



AN INTRODUCTION TO LOCAL VOLATILITY MODELS

Jan-Frederik Mai
XAIA Investment GmbH
Sonnenstraße 19, 80331 München, Germany
jan-frederik.mai@xaia.com

First Version: February 4, 2015

This Version: July 26, 2017

Abstract The fundamental mathematical techniques underpinning the theory of local volatility models are reviewed and explained in simple terms. Moreover, it is outlined how the continuous-time models are put into practice via numerically solving pricing PDEs. It is explained how different pricing approaches relying on either backward or forward PDEs are implicitly interrelated in continuous and discrete time.

1 Introduction We denote by $\{F_t\}_{t \in [0, T]}$ the evolution of a forward strike price process with maturity T on some underlying asset. Moreover, we assume that interest rates are modeled deterministically and denote by $DF(t, T)$ the value at time t of a zero coupon bond with maturity T . Postulating absence of arbitrage, classical theory claims the existence of some risk-neutral pricing measure under which the stochastic process $\{F_t\}_{t \in [0, T]}$ is a local martingale. A local volatility model is a common choice for this local martingale. The forward price process under some risk-neutral measure, which we fix henceforth, is modeled as

$$dF_t = \sigma(t, F_t) dW_t, \quad (1)$$

where $\{W_t\}_{t \in [0, T]}$ denotes standard Brownian motion and σ is a bivariate, deterministic function, called the *local volatility function*. Throughout, we assume σ is nice enough so that the stochastic differential equation (1) defines a unique strong and non-negative solution and that F_t admits a density with respect to Lebesgue measure. The natural filtration of the Brownian motion is assumed to coincide with the market filtration, i.e. the observable information flow in the marketplace. At time $t \in [0, T]$ the value of a European derivative on the underlying asset with payoff $f(F_T)$ at time T is well-known by arbitrage pricing theory to equal the expected value of $DF(t, T) f(F_T)$ conditioned on the information available in the marketplace at time t . Since Brownian motion is a Markov process and σ is deterministic, the forward price process is Markovian. Consequently, the market does not need the full information available at time t in order to determine the expected value of $DF(t, T) f(F_T)$, but only the information about the current level F_t of the forward. This implies that the value V of the European derivative at time t is a function of two variables, time and forward, i.e.

$$V(t, F_t) := DF(t, T) \mathbb{E}[f(F_T) | F_t].$$



Our goal is to determine today's value $V(0, F_0)$ of the derivative, which is the (unconditional) expected value of $f(F_T)$, discounted back into time $t = 0$. The problem with the computation of this expected value is that, in general, the distribution of F_T is unknown, because it depends on the function σ . On the one hand, one could now opt for a specific parametric form of σ and then try to derive the probability law of F_T in closed form. Popular examples with known and tractable distributions of F_T comprise the classical lognormal Black model ($\sigma(t, F_t) = \sigma F_t$), a mixture of lognormal models (σ can be given explicitly in this case, cf. Brigo, Mercurio (2002)), and the constant elasticity of variance model ($\sigma(t, F_t) = \sigma F_t^\rho$ with $\rho \in (0, 1)$, cf. Delbaen, Shirakawa (2002)). On the other hand, one could stick with the general definition (1) for arbitrary σ and work non-parametrically or semi-parametrically. For example, one may try to reengineer the functional form of σ from observable prices of European stock derivatives. This is possible, since already for the general model (1) there is quite some mathematical theory available which helps to compute $V(0, F_0)$ efficiently by numerical methods. Pioneered by the seminal work of Dupire (1994), such an approach is outlined, e.g., in Andreasen, Huge (2011a). The present article surveys the mathematical concepts underpinning local volatility models. For further reading, the interested reader is referred to the standard textbook Gatheral (2006).

The remaining article is organized as follows. Section 2 discusses three famous partial differential equations (PDEs) in the context of local volatility models, their coherences and their motivations. Section 3 explains how the PDEs of Section 2 are put into practice by means of discretization, and how continuous-time coherences translate into discrete-time coherences via matrix algebra. Finally, Section 4 concludes.

2 Three useful PDEs We denote by $x \mapsto p(t, x)$ the density of the random variable F_t , for each $t \in (0, T]$. Since for $t = 0$ we know the value F_0 , the density $x \mapsto p(0, x)$ does not exist but the law of F_0 is a Dirac measure at the known value F_0 , denoted by δ_{F_0} in the sequel. Recall that we are interested in the computation of the value $V(0, F_0)$. Since the discount factor $DF(0, T)$ in $V(0, F_0)$ is assumed to be given, we can without loss of generality concentrate on the computation of the value $v(0, F_0)$, where we define $v(t, F_t) := V(t, F_t)/DF(t, T)$. The latter value can be written with arbitrary $t \in (0, T]$ as

$$\begin{aligned} v(0, F_0) &= \mathbb{E}[f(F_T)] = \mathbb{E}[\mathbb{E}[f(F_T) | F_t]] \\ &= \mathbb{E}[v(t, F_t)] = \int_0^\infty v(t, x) p(t, x) dx, \end{aligned} \quad (2)$$

where the tower property of conditional expectation has been used in the second equality. The last equation unravels a duality between v and p : for $t \searrow 0$, the probability law $p(t, x) dx$ approaches the one point mass $\delta_{F_0}(dx)$ at F_0 , i.e. for small t the function p becomes more and more "known". The situation is reversed for $t \nearrow T$, when the function $v(t, x)$ approaches the known function $v(T, x) = f(x)$, but the density p becomes more and more "unknown". Consequently, there are two conjugate strategies for



determining the desired value $v(0, F_0)$, either working forward in time and learning more and more about p , or working backward in time and learning more and more about v . These strategies are made precise by the following forward and backward PDEs for p and v . First, Subsection 2.1 treats the Fokker-Planck equation, which is a forward PDE for p . Second, Subsection 2.2 treats the classical backward pricing PDE for v . Third, Subsection 2.3 treats Dupire's forward PDE for the special case of a call option on the underlying, which differs from the backward pricing PDE in the sense that the underlying argument is replaced by the call strike price.

2.1 The Fokker–Planck equation

The Fokker–Planck equation is of paramount interest in physics, see, e.g., Risken (1989) for a comprehensive treatment. For our purpose, a special case of it states that

$$\frac{\partial}{\partial t} p(t, x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(t, x) p(t, x)). \quad (3)$$

Notice that Equation (3) can be solved forwards in time together with the boundary condition $p(t, x) dx \searrow \delta_{F_0}(dx)$. When solving the equation numerically in discrete time, each continuous measure $p(t, x) dx$ must be approximated by a discrete measure anyway, so that the one-point mass start value is notationally more convenient than in the above continuous-time formulation. In particular, each such family of discrete approximations for the measures $p(t, x) dx$ defines a discrete-time stochastic process approximating the continuous-time process $\{F_t\}_{t \in [0, T]}$. Viewed this way, finding $v(0, F_0)$ via *tree pricing* is closely connected to the procedure of discretizing the Fokker–Planck equation. A rigorous derivation of the Fokker–Planck equation is technical, but a heuristic explanation can be given as follows: fix an arbitrary Borel set B and heuristically¹ apply Itô's formula to the indicator function $I(F_t) := 1_{\{F_t \in B\}}$. This yields

$$dI(F_t) = I'(F_t) \sigma(t, F_t) dW_t + \frac{1}{2} \sigma^2(t, F_t) I''(F_t) dt.$$

Further assuming that we may interchange expectation and differentiation, we obtain

$$d \mathbb{E}[I(F_t)] = \mathbb{E}[dI(F_t)] = \frac{1}{2} \mathbb{E}[\sigma^2(t, F_t) I''(F_t)] dt,$$

or, put differently,

$$\begin{aligned} \int_B \frac{\partial}{\partial t} p(t, x) dx &= \frac{\partial}{\partial t} \int_B p(t, x) dx = \frac{\partial}{\partial t} \mathbb{E}[I(F_t)] \\ &= \frac{1}{2} \int_0^\infty I''(x) \sigma^2(t, x) p(t, x) dx. \end{aligned}$$

Applying now twice integration by parts and assuming that the integrand vanishes at the boundaries $x = 0$ and $x = \infty$, we

¹Of course, I is not twice differentiable, so in order to carry out this argument rigorously, one would need to approximate I by smooth functions and verify the convergence of the approximations.



observe

$$\begin{aligned} \int_0^\infty I''(x) \sigma^2(t, x) p(t, x) dx &= - \int_0^\infty I'(x) \frac{\partial}{\partial x} (\sigma^2(t, x) p(t, x)) dx \\ &= \int_0^\infty I(x) \frac{\partial^2}{\partial x^2} (\sigma^2(t, x) p(t, x)) dx = \int_B \frac{\partial^2}{\partial x^2} (\sigma^2(t, x) p(t, x)) dx. \end{aligned}$$

Putting together the pieces, we conclude for an arbitrary Borel set B that

$$\int_B \frac{\partial}{\partial t} p(t, x) - \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(t, x) p(t, x)) dx = 0.$$

2.2 The backward pricing PDE

As a direct extension of the classical Black–Scholes PDE², the backward pricing PDE states that

$$\frac{\partial}{\partial t} v(t, x) = -\frac{1}{2} \sigma^2(t, x) \frac{\partial^2}{\partial x^2} v(t, x). \quad (4)$$

Notice that Equation (4) can be solved backwards in time together with the boundary condition $v(T, x) = f(x)$.

The derivation of Formula (4) follows directly from Itô's formula, which states that

$$dv(t, F_t) = \left(\frac{\partial}{\partial t} v(t, F_t) + \frac{1}{2} \sigma^2(t, F_t) \frac{\partial^2}{\partial x^2} v(t, F_t) \right) dt + \dots dW_t.$$

By definition, the stochastic process $t \mapsto v(t, F_t) = \mathbb{E}[f(F_T) | F_t]$ is a (closed) martingale, implying that the drift term in the last stochastic differential equation vanishes. This yields Equation (4).

2.3 Dupire's forward PDE

Dupire's forward PDE, which is named after the seminal reference Dupire (1994), is specifically designed for European call options, since these are the best-understood and most liquid options. To this end, we denote

$$c(t, k) := \mathbb{E}[(F_t - k)_+] = \int_k^\infty (x - k) p(t, x) dx$$

and mention that $DF(0, T) c(T, k)$ equals the desired value of a call option on the underlying with maturity T and strike price k . Dupire's forward PDE states that

$$\frac{\partial}{\partial t} c(t, k) = \frac{1}{2} \sigma^2(t, k) \frac{\partial^2}{\partial k^2} c(t, k). \quad (5)$$

Notice that Equation (5) can be solved forwards in time together with the boundary condition $c(0, k) = (F_0 - k)_+$. The core application of Dupire's forward PDE is that it might be used in order to bootstrap the function σ from observed market prices for call options. The article Andreasen, Huge (2011a) provides a recipe for how this might be done in a numerically efficient and arbitrage-consistent way.

²Notice that the classical Black–Scholes PDE involves also an interest rate, which we conveniently got rid of by concentrating on the forward process rather than the stock price process.



The derivation of Dupire's PDE works as follows. Using the parameter integral differentiation rule, it is readily observed that $\frac{\partial^2}{\partial k^2} c(t, k) = p(t, k)$. Using the Fokker-Planck equation (3) and a heuristic interchange of integration and differentiation, the derivative of $c(t, k)$ with respect to t may be written as

$$\begin{aligned} \frac{\partial}{\partial t} c(t, k) &= \int_k^\infty (x - k) \frac{\partial}{\partial t} p(t, x) dx \\ &= \int_k^\infty (x - k) \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(t, x) p(t, x)) dx = (*). \end{aligned}$$

Via integration by parts and using that $\sigma^2(t, x) p(t, x)$ and its x -derivative vanish rapidly as $x \rightarrow \infty$, we compute further that

$$\begin{aligned} (*) &= -\frac{1}{2} \int_k^\infty \frac{\partial}{\partial x} (\sigma^2(t, x) p(t, x)) dx \\ &= \frac{1}{2} \sigma^2(t, k) p(t, k) = \frac{1}{2} \sigma^2(t, k) \frac{\partial^2}{\partial k^2} c(t, k). \end{aligned}$$

3 Implementation in discrete time

When actually computing $v(0, F_0)$ in practice using one of the PDEs mentioned in the previous section, required is a discretization of the respective continuous-time PDE. To this end, standard approaches rely on the replacement of the appearing differential operators by their respective discrete-time difference operators. This typically leads to iterative linear equation systems involving well-behaved tridiagonal matrices, for which numerically efficient algorithms are known. On the one hand, any discretization of the Fokker-Planck PDE naturally leads to a trinomial tree approximation of the continuous-time financial model (1). This has the advantage that the approach is rigorously embedded into arbitrage pricing theory and the induced discrete-time model may consistently be used for exotic derivatives, e.g. as the basis for a Monte Carlo simulation of the model. On the other hand, a discretization of the backward pricing PDE does not explicitly involve a discrete-time financial model, but only implicitly via the duality between Fokker-Planck and the backward PDE that follows from (2). Andreasen, Huge (2011b) describe a discretization procedure which achieves consistency of both forward and backward discretization schemes. This is important, especially when the same model is applied to the pricing of different types of derivatives and, depending on the specific claim, either one or the other approach is advantageous. It is educational to once work through the discrete math and make explicit the discrete-time financial model that is implicitly used when discretizing the backward pricing PDE. This is done in the sequel.

3.1 Discretization of backward PDE

Discretization requires to introduce some notation, so we consider a finite grid $0 = t_0 < t_1 < t_2 < \dots < t_m = T$ for the time variable as well as a finite, equidistant grid $0 = x_0 < x_1 < x_2 < \dots < x_n$ for the state variable, where x_n should be sufficiently large. Our functions c, v and p in two variables then become $(n + 1) \times (m + 1)$ -matrices. All differential operators appearing in the PDEs above are replaced with their discretized counterparts. More precisely, for all indices $i = 0, \dots, m - 1$



and $j = 0, \dots, n - 1$ we denote $\Delta x := x_{j+1} - x_j = x_1$ and $\Delta t_i := t_{i+1} - t_i$. Then we introduce the row vectors

$$\delta_{xx}^{(j)} := \frac{1}{(\Delta x)^2} \left(\underbrace{0, \dots, 0}_{0, \dots, j-2}, \underbrace{1}_{j-1}, \underbrace{-2}_j, \underbrace{1}_{j+1}, \underbrace{0, \dots, 0}_{j+2, \dots, n} \right)$$

for $j = 1, \dots, n - 1$. For an arbitrary function h in x we denote by $\vec{h} := (h(x_0), \dots, h(x_n))'$, so that

$$\delta_{xx}^{(j)} \vec{h} = \frac{h(x_{j-1}) - 2h(x_j) + h(x_{j+1}))}{(\Delta x)^2}$$

provides a discrete analog of $h''(x_j)$ according to our grid. We are going to apply this with the functions $h(x) = v(t_i, x)$ below. For $j = 0$ and $j = n$, the definitions of $\delta_{xx}^{(0)}$, respectively of $\delta_{xx}^{(n)}$, requires special attention, because one has to define boundary conditions which depend on the specific derivative in concern (because the required values x_{-1} and x_{n+1} do not exist). For example, if a call option is considered, one might³ set $\delta_{xx}^{(0)} = \delta_{xx}^{(n)} = (0, \dots, 0)$. This intuitively means that for very small forward values (when the call price equals effectively zero) and for very large forward values (when the call price is effectively linear in the forward) the price of the call option is well approximated by linear functions. However, depending on the derivative in concern, one might not be able to avoid careful deliberations on how to impose boundary conditions. Finally, we also write δ_{xx} for the $(n + 1) \times (n + 1)$ -matrix with rows $\delta_{xx}^{(j)}$, $j = 0, \dots, n$.

We now solve the backward pricing PDE for $\vec{v}(t_0)$ in discrete time. To this end, we work backwards in time from the starting condition $\vec{v}(t_m) = \vec{f}$. For each $i = 0, \dots, m$ we denote by $\Sigma(t_i) \in \mathbb{R}^{(n+1) \times (n+1)}$ a matrix with all off-diagonal elements zero and diagonal being given by the vector $(\sigma(t_i, x_0), \dots, \sigma(t_i, x_n))$. Discretization of the PDE (4) applying the second difference operator matrix δ_{xx} yields the equations⁴

$$\vec{v}(t_{i+1}) = \vec{v}(t_i) - \Delta t_i \frac{1}{2} \Sigma^2(t_i) \delta_{xx} \vec{v}(t_i), \quad i = 0, \dots, m - 1.$$

Since we need to work backwards in time starting from $\vec{v}(t_m)$ it is appropriate to rearrange the last equation for $\vec{v}(t_i)$ yielding

$$\vec{v}(t_i) = \left(I_{n+1} - \Delta t_i \frac{1}{2} \Sigma^2(t_i) \delta_{xx} \right)^{-1} \vec{v}(t_{i+1}), \quad i = 0, \dots, m - 1, \quad (6)$$

with $I_{n+1} \in \mathbb{R}^{(n+1) \times (n+1)}$ denoting the identity matrix. By construction, the matrices

$$A(t_i) := I_{n+1} - \Delta t_i \frac{1}{2} \Sigma^2(t_i) \delta_{xx}$$

³It is more common to impose a so-called Dirichlet boundary condition at x_0 , because the value of the call option is known to equal zero in that case. However, these deliberations (and the resulting more involved notation) is beyond the scope of this introduction.

⁴The presented discretization corresponds to a fully implicit scheme, since $\vec{v}(t_i)$ is the unknown. Alternatively, one might also use a fully explicit or a convex mixture of fully implicit and explicit method, e.g. the Crank-Nicholson method, in order to improve numerical stability.



are tridiagonal and have an inverse, since they are obviously diagonally dominant. Furthermore, $A(t_i)^{-1}$ is a probability transition matrix, i.e. has non-negative entries and unit row sums. Indeed, since $A(t_i)$ is a tridiagonal matrix whose row sums equal one by construction, i.e. $A(t_i) \vec{1} = \vec{1}$, the same is valid for its inverse, i.e. $A(t_i)^{-1} \vec{1} = \vec{1}$. That $A(t_i)^{-1}$ has only non-negative entries follows immediately from one of the characterizations in Plemmons (1977), e.g. characterization K_{34} with $\vec{x} = \vec{1}$.

3.2 Implicit discrete-time model

Next, we explain how the matrices $A(t_i)$ from the backward PDE solution induce a discrete-time tree approximation for $\{F_t\}_{t \in [0, T]}$ via the conjugate Fokker–Planck forward equation. Conveniently assume that the current forward value F_0 is contained in our grid for the state space⁵, say $x_k = F_0$. Then $\vec{p}(t_0)$ equals the $(n+1)$ -dimensional vector with entries $1_{\{j=k\}}$, $j = 0, \dots, n$, i.e. is a unit vector. We now claim that a tree approximation for the forward process is given by defining the discretization of p iteratively via

$$\vec{p}(t_{i+1}) := (A(t_i)^{-1})' \vec{p}(t_i), \quad i = 0, \dots, m-1. \quad (7)$$

Recall that $A(t_i)^{-1}$ is a probability transition matrix, which implies that each $\vec{p}(t_i)$ defines a discrete probability distribution on the states $\{x_0, \dots, x_n\}$, so that the sequence $\vec{p}(t_0), \dots, \vec{p}(t_m)$ defines a discrete-time stochastic process $\{F_{t_i}^{(d)}\}_{i=0, \dots, m}$. Using the notation \vec{x} for $(x_0, \dots, x_n)'$ and \vec{e}_k for the unit vector with components $1_{\{j=k\}}$, $j = 0, \dots, n$, it follows for $j = 0, \dots, n$ that

$$\mathbb{E}[F_{t_{i+1}}^{(d)} | F_{t_i}^{(d)} = x_j] = \vec{x}' \underbrace{((A(t_i)^{-1})' \vec{e}_j)}_{\substack{\text{1-step transition} \\ \text{probabilities}}} = \vec{e}_j' A(t_i)^{-1} \vec{x} = x_j,$$

showing that $\{F_{t_i}^{(d)}\}_{i=0, \dots, m}$ defines a martingale. Notice that the last equality uses (6) with payoff function $f(x) = x$, i.e. $\vec{v}(t_i) = \vec{x}$ for arbitrary i . In other words, \vec{x} is an Eigenvalue for $A(t_i)^{-1}$. Furthermore, observe for $i = 0, \dots, m-1$ that

$$\vec{p}(t_i)' \vec{v}(t_i) = \vec{p}(t_{i+1})' A(t_i) \vec{v}(t_i) = \vec{p}(t_{i+1})' \vec{v}(t_{i+1}),$$

which is the discrete-time analog of (2). Consequently, (7) defines a valid discrete-time financial market model for the forward process. Does it approximate the continuous-time forward model (1) as the mesh of the x - and t -grids approach zero? Indeed, this can be seen as follows: consider as payoff f an arbitrary continuous and bounded function. Then by construction we have that

$$\mathbb{E}[f(F_T)] = v(0, F_0) \approx \sum_{j=0}^n p(t_m, x_j) f(x_j) = \mathbb{E}[f(F_T^{(d)})],$$

where the approximation becomes an equality in the limit as the mesh of state- and time-grid become zero, because the difference operators tend to their respective differential operators. This proves distributional equality in the limit of F_T and $F_T^{(d)}$.

⁵If not, choose the nearest grid point. This inaccuracy is resolved in the limit as the mesh of the grid becomes finer and finer.



Since the process $\{F_t\}_{t \in [0, T]}$ is a closed martingale with terminal value F_T , and the same holds for the discrete-time martingale $\{F_{t_i}^{(d)}\}_{i=0, \dots, m}$, considering only the final time horizon T is sufficient for distributional equality on the process level.

Summarizing, we have seen in this subsection that the aforementioned discrete solution of the backward pricing PDE implicitly relies on a discrete-time financial model, where the respective discrete approximation of the forward process can be obtained from the matrices $A(t_i)$ introduced above via the iteration (7). The latter may itself be viewed as a discretization of the Fokker–Planck forward PDE which is by definition fully consistent with the discretization of the backward pricing PDE.

4 Conclusion Three fundamental pricing PDEs in the context of local volatility models have been introduced, and their derivations have been sketched. Finally, it has been explained how a numerical scheme for solving a backward pricing PDE implicitly induces a discrete-time financial model, which can be written down explicitly using matrix algebra.

- References**
- J. Andreasen, B. Høge, Volatility interpolation, *Risk magazine* (March 2011) pp. 76–79.
 - J. Andreasen, B. Høge, Random grids, *Risk magazine* (July 2011) pp. 66–71.
 - D. Brigo, F. Mercurio, Lognormal-mixture dynamics and calibration to market volatility smiles, *International Journal of Theoretical and Applied Finance* **5** (2002) pp. 427–446.
 - F. Delbaen, H. Shirakawa, A note on option pricing for the constant elasticity of variance model, *Asia-Pacific Financial Markets* **9:2** (2002) pp. 85–99.
 - B. Dupire, Pricing with a smile, *Risk magazine* (January 1994) pp. 18–20.
 - J. Gatheral, The volatility surface – a practitioner’s guide, *Wiley & Sons, Inc., Hoboken, New Jersey* (2006).
 - R.J. Plemmons, M-matrix characterizations. I - nonsingular M-matrices, *Linear Algebra and its Applications* **18** (1977) pp. 175–188.
 - H. Risken, The Fokker–Planck equation – methods of solution and applications, 2nd edition, 3rd printing *Springer-Verlag, Berlin, Heidelberg, New York* (1989).