with $\mathbb{E} \int_0^T \theta_s^2 ds < \infty$, define

$$Z_t = \exp\left(\int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds\right).$$

By Itô’s formula, $dZ_t = Z_t \theta_t dB_t$, so Z is a positive local martingale with $Z_0 = 1$. It is not automatic that Z is a *true* martingale. This requires e.g. **Novikov’s condition**

$$\mathbb{E} \exp\left(\frac{1}{2} \int_0^T \theta_s^2 ds\right) < \infty.$$

Under Novikov, $\mathbb{E}Z_T = 1$, so Z_T can serve as a Radon–Nikodym density.

6.2 Girsanov’s Theorem

Theorem 6.1 (Girsanov). *Assume Novikov’s condition. Define a new probability measure \mathbb{Q} on (Ω, \mathcal{F}_T) by*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_T} = Z_T.$$

Then the process

$$\tilde{B}_t = B_t - \int_0^t \theta_s ds$$

is a standard Brownian motion under \mathbb{Q} on $[0, T]$.

Idea. The condition for \tilde{B} to be a \mathbb{Q} -Brownian motion is that \tilde{B} is continuous (clear), has quadratic variation t (clear, since the bounded-variation drift contributes nothing), and that \tilde{B} is a \mathbb{Q} -martingale (the key claim). The last is a Bayes-type computation: for $s < t$ and $A \in \mathcal{F}_s$,

$$\mathbb{E}_{\mathbb{Q}}[\tilde{B}_t \mathbf{1}_A] = \mathbb{E}_{\mathbb{P}}[Z_T \tilde{B}_t \mathbf{1}_A] = \mathbb{E}_{\mathbb{P}}[Z_t \tilde{B}_t \mathbf{1}_A]$$

(using the tower property and that Z is a \mathbb{P} -martingale). One checks by Itô that $Z_t \tilde{B}_t$ is itself a \mathbb{P} -martingale, this comes down to the quadratic-covariation identity $d[Z, \tilde{B}]_t = Z_t \theta_t dt$ canceling the drift in \tilde{B} . Hence the right side equals $\mathbb{E}_{\mathbb{P}}[Z_s \tilde{B}_s \mathbf{1}_A] = \mathbb{E}_{\mathbb{Q}}[\tilde{B}_s \mathbf{1}_A]$. \square

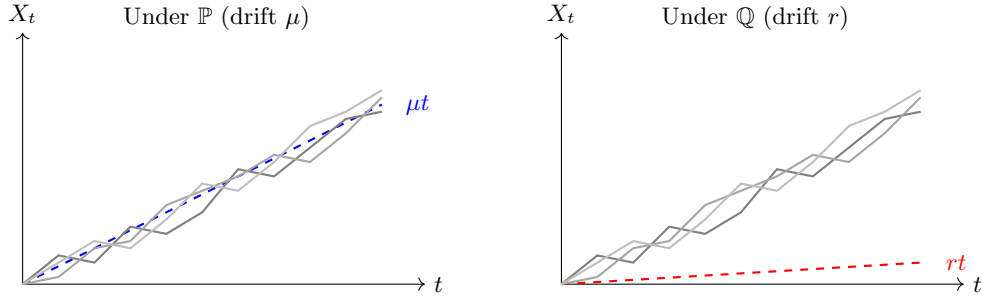


Figure 4: Girsanov in pictures. The two panels show the *same* three sample paths. What changes is the probability weighting: under \mathbb{P} these paths look typical (clustering around the trend line μt); under \mathbb{Q} the same paths look like atypical upward excursions (the trend line is now the much shallower rt). The pathwise diffusion is identical — only the location of the “expected trend” shifts. Hence the diffusion coefficient b is a measure-theoretic invariant, while the drift is not.

The interpretation is striking. Under \mathbb{P} , the SDE

$$dX_t = a(X_t) dt + b(X_t) dB_t$$

has drift $a(X_t)$. Choose θ_t so that $b(X_t)\theta_t = \tilde{a}(X_t) - a(X_t)$ for a target drift \tilde{a} . Then under \mathbb{Q} ,

$$dX_t = a(X_t) dt + b(X_t)(d\tilde{B}_t + \theta_t dt) = \tilde{a}(X_t) dt + b(X_t) d\tilde{B}_t.$$

The diffusion coefficient is invariant under the change of measure, only the drift moves. This is a strong statement: it says the “law of the diffusion coefficient” (in the sense of the quadratic variation $[X]_t = \int_0^t b^2(X_s) ds$) is a measure-theoretic invariant, while the drift can be tuned by reweighting paths.

6.3 Application: Risk-Neutral Pricing

In the Black–Scholes model under the real-world measure \mathbb{P} , the stock satisfies $dS_t = \mu S_t dt + \sigma S_t dB_t$. The drift μ reflects the actual expected return, which is generally different from the risk-free rate r . Choose

$$\theta = \frac{r - \mu}{\sigma}$$

(the **market price of risk**). Under the corresponding \mathbb{Q} ,

$$dS_t = r S_t dt + \sigma S_t d\tilde{B}_t.$$

This is the **risk-neutral** dynamics, and option prices are computed as discounted expectations under \mathbb{Q} rather than \mathbb{P} . The mathematical content is that the same stochastic process, the same set of paths, carries two perfectly valid probabilistic interpretations; the choice of measure depends on the financial question being asked.

7 Coda: The Big Picture

Stepping back, the SDE–PDE bridge developed in this article forms a diagram:

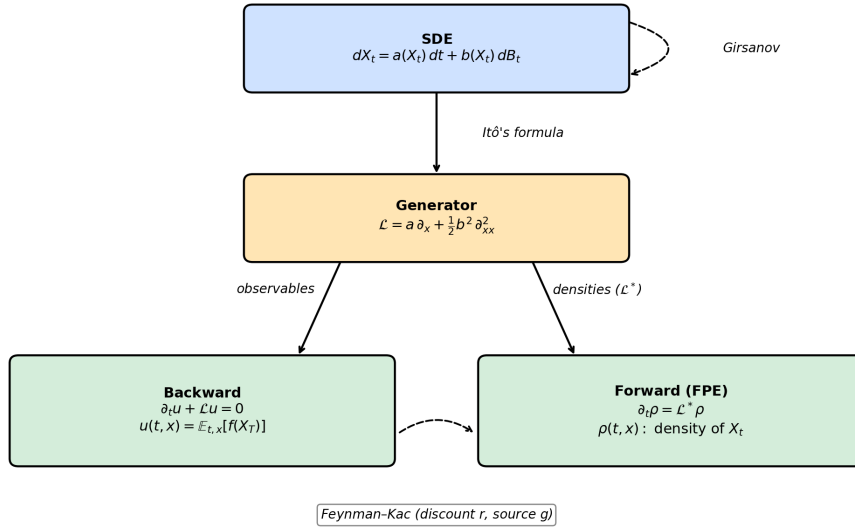


Figure 5: Caption

Existence/uniqueness tells us the SDE is well-posed; the Markov property says its solution carries no memory beyond its current state; the generator distills the SDE to a single linear operator; the two Kolmogorov equations translate that operator into PDEs for observables and densities; Feynman–Kac provides stochastic solutions to a wider class of parabolic equations; Girsanov tells us that drift is malleable while diffusion is not.

In the next installment we will use this machinery to study **specific** families of SDEs: linear SDEs, the CIR process, and the Heston stochastic volatility model, and develop tools (the change of variables, time changes, the Lamperti transform) that turn nonlinear SDEs into tractable ones.

A Gronwall’s Inequality

The proofs in this article (and in [II]) all reduce to controlling a function $\phi(t) \geq 0$ that satisfies an integral inequality of the form

$$\phi(t) \leq \alpha + \beta \int_0^t \phi(s) ds.$$

Gronwall’s lemma (integral form) states that any such ϕ , when $\alpha, \beta \geq 0$ are constants, satisfies

$$\phi(t) \leq \alpha e^{\beta t}.$$

The proof is two lines: let $\Phi(t) = \alpha + \beta \int_0^t \phi$, then $\Phi' = \beta\phi \leq \beta\Phi$, so $(\Phi e^{-\beta t})' \leq 0$, giving $\Phi(t) \leq \Phi(0)e^{\beta t} = \alpha e^{\beta t}$, and $\phi \leq \Phi$ closes the argument.

For SDE applications one often needs a slightly more general version where α is a non-decreasing function of t : if $\phi(t) \leq \alpha(t) + \beta \int_0^t \phi(s) ds$, then $\phi(t) \leq \alpha(t)e^{\beta t}$. The proof is identical.

B Details of the Existence/Uniqueness Theorem

We sketch the bookkeeping for Step 2 of the Picard iteration. Define the difference $D_t^{(n)} = X_t^{(n+1)} - X_t^{(n)}$. Then

$$D_t^{(n)} = \int_0^t [a(s, X_s^{(n)}) - a(s, X_s^{(n-1)})] ds + \int_0^t [b(s, X_s^{(n)}) - b(s, X_s^{(n-1)})] dB_s.$$

Apply the elementary inequality $(p + q)^2 \leq 2p^2 + 2q^2$, take the supremum over $s \leq t$ inside the expectation, and use **Doob's L^2 inequality** on the martingale piece:

$$\mathbb{E} \sup_{s \leq t} \left| \int_0^s [b(r, X_r^{(n)}) - b(r, X_r^{(n-1)})] dB_r \right|^2 \leq 4 \mathbb{E} \int_0^t |b(r, X_r^{(n)}) - b(r, X_r^{(n-1)})|^2 dr.$$

For the drift piece, Cauchy–Schwarz gives a similar bound with a t in front. Using the Lipschitz condition (L),

$$\mathbb{E} \sup_{s \leq t} |D_s^{(n)}|^2 \leq K_1 \int_0^t \mathbb{E} \sup_{r \leq s} |D_r^{(n-1)}|^2 ds,$$

where $K_1 = 2L^2(t + 4)$ can be bounded uniformly by $K_1 = 2L^2(T + 4)$ on $[0, T]$. Iterating n times yields

$$\mathbb{E} \sup_{s \leq T} |D_s^{(n)}|^2 \leq \frac{(K_1 T)^n}{n!} \mathbb{E} \sup_{s \leq T} |D_s^{(0)}|^2,$$

and the factorial denominator wins against the exponential numerator, giving summability.

Step 3 (uniqueness) is the same computation with $D = X - Y$ for two purported solutions: one finds $\mathbb{E} \sup_{s \leq t} |D_s|^2 \leq K_1 \int_0^t \mathbb{E} \sup_{r \leq s} |D_r|^2 ds$, and Gronwall (with $\alpha = 0$) closes the loop.

C Quick Reference: Operators and Their Adjoints

For the operator $\mathcal{L} = a(x)\partial_x + \frac{1}{2}b^2(x)\partial_{xx}^2$ acting on smooth functions, integration by parts twice yields the formal adjoint

$$\mathcal{L}^* \rho = -\partial_x(a\rho) + \frac{1}{2}\partial_{xx}^2(b^2\rho),$$

satisfying $\int (\mathcal{L}u)\rho dx = \int u(\mathcal{L}^*\rho) dx$ whenever boundary terms vanish. The asymmetry a remains inside a single derivative while b^2 sits under two, reflects the different parities of the drift and diffusion contributions and is essentially the same fact, viewed from a different angle, as why the generator has $a\partial_x$ (first order) and $\frac{1}{2}b^2\partial_{xx}^2$ (second order).

In multivariate form ($X_t \in \mathbb{R}^d$, $a : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $b : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$, $\sigma = bb^\top$),

$$\mathcal{L} = \sum_i a_i \partial_i + \frac{1}{2} \sum_{i,j} \sigma_{ij} \partial_{ij}^2, \quad \mathcal{L}^* \rho = - \sum_i \partial_i(a_i \rho) + \frac{1}{2} \sum_{i,j} \partial_{ij}^2(\sigma_{ij} \rho).$$

All of the theorems in this article extend mutatis mutandis to the multivariate case; the proofs are notationally heavier but conceptually identical.