Chapter 3

The Generator

One operator runs through most of what follows, and it repays meeting on its own before it appears in disguise four times. For a diffusion it is $\mathcal{L}f = \mu f' + \frac{1}{2}\sigma^2 f''$, and Itô's lemma says precisely that $\mathcal{L}$ is the drift of $f(X)$. Everything else in this chapter is a consequence: pricing runs $\mathcal{L}$ backwards on functions of the state, densities run its adjoint forwards, and the two are related by integration by parts and nothing more. Later chapters derive both equations separately and at length. They are the same object, and knowing that in advance turns four derivations into one.

3.1 Itô's Lemma, Read as a Definition

Chapter 2 established that continuous martingale noise has only one source: a continuous local martingale is an integral against Brownian motion, and there is nothing else on offer. What it did not establish is that a *price* may be written that way at all — that needs the price to be a semimartingale, which follows from no-arbitrage rather than from continuity, and chapter 4 is where it is obtained. Taking both, together with a drift that accumulates at a rate, a continuous arbitrage-free price is

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

and this chapter takes that form as given. Itô's lemma then says that for smooth f,

$$df(X_t) = \Big(\mu\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}\Big)dt + \sigma\frac{\partial f}{\partial x}\,dW_t. \tag{3.1}$$

The bracket has a name.

Definition 3.1 (Infinitesimal generator)**.** The generator of the diffusion is the operator

$$\mathcal{L}f = \mu(x,t)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 f}{\partial x^2}. \tag{3.2}$$

With that, (3.1) reads $df(X_t) = \mathcal{L}f\,dt + \sigma f'\,dW_t$, and since the stochastic integral has no drift:

Remark (The one sentence to remember)**.** $\mathcal{L}f$ is the drift of $f(X)$.

Equivalently, and this is the definition that does not presuppose Itô's lemma,

$$\mathcal{L}f(x) = \lim_{t \downarrow 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t}, \tag{3.3}$$

the rate at which the expectation of f starts to move from a state x. Itô's lemma is then the statement that for a diffusion this limit is (3.2), which is a computation rather than a definition, and once done it need not be done again.

Remark (The limit is a definition only where it exists)**.** Equation (3.3) is written as though it always makes sense, and it does not. The set of f for which the limit exists — uniformly, in the sense the theory requires — is called the *domain* of the generator, written $\mathcal{D}(\mathcal{L})$, and it is a genuine restriction rather than a technicality to be waved past. Two examples, failing in different ways.

Roughness. Take standard Brownian motion and $f(x) = |x|$, and evaluate (3.3) at the origin. Since $\mathbb{E}|W_t| = \sqrt{2t/\pi}$,

$$\frac{\mathbb{E}^0[f(W_t)] - f(0)}{t} = \frac{1}{t}\sqrt{\frac{2t}{\pi}} = \sqrt{\frac{2}{\pi t}} \longrightarrow \infty.$$

The limit does not exist, so $|x|$ is not in the domain, and $\mathcal{L}|x|$ is not a hard quantity to evaluate at the origin — it does not exist.[1] This is not a curiosity, because a call payoff $(S-K)^+$ has exactly this kink at the strike. What saves the pricing equation of chapter 5 is that it never asks for $\mathcal{L}$ applied to the payoff. The payoff enters as the terminal condition $V(T,\cdot) = (S-K)^+$, and the equation $\partial_t V + \mathcal{L}V = 0$ is solved backwards from it. For every $t < T$ the solution is $V(t,\cdot) = P_{T-t}(S-K)^+$, a bounded function convolved with a Gaussian density, and that is infinitely differentiable however rough the function was: the semigroup smooths instantly, and the kink exists only at the single instant T. So $\mathcal{L}$ is applied to $V(t,\cdot)$, which is in its domain, and never to the data that started it off.

The one place the kink is still visible is numerical, where the payoff is evaluated on a finite set of points rather than integrated against a density. The grid quadrature measured in chapter 19 wobbles for exactly this reason as its node count changes: the strike falls in a different place between nodes each time, and there is no smoothing left to hide it.

Growth. Now take $f(x) = e^{x^2}$, which is smooth everywhere, so the formal expression (3.2) is perfectly well defined at every point. What fails is the expectation:

$$\mathbb{E}^0\big[e^{W_t^2}\big] = \frac{1}{\sqrt{1-2t}} \quad \text{for } t < \tfrac{1}{2}, \qquad \text{and } +\infty \text{ beyond.}$$

So $P_t f$ does not exist for $t \geq \frac{1}{2}$, and no amount of care with the limit at zero repairs it.[2] The domain therefore has to control growth as well as smoothness, which is why the natural space to work in is bounded continuous functions rather than all of C^2.

The domain is also where boundary behaviour lives.

Definition 3.2 (Absorbing and reflecting barriers)**.** Let X be a diffusion on $[0,\infty)$ and $\tau = \inf\{t : X_t = 0\}$. The barrier at 0 is *absorbing* if the path stops on arrival, $X_t = 0$ for all $t \geq \tau$, and

[1] `both the closed form and the divergence, checked.`

[2] `checked, with the exponents combined --- evaluating` e^{x^2} `and the density separately overflows one and underflows the other.`

reflecting if the path is turned back instantly, spending zero time at the origin and immediately re-entering $(0, \infty)$.

Both occur. A default intensity or an equity price that is wiped out has an absorbing zero — chapter 16 is built on it, since a defaulted name does not recover — and so does a knocked-out barrier option, whose value is zero from the moment the barrier is touched. A short rate floored at zero which is meant to keep trading has a reflecting one, as does any process constrained to a range for reasons that do not stop it.

Remark (The same $\mathcal{L}$, two domains)**.** The coefficients cannot tell these apart: μ and σ are the same function of x in both cases, and the difference is entirely a condition at the boundary — which is to say, a domain.

For a reflecting barrier the condition is $f'(0) = 0$. A one-sided derivative at the origin would be multiplied by a drift the path is not allowed to follow, and requiring it to vanish is what makes the two sides of the reflection agree.

For an absorbing barrier the process at 0 does not move, so $\mathbb{E}^0[f(X_t)] = f(0)$ for every t and (3.3) gives $\mathcal{L}f(0) = 0$: the condition is on $\mathcal{L}f$, not on f. If instead the process is *killed* at the boundary — sent to a cemetery state where every payoff is worth zero, which is what a knock-out does — the natural space is functions with $f(0) = 0$. The two coincide whenever the payoff is worth nothing at the barrier, which in the financial applications it is, and that is why the distinction is easy to miss.

Every statement below is a statement about $f \in \mathcal{D}(\mathcal{L})$, and we will stop saying so.

Example 3.1 (Two generators worth recognising)**.** Geometric Brownian motion, $dS = rS\,dt + \sigma S\,dW$, has

$$\mathcal{L}f = rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2},$$

which is every term of the Black-Scholes equation except the $\partial f/\partial t$ and the discounting. Chapter 5 will assemble it.

Theorem 3.3 (Dynkin's formula)**.** *Let $f \in \mathcal{D}(\mathcal{L})$ and let τ be a stopping time with $\mathbb{E}^x[\tau] < \infty$. Then*

$$\mathbb{E}^x\big[f(X_\tau)\big] = f(x) + \mathbb{E}^x\left[\int_0^\tau \mathcal{L}f(X_s)\,ds\right]. \tag{3.4}$$

Proof. Integrate $df(X_t) = \mathcal{L}f\,dt + \sigma f'\,dW_t$ from 0 to τ:

$$f(X_\tau) = f(x) + \int_0^\tau \mathcal{L}f(X_s)\,ds + \int_0^\tau \sigma(X_s)f'(X_s)\,dW_s.$$

The last term has zero expectation, and that is the only step needing an argument. A stochastic integral is a local martingale always and a martingale only when its integrand is square integrable over the interval; with f in the domain, $\sigma f'$ is bounded, so

$$\mathbb{E}^x \int_0^\tau \sigma^2 (f')^2\,ds \;\le\; \|\sigma f'\|_\infty^2\,\mathbb{E}^x[\tau] \;<\; \infty$$

by the hypothesis on τ, and the Itô isometry of chapter 1 turns that into a finite second moment. So the integral is a true martingale, stopped at τ it still has expectation zero, and (3.4) follows. □

Remark (Where the hypothesis on τ earns its keep)**.** Dropping $\mathbb{E}[\tau] < \infty$ does not merely weaken (3.4); it makes it false. Let X be standard Brownian motion, $f(x) = x$, so $\mathcal{L}f = 0$, and let τ be the first time X reaches 1. Then τ is finite almost surely, $f(X_\tau) = 1$ always, and (3.4) would read $1 = 0 + 0$.

What has gone wrong is exactly the step above: $\mathbb{E}[\tau] = \infty$ for that stopping time, the local martingale is not a martingale, and its expectation is not preserved. This is the same failure that appears in chapter 14 as a Riccati solution exploding and in chapter 4 as a doubling strategy.

Almost every expectation computed in these notes is (3.4) with a particular f chosen to make $\mathcal{L}f$ simple — often zero, in which case $f(X)$ is a martingale and the expectation is just $f(x)$.

Remark (Finding a martingale is choosing f with $\mathcal{L}f = 0$)**.** This converts a search into an equation. A function of the state is a martingale exactly when it is in the kernel of the generator, so hunting for martingales — which is what pricing does — is solving $\mathcal{L}f = 0$.

Chapter 4's risk neutral measure will be the statement that discounted prices are martingales, which is the statement that they satisfy $\mathcal{L}f = 0$ for the appropriately adjusted $\mathcal{L}$. That is the whole of the pricing equation, and it is already visible here.

3.2 The Semigroup

Fix f and let the starting point vary. Define

$$(P_t f)(x) = \mathbb{E}^x\big[f(X_t)\big]. \tag{3.5}$$

Proposition 3.4 (The semigroup property is the Markov property)**.** *For bounded measurable f and $s, t \geq 0$, $P_{t+s} = P_s P_t$.*

Proof. By the tower property, conditioning on $\mathcal{F}_s$,

$$(P_{t+s} f)(x) = \mathbb{E}^x\big[f(X_{t+s})\big] = \mathbb{E}^x\Big[\mathbb{E}\big[f(X_{t+s}) \,\big|\, \mathcal{F}_s\big]\Big].$$

The Markov property is the statement that the inner conditional expectation depends on $\mathcal{F}_s$ only through X_s, and equals what the same expectation would be for a process started afresh at X_s and run for time t — that is, $\mathbb{E}[f(X_{t+s}) \mid \mathcal{F}_s] = (P_t f)(X_s)$. Substituting,

$$(P_{t+s} f)(x) = \mathbb{E}^x\big[(P_t f)(X_s)\big] = \big(P_s(P_t f)\big)(x). \qquad \square$$

The middle step is the whole content. The semigroup identity is not a consequence of being Markov that one derives; it is what being Markov *says*, rewritten with the conditioning replaced by composition of operators. Everything that follows in this chapter is available for Markov processes and unavailable otherwise, and chapter 10's rough volatility models are outside it for precisely this reason.

Proposition 3.5 (The backward equation)**.** *For $f \in \mathcal{D}(\mathcal{L})$,*

$$\frac{\partial}{\partial t} P_t f = \mathcal{L} P_t f = P_t \mathcal{L} f. \tag{3.6}$$

Proof. Both equalities come from proposition 3.4 and (3.3), differing only in which side the increment is taken on. Writing the difference quotient with the extra time at the front,

$$\frac{P_{t+h}f - P_t f}{h} = \frac{P_h(P_t f) - P_t f}{h} \longrightarrow \mathcal{L}\, P_t f$$

as $h \downarrow 0$, by the definition of $\mathcal{L}$ applied to the function $P_t f$. Taking it at the back instead,

$$\frac{P_{t+h}f - P_t f}{h} = P_t\Big(\frac{P_h f - f}{h}\Big) \longrightarrow P_t\mathcal{L}f,$$

the limit passing through P_t because P_t is a contraction on bounded functions,

$$\|P_t g\|_\infty = \sup_x \big|\mathbb{E}^x[g(X_t)]\big| \le \sup_x \mathbb{E}^x\big[|g(X_t)|\big] \le \sup_y |g(y)| = \|g\|_\infty,$$

since an average of values of g cannot exceed the largest value g takes anywhere. So uniform convergence of the argument gives uniform convergence of the image. □

The two readings of (3.6) are the two ways of computing an expectation, and the chapter's last section is about the difference between them. The first evolves the *function*; the second evolves the *starting point*.

Structure (The generator is a velocity and the semigroup is a flow)**.** Equation (3.6) is an ordinary differential equation for an operator-valued function of time, with initial condition $P_0 = I$. Its solution is written

$$P_t = e^{t\mathcal{L}},$$

and the notation is more than decoration even though it is formal.

The family $\{P_t\}_{t\geq 0}$ is a one-parameter semigroup of operators, and $\mathcal{L}$ is its derivative at the identity. That is exactly the relationship between a Lie group and its algebra: a rotation by angle θ is $e^{\theta A}$ for an angular velocity A, and knowing A — one matrix, at one instant — determines the whole family of rotations. Here knowing $\mathcal{L}$ — coefficients, at one instant — determines the whole evolution. This is why a chapter on an operator is a chapter about a process, and it is the same identification chapter 12 needs when the criterion for a finite dimensional realisation turns out to be a condition on a Lie algebra.

Where the analogy strains is that $\mathcal{L}$ is *unbounded*: it involves two derivatives, and differentiation makes functions larger without limit. So the power series for $e^{t\mathcal{L}}$ does not converge in any straightforward sense, and the identification cannot be established by writing it down.

Remark (What makes the exponential legitimate)**.** The theorem repairing that is the Hille-Yosida theorem.

The difficulty is that $e^{t\mathcal{L}}$ cannot be defined by its series. Hille-Yosida instead characterises which unbounded operators generate a semigroup, and does it through the *resolvent* $(\lambda I - \mathcal{L})^{-1}$ — an object that is bounded where $\mathcal{L}$ is not, because inverting a differential operator is integrating, and integrating smooths. The statement is that $\mathcal{L}$ generates a contraction semigroup precisely when it is densely defined, closed, and its resolvent satisfies $\|(\lambda I - \mathcal{L})^{-1}\| \le 1/\lambda$ for every $\lambda > 0$.

Two things follow. The exponential notation is a theorem rather than an abbreviation — something has to be checked, and for the diffusions in these notes it holds. And the resolvent, which appears here as a technical device, is the Laplace transform of the semigroup,

$$(\lambda I - \mathcal{L})^{-1} f = \int_0^\infty e^{-\lambda t}\, P_t f\, dt,$$

which is to say it is the price of a perpetual cashflow discounted at rate λ. Perpetuities are the objects on which an unbounded generator becomes bounded, and that is not a coincidence: discounting is what makes an infinite horizon finite.

Remark (What this buys)**.** Two things immediately. The first equality in (3.6) is the backward equation and the second is a statement that the generator commutes with its own semigroup, which is why one may evolve either the function or the starting point and get the same answer — the subject of the next section.

The second is a way of reading the model classes to come. If $\mathcal{L}$ maps some finite-dimensional family of functions into itself, then $e^{t\mathcal{L}}$ does too, and computing $P_t f$ for f in that family reduces to a finite system of ordinary differential equations. That single observation is why affine models are solvable, and chapter 14 is devoted to it.

3.3 The Adjoint, and Why There Are Two Equations

So far $\mathcal{L}$ has acted on functions of the state. Densities are not functions of the state in the same sense — they are the things one integrates functions against — and the operator that moves them is not $\mathcal{L}$.

Definition 3.6 (Adjoint)**.** Under the pairing $\langle f, p\rangle = \int f(x)p(x)\,dx$, the adjoint $\mathcal{L}^*$ is the operator satisfying

$$\langle \mathcal{L} f, p\rangle = \langle f, \mathcal{L}^* p\rangle \tag{3.7}$$

for every f in the domain of $\mathcal{L}$ and every p in the domain of $\mathcal{L}^*$.

Calculation 3.7 (Computing it, boundary terms included)**.** Integrate by parts on an interval $[a, b]$, keeping every term. Once for the first order piece,

$$\int_a^b \mu p \frac{\partial f}{\partial x}\, dx = \Big[\mu p f\Big]_a^b - \int_a^b f \frac{\partial}{\partial x}[\mu p]\, dx,$$

and twice for the second, which leaves two boundary terms because each integration produces one:

$$\int_a^b \tfrac{1}{2}\sigma^2 p \frac{\partial^2 f}{\partial x^2}\, dx = \Big[\tfrac{1}{2}\sigma^2 p \frac{\partial f}{\partial x} - f\frac{\partial}{\partial x}\big[\tfrac{1}{2}\sigma^2 p\big]\Big]_a^b + \int_a^b f \frac{\partial^2}{\partial x^2}\Big[\tfrac{1}{2}\sigma^2 p\Big]\, dx.$$

Collecting, (3.7) holds with

$$\mathcal{L}^* p = -\frac{\partial}{\partial x}[\mu p] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[\sigma^2 p] \tag{3.8}$$

provided the boundary terms cancel:

$$\left[\underbrace{\mu p f - f\frac{\partial}{\partial x}\Big[\tfrac{1}{2}\sigma^2 p\Big]}_{-f\,\cdot\, J[p]} + \tfrac{1}{2}\sigma^2 p \frac{\partial f}{\partial x}\right]_a^b = 0. \tag{3.9}$$

The bracket in (3.9) is not clutter, and the quantity named under it is the reason.

Definition 3.8 (Probability flux)**.** The flux of a density p is

$$J[p] = \mu p - \frac{\partial}{\partial x}\left[\tfrac{1}{2}\sigma^2 p\right], \tag{3.10}$$

so that $\mathcal{L}^* p = -\partial_x J[p]$.

Remark (What the boundary terms are for)**.** Writing $\mathcal{L}^*$ as the derivative of a flux makes the forward equation a conservation law: $\partial_t p = -\partial_x J$ says probability is neither created nor destroyed, only transported, and J is the rate at which it moves past a point. Chapter 9 makes use of exactly this form.

The boundary terms then have a reading rather than being a nuisance to assume away. On the whole line with p decaying and f bounded, both terms vanish and (3.8) holds as stated. On a bounded interval they do not vanish by themselves, and requiring them to cancel is precisely a boundary condition — $J = 0$ at the ends for a reflecting barrier, $p = 0$ for an absorbing one. So the pair $(\mathcal{L}, \mathcal{L}^*)$ carries the boundary behaviour in the condition that makes them adjoint, which is the concrete form of the remark above that a domain, not a formula, distinguishes two processes with the same coefficients.

The flux says what the two barriers of the previous section are doing, in one quantity. Integrating the forward equation over $[0, \infty)$,

$$\frac{d}{dt}\int_0^\infty p(t,x)\,dx = -\big[J\big]_0^\infty = J(t,0)$$

when J vanishes at infinity. A reflecting barrier has $J(t,0) = 0$ and the total is conserved: nothing crosses. An absorbing barrier has $J(t,0) < 0$, probability leaving through the origin and not returning, so p is the density of the paths that have survived and integrates to less than one. And since the mass lost by time t is exactly $\mathbb{P}(\tau \le t)$, the outward flux at the barrier *is* the density of the first passage time — the object chapter 16 calls the default density and prices from.

Chapter 14 needs the zero-flux case specifically. Setting $J[p] = 0$ pointwise and solving is what produces the stationary density there, and the reason it can be done in one dimension is visible in (3.10): a flux with zero divergence on a line must be constant, and a constant flux that vanishes at either end vanishes everywhere. In two dimensions a divergence-free flux may circulate instead, which is the whole of that chapter's obstruction to reversibility.

Structure (It is the same flux as in physics)**.** The word is not an analogy. $\partial_t p + \partial_x J = 0$ is the continuity equation, the same statement that carries charge in electromagnetism and mass in fluid dynamics, and it is derived the same way: whatever leaves a region has to cross its boundary, which in higher dimensions is the divergence theorem and in one dimension is the $[\,\cdot\,]_a^b$ that integration by parts produced above. Quantum mechanics calls the corresponding object for $|\psi|^2$ the probability current, and means by it exactly what (3.10) means.

The two terms are the two classical fluxes. The drift term μp is convective transport, density carried along by a velocity field, the ρv of a fluid. The second term is diffusive transport down a gradient, which is Fick's law for a solute and Fourier's law for heat, with $\frac{1}{2}\sigma^2$ in the role of the diffusion constant. A Fokker-Planck equation is an advection-diffusion equation, and a Feynman-Kac problem with a discount rate is one with a sink term.

The boundaries transfer too. A reflecting barrier is an insulated wall, no heat crossing it, a Neumann condition. An absorbing barrier is a perfect sink held at zero concentration, a Dirichlet condition, and the flux at it is the rate of arrival — the same computation that gives a first passage density gives the heat flowing into a cold wall.

Two places in the dictionary need care. The first is that $\frac{1}{2}\sigma^2$ sits *inside* the derivative, where the physical Fick's law puts the diffusion constant outside. For constant σ there is no difference; for state-dependent σ there is, and it is the same question as the Itô against Stratonovich choice of chapter 1 — physics knows it as the Itô-Stratonovich dilemma and has to settle it by asking what the noise is a model of. The next remark is about that difference from the operator side.

The second is the two-dimensional obstruction just mentioned. In stationarity $\partial_t p = 0$, so the flux is divergence-free — and a divergence-free field need not vanish. It may circulate, exactly as a magnetostatic field does around a current. Reversibility is the condition that the stationary flux is not merely divergence-free but zero; what stands between the two is the circulating part that the Helmholtz decomposition isolates, and on a line there is nowhere for it to live. Above one dimension there is, and an irreversible stationary diffusion is one with a vortex in it.

Remark (Inside or outside is the whole difference)**.** Put (3.2) and (3.8) side by side. The coefficients are the same, the derivatives are the same, and the only difference is that in $\mathcal{L}$ they sit outside the derivatives and in $\mathcal{L}^*$ they sit inside. Integration by parts is the operation that moves them, and that is the entire content of the relationship.

The difference is not cosmetic, because μ and σ depend on x: $\partial_x[\mu p] = \mu\,\partial_x p + p\,\partial_x \mu$, and the second term is real. Getting it wrong gives an equation that does not conserve probability, and `a test` makes the failure explicit rather than leaving it as a warning.

That is all the adjoint is. And it immediately gives the equation the density satisfies, at no further cost.

Calculation 3.9 (The forward equation, in two lines)**.** Let p_t be the density of X_t. Differentiating the pairing two ways and using (3.7),

$$\left\langle f, \frac{\partial p_t}{\partial t} \right\rangle = \frac{d}{dt}\langle f, p_t\rangle = \frac{d}{dt}\,\mathbb{E}[f(X_t)] = \langle \mathcal{L}f, p_t\rangle = \langle f, \mathcal{L}^* p_t\rangle,$$

the third equality being the definition of the generator. This holds for every test function f, so the two right hand arguments agree:

$$\frac{\partial p}{\partial t} = \mathcal{L}^* p. \tag{3.11}$$

Under its own name this is the Fokker-Planck, or forward Kolmogorov, equation. Chapter 9 puts it to work; there is nothing left to derive there.

The two equations that occupy later chapters are now both visible, and are one statement seen from two sides:

	Acts on	**Equation**
Backward, $\mathcal{L}$	functions of the state	$\partial u/\partial t + \mathcal{L}u = 0$
Forward, $\mathcal{L}^*$	densities	$\partial p/\partial t = \mathcal{L}^* p$

The first is Feynman-Kac, derived in chapter 5 and used to price. The second is (3.11), which chapter 9 uses to obtain Dupire's formula. Neither derivation needs to be read as new once this table is understood; each is (3.7) with the specific coefficients substituted.

Remark (The duality, in words that a desk uses)**.** An expectation $\mathbb{E}[f(X_T)]$ can be computed in two ways. Evolve the payoff backwards to today and pair it with today's density, or evolve the density forwards to expiry and pair it with the payoff. By (3.7) these agree exactly, and the choice is one of convenience:

$$\langle f, p_T \rangle = \langle P_T f, p_0 \rangle. \tag{3.12}$$

Which is more convenient depends on what is being varied. One payoff and many starting points — an option, revalued as the market moves — wants the backward equation. Many payoffs and one starting point — a whole strike surface, today — wants the forward one, since a single forward solve prices every strike at once. That is why chapter 9 reaches for the forward equation and chapter 5 for the backward.

3.4 What to Carry Forward

Three things, and they are the reason this chapter comes before the pricing rather than after it.

First, $\mathcal{L}f$ is the drift of $f(X)$, so a martingale is a solution of $\mathcal{L}f = 0$. Every pricing equation in these notes is that sentence with discounting attached.

Second, there are not two theories, one for prices and one for densities. There is one operator and its adjoint, and integration by parts is the only thing separating them. When chapter 9 derives the Fokker-Planck equation from scratch, notice that the derivation is three lines of integration by parts and nothing else.

Third, and least obvious now, tractability is a property of $\mathcal{L}$. A model is solvable when its generator preserves a small family of functions, so that an infinite-dimensional evolution collapses onto a finite one. Affine models, quadratic models and polynomial models are three answers to the same question about $\mathcal{L}$, and chapter 14 asks it properly.

References

- Oksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*, 6th ed. Springer.

- Ethier, S. N., & Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence.* Wiley.

- Pavliotis, G. A. (2014). *Stochastic Processes and Applications.* Springer.

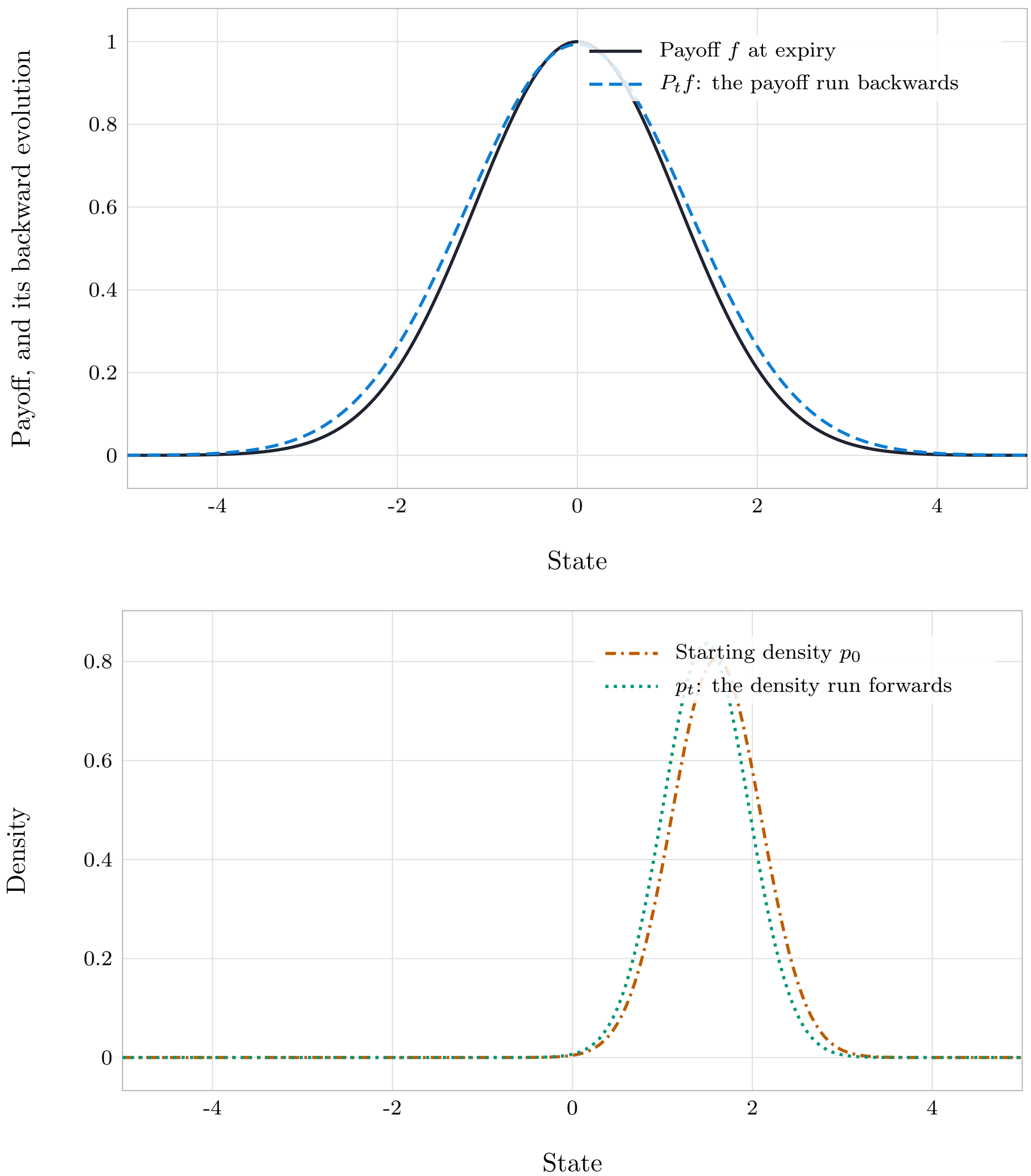

Figure 3.1: The same expectation computed from both ends, for a mean reverting process. The upper panel shows a payoff and what it becomes when run backwards under $\mathcal{L}$; the lower shows a starting density and what it becomes when run forwards under $\mathcal{L}^*$. Pairing the original payoff with the evolved density gives 0.4437028414; pairing the evolved payoff with the original density gives 0.4437028414. The two routes have no numerical step in common beyond the grid, and (3.12) says they must agree.

Drawn by `Diffusion::generator` and `Diffusion::adjoint`.