Chapter 14

Solvable Models

Chapter 3 ended on a promise: tractability is a property of the generator. This chapter cashes it. A model is solvable when $\mathcal{L}$ preserves a small family of functions, so that an evolution on an infinite-dimensional space collapses onto a finite one — and the three classical families, exponential-affine, exponential-quadratic and polynomial, are three answers to that one question. We separate that collapse from a different one it is often confused with, being Markov, since a model can have either without the other. We take the Riccati equation apart to see why it is solvable rather than merely quadratic. And we look at the spectral route, which is how physics solves such equations and which works here for exactly the diffusions whose eigenfunctions are the classical orthogonal polynomials. The chapter ends with the recipe read backwards: how to look for a tractable class that nobody has written down.

14.1 Two Collapses, Often Confused

Pricing begins as an infinite-dimensional problem. The state is a path, the expectation is over path space. Every tractable model is the result of two separate reductions that are independent.

	Reduction	**Bought by**
Markov	path space $\to$ a finite state	a state whose future ignores its past
Solvable	a PDE $\to$ finitely many ODEs	$\mathcal{L}$ preserving a small family

The first turns a functional integral into a partial differential equation in a handful of variables. The second turns that equation into a system one can solve with a pen. They are not the same thing, and much confusion follows from treating them as one.

The second needs a definition, since the whole chapter is named for it and the word is used loosely in the literature — sometimes for a closed-form formula, sometimes for anything faster than simulation.

Definition 14.1 (Solvable)**.** A model is *solvable* for a class of payoffs if the price of any payoff in the class can be computed to arbitrary accuracy by

(i) solving finitely many ordinary differential equations, the number fixed by the model and not by the accuracy demanded; and

(ii) evaluating finitely many one-dimensional integrals.

Three features of definition 14.1 are deliberate.

It does not ask for a closed-form formula. Almost nothing is solvable in that sense — Heston is not, and neither is Vasicek's swaption price without the normal distribution function — and the distinction between an expression built from special functions and one built from a Riccati solution is a fact about which functions have names, not about the model.

It does say "the number fixed by the model". A grid also produces a system of ordinary differential equations, one per node, but the number grows as the grid is refined and grows exponentially in the dimension; refining a grid is buying accuracy with state variables. A solvable model buys accuracy with nothing, because the same two Riccati equations serve every accuracy and every strike.

And it is relative to a *class of payoffs*. Solvability is not a property of a model alone. The same model appears in both columns of the table below depending on what is asked of it, which is why chapter 19 exists.

Definition 14.1 is the general form, framed this way so that it can be set against a grid on the same axis. In a fixed income setting it has a concrete and more demanding reading, and that reading is what practitioners mean by the word.

Definition 14.2 (Solvable, as a rates desk means it)**.** A term structure model is solvable if

(i) the discount bond $P(t,T)$ is an explicit function of the state, for every t and T and at *every* value the state can take; and

(ii) European swaption prices follow in closed form or by one quadrature.

The first clause does the work, and the phrase "at every value the state can take" does all of it. Definition 14.1 asks for a number: today's price of one payoff. Definition 14.2 asks for a *field* — the whole curve, as a function on the state space, available wherever one cares to evaluate it. That is strictly more, and it is what a desk needs, for two reasons.

- *Assembling instruments.* A swap, a cap, a range accrual and an amortising callable are all built from discount factors at many dates. If $P(t,T)$ is a function of the state then so is every one of them, and no separate solution is needed for each.
- *Backward induction.* This is the decisive one. A Bermudan swaption needs, at every exercise date and every node, the value of exercising — which is a swap, and so needs the whole curve at that node. A model delivering only a time-zero transform cannot supply it. Chapter 12's reconstruction formula, which writes every bond as an explicit exponential in the two state variables, is exactly clause (i) being met, and it is why that chapter's model can price callables while an equity model with the same transform machinery cannot.

So the two definitions separate the two things solvability is wanted for. Definition 14.1 is what makes a calibration cheap: fit European prices without a grid. Definition 14.2 is what makes an exotic possible: evaluate the model anywhere and let backward induction do the rest. A model can satisfy the first and fail the second — Heston does — and the failure is not a matter of degree, since no amount of extra computation turns a transform into a value function.

Structure. Read definition 14.1 against the semigroup of chapter 3. Pricing is evaluating $P_T f$, and $P_T = e^{T\mathcal{L}}$ is the exponential of an operator on an infinite-dimensional space. Nothing can be done with that directly.

A grid makes the space finite by force: replace the functions on $\mathbb{R}^d$ by their values at n^d nodes, and $\mathcal{L}$ by a matrix of that size. The exponential is then computable, and the cost of accuracy is the cost of a finer discretisation.

Solvability makes it finite by *restriction* instead. If some small family is carried into itself by the flow, then on that family P_T is an operator of low dimension exactly — no approximation — and the accuracy is limited only by how well the ODEs are integrated, which is a one-dimensional problem. Both routes reduce an infinite-dimensional exponential to a finite computation. One truncates the space and the other finds a piece of it that is already closed.

Remark (Markov is necessary and nowhere near sufficient)**.** Being Markov buys a great deal — it is the difference between a problem with a differential equation and a problem without one — and it buys nothing about closed forms.

The clearest case is chapter 9's local volatility model. It is perfectly Markov: the state is the spot, one dimension, and the pricing equation is an ordinary parabolic PDE. It has no closed-form solution for a general local volatility function, and nobody expects one. Markov gave the PDE; nothing gave the formula.

The clearest case in the other direction is what chapter 12 had to do. The Heath-Jarrow-Morton framework is not Markov in any finite state at all: the state is the entire forward curve, and the drift condition makes the evolution depend on the whole of it. Quasi-Gaussian models are the response, and they buy a finite Markov state by insisting the volatility structure decay exponentially — a restriction imposed for no reason except that it makes the first reduction possible. Rough volatility, in chapter 10, refuses the same restriction and pays for it by having no finite Markov state and therefore no PDE.

Model	**Markov**	**Solvable**	**What one gets**
Black-Scholes	yes	yes	a formula
Vasicek, Hull-White	yes	yes	bonds and swaptions in closed form
Heston	yes	yes	a characteristic function
Local volatility	yes	no	a PDE, solved on a grid
SABR	yes	no	an asymptotic expansion
Quasi-Gaussian	yes	partly	a low-dimensional grid
General HJM	no	no	a simulation
Rough volatility	no	no	a simulation, and an awkward one

Remark (Which one a difficulty belongs to)**.** Consult the table when a model is being hard to work with, because the two failures have different remedies. A model that is not Markov cannot be put on a grid at all, and the fix — if there is one — is to enlarge the state until it is, as chapter 12 does. A model that is Markov but not solvable is merely slow, and the fix is chapter 19's: a better numerical method, or a machine.

Confusing them wastes time in a predictable direction. One cannot make rough volatility affine, and one does not need a new model class to price a local volatility barrier.

14.2 Invariant Families

Now the second reduction, properly.

Definition 14.3 (Invariant family)**.** A family $\mathcal{F}$ of functions is invariant if the evolution keeps a solution inside it: whenever $u(0,\cdot) \in \mathcal{F}$ and $\partial_t u = \mathcal{L}u$, we have $u(t,\cdot) \in \mathcal{F}$ for every t.

If $\mathcal{F}$ is finite dimensional and invariant, a function that starts inside it stays inside it, and its evolution is described by however many numbers it takes to say *which* member of $\mathcal{F}$ it is. An infinite dimensional flow has collapsed onto a finite dimensional one. That is the whole of solvability. Two examples make this concrete before anything is said in general — one where invariance is the easiest thing in the world to arrange, and one where it is not.

14.2.1 A Subspace: Polynomials

Take $\mathcal{F}$ to be the polynomials of degree at most n — a linear subspace of dimension $n+1$, with the obvious basis $1, x, x^2, \dots, x^n$. A member is specified by its coefficient vector, so "which member" is $n+1$ numbers, and since $\mathcal{L}$ is a linear operator its action on those numbers is a matrix.

Calculation 14.4 (The matrix, for an Ornstein-Uhlenbeck process)**.** Take $\mathcal{L} = -\kappa x\,\partial_x + \frac{1}{2}\sigma^2\partial_{xx}$. Applying it to a single power,

$$\mathcal{L}\,x^n \;=\; -\kappa n\,x^n \;+\; \tfrac{1}{2}\sigma^2 n(n-1)\,x^{n-2}, \tag{14.1}$$

which is a polynomial of degree n: the family is invariant, and visibly so. Writing that out on $1, x, x^2, x^3$ gives

$$\mathcal{L} \;\longleftrightarrow\; \begin{pmatrix} 0 & 0 & \sigma^2 & 0 \\ 0 & -\kappa & 0 & 3\sigma^2 \\ 0 & 0 & -2\kappa & 0 \\ 0 & 0 & 0 & -3\kappa \end{pmatrix},$$

reading columns as inputs. It is triangular, because (14.1) can lower a degree but never raise one — the drift takes one power off and gives one back, the diffusion takes two off. And its diagonal is $-\kappa n$, so those are the generator's eigenvalues on this subspace; they are the same numbers the spectral section below reaches by a different route.

Now use it. Let $m_n(t) = \mathbb{E}[x_t^n]$. Since $\frac{d}{dt}\mathbb{E}[f(x_t)] = \mathbb{E}[\mathcal{L}f(x_t)]$ by Dynkin, taking expectations of (14.1) gives

$$\frac{dm_n}{dt} = -\kappa n\,m_n \;+\; \tfrac{1}{2}\sigma^2 n(n-1)\,m_{n-2}, \tag{14.2}$$

a linear system, triangular, and therefore solvable one rung at a time. The first three rungs:

- $m_0 = 1$, since a probability is a probability.
- $m_1' = -\kappa m_1$, so $m_1 = x_0 e^{-\kappa t}$: the mean decays at κ.
- $m_2' = -2\kappa m_2 + \sigma^2$, so $m_2 = x_0^2 e^{-2\kappa t} + \dfrac{\sigma^2}{2\kappa}\big(1 - e^{-2\kappa t}\big)$, which is the familiar variance plus the squared mean.

Every higher moment follows by the same substitution, with no simulation, no grid, and no distributional assumption — (14.2) never mentions normality. Integrating it numerically and comparing against the exact Gaussian moments confirms both the matrix and the ladder.[1]

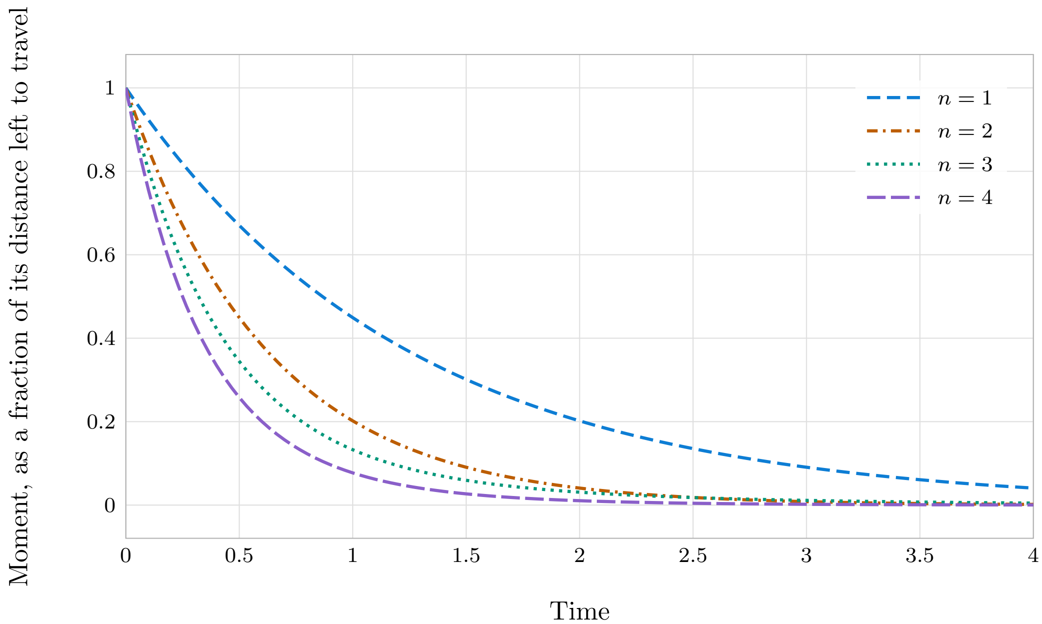

Figure 14.1: Each moment of an Ornstein-Uhlenbeck process relaxing towards its stationary value, as a fraction of the distance it had to travel. The curves are solved from (14.2) and nothing else — no simulation, no grid, no assumption of normality. The ordering is the point: the fourth moment settles first and the first moment last, because rung n of the ladder relaxes at κn, which is the diagonal of the matrix above. Read the other way, the slowest thing about a mean-reverting process is its mean.

Drawn by `polynomial_generator_matrix` and `polynomial_moments`.

This is the pattern for a subspace generally: pick a basis, write $\mathcal{L}$ as a matrix, and the flow is $\dot{a} = Ma$ on coefficients, with P_t the matrix exponential e^{tM}. Linear, always, because $\mathcal{L}$ is linear and a subspace is where linearity has somewhere to act.

Remark (Two ways to be invariant)**.** Definition 14.3 is about the flow rather than about $\mathcal{L}$, and the distinction is not pedantry — it is what the rest of this section turns on.

For the polynomials just treated, a linear *subspace*, the two are the same thing. $\mathcal{L}$ maps $\mathcal{F}$ into itself if and only if $P_t = e^{t\mathcal{L}}$ does, because the exponential of an operator preserving a subspace preserves it too. That is why writing $\mathcal{L}$ as a matrix was enough.

The next family is a *manifold*, and there the two part company: the flow condition is genuinely weaker. Take the family of exponentials below: $\mathcal{L}e^{\phi+\psi x}$ is some function of x multiplied by $e^{\phi+\psi x}$, and that product is not an exponential-affine function at all. The family is not preserved by $\mathcal{L}$. What is required instead is that the multiplier be something the *parametrisation* can absorb, and that phrase has an exact meaning.

[1] `generator::polynomial_moments`.

Definition 14.5 (What a parametrisation absorbs)**.** Write a family in exponential form with a fixed set of functions in the exponent,

$$u \;=\; \exp\left(\sum_{i=1}^{n} \theta_i \, g_i(x)\right), \tag{14.3}$$

the θ_i being the parameters and the g_i fixed. Differentiating in time,

$$\frac{\partial_t u}{u} \;=\; \partial_t \log u \;=\; \sum_{i=1}^{n} \theta_i'(t)\, g_i(x),$$

so the *only* multipliers the parameters are capable of producing are combinations of $g_1, \ldots, g_n$. That span is what the parametrisation can absorb.

The family is therefore invariant exactly when $\mathcal{L}u/u$ lies in $\operatorname{span}\{g_1, \ldots, g_n\}$ for every member; and when it does, the θ_i' are read off as its coordinates in that span, which is where the ODEs come from. If the multiplier lands outside the span there is no choice of θ_i' that can match it, and the flow leaves the family at once.

For the exponential-affine family the exponent is $\phi \cdot 1 + \psi \cdot x$, so $g_1 = 1$, $g_2 = x$, and the absorbable multipliers are the affine functions of x. That is the whole origin of the word "affine" in this subject.

Example 14.1 (The rule applied to a quadratic exponent)**.** Definition 14.5 is a calculation and not just a slogan, so run it on $\mathcal{F} = \{e^{q(x)}\}$ with q quadratic — the family behind the quadratic-Gaussian models. Here $g = (1, x, x^2)$, so quadratics are absorbable. Since $\partial_x e^q = q' e^q$ and $\partial_{xx} e^q = (q'' + (q')^2)e^q$,

$$\frac{\mathcal{L}u}{u} \;=\; \mu\, q' \;+\; \tfrac{1}{2} v \left((q')^2 + q''\right),$$

and with q quadratic, q' is linear and q'' constant. Counting degrees: $\mu q'$ has degree $\deg \mu + 1$ and $v(q')^2$ has degree $\deg v + 2$. Both must be at most two, so

$$\deg \mu \le 1, \qquad \deg v = 0.$$

A quadratic exponent demands a *constant* variance — strictly more than the affine family asks of it, which is why the quadratic-Gaussian class is built on a Gaussian state. Give the variance even a linear term and $v(q')^2$ turns cubic, leaving the span; the multiplier then has a component no choice of ϕ', ψ' or γ' can match.[2]

14.2.2 A Manifold: Exponential-Affine Functions

Now the family that motivated the remark above. Take

$$\mathcal{F} = \left\{ x \mapsto e^{\phi + \psi x} \;:\; \phi, \psi \in \mathbb{R} \right\}.$$

It is two dimensional in the sense that two numbers name a member, but it is not a linear subspace: the sum of two exponentials is not an exponential. There is no basis, and no coefficient vector

[2] `both halves measured`.

to move. What moves instead is the *label* (ϕ, ψ) — the family is a surface in function space, parametrised by two coordinates, and a flow that preserves the surface is a curve traced out in those coordinates. That is what "the induced flow is on the parameters" means, and the consequence is that nothing forces it to be linear, because (ϕ, ψ) are not coordinates in a basis and $\mathcal{L}$ is not acting on them by matrix multiplication.

Before solving anything, the invariance has to be checked, and by the remark distinguishing the two invariances it is the weaker condition that applies. Applying the generator to a member,

$$\mathcal{L}\, e^{\phi+\psi x} \;=\; \Big[\mu(x)\,\psi + \tfrac{1}{2}v(x)\,\psi^2\Big]\, e^{\phi+\psi x}, \qquad v = \sigma^2, \tag{14.4}$$

using $\partial_x e^{\phi+\psi x} = \psi\, e^{\phi+\psi x}$ and $\partial_{xx} e^{\phi+\psi x} = \psi^2 e^{\phi+\psi x}$. The bracket is a function of x, so the right-hand side is *not* an exponential-affine function and $\mathcal{L}$ does not map the family into itself. By definition 14.5 what matters is whether the bracket lies in $\operatorname{span}\{1, x\}$, and it does exactly when μ and v are affine. So the family is invariant in the sense of definition 14.3 precisely for affine drift and affine variance, which is the definition of an affine model, arrived at rather than assumed.

That is the structural claim. Chapter 12 already derived a Riccati equation for a particular model. The point here is not to derive it again but to say where its *shape* comes from: not from the model, and not from any nonlinearity in $\mathcal{L}$, but from the family being a manifold whose parameters sit in an exponent.

Calculation 14.6 (Where the Riccati equation comes from)**.** Write $u(t, x) = e^{\phi(t)+\psi(t)x}$ and ask that it satisfy $\partial_t u = \mathcal{L}u$. Dividing both sides by u makes the left

$$\frac{\partial_t u}{u} \;=\; \phi'(t) + \psi'(t)\, x,$$

and the right is (14.4).

Put $\mu(x) = a + bx$ and $v(x) = c + dx$, expand, and match the two powers of x separately:

$$\phi' = a\,\psi + \tfrac{1}{2}c\,\psi^2, \tag{14.5}$$

$$\psi' = b\,\psi + \tfrac{1}{2}d\,\psi^2. \tag{14.6}$$

Equation (14.6) is a Riccati equation — quadratic in the unknown — and (14.5) is then an integral, since its right side is known once ψ is. Two ordinary differential equations in place of a partial one.

Checking (14.4) against a numerically differenced generator confirms both the identity and that its right side really is affine in x, which is the step that lets the coefficients be read off one at a time.[3]

So "invariant" is one condition on $\mathcal{L}$ for a subspace — that it map $\mathcal{F}$ into itself — and a different, weaker one for a manifold, that $\mathcal{L}u/u$ fall in the exponent's span. Both examples have so far shown the condition *necessary*: if the flow stays in the family, the multiplier has nowhere else to be. That is the easy half and not the useful one. What a model needs is the converse — that the condition *delivers* a solution — and it does not follow for free, because an ODE system is not obviously solvable and a function satisfying a partial differential equation is not obviously the price of anything.

[3] `the test`.

Theorem 14.7 (Invariance gives solvability)**.** *Let $\mathcal{F} = \big\{\exp\big(\sum_{i=1}^{n}\theta_i g_i\big)\big\}$ with the g_i linearly independent, and suppose that for every member*

$$\frac{\mathcal{L}u}{u} \;\in\; \mathrm{span}\{g_1,\dots,g_n\}, \tag{14.7}$$

so that the coordinates of the multiplier define a map $\theta \mapsto V(\theta) \in \mathbb{R}^n$, locally Lipschitz. Then for each $\theta(0)$ there is a $T^* > 0$ and a unique

$$\theta'(t) = V\big(\theta(t)\big), \qquad t \in [0, T^*), \tag{14.8}$$

and on that interval

$$\mathbb{E}\Big[\exp\Big(\textstyle\sum_i \theta_i(0)\, g_i(X_T)\Big)\,\Big|\, X_0 = x\Big] \;=\; \exp\Big(\textstyle\sum_i \theta_i(T)\, g_i(x)\Big), \tag{14.9}$$

provided the left side is finite. The model is therefore solvable, in the sense of definition 14.1, for every payoff whose transform is a member of $\mathcal{F}$: the n equations of (14.8) *are the ODEs, and n does not depend on the accuracy wanted.*

Note that (14.9) *delivers more than definition 14.1 asks. Its right-hand side is a function of x and not a number — the expectation from* any *starting state, at once, for the price of the same n equations. So an invariant family meets clause (i) of definition 14.2 as well, and that is why the affine construction is what rates modelling is built on rather than merely a way of pricing one European option.*

Proof. Three steps.

Existence and uniqueness of the parameter flow. Condition (14.7) says the multiplier lies in the span; since the g_i are independent, its coordinates in that span are determined, which is what makes V a well-defined function of θ rather than a relation. With V locally Lipschitz, Picard-Lindelöf gives a unique solution of (14.8) on some interval around zero. It need not be all of $[0,\infty)$: V is typically quadratic, and a quadratic vector field can reach infinity in finite time. That is the explosion the Riccati section returns to, and T^* is where it happens.

The candidate satisfies the equation. Define $u(t,x) = \exp(\sum_i \theta_i(t) g_i(x))$ using that solution. Then

$$\frac{\partial_t u}{u} = \sum_i \theta_i'(t)\, g_i = \sum_i V_i(\theta)\, g_i = \frac{\mathcal{L}u}{u},$$

the last equality by the definition of V as the coordinates of $\mathcal{L}u/u$. So $\partial_t u = \mathcal{L}u$, with $u(0,\cdot)$ the payoff transform. This step is pure bookkeeping — it is the reason the construction was set up this way — and it is also where the argument would break if the multiplier had a component outside the span, since then no choice of θ' could match it.

The candidate is the expectation. This is the step with content, and it does not follow from solving the equation: many functions satisfy a parabolic equation and only one is the conditional expectation. Apply Itô's lemma to $u(T-t, X_t)$, which is legitimate since u is smooth in x and C^1 in t:

$$d\,u(T-t, X_t) = \Big(-\partial_t u + \mathcal{L}u\Big)dt + \partial_x u \cdot \sigma(X_t)\, dW_t = \partial_x u \cdot \sigma(X_t)\, dW_t,$$

the drift vanishing by the previous step. So $u(T-t, X_t)$ is a local martingale. If it is a *true* martingale, taking expectations between 0 and T gives (14.9) directly, since $u(0, X_T)$ is the payoff. Integrability of the left side of (14.9) is what upgrades the local martingale to a martingale, by dominated convergence along a localising sequence. □

Structure (Move the point or move the label)**.** Read equation (14.9) slowly, because the two sides do different things.

On the left the state moves. It starts at x, runs to X_T, and the label $\theta(0)$ is held fixed while an expectation is taken over every path it might have taken. On the right the state does not move at all — it is still x — and the *label* has run from $\theta(0)$ to $\theta(T)$ instead. The same number is obtained by moving the point or by moving the label, and never by moving both.

Which of the two is cheap is the point. Moving the point means an expectation over paths: a partial differential equation in as many dimensions as the state has, or a simulation. Moving the label means the n ordinary differential equations of (14.8), and n is fixed by the family rather than by the dimension of the state or the length of the horizon.

The duality is not a coincidence of the algebra; it is what invariance says. The semigroup $P_T = e^{T\mathcal{L}}$ acts on functions, and $\mathcal{F}$ is invariant, so P_T maps $\mathcal{F}$ into itself. A map of a manifold into itself, read in a chart, is a map of coordinates — and the coordinates here are the θ. So (14.8) is not an approximation to the semigroup. It *is* the semigroup, written in the only coordinates the invariant family has. The remark above on the two ways to be invariant is what licenses that reading: the vector field has to be tangent to the family, or there would be no induced motion on the coordinates to write down.

Notice too that the label runs the opposite way to the state. It is $\theta(0)$ that sits inside the expectation, against the terminal state, and $\theta(T)$ that sits outside, against the initial one. That reversal is the signature of a pairing: the g_i pair a point with a label, and moving one of the two is the same as moving the other backwards. It is the same reversal that separates the forward equation, which pushes densities from today into the future, from the backward equation, which pulls payoffs from the future back to today — and this theorem is the backward equation with the function replaced by its coordinates.

The manoeuvre generalises well beyond this chapter. Substituting an exponential ansatz and finding that a linear equation for the function has become a nonlinear equation for the exponent is the pattern shared by the method of characteristics, by Hamilton-Jacobi theory, and by the large deviation asymptotics of these same models — and, in chapter 12, by the Riccati system that appeared there for what looked like an unrelated reason. In each the exponent's coefficients are coordinates dual to the state, and solving the problem means flowing them rather than the state.

Remark (Where the proviso bites)**.** The integrability proviso is not decoration. In an affine model the transform $\mathbb{E}[e^{\psi_0 X_T}]$ is finite only for ψ_0 in a strip, and outside it the moment does not exist — which is the same statement as the Riccati solution exploding before T. So the three failure modes of theorem 14.7 line up: the ODE reaching infinity, the moment being infinite, and the local martingale not being a martingale are one phenomenon seen from three sides.

It matters in practice because the strip is where the damping parameter of a Fourier inversion has to live. A transform method evaluates the characteristic function at a complex argument, and choosing that argument outside the strip returns a number the model does not have. Implementations report this as an overflow in the Riccati solver; it is really a statement about which moments exist.

14.2.3 What Each One Actually Gives You

Both calculations end with an ordinary differential equation, but what one does with the answer differs, and the two entries in the table below are opaque without saying how.

Remark (The transform, and inverting it)**.** Solving (14.5) and (14.6) out to T, starting from $\phi(0) = 0$ and $\psi(0) = \psi_0$, gives

$$\mathbb{E}\Big[e^{\psi_0 x_T}\Big] \;=\; e^{\phi(T)+\psi(T)x_0}, \tag{14.10}$$

since $u(t,x) = \mathbb{E}[e^{\psi_0 x_T} \mid x_t = x]$ is the solution of the backward equation with that terminal condition. Setting $\psi_0 = iw$ makes the left side the *characteristic function* of x_T, and that is what "transform" means here: not a rearrangement of the problem but a specific function of one variable w, delivered in closed form by two ODEs, one w at a time.

A characteristic function determines a distribution, so it determines every European price. Recovering the price is an integral over w — Fourier *inversion.*

Calculation 14.8 (Inverting the transform)**.** Write $s_T = \ln S_T$, let $k = \ln K$ be the log-strike, and take zero rates so that

$$C(k) = \mathbb{E}\big[(e^{s_T} - e^k)^+\big].$$

The natural move is to Fourier transform in k, and it fails immediately: as $k \to -\infty$ the call price tends to S_0, not to zero, so C is not integrable and has no transform.

Damp it. Multiply by $e^{\alpha k}$ for some $\alpha > 0$. The left tail is now killed by the exponential, and the right tail was already fine, so $c(k) = e^{\alpha k} C(k)$ is integrable and

$$\zeta(w) = \int_{-\infty}^{\infty} e^{iwk}\, c(k)\, dk = \int_{-\infty}^{\infty} e^{zk}\, C(k)\, dk, \qquad z = \alpha + iw,$$

exists.

Exchange the order of integration. Writing q for the density of s_T and using that the payoff vanishes for $k > s$,

$$\zeta(w) = \int_{-\infty}^{\infty} q(s) \int_{-\infty}^{s} \big(e^s - e^k\big) e^{zk}\, dk\, ds.$$

The inner integral is elementary, and needs $\operatorname{Re} z = \alpha > 0$ for its lower limit to converge:

$$\int_{-\infty}^{s} e^s e^{zk} dk - \int_{-\infty}^{s} e^{(z+1)k} dk = e^{(z+1)s}\left(\frac{1}{z} - \frac{1}{z+1}\right) = \frac{e^{(z+1)s}}{z(z+1)}.$$

Recognise what is left. The outer integral is now $\int q(s) e^{(z+1)s} ds = \mathbb{E}[e^{(z+1)s_T}]$, which is the characteristic function at the complex argument $-i(z+1)$. So

$$\zeta(w) \;=\; \frac{\varphi\big(w - i(\alpha+1)\big)}{(\alpha + iw)(\alpha + 1 + iw)}, \tag{14.11}$$

with $\varphi(u) = \mathbb{E}[e^{ius_T}]$. Everything on the right is available from the Riccati equations, one w at a time.

Invert. Finally $c(k) = \frac{1}{2\pi}\int e^{-iwk}\zeta(w)\, dw$, and since C is real, $\zeta(-w) = \overline{\zeta(w)}$ and the two halves of the line fold together:

$$C(k) \;=\; \frac{e^{-\alpha k}}{\pi} \int_0^{\infty} \operatorname{Re}\Big[e^{-iwk}\,\zeta(w)\Big]\, dw. \tag{14.12}$$

One integral per strike, with an integrand in closed form. Applied to a lognormal, whose characteristic function is immediate, it returns the Black-Scholes price to ten decimals.[4]

[4] `checked here`, against a formula the routine knows nothing about.

Remark (What α has to satisfy, and why it is the same condition as before)**.** Two constraints have appeared, and the second is the interesting one.

The damping must be strictly positive, $\alpha > 0$, or the inner integral above diverges. Equivalently: at $\alpha = 0$ the denominator of (14.11) vanishes at $w = 0$, so the undamped transform has a pole at the origin. That pole is the statement that a call price is not integrable in log-strike.[5]

The numerator of (14.11) evaluates φ at $w - i(\alpha + 1)$, which is to say it evaluates

$$\mathbb{E}\big[S_T^{1+\alpha}\big].$$

So α is admissible only if the moment of order $1 + \alpha$ exists — and that is exactly the proviso of theorem 14.7, arriving from the other side. There the condition appeared as the integrability that upgrades a local martingale; here it appears as the strip in which a damping parameter may be chosen. They are one condition.

The practical consequence is that in a model whose moments fail above some order — Heston with a large enough correlation, or any model with a heavy right tail — choosing α beyond that order does not fail loudly. The Riccati solver returns something, the integral evaluates, and the answer is not a price.[6] Within the strip, the answer is insensitive to α, as it must be for a device that is not part of the problem.[7]

Remark (Two results, one pair of authors)**.** Carr and Madan appear twice in these notes and the results are unrelated, which deserves flagging since the names collide. The inversion above is their 1999 transform method. Chapter 15's static replication — writing a payoff as a portfolio of options with weights $g''(K)$ — is their 2001 positioning result, and it is model-free, involves no transform, and needs no characteristic function. One turns a model into prices; the other turns prices into other prices without a model.

Remark (Moments, and what they will not give you)**.** The polynomial side delivers something more elementary and more limited. Equation (14.2) gives $\mathbb{E}[x_T^n]$ for as many n as wanted, exactly — the ladder is a finite triangular linear system, so its solution is a finite combination of exponentials in T with no discretisation error and no sampling error — and cheaply, since an $(n+1)$-dimensional matrix exponential is the whole of the work.

What it does not give is a price. Moments do not determine a distribution in any usable numerical sense: reconstructing a density from its moments is severely ill-conditioned, and the error in a price built that way is not controlled by the error in the moments. So a polynomial family is the right tool when the moments *are* the answer — variance swaps and other quadratic payoffs, method-of-moments estimation, or the covariance structure a hedge ratio needs — and the wrong one when a strike is involved.

That asymmetry is the honest reason affine models are the ones on trading desks and polynomial models are not, despite the polynomial condition being the weaker of the two.

14.2.4 The Distinction, and What It Decides

The two calculations differ in one structural respect and it accounts for everything else.

[5] `the pole is visible in the scaling`: undamped, shrinking w by a thousand grows the transform by a thousand; damped, it does not grow at all.

[6] `demonstrated by capping the moments by hand`.

[7] `checked across a range`.

Structure. A linear *subspace* has a basis, a member is a coefficient vector, and a linear operator acts on coefficient vectors by a matrix. So the induced flow is $\dot{a} = Ma$ and is linear, necessarily and always.

A *manifold* has no basis. A member is named by parameters, and those parameters enter the function nonlinearly — here they sit in an exponent. Applying $\mathcal{L}$ therefore does not act on them linearly; the second derivative brings ψ down twice, and the induced flow acquires a ψ^2. It is a vector field on the parameter space, and its degree is inherited from how the parameters were used to build the family, not from any nonlinearity in $\mathcal{L}$ — which remains a perfectly linear operator throughout.

So the shape of the family fixes the shape of the equations. Subspaces give linear systems; the exponential manifold gives Riccati. Neither is an accident and neither is a fact about the model, only about the family one chose to preserve.

Family	**Shape**	**Equations**	**What it computes**
$e^{\phi+\psi x}$	manifold	Riccati	the characteristic function, and so any European price, by one Fourier integral per strike
$e^{\phi+\psi x+x^\top Cx}$	manifold	matrix Riccati	the same, with a quadratic state
polynomials of degree $\leq n$	subspace	linear	every moment up to n, in closed form; no prices without one

14.3 Reading the Recipe Backwards

The definition can be used as a search rather than as a description, and this is how one would look for a class that is not in the table.

Calculation 14.9 (What preserving polynomials forces)**.** Suppose we want $\mathcal{L}$ to map polynomials of degree at most n into polynomials of degree at most n, for every n. Apply it to x^n:

$$\mathcal{L}x^n = n\,\mu(x)\,x^{n-1} + \tfrac{1}{2}n(n-1)\,\sigma^2(x)\,x^{n-2}.$$

The first term has degree $n-1+\deg\mu$ and the second $n-2+\deg\sigma^2$. For both to be at most n,

$$\deg\mu \leq 1, \qquad \deg\sigma^2 \leq 2.$$

So the drift must be affine and the variance quadratic — and that is the definition of a polynomial process, obtained rather than assumed.

Remark (The same question, asked of other families)**.** That calculation is a template. Choose the family one wants preserved, apply $\mathcal{L}$ to a general member, and demand the result stay inside. What comes back is a condition on μ and σ, and it is usually restrictive enough to name a model class.

Asking it of $e^{\psi x}$ gives affine drift and affine variance — chapter 12's structure note is that calculation. Asking it of $e^{\psi x+cx^2}$ gives the quadratic-Gaussian class. Asking it of polynomials gives the above. These are not three discoveries; they are one question asked three times.

It also says when to stop. If the condition that comes back is that μ and σ are constant, the family was too small to be interesting. If nothing comes back, it was too large.

14.4 Why the Riccati Equation Is Solvable

Chapter 12 derived $\psi' = \beta\psi + \frac{1}{2}\alpha\psi^2$ and solved it in the Vasicek case by noticing that the quadratic term vanishes. The general case is solvable too because it is the reason Riccati equations are solvable anywhere in mathematics.

Calculation 14.10 (A Riccati equation is a linear system in disguise)**.** Substitute $\psi = -\dfrac{2}{\alpha}\dfrac{w'}{w}$. Then

$$\psi' = -\frac{2}{\alpha}\left(\frac{w''}{w} - \frac{(w')^2}{w^2}\right) = -\frac{2}{\alpha}\frac{w''}{w} + \frac{\alpha}{2}\psi^2,$$

and substituting into the Riccati equation, the quadratic terms cancel identically, leaving

$$w'' + \beta w' = 0, \tag{14.13}$$

a *linear* second order equation with constant coefficients.

Structure (What the substitution means)**.** The cancellation looks like luck and is not. A Riccati equation is what a linear flow looks like after projection.

Take the linear system

$$\begin{pmatrix} u \\ v \end{pmatrix}' = M \begin{pmatrix} u \\ v \end{pmatrix}, \qquad M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

and follow only the ratio $\psi = v/u$. The quotient rule gives two terms,

$$\psi' = \frac{v'u - vu'}{u^2} = \underbrace{\frac{v'}{u}}_{\text{numerator moves}} - \underbrace{\psi\,\frac{u'}{u}}_{\text{denominator moves}},$$

and both fractions become affine in ψ once the linear dynamics are divided through by u: $v'/u = c + d\psi$ and $u'/u = a + b\psi$. Substituting,

$$\psi' = (c + d\psi) - \psi\,(a + b\psi) = c + (d - a)\,\psi - b\,\psi^2. \tag{14.14}$$

There is the quadratic, and it comes from one place: the second term, in which the ratio multiplies its own denominator's growth rate. Were the denominator held fixed, $u' = 0$, the equation would be affine and there would be nothing to solve. The nonlinearity is the price of the denominator being alive.[8]

So ψ satisfies a Riccati equation — and conversely, every Riccati equation arises this way, since (14.14) can be matched to any given coefficients. The substitution above is that projection run backwards: it lifts a nonlinear flow on a line to a linear flow on a plane, where it can be solved by exponentiating a matrix.

Two consequences follow immediately. Riccati equations are solvable exactly when the linear system above them is, which is why affine models have closed forms at all. And a Riccati solution can blow up in finite time even though the linear system never does — the ratio v/u runs to infinity when u passes through zero, while nothing has happened upstairs. That is not a numerical failure and cannot be fixed by a smaller step; it is a feature of the projection, and it is why affine model

[8] the test that checks (14.14) against an integrated linear flow.

calibrations sometimes report an explosion at a maturity for no financial reason. The cleanest illustration is a rotation, $M = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$: the pair (u, v) walks around the unit circle forever, perfectly bounded, while $\psi = \tan t$ runs to infinity at $t = \pi/2$ and comes back from the other side.[9]

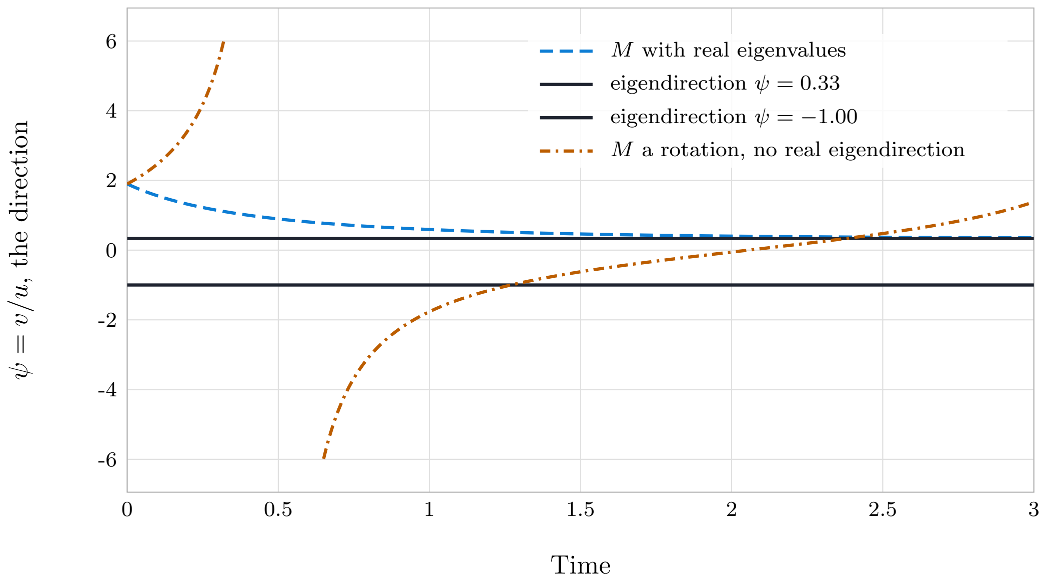

Figure 14.2: The Riccati flow, plotted as the direction $\psi = v/u$ of a linear flow in the plane, for two matrices from the same starting direction. Where M has real eigenvalues the two horizontal lines are the corresponding eigendirections, and the trajectory is trapped between them and settles on one — a fixed point of the projected flow is a direction the linear flow preserves. Where M is a rotation there is no real eigendirection, nothing traps the trajectory, and it runs off the top of the chart and re-enters from the bottom, passing through infinity at $t = \pi/2 - \arctan 1.9 \approx 0.49$: that is the blow-up, and the figure shows what it is, namely the direction passing through $u = 0$ where the coordinate ψ stops covering the line. Nothing happens to the flow there. The curve is computed in the angle $\theta = \arctan \psi$, which covers the whole line and in which the field is bounded, so the integrator passes the blow-up without noticing — one working in ψ could not, and would stop at the edge.

Drawn by `the_ratio_of_a_linear_flow_obeys_a_riccati_equation`.

Remark (The same construction as a blow-up)**.** A reader who has met algebraic geometry will recognise this projection, and the resemblance is exact rather than atmospheric.

The linear flow lives on the plane and vanishes at the origin, so the origin is a singular point of the family of trajectories — every trajectory approaches or leaves it, and no single direction is defined there. *Blowing up* the origin is the standard repair: replace that one point by the projective line $\mathbb{P}^1$ of directions through it. A point of $\mathbb{P}^1$ is a line through the origin, which is to say a value of the ratio $\psi = v/u$, together with the one direction $u = 0$. The Riccati equation is exactly the flow induced on that line.

Two things follow, and both are visible in what the equation does.

[9] `measured here`.

The singularity is resolved into eigendirections. The trajectories' bad point at the origin becomes finitely many bad points on the exceptional line, and they are the fixed points of the Riccati flow, $c + (d-a)\psi - b\psi^2 = 0$. Those are the eigendirections of M: substituting $\psi = (\lambda - a)/b$ turns the quadratic into the characteristic polynomial $(\lambda - a)(\lambda - d) - bc$, so it vanishes exactly when λ is an eigenvalue.[10] One badly behaved point has become two simple ones whose structure can be read off, which is what resolving a singularity buys.

The solution map is a Möbius transformation. A linear map of the plane descends to a fractional linear map of the line of ratios, so with $e^{TM} = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$,

$$\psi(T) = \frac{C + D\,\psi_0}{A + B\,\psi_0}. \tag{14.15}$$

That is a strong statement about a nonlinear differential equation, and it has a testable consequence: Möbius maps are precisely those preserving the cross-ratio, so four solutions started from four different points keep their cross-ratio for all time.[11]

The blow-up in finite time is then a chart artefact rather than an event. $\mathbb{P}^1$ is compact and the flow on it is perfectly well behaved; the coordinate $\psi = v/u$ merely fails to cover the direction $u = 0$, so a trajectory passing through that direction leaves the chart at $+\infty$ and re-enters at $-\infty$.

Remark (Why the ceiling is quadratic)**.** Everything above explains what a quadratic buys. It does not explain why the story stops there — why a cubic term would not buy something further — and the answer is not that nobody has tried.

Take vector fields on a line, $f(z)\,\partial_z$, which is what the right-hand side of (14.14) is. Two of them bracket as

$$[\,f\partial,\ g\partial\,] \;=\; (fg' - gf')\,\partial, \qquad \text{so} \qquad [\,z^i\partial,\ z^j\partial\,] \;=\; (j-i)\,z^{\,i+j-1}\,\partial.$$

On the span of $\{1, z, z^2\}$ that produces nothing new: the highest degree it can reach is $i+j-1 = 2$, attained by $[z\partial, z^2\partial]$. So those three fields close into a three dimensional Lie algebra — $\mathfrak{sl}_2$, whose group is the one acting on $\mathbb{P}^1$ by the Möbius maps of (14.15). The quadratic is exactly what is needed to fill it and exactly what it can hold.

Admit a cubic and it escapes immediately. $[z^2\partial, z^3\partial]$ is a quartic, which brackets to a quintic, and the algebra generated is infinite dimensional. There is then no finite dimensional group acting on the solutions, so there is no solution map of finitely many functions of time, no invariant to inherit, and no linear system of fixed size overhead to lift to. The failure is structural rather than a gap in anyone's ingenuity.

It is also visible. Integrating a cubic field with the same scheme and the same four starting points, the cross-ratio that a Riccati preserves to machine precision drifts by three percent at a modest cubic coefficient and seven at a larger one.[12]

So the boundary of solvability sits at the quadratic because the line has a three dimensional symmetry algebra and no more. That is the same criterion chapter 12 applies one level up, where a finite dimensional realisation exists exactly when the Lie algebra generated by the drift and volatility vector fields is finite dimensional. There the algebra lives on the curve and here on the

[10] `checked here.`
[11] `measured, and it does.`
[12] `measured.`

parameter, and in both places finite dimensionality of an algebra is what finiteness of a state or a solution formula comes down to.

One thing not to run together with this. Chapter 12 also warns that a *quadratic diffusion coefficient* can carry the state to infinity in finite time, and that quadratic is a different one — a polynomial in the state, not in the parameter, and the two ceilings have nothing to do with each other. That one is a question about whether an SDE has a solution; this one is about whether an ODE has a symmetry group.

Remark (Where physics met finance)**.** The same substitution turns the Riccati equation of a stochastic control problem into a Schrödinger equation, and it is the standard route in both directions. The historical route by which affine term structure models were found runs through this correspondence rather than through finance: the equations were recognised as ones already solved elsewhere.

This is the general lesson of the chapter and the reason for the emphasis. New model classes have rarely been invented by writing down dynamics and hoping. They have been found by noticing that a tractable structure from another subject can be mapped onto a financial one, and the mapping is almost always at the level of the generator.

14.5 The Spectral Route

There is a second way to solve $\partial u/\partial t = \mathcal{L}u$, and it is the one a physicist reaches for first: diagonalise the operator.

If there are functions φ_n and numbers λ_n with

$$\mathcal{L}\varphi_n = -\lambda_n \varphi_n, \tag{14.16}$$

then expanding an initial condition in the φ_n reduces the evolution to scalar multiplication:

$$P_t f = \sum_n c_n e^{-\lambda_n t} \varphi_n, \qquad f = \sum_n c_n \varphi_n. \tag{14.17}$$

Each mode decays at its own rate and nothing mixes. This is an invariant family again — each φ_n spans a one-dimensional one — but arranged so the whole operator is diagonal at once rather than merely block-finite.

The expansion needs the eigenfunctions to form a basis, and that needs $\mathcal{L}$ to be self-adjoint. Not in the plain pairing of chapter 3, though — in a weighted one.

Definition 14.11 ($L^2(\pi)$)**.** Let π be the stationary density. The space $L^2(\pi)$ is the functions square integrable against it, with inner product

$$\langle f, g \rangle_\pi = \int f(x)\, g(x)\, \pi(x)\, dx. \tag{14.18}$$

The weighting says that a function's size is measured where the process actually spends its time. For an Ornstein-Uhlenbeck process with $\pi = N(0, \sigma^2/2\kappa)$, a function that grows wildly at $x = 10$ standard deviations is small in $L^2(\pi)$, because the process is never there. In the flat pairing it would be enormous. The financial reading is direct: what matters about a payoff is its behaviour over the states the model visits.

Definition 14.12 (Reversible)**.** A stationary diffusion is reversible if its law is unchanged by running time backwards — equivalently, if the probability flowing from x to y in time t equals that flowing from y to x, $\pi(x)\, p_t(x,y) = \pi(y)\, p_t(y,x)$.

Concretely: film the process in stationarity, show the film to someone, and ask which way it is running. If they cannot tell, the process is reversible. A mean-reverting rate wandering around its long-run level passes that test; it looks the same either way. The theorem is that $\mathcal{L}$ is self-adjoint in $L^2(\pi)$ exactly when the process passes it.

Calculation 14.13 (Why one dimension is free)**.** A stationary density carries no net probability flux. In one dimension there is nowhere for a flux to circulate to, so it must vanish *pointwise*, which forces $\mu\pi = \frac{1}{2}(\sigma^2\pi)'$ and therefore lets the generator be rewritten as a divergence:

$$\mathcal{L}f = \frac{1}{\pi}\left(\tfrac{1}{2}\sigma^2\pi\, f'\right)'. \tag{14.19}$$

Now pair it. The π in (14.18) cancels the $1/\pi$ in (14.19), and one integration by parts gives

$$\langle \mathcal{L}f, g\rangle_\pi = \int \left(\tfrac{1}{2}\sigma^2\pi f'\right)' g\, dx = -\int \tfrac{1}{2}\sigma^2\pi\, f'\, g'\, dx,$$

which is symmetric in f and g on its face. So $\langle \mathcal{L}f, g\rangle_\pi = \langle f, \mathcal{L}g\rangle_\pi$ for *every* scalar diffusion, whatever its drift and volatility — not merely for the classical ones.[13]

Example 14.2 (A two-dimensional process that fails)**.** Above one dimension probability can circulate, and then it generally does. Take a linear model in the plane,

$$dZ_t = -AZ_t\, dt + dW_t, \qquad A = \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix},$$

whose antisymmetric part is a rotation. Its stationary covariance solves $A\Sigma + \Sigma A^\top = I$, and the rotation drops out of that equation entirely: $\Sigma = \frac{1}{2}I$, the same isotropic Gaussian a symmetric A would give. A snapshot of this process is indistinguishable from a reversible one.

The film is not. The process circulates around the origin at rate θ, so running it backwards makes it circulate the other way, and one can tell. Its generator is correspondingly not self-adjoint: on linear functions $\mathcal{L}$ acts as $-A^\top$, so

$$\langle \mathcal{L}f, g\rangle_\pi = -\alpha^\top A\Sigma\beta, \qquad \langle f, \mathcal{L}g\rangle_\pi = -\alpha^\top \Sigma A^\top \beta,$$

and these agree for all α, β only when $A\Sigma$ is symmetric, which the swirl prevents. The spectrum moves off the real line to match — the eigenvalues of $-A$ are $-1 \mp i\theta$ — so there is no real orthogonal expansion to be had.[14]

So the spectral method is a one-dimensional tool in practice, and the reason is structural rather than technical: reversibility is nearly automatic on a line and generically false in a plane. Most multi-factor models of interest have a swirl of exactly this kind, because correlated factors with different mean reversion rates produce one. That is the honest boundary of the technique, and it is the same shape as the boundary of every other method here: a structural condition, cheap in low dimension and expensive above it.

[13] `checked numerically for an Ornstein-Uhlenbeck process, a double well and a state-dependent volatility`.

[14] `measured, including that the stationary covariance is unchanged`.

Example 14.3 (The three classical cases)**.** Where the method works completely, the eigenfunctions are the classical orthogonal polynomials, and the correspondence is not a coincidence but the polynomial invariance of the previous section seen through a self-adjoint lens:

Process	**Stationary law**	**Eigenfunctions**	**Eigenvalues**
Ornstein-Uhlenbeck	Gaussian	Hermite	$-n\kappa$
Square root (CIR)	Gamma	Laguerre	$-n\kappa$
Jacobi	Beta	Jacobi	$-n(\cdots)$

These are the Pearson diffusions, and they are the intersection of the two ideas in this chapter: their generators preserve polynomials, and they are reversible, so the polynomial invariant subspace can be diagonalised rather than merely triangularised.

Calculation 14.14 (The Ornstein-Uhlenbeck case, checked)**.** For $dX = -\kappa X\,dt + \sigma\,dW$ with stationary variance $s^2 = \sigma^2/2\kappa$, the eigenfunctions are the Hermite polynomials $He_n(x/s)$ with $\mathcal{L}He_n = -n\kappa\,He_n$. The case $n = 2$ shows the mechanism: with $f = (x/s)^2 - 1$,

$$\mathcal{L}f = -\kappa x \cdot \frac{2x}{s^2} + \frac{1}{2}\sigma^2 \cdot \frac{2}{s^2} = -\frac{2\kappa x^2}{s^2} + 2\kappa = -2\kappa f,$$

using $\sigma^2/s^2 = 2\kappa$. The drift term lowers the degree and the diffusion term supplies exactly the constant needed to close the polynomial back on itself.

`quant/src/generator.rs` verifies this for the first five eigenfunctions on a grid, and then verifies the expansion (14.17) end to end against the exact Gaussian transition law integrated numerically — two routes sharing nothing but the definition of the process.

Remark (What the spectrum is good for)**.** Three things, in increasing order of how much they matter.

The expansion is cheap across maturities. The eigenfunctions and coefficients do not depend on t, so once computed, every maturity costs one exponential per mode. A model that must be re-solved for each expiry becomes one that is solved once.

It gives an honest error. Truncating after N modes leaves an error bounded by the next eigenvalue's decay, which is a statement one can make in advance — unlike a Monte Carlo standard error, which is a statement about the run one happened to do.

And the spectrum has a financial reading. The gap between the first two eigenvalues is the rate at which the process forgets where it started, so for an Ornstein-Uhlenbeck process the spectral gap *is* the mean reversion rate and $\ln 2/\lambda_1$ is the half-life. Chapter 20 spends a section on estimating that number badly. In this language it is estimating a spectral gap from a finite sample, and the difficulty is the usual one: the slowest mode is the one a short window sees least of.

Remark (The gap is an asymptotic rate)**.** A deviation from stationarity is a sum over modes,

$$P_t f - \mathbb{E}_\pi[f] \;=\; \sum_{n\geq 1} c_n\, e^{-\lambda_n t}\,\varphi_n,$$

and every term of it is decaying, the higher ones faster. So the total falls off *more quickly* than $e^{-\lambda_1 t}$ at first, and only settles to that rate once the higher modes have spent themselves. The gap

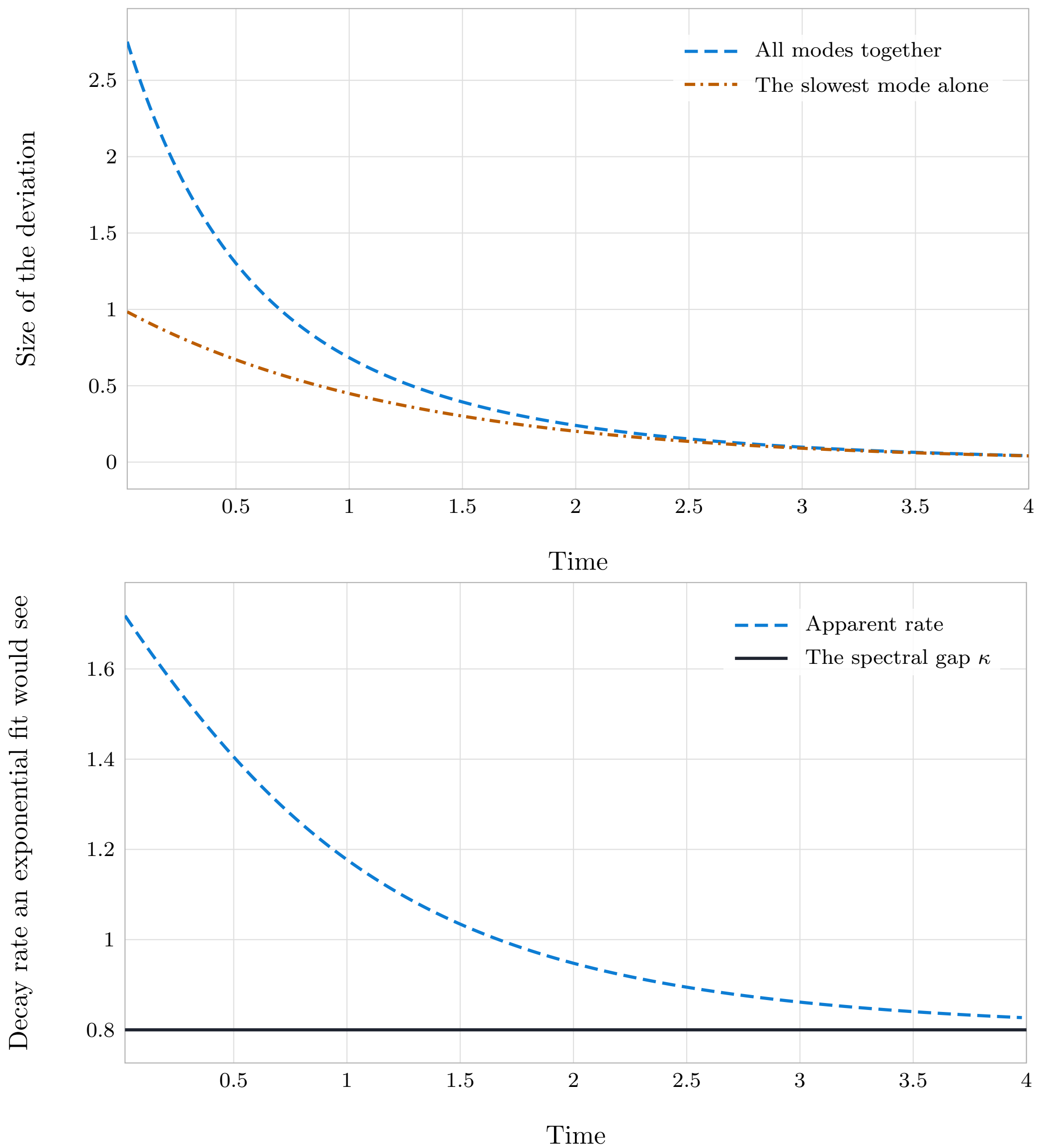

Figure 14.3: Above: a deviation from stationarity with weight on four modes, against the slowest mode alone. The total starts below the slowest mode and approaches it, because the higher modes are still alive early and decaying faster. Below: the decay rate an exponential fit would report if it began at that time, against the spectral gap it is supposed to be estimating. Early windows see a rate well above κ and the bias is one-sided — which is the same overstatement of mean reversion chapter 20 finds from the estimator's side, arriving here from the spectrum's.

Drawn by `ou_spectral`.

is the rate the decay approaches, not the rate at which the first halving happens; the first halving is quicker.

That distinction has teeth when a half-life is estimated from data. A sample of a given length contains the fast early relaxation as well as the slow tail, and fitting one exponential to the whole of it splits the difference — returning a rate above λ_1, which is to say overstating the mean reversion. It is a second and independent route to the bias chapter 20 measures, arriving from the spectrum rather than from the estimator.

14.6 What This Chapter Is For

When a model resists, ask which reduction has failed. If the state is infinite, no grid exists and the state must be enlarged or the model abandoned. If the state is finite but no family is preserved, the model is fine and merely needs computing, which is chapter 19. If a family is preserved, find out whether it is a subspace or a manifold, because that determines whether the equations are linear or Riccati.

References

- Duffie, D., Filipovic, D., & Schachermayer, W. (2003). Affine processes and applications in finance. *Annals of Applied Probability*, 13(3), 984–1053.

- Cuchiero, C., Keller-Ressel, M., & Teichmann, J. (2012). Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16(4), 711–740.

- Linetsky, V. (2004). The spectral decomposition of the option value. *International Journal of Theoretical and Applied Finance*, 7(3), 337–384.

- Forman, J. L., & Sorensen, M. (2008). The Pearson diffusions: a class of statistically tractable diffusion processes. *Scandinavian Journal of Statistics*, 35(3), 438–465.

- Reid, W. T. (1972). *Riccati Differential Equations.* Academic Press.