Chapter 7 Beyond Black and Scholes

The Black–Scholes (BS) model for the value $V(S,t)$ of a vanilla option is based on some assumptions on the market. In particular, the BS model assumes the price S_t of the asset on which the option is written to follow a geometric Brownian motion with a constant volatility σ. Further, transaction costs are neglected, and trading of the underlying is supposed to have no influence on the price S_t. As has been discussed extensively, the value function $V(S,t)$ for standard options ("plain vanilla") of the European type, satisfies the Black–Scholes equation (1.2),

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0\,. \tag{BSE}$$

Solutions of this linear equation are subject to the terminal condition $V(S,T) = \Psi(S)$, where Ψ defines the payoff.

The BS-model is the core example of a complete market. In these idealized markets, the risk exposure to variations in the underlying can be hedged away. The corresponding risk strategy is unique. Hence vanilla options modeled by Assumption 1.2 have a unique price, given by the costs of the replication strategy (⟶ Appendix A4). Essentially, Chapters 4 through 6 have applied numerical methods to complete markets.

For the more realistic incomplete markets, there are no perfect hedges, and a risk remains. Each hedging strategy leads to a specific model with its own price [ConT04]. The hedger compensates the remaining risk in incomplete markets by charging an additional risk premium. Hence the value function or expected value is not the price for which the option is sold. Depending on the way how the comfortable assumption of completeness of the BS-market is lost, different models are set up, calling for different numerical approaches. This Chapter 7 is devoted to computational tools for incomplete markets.

Relaxing several of the assumptions of the Black–Scholes market, *nonlinear* extensions of the BS equation can be derived. These "nonlinear Black–Scholes type equations" are of the form

$$\frac{\partial V}{\partial t} + \frac{1}{2}\left[\hat{\sigma}(S,t,\frac{\partial^2 V}{\partial S^2})\right]^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0\,. \tag{7.1}$$

In this class of models, the volatility $\hat{\sigma}$ is a function that may incorporate several types of nonlinearity. The standard PDE (BSE) is included for $\hat{\sigma} \equiv \sigma$.

R.U. Seydel, *Tools for Computational Finance*, Universitext,
DOI 10.1007/978-1-4471-2993-6_7, © Springer-Verlag London Limited 2012

In Section 7.1 we describe three scenarios leading to three different functions $\hat{\sigma}$ of the volatility. A nonlinear PDE as (7.1) requires special numerical treatment, which will be the focus of Section 7.2.

Another stream of research beyond Black and Scholes is devoted to jump processes (Section 7.3). One of the numerical approaches is based on partial integro-differential equations (PIDE). Some highly efficient methods apply the Fourier transform; a basic approach is discussed in Section 7.4.

7.1 Nonlinearities in Models for Financial Options

In this section we briefly discuss three sources of nonlinearity in (7.1). We start with transaction costs based on Leland's approach [Lel85], and touch the more sophisticated model of Barles and Soner [BaS98]. Then we turn to specifying ranges of volatility. Finally we address the feedback by market illiquidity.

7.1.1 Leland's Model of Transaction Costs

Basic for the Black–Scholes model is the idea of rebalancing the portfolio continuously. But in financial reality this continuous trading would cause arbitrarily high trading costs. Keeping transaction costs low forces to abandon the optimal Black–Scholes hedging. But without the ideal BS hedging, the model suffers from hedging errors. To compromise, the hedger searches a balance between keeping both the transaction costs low *and* the hedging errors low.

Suppose that instead of rebalancing continuously, trading is only possible at discrete time instances with time step Δt apart (Δt fixed and finite). We assume a transaction cost rate proportional to the trading volume νS:

$$\text{trading } \nu \text{ assets costs the amount } c|\nu|S$$

for some cost parameter c.

Here we sketch a heuristic derivation of a model due to [Lel85], [HoWW94]. The discussion of this model parallels that for the Black–Scholes model, now adapted to the discrete scenario.[1] The stochastic changes of the asset with price S and of a riskless bond with price B are

$$\begin{aligned} \Delta S &= \mu S \Delta t + \sigma S \Delta W \\ \Delta B &= r B \Delta t \,. \end{aligned}$$

The portfolio with value Π is taken in the form

[1] All other BS-assumptions remain untouched [Kwok98]. The following analysis uses or modifies Appendix A4 with (A4.1), (A4.3), (A4.8). Δ means the increment, and *not* the greek $\frac{\partial V}{\partial S}$.

$$\Pi = \alpha S + \beta B\,,$$

with α units of the asset and β units of the bond. Suppose the portfolio is self-financing in the sense $S\Delta\alpha + B\Delta\beta = 0$, which is sufficient for $\Delta\Pi = \alpha\Delta S + \beta\Delta B$. Further assume that trading is such that the portfolio Π replicates the value of the option.

By definition, $\nu = \Delta\alpha$. After one time interval, $\nu = \Delta\alpha$ assets are traded, with transaction costs $cS|\Delta\alpha|$. The change in the value of the portfolio is

$$\begin{aligned}\Delta\Pi &= \alpha\Delta S + \beta\Delta B - cS|\Delta\alpha| \\ &= (\alpha\mu S + \beta rB)\Delta t + \alpha\sigma S\Delta W - cS|\Delta\alpha|\,.\end{aligned}$$

Let V be the value function of the option. Itô's lemma adapted to the discrete scenario gives

$$\Delta V = \frac{\partial V}{\partial S}\,\Delta S + \left(\frac{\partial V}{\partial t} + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2}\right)\Delta t\,.$$

By the no-arbitrage principle $\Delta V = \Delta\Pi$ holds for the replicating and self-financing portfolio. And coefficient matching will give further information. But first let us approximate the $\Delta\alpha$-term.

From BS theory we expect $\alpha \approx \frac{\partial V}{\partial S}$. So $\nu = \Delta\alpha$ will be approximated by

$$\begin{aligned}&\frac{\partial V(S+\Delta S, t+\Delta t)}{\partial S} - \frac{\partial V(S,t)}{\partial S} \\ &= \frac{\partial^2 V(S,t)}{\partial S^2}\,\Delta S + \frac{\partial^2 V(S,t)}{\partial S\,\partial t}\,\Delta t + \text{t.h.o.}\,,\end{aligned}$$

invoking Taylor's expansion. After substituting ΔS we realize that the term of lowest order is

$$\sigma S\frac{\partial^2 V(S,t)}{\partial S^2}\,\Delta W\,.$$

In summary, the transaction costs in $\Delta\Pi$ can be approximated by

$$-cS|\Delta\alpha| = -c\sigma S^2\left|\frac{\partial^2 V(S,t)}{\partial S^2}\right|\,|\Delta W| + \text{t.h.o.}\,,$$

which is path-dependent. Leland [Lel85] boldly suggested to approximate $|\Delta W| \approx \mathsf{E}(|\Delta W|)$. Exercise 7.1 tells

$$\mathsf{E}(|\Delta W|) = \sqrt{\Delta t}\,\sqrt{\frac{2}{\pi}}\,.$$

In this way, the trading cost term $-cS\,|\Delta\alpha|$ is approximated by the deterministic expression

$$-c\sigma S^2\left|\frac{\partial^2 V(S,t)}{\partial S^2}\right|\,\sqrt{\Delta t}\,\sqrt{\frac{2}{\pi}}\,. \tag{7.2}$$

This may be seen as further assumption, motivated by the above arguing. The approximation (7.2) of the transaction costs and its artificial parameter $\sqrt{2/\pi} \approx 0.8$ reflect the lack of a unique price in incomplete markets.

With this somewhat artificial approximation (7.2) of the trading costs $-cS|\alpha|$, coefficient matching of $\Delta V = \Delta\Pi$ leads to match the remaining stochastic terms,

$$\alpha\sigma S\Delta W = \sigma S\frac{\partial V}{\partial S}\,\Delta W\,,$$

or $\alpha = \frac{\partial V}{\partial S}$, which is the famous "delta hedging," consistent with the modeling of $\Delta\alpha$ above. The remaining terms are deterministic. Use $\beta B + S\frac{\partial V}{\partial S} = \Pi = V$ to obtain

$$\begin{aligned}&\left(\mu S\frac{\partial V}{\partial S} + rV - rS\frac{\partial V}{\partial S}\right)\,\Delta t - cS|\Delta\alpha| \\ &= \left(\frac{\partial V}{\partial t} + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2} + \mu S\frac{\partial V}{\partial S}\right)\,\Delta t\end{aligned} \tag{7.3}$$

The μ-terms cancel out. (7.3) with transaction costs replaced by (7.2) lead to the variant of the Black–Scholes equation. With the coefficient

$$\gamma := \sqrt{\frac{2}{\pi}}\left(\frac{2c}{\sigma\sqrt{\Delta t}}\right) \tag{7.4}$$

the resulting equation is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\sigma^2 S^2\gamma\left|\frac{\partial^2 V}{\partial S^2}\right| + rS\frac{\partial V}{\partial S} - rV = 0 \tag{7.5}$$

Formally, this becomes the standard Black–Scholes equation with a modified volatility

$$\hat{\sigma}^2(\Gamma) := \sigma^2[1 + \gamma\ \mathrm{sign}(\Gamma)]\,, \tag{7.6}$$

with $\Gamma := \frac{\partial^2 V}{\partial S^2}$. For convex payoff, this amounts to augment the volatility to a constant $\hat{\sigma} > \sigma$ (Leland's scenario). In this case Γ does not change sign, and the PDE (7.5) is again linear. But note that for instance for barrier options, Γ does change sign, and the PDE is nonlinear and of the general type of equation (7.1). For $c = 0$ (no transaction costs) (7.5) specializes to the BS-equation. To have a well-posed PDE, Δt must be such that $\gamma < 1$. In particular, $\Delta t \to 0$ does not make sense.

7.1.2 The Barles and Soner Model of Transaction Costs

Barles and Soner [BaS98] assume a price process $\mathrm{d}S_t = S_t(\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t)$, with constant volatility σ, $0 \le t \le T$, and model transactions using the following variables:

α_t shares of the asset with price S_t,
β_t shares of the bond,

L_t cumulative transfer form cash to stock, nondecreasing, $L(0) = 0$,
M_t cumulative transfer from stock to cash, nondecreasing, $M(0) = 0$.

Consequently,

$$\begin{aligned}\alpha_t &= \alpha_0 + L_t - M_t \\ \beta_t &= \beta_0 - \int_0^t S_\tau \cdot (1+c)\,\mathrm{d}L_\tau + \int_0^t S_\tau \cdot (1-c)\,\mathrm{d}M_\tau + \int_0^t r\beta_\tau\,\mathrm{d}\tau\end{aligned} .$$

That is, in both cases buying and selling of stocks, transaction costs $\int S_\tau c$ are charged to β, where c again denotes proportional transaction costs. The further derivation of [BaS98] is based on a utility function. The final result is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \cdot \left[1 + f\left(\mathrm{e}^{r(T-t)} a^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\right] + rS\frac{\partial V}{\partial S} - rV = 0 \quad (7.7)$$

where a is a parameter representing proportional transaction costs and risk aversion. The function f is the unique solution of the ODE

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{f(x)+1}{2\sqrt{xf(x)} - x} \quad \text{with} \quad f(0) = 0 .$$

The resulting function f is singular at $x = 0$ ($\longrightarrow$ Exercise 7.2). Figure 7.1 shows the difference between the BS-solution and the solution of the corresponding nonlinear model (7.7).

7.1.3 Specifying a Range of Volatilities

The two above models of transaction costs come up with a nonlinear volatility function $\hat{\sigma}(\Gamma)$. Usually this function is not known, and is subject to speculation (modeling). It will be easier to specify a *range* of volatility, assuming that $\hat{\sigma}$ lies within an interval or band

$$0 < \sigma_{\min} \le \sigma \le \sigma_{\max} < 1 .$$

This is the uncertain-volatility model of [AvP94], [AvLP95], [Lyo95].

The derivation starts as above, leading to (7.3) with $c = 0$. (Transaction costs are not considered here.) Formally, the result is the Black–Scholes equation (BSE), except that σ is no constant, but is considered as a stochastic variable $\sigma(t)$:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma(t)^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 .$$

This is a PDE with stochastic control parameter $\sigma(t)$. There is an ambitious theory for such controlled diffusion processes, see the monograph [Kry80]. To avoid the use of this methodology, we adopt a simplified arguing, similar as in [Wil98].

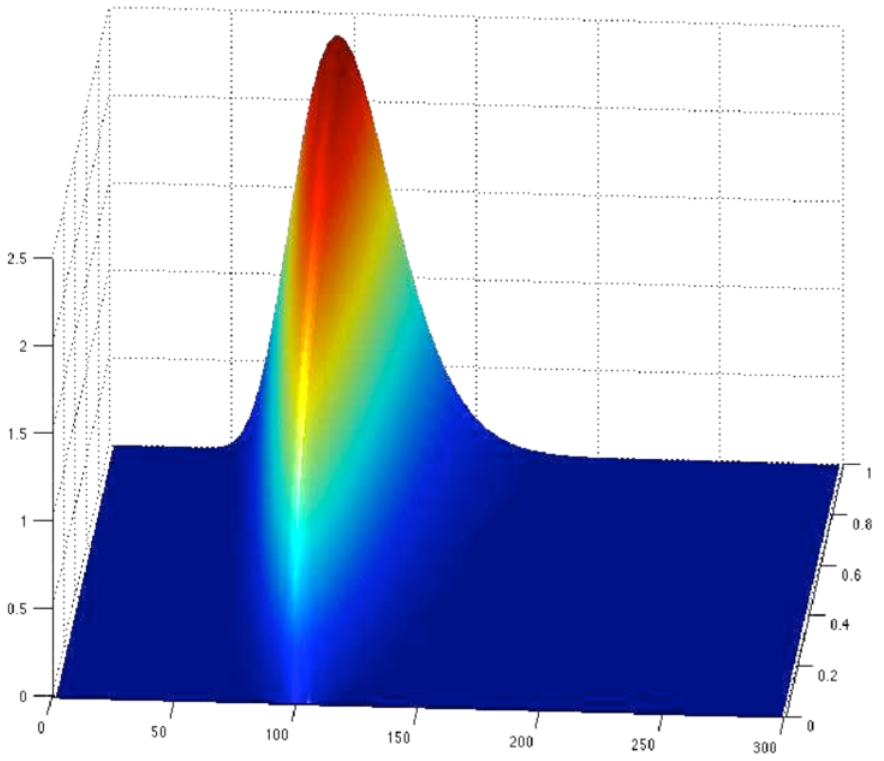

Fig. 7.1. $V(S, T-t)$: difference between solutions of (BSE) and (7.7); $K = 100, r = 0.1, \sigma = 0.2, a = 0.02, T = 1$. With kind permission of Pascal Heider.

Using an argumentation of Black and Scholes, we construct a portfolio of one option (value V), and hedge it with $-\alpha$ units of the underlying asset,

$$\Pi = V - \alpha S \, .$$

Assuming a change in the value of this portfolio in the form $\Delta\Pi = \Delta V - \alpha \Delta S$, we have as above

$$\Delta\Pi = \frac{\partial V}{\partial S}\Delta S + \left(\frac{\partial V}{\partial t} + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2}\right)\Delta t - \alpha\Delta S \, .$$

The choice $\alpha = \frac{\partial V}{\partial S}$ eliminates the risk represented by the ΔW-terms. This results in

$$\Delta\Pi = \left(\frac{\partial V}{\partial t} + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2}\right)\Delta t \, .$$

Note that the return $\Delta\Pi$ of the portfolio still depends on the unknown stochastic $\sigma(t)$.

We now define artificially two specific functions $\sigma^+(t)$ and $\sigma^-(t)$ chosen such that the return $\Delta\Pi(\sigma)$ increases by the maximum amount, or by the least amount:

$\sigma^+(t)$ chosen such that $\Delta\Pi(\sigma^+)$ is a maximum,
$\sigma^-(t)$ chosen such that $\Delta\Pi(\sigma^-)$ is a minimum.

These returns reflect the best case and the worst case as seen by the holder. For every function $\sigma(t)$ the no-arbitrage principle holds. Hence both cases $\sigma^+(t)$ and $\sigma^-(t)$ result in a return $\Delta\Pi = r\Pi\Delta t$. This can be summarized as

$$\begin{aligned} \sigma^+ \text{ maximizes} \quad & \max_{\sigma_{\min}\le\sigma\le\sigma_{\max}} \Delta\Pi(\sigma) = r\Pi\Delta t \\ \sigma^- \text{ minimizes} \quad & \min_{\sigma_{\min}\le\sigma\le\sigma_{\max}} \Delta\Pi(\sigma) = r\Pi\Delta t \end{aligned}$$

In view of the expression for $\Delta\Pi(\sigma)$, the two artificial functions σ^+, σ^- enter via the term

$$\sigma^2 \frac{\partial^2 V}{\partial S^2}$$

For $\Delta\Pi$ to become a maximum or minimum, σ^+ (or σ^-) will equal $\sigma_{\min}$ or $\sigma_{\max}$, depending on the sign of $\Gamma = \frac{\partial^2 V}{\partial S^2}$. To become a maximum, set

$$\sigma^+(\Gamma) := \begin{cases} \sigma_{\max} & \text{if } \Gamma \ge 0 \\ \sigma_{\min} & \text{if } \Gamma < 0 \end{cases} . \tag{7.8a}$$

And to become a minimum, set

$$\sigma^-(\Gamma) := \begin{cases} \sigma_{\max} & \text{if } \Gamma < 0 \\ \sigma_{\min} & \text{if } \Gamma \ge 0 \end{cases} . \tag{7.8b}$$

This defines two specific control functions σ, which after substitution into the PDE $\Delta\Pi(\sigma) = r\Pi\Delta t$ yields *two* nonlinear PDEs

$$\frac{\partial V}{\partial t} + \frac{1}{2}\hat{\sigma}(\Gamma)^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 , \tag{7.9}$$

with $\hat{\sigma} = \sigma^+$ and $\hat{\sigma} = \sigma^-$ from (7.8). Let us denote the corresponding solutions V^+ and V^-. Since σ^+ yields the maximum return, we expect $V \le V^+$, and similarly, $V^- \le V$. This provides the range $V^- \le V \le V^+$ for the option price.

In the special case of vanilla options, the convexity of $V(S,.)$ implies $\Gamma \ge 0$ and hence $\sigma^+ = \sigma_{\max}$ and $\sigma^- = \sigma_{\min}$; the nonlinearity is not effective then. The monotonicity of V with respect to σ is clear for vanilla options, but is not valid, for example, for barrier options. And convexity of $V(S,.)$ is lost for barrier options, butterfly spreads, digital options, and many other options [PoFV03]. The great potential of the uncertain-volatility model is illustrated by Figure 7.2. For the example of a butterfly option, and an uncertainty interval $0.15 \le \sigma \le 0.25$ we show the band $V^- \le V \le V^+$, with two

Black–Scholes curves therein. The payoff of a butterfly spread is illustrated schematically in Figure 1.24(d), see also Exercise 7.4. The functions V^-, V^+ were calculated with the methods explained in Section 7.2. For barrier options, the success of the method is doubtful because of the high sensitivity w.r.t. σ close to the barrier. Then the bandwidth may be so large that it is not of practical use. Such an example is shown in Figure 7.3.

7.1.4 Market Illiquidity

As pointed out by [FrS97], [ScW00], [FrP02], the assumption that a big investor can trade large amounts of an asset without affecting its price, is not realistic. There will be a feedback, and the assumption of an infinite market liquidity may fail. [FrS97], [ScW00] introduce a market liquidity parameter λ, with $0 \le \lambda \le 1$, and derive the nonlinear PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\sigma^2 S^2}{(1-\lambda\frac{\partial^2 V}{\partial S^2})^2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0\,. \tag{7.10}$$

Here we do not discuss further details. Note that this model is also of the form of equation (7.1).

7.2 Numerical Solution of Nonlinear Black–Scholes Equations

All the nonlinear PDEs of Section 7.1 fall under the general type of equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\hat{\sigma}^2(S,t,\frac{\partial^2 V}{\partial S^2})S^2\frac{\partial^2 V}{\partial S^2} + (r-\delta)S\frac{\partial V}{\partial S} - rV = 0\,, \tag{7.11}$$

which we are going to solve next. In this form, equation (7.11) represents the value of a European-style option. There is no analytical solution known for (7.11), so a numerical approach is needed also in the European case.

For an American-style option, a penalization can be applied, and an additional nonlinear term appears in (7.11). A classical penalty approach (e.g., [ElO82], [FoV02]) is to add the penalty $\hat{p}\max(\Psi - V, 0)$, where Ψ denotes the payoff, and the penalty parameter $\hat{p}$ is chosen large, say, $\hat{p} = 10^6$. The resulting PDE is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\hat{\sigma}^2(S,t,\frac{\partial^2 V}{\partial S^2})S^2\frac{\partial^2 V}{\partial S^2} + (r-\delta)S\frac{\partial V}{\partial S} - rV + \hat{p}\max(\Psi - V, 0) = 0\,. \tag{7.12}$$

In the continuation region, for $V \ge \Psi$, the penalty term is zero, and (7.11) results. For $\hat{p} \to \infty$, think of dividing the equation by $\hat{p}$ to be convinced that V sticks close to Ψ. In Chapter 4, we could conserve the linear equation by the elegant complementarity approach. In (7.12) the PDE is nonlinear by

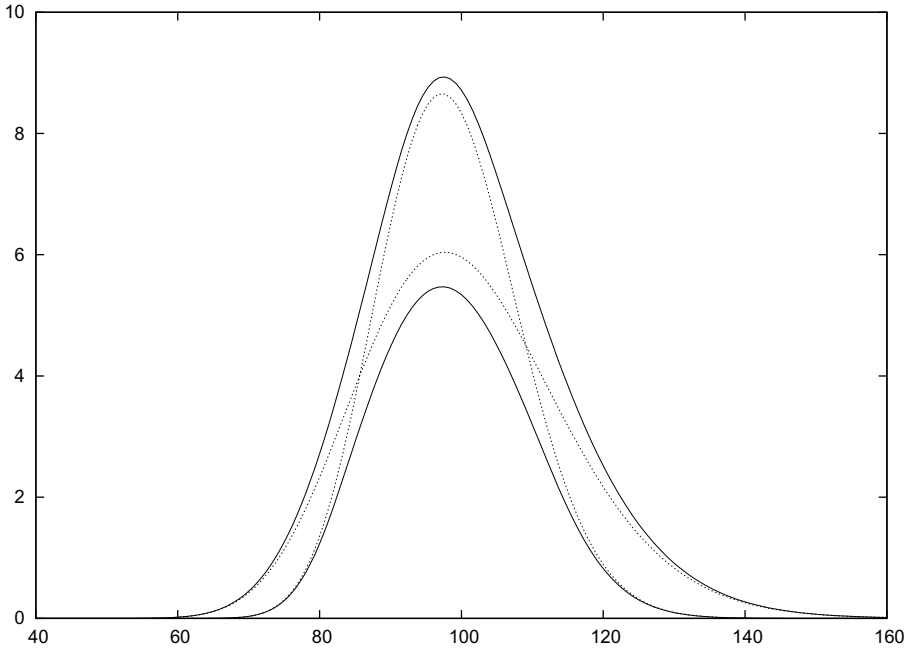

Fig. 7.2. $V(S,0)$ of a European butterfly spread, uncertain-volatility model of Avellaneda at al., Section 7.1.3; with $K = 100$, $K_1 = 85$, $K_2 = 115$, $r = 0.13$, $\sigma_{\min} = 0.15$, $\sigma_{\max} = 0.25$, $\delta = 0.03$, $T = 0.27$. Four curves are shown: the bounding functions V^+ and V^- as heavy lines, and V with dotted curves of the standard Black–Scholes model with constant volatilities $\sigma = 0.15$ (the steeper curve) and $\sigma = 0.25$ (the lower profile).

the volatility function $\hat{\sigma}$, and thus the nonlinear penalty term does not cause further harm.

7.2.1 Transformation

The transformation (4.3) of Chapter 4 is not valid here, because the volatility $\hat{\sigma}$ is no longer constant. But assuming constant r, δ, the independent variables S, t can be transformed similarly. The transformation from variables S, t, V to x, τ, u is

$$x := \log \frac{S}{K}\,, \quad \tau := \frac{1}{2}\sigma_0^2 \cdot (T - t)\,, \quad u(x,\tau) := \mathrm{e}^{-x}\frac{V(S,t)}{K}\,. \tag{7.13}$$

σ_0 is a scaling parameter. As a result of the transformation, $V_S = u + u_x$ and $SV_{SS} = u_x + u_{xx}$. Here we use the notations $V_S, V_{SS}, u_\tau, u_x, u_{xx}$ for partial derivatives. And (7.11) becomes

$$- u_\tau + \tilde{\sigma}^2(x,\tau,u_x,u_{xx})(u_x+u_{xx}) + \frac{2(r-\delta)}{\sigma_0^2}u_x - \frac{2\delta}{\sigma_0^2}u = 0$$

$$\text{with} \quad \tilde{\sigma} := \frac{1}{\sigma_0}\,\hat{\sigma}\left(S,t,\frac{\partial^2 V}{\partial S^2}\right) = \frac{1}{\sigma_0}\,\hat{\sigma}\left(K\mathrm{e}^x,\, T-\frac{2\tau}{\sigma_0^2},\, \frac{\mathrm{e}^{-x}}{K}(u_x+u_{xx})\right). \tag{7.14}$$

(Transform (7.12) in Exercise 7.3.) For example, for Leland's model,

$$\tilde{\sigma}^2 = 1 + \gamma\,\mathrm{sign}(u_x+u_{xx})\,.$$

For all of the models of Section 7.1 the nonlinearity is of the type

$$\tilde{\sigma}^2(x,\tau,s)\cdot s \quad \text{with} \quad s := u_x + u_{xx}\,, \tag{7.15}$$

with $\tilde{\sigma}$ from (7.14).

The payoffs Ψ of the options are transformed as well. Let u^* denote the transformed payoff. For the payoff of a vanilla put,

$$V(S,T) = Ke^x u(x,0) = (K-S)^+ = K(1-\mathrm{e}^x)^+$$

and hence

$$u(x,0) = u^*(x) := (\mathrm{e}^{-x}-1)^+\,.$$

Similarly, for a vanilla call,

$$u(x,0) = u^*(x) := (1-\mathrm{e}^{-x})^+\,.$$

This is similar for exotic options ($\longrightarrow$ Exercise 7.4).

Finally, boundary conditions are chosen (as in Section 4.4) and transformed. For example, applying (4.18) for a vanilla call of the European type,

$$\begin{aligned} u(x_{\max},\tau) &= \frac{\mathrm{e}^{-x_{\max}}}{K}V(S_{\max},t) \\ &= \frac{\mathrm{e}^{-x_{\max}}}{K}(S_{\max}\mathrm{e}^{-\delta(T-t)} - K\mathrm{e}^{-r(T-t)}) \\ &= \mathrm{e}^{-\delta(T-t)} - \exp(-r(T-t)-x_{\max}) \\ &= \exp(-\tau\frac{2\delta}{\sigma_0^2}) - \exp(-\tau\frac{2r}{\sigma_0^2} - x_{\max}) \\ u(x_{\min},\tau) &= 0\,. \end{aligned}$$

For a vanilla put and $S_{\min} \approx 0$ one may choose

$$\begin{aligned} u(x_{\min},\tau) &= \frac{1}{K}\mathrm{e}^{-x_{\min}}K\mathrm{e}^{-r(T-t)} = \exp(-\tau\frac{2r}{\sigma_0^2} - x_{\min}) \\ u(x_{\max},\tau) &= 0\,. \end{aligned}$$

For vanilla American-style options with penalty formulation (7.12), the nonzero boundary conditions are just that u is in contact with the payoff,

$$\begin{aligned} u(x_{\min}) = u^*(x_{\min}) &= \mathrm{e}^{-x_{\min}} - 1 \qquad \text{for a put, and} \\ u(x_{\max}) = u^*(x_{\max}) &= 1 - \mathrm{e}^{-x_{\max}} \qquad \text{for a call.} \end{aligned}$$

7.2.2 Discretization

Finite differences in a standard fashion as in Chapter 4, with the same grid, lead to nonlinear equations for the vector $w^{(\nu)}$ of approximate values at time level $\tau_\nu = \tau_{\nu-1} + \Delta\tau$. The equidistant x-spacing with mesh size Δx consists of m subintervals, see Section 4.2.2. The components w_0 and w_m are defined by boundary conditions. The finite differences include

$$\delta_x w_{i,\nu} := \frac{w_{i+1,\nu} - w_{i-1,\nu}}{2\Delta x}$$
$$\delta_{xx} w_{i,\nu} := \frac{w_{i+1,\nu} - 2w_{i,\nu} + w_{i-1,\nu}}{\Delta x^2}$$

where Δx^2 is understood as $(\Delta x)^2$. For the discretization replace s by $\bar{s}$ with

$$\bar{s}_{i,\nu} := (\delta_x + \delta_{xx}) w_{i,\nu} = \frac{w_{i+1,\nu} - w_{i-1,\nu}}{2\Delta x} + \frac{w_{i+1,\nu} - 2w_{i,\nu} + w_{i-1,\nu}}{\Delta x^2}$$

Substituting into the PDEs is the next step. Here we confine ourselves to the European case (7.11); the discretization of (7.12) is analogous and left to the reader. Define

$$\begin{aligned}\mathcal{L}_{i,\nu} :=& \tilde{\sigma}^2(x_i, \tau_\nu, \delta_x w_{i,\nu}, \delta_{xx} w_{i,\nu})(\delta_x w_{i,\nu} + \delta_{xx} w_{i,\nu}) \\ &+ \frac{2(r-\delta)}{\sigma_0^2} \delta_x w_{i,\nu} - \frac{2\delta}{\sigma_0^2} w_{i,\nu}\end{aligned}$$

to arrive at the θ-approach

$$\frac{-w_{i,\nu+1} + w_{i,\nu}}{\Delta\tau} + \theta \mathcal{L}_{i,\nu+1} + (1-\theta)\mathcal{L}_{i,\nu} = 0\,. \tag{7.16}$$

Recall that this includes Crank–Nicolson for $\theta = \frac{1}{2}$, and for $\theta = 1$ the fully implicit Euler (BDF). The $\tilde{\sigma}$ of the above examples is represented by the discretization $\tilde{\sigma}(x_i, \tau_\nu, \bar{s}_{i,\nu})$ with

$$\begin{aligned}\bar{s}_{i,\nu} &= w_{i-1,\nu}\left(-\frac{1}{2\Delta x} + \frac{1}{\Delta x^2}\right) - \frac{2}{\Delta x^2} w_{i,\nu} + w_{i+1,\nu}\left(\frac{1}{2\Delta x} + \frac{1}{\Delta x^2}\right) \\ &= \alpha\, w_{i-1,\nu} - \frac{2}{\Delta x^2}\, w_{i,\nu} + \beta\, w_{i+1,\nu}\end{aligned} \tag{7.17a}$$

where we denote

$$\alpha := -\frac{1}{2\Delta x} + \frac{1}{\Delta x^2}, \quad \beta := \frac{1}{2\Delta x} + \frac{1}{\Delta x^2} \tag{7.17b}$$

and reuse the notation $\tilde{\sigma}$ for the three-argument version. The discretized version of the operator $\mathcal{L}_{i,\nu}$ is now

$$\mathcal{L}_{i,\nu} = \tilde{\sigma}^2(x_i, \tau_\nu, \bar{s}_{i,\nu})\bar{s}_{i,\nu} + \frac{r-\delta}{\sigma_0^2 \Delta x}(w_{i+1,\nu} - w_{i-1,\nu}) - \frac{2\delta}{\sigma_0^2} w_{i,\nu} \tag{7.18}$$

and the θ-method reads

$$-w_{i,\nu+1} + w_{i,\nu} + \theta\Delta\tau\mathcal{L}_{i,\nu+1} + (1-\theta)\Delta\tau\mathcal{L}_{i,\nu} = 0\,. \tag{7.19}$$

With the vector notation $w^{(\nu)}$ as in Chapter 4 this is written

$$F(w^{(\nu+1)}, w^{(\nu)}) = 0\,.$$

For the fully implicit BDF method ($\theta = 1$), the ith equation of the vector equation $F = 0$ reads

$$\begin{aligned} F_i = &- w_i^{(\nu+1)} + w_i^{(\nu)} \\ &+ \Delta\tau\Big[\tilde{\sigma}^2(x_i, \tau_{\nu+1}, \alpha w_{i-1}^{(\nu+1)} - \frac{2}{\Delta x^2} w_i^{(\nu+1)} + \beta w_{i+1}^{(\nu+1)})\cdot \\ &\qquad (\alpha w_{i-1}^{(\nu+1)} - \frac{2}{\Delta x^2} w_i^{(\nu+1)} + \beta w_{i+1}^{(\nu+1)}) \\ &- \frac{r-\delta}{\sigma_0^2 \Delta x} w_{i-1}^{(\nu+1)} - \frac{2\delta}{\sigma_0^2} w_i^{(\nu+1)} + \frac{r-\delta}{\sigma_0^2 \Delta x} w_{i+1}^{(\nu+1)}\Big] = 0 \end{aligned} \tag{7.20a}$$

For $i = 0$ and $i = m$, boundary conditions enter. Their basic structure is

$$\begin{aligned} F_0^{(\nu)} &:= u(x_{\min}, \tau_\nu) - w_0^{(\nu)} \\ F_m^{(\nu)} &:= u(x_{\max}, \tau_\nu) - w_m^{(\nu)}\,. \end{aligned} \tag{7.20b}$$

For the θ-method (7.19) they enter in the form $\theta F^{(\nu+1)} + (1-\theta)F^{(\nu)}$. The nonlinear equation $F(w^{(\nu+1)}, w^{(\nu)}) = 0$ with components defined by (7.20) represents a discretization of (7.11). It is solved iteratively by Newton's method.

7.2.3 Convergence of the Discrete Equations

The above numerical scheme is of the form

$$F(\Delta\tau, \Delta x, \nu, i, w_{i,\nu}, \tilde{w}) = 0$$

where $\tilde{w}$ stands for the vector of all $w_{k,l}$. For such a scheme convergence to the unique viscosity solution (⟶ Appendix C5) can be proved, provided F satisfies three conditions [BaDR95], namely,

- stability,
- consistency, and
- monotonicity.

Not for the numerical scheme but for the equation an additional property must be assumed, namely, the strong uniqueness. For the uniqueness we refer to the special literature [CrIL92].

The proof that for a particular scheme all of these three criteria are satisfied, can be quite involved [PoFV03], [Hei10], [HeS10]. Checking stability

and consistency is rather standard. Here we concentrate on the monotonicity of the scheme, which is a new aspect as compared to the investigations for the linear equation in Chapter 4.

Definition 7.1 (monotone scheme)
A discretization $F(w^{(\nu+1)}, w^{(\nu)})$ is *monotone* if for all $i = 0, \ldots, m$

(a) $$F_i(w^{(\nu+1)} + \epsilon^{(\nu+1)},\, w^{(\nu)} + \epsilon^{(\nu)}) \geq F_i(w^{(\nu+1)}, w^{(\nu)})$$

for all

$$\epsilon^{(\nu+1)} = (0, \ldots, 0, \epsilon_{i-1}^{(\nu+1)}, 0, \epsilon_{i+1}^{(\nu+1)}, 0, \ldots, 0) \geq 0$$
$$\epsilon^{(\nu)} = (0, \ldots, 0, \epsilon_{i-1}^{(\nu)}, \epsilon_{i}^{(\nu)}, \epsilon_{i+1}^{(\nu)}, 0, \ldots, 0) \geq 0\,,$$

and

(b) $$F_i(w^{(\nu+1)} + \epsilon^{(\nu+1)},\, w^{(\nu)}) \leq F_i(w^{(\nu+1)}, w^{(\nu)})$$

for all

$$\epsilon^{(\nu+1)} = (0, \ldots, 0, \epsilon_{i}^{(\nu+1)}, 0, \ldots, 0) \geq 0\,.$$

Translated into the fully implicit scheme (7.20), the condition (a) of monotonicity reads

$$F_i(w_i^{(\nu+1)},\, w_{i-1}^{(\nu+1)} + \epsilon_1,\, w_{i+1}^{(\nu+1)} + \epsilon_2,\, w_i^{(\nu)} + \epsilon_3) \geq$$
$$F_i(w_i^{(\nu+1)},\, w_{i-1}^{(\nu+1)},\, w_{i+1}^{(\nu+1)},\, w_i^{(\nu)})$$

for scalar $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon$. Because of transitivity, it suffices to show separately

(a1) $F_i(w_i^{(\nu+1)},\, w_{i-1}^{(\nu+1)} + \epsilon,\, w_{i+1}^{(\nu+1)},\, w_i^{(\nu)}) \geq F_i(w_i^{(\nu+1)}, w_{i-1}^{(\nu+1)}, w_{i+1}^{(\nu+1)}, w_i^{(\nu)})$
(a2) $F_i(w_i^{(\nu+1)},\, w_{i-1}^{(\nu+1)},\, w_{i+1}^{(\nu+1)} + \epsilon,\, w_i^{(\nu)}) \geq F_i(w_i^{(\nu+1)}, w_{i-1}^{(\nu+1)}, w_{i+1}^{(\nu+1)}, w_i^{(\nu)})$
(a3) $F_i(w_i^{(\nu+1)},\, w_{i-1}^{(\nu+1)},\, w_{i+1}^{(\nu+1)},\, w_i^{(\nu)} + \epsilon) \geq F_i(w_i^{(\nu+1)}, w_{i-1}^{(\nu+1)}, w_{i+1}^{(\nu+1)}, w_i^{(\nu)})$

for (a) to hold, and for (b)

$$F_i(w_i^{(\nu+1)} + \epsilon,\, w_{i-1}^{(\nu+1)},\, w_{i+1}^{(\nu+1)},\, w_i^{(\nu)}) \leq F_i(w_i^{(\nu+1)},\, w_{i-1}^{(\nu+1)},\, w_{i+1}^{(\nu+1)},\, w_i^{(\nu)})\,.$$

Next we check under which conditions the scheme (7.20) is monotone. [Hei10] has shown that the scheme converges whenever the nonlinear term $\tilde{\sigma}^2(x, \tau, s)s$ satisfies conditions (i)–(iii):

Theorem 7.2
Assume $\tilde{\sigma}^2(x, \tau, u_x, u_{xx})$ in the form $\tilde{\sigma}^2(x, \tau, s)$, with s from (7.15), and
(i) $\tilde{\sigma}^2(x, \tau, s)s$ is continuous and monotone increasing in s,
(ii) there exists a constant $c_+ > 0$ such that for all s and $\epsilon > 0$

$$\tilde{\sigma}^2(x, \tau, s + \epsilon) \cdot (s + \epsilon) \geq \tilde{\sigma}^2(x, \tau, s) \cdot s + c_+\epsilon$$

(iii) Δx is small enough such that

$$c_+\frac{2-\Delta x}{\Delta x}-\frac{2(r-\delta)}{\sigma_0^2}\geq 0 \quad\text{and}\quad c_+\frac{2+\Delta x}{\Delta x}+\frac{2(r-\delta)}{\sigma_0^2}\geq 0$$

Then the fully implicit BDF scheme (7.20) converges to the viscosity solution of (7.11).

Proof: Here we confine ourselves to the proof of monotonicity. As noted above, we can proceed componentwise and check (a1), (a2), (a3), and (b) separately. We begin with $0<i<m$.

To show (a1), perturb $w_{i-1}^{(\nu+1)}\to w_{i-1}^{(\nu+1)}+\epsilon$ for $\epsilon>0$. Then $\bar{s}_{i,\nu}\to\bar{s}_{i,\nu}+\alpha\epsilon$, and

$$\begin{aligned}
&F_i(w_i^{(\nu+1)},\, w_{i-1}^{(\nu+1)}+\epsilon,\, w_{i+1}^{(\nu+1)},\, w_i^{(\nu)}) = \\
&\quad -w_i^{(\nu+1)}+w_i^{(\nu)}+\Delta\tau\Big[\tilde{\sigma}^2(x_i,\tau_{\nu+1},\bar{s}_{i,\nu}+\alpha\epsilon)(\bar{s}_{i,\nu}+\alpha\epsilon) \\
&\quad -\frac{r-\delta}{\sigma_0^2\Delta x}(w_{i-1}^{(\nu+1)}+\epsilon)-\frac{2\delta}{\sigma_0^2}w_i^{(\nu+1)}+\frac{r-\delta}{\sigma_0^2\Delta x}w_{i+1}^{(\nu+1)}\Big] \\
&\geq -w_i^{(\nu+1)}+w_i^{(\nu)}+\Delta\tau\Big[\tilde{\sigma}^2(x_i,\tau_{\nu+1},\bar{s}_{i,\nu})\bar{s}_{i,\nu}+c_+\epsilon\alpha \\
&\quad -\frac{r-\delta}{\sigma_0^2\Delta x}w_{i-1}^{(\nu+1)}-\frac{2\delta}{\sigma_0^2}w_i^{(\nu+1)}+\frac{r-\delta}{\sigma_0^2\Delta x}w_{i+1}^{(\nu+1)}-\frac{r-\delta}{\sigma_0^2\Delta x}\epsilon\Big]
\end{aligned}$$

where the inequality is due to (ii). Compare with F_i in (7.20) and realize two extra terms. By (iii), with α from (7.17b), they are

$$c_+\epsilon\alpha-\frac{r-\delta}{\sigma_0^2\Delta x}\epsilon=\frac{\epsilon}{2\Delta x}\left[c_+\frac{2-\Delta x}{\Delta x}-\frac{2(r-\delta)}{\sigma_0^2}\right]\geq 0\,.$$

So we have shown (a1), the first of the four criteria of monotonicity.

To show (a2), perturb $w_{i+1}^{(\nu+1)}\to w_{i+1}^{(\nu+1)}+\epsilon$. Then $\bar{s}_{i,\nu}\to\bar{s}_{i,\nu}+\epsilon\beta$ and the perturbed F_i is

$$\begin{aligned}
&-w_i^{(\nu+1)}+w_i^{(\nu)}+\Delta\tau\Big[\tilde{\sigma}^2(x_i,\tau_{\nu+1},\bar{s}_{i,\nu}+\beta\epsilon)(\bar{s}_{i,\nu}+\beta\epsilon) \\
&-\frac{r-\delta}{\sigma_0^2\Delta x}w_{i-1}^{(\nu+1)}-\frac{2\delta}{\sigma_0^2}w_i^{(\nu+1)}+\frac{r-\delta}{\sigma_0^2\Delta x}w_{i+1}^{(\nu+1)}+\epsilon\frac{r-\delta}{\sigma_0^2\Delta x}\Big]\,.
\end{aligned}$$

Again we obtain a lower bound by (ii), and arrive at the sum of two extra terms

$$c_+\epsilon\beta+\epsilon\frac{r-\delta}{\sigma_0^2\Delta x}\,,$$

which is ≥ 0 by (iii). So the perturbed F_i is larger or equal the unperturbed F_i, and (a2) is satisfied.

The assertion (a3) is clearly satisfied since the perturbation $w_i^{(\nu)} \to w_i^{(\nu)} + \epsilon$ only affects the term outside the brackets.

To show (b), perturb $w_i^{(\nu+1)} \to w_i^{(\nu+1)} + \epsilon$. Then $\bar{s}_{i,\nu} \to \bar{s}_{i,\nu} - \frac{2\epsilon}{\Delta x^2}$, and F_i is perturbed to

$$
\begin{aligned}
&- w_i^{(\nu+1)} - \epsilon + w_i^{(\nu)} + \Delta\tau \Bigg[\tilde{\sigma}^2(x_i, \tau_{\nu+1}, \bar{s}_{i,\nu} - \epsilon \frac{2}{\Delta x^2})(\bar{s}_{i,\nu} - \epsilon \frac{2}{\Delta x^2}) \\
&- \frac{r-\delta}{\sigma_0^2 \Delta x} w_{i-1}^{(\nu+1)} - \frac{2\delta}{\sigma_0^2} w_i^{(\nu+1)} - \frac{2\delta}{\sigma_0^2}\epsilon + \frac{r-\delta}{\sigma_0^2 \Delta x} w_{i+1}^{(\nu+1)} \Bigg] .
\end{aligned}
$$

By the monotonicity (i) and by $\epsilon > 0$, $\delta \geq 0$, the above is smaller or equal to the unperturbed F_i —that is, (b) holds true.

Finally, monotonicity must be checked for F_0 and F_m. For $\theta = 1$, F_0 depends on $w_0^{(\nu+1)}$ and F_m depends on $w_m^{(\nu+1)}$. Hence only (b) needs to be checked, which is clearly satisfied.

This ends the proof that the conditions (i), (ii), (iii) imply monotonicity of the fully implicit scheme.

Example 7.3 (Leland's model)
Let us inspect whether the criteria (i), (ii), (iii) of Theorem 7.2 are satisfied for Leland's model of transaction costs. For (i) we require $|\gamma| < 1$. With some simple manipulations, one shows that (ii) is satisfied with $c_+ = 1 - \gamma$. And for (iii) to hold, the grid size Δx must be small enough. ($\longrightarrow$ Exercise 7.5). Specifically, for zero dividend rate $\delta = 0$ the θ-method is

$$
\begin{aligned}
&- w_i^{(\nu+1)} + w_i^{(\nu)} + \Delta\tau \cdot \theta \, [\tilde{\sigma}^2(\bar{s}_i^{(\nu+1)})\bar{s}_i^{(\nu+1)} + \frac{2r}{\sigma_0^2}\delta_x w_i^{(\nu+1)}] \\
&+ \Delta\tau(1-\theta) \, [\tilde{\sigma}^2(\bar{s}_i^{(\nu)})\bar{s}_i^{(\nu)} + \frac{2r}{\sigma_0^2}\delta_x w_i^{(\nu)}] = 0 \, .
\end{aligned}
$$

□

Sufficient conditions for the Crank–Nicolson scheme ($\theta = 1/2$) to converge include (i), (ii), (iii), and in addition (iv) and (v):

(iv) There exists a constant $c_- > 0$ such that for all $\epsilon > 0$ and s

$$
\tilde{\sigma}^2(x, \tau, s - \epsilon)(s - \epsilon) \geq \tilde{\sigma}^2(x, \tau, s)s - c_-\epsilon
$$

(v)

$$
\Delta\tau \leq \frac{\Delta x^2}{c_-} \frac{\sigma_0^2}{\sigma_0^2 + \Delta x \, \delta} \, ,
$$

see [Hei10], [HeS10]. Condition (iv) holds for Leland's model with $c_- = 1 + \gamma$, and for the uncertain volatility model with $c_- = \sigma_{\max}^2$. Conditions (iii) and (iv) amount to stability bounds. We emphasize that in the case of nonlinear models, unconditional stability does *not* hold!

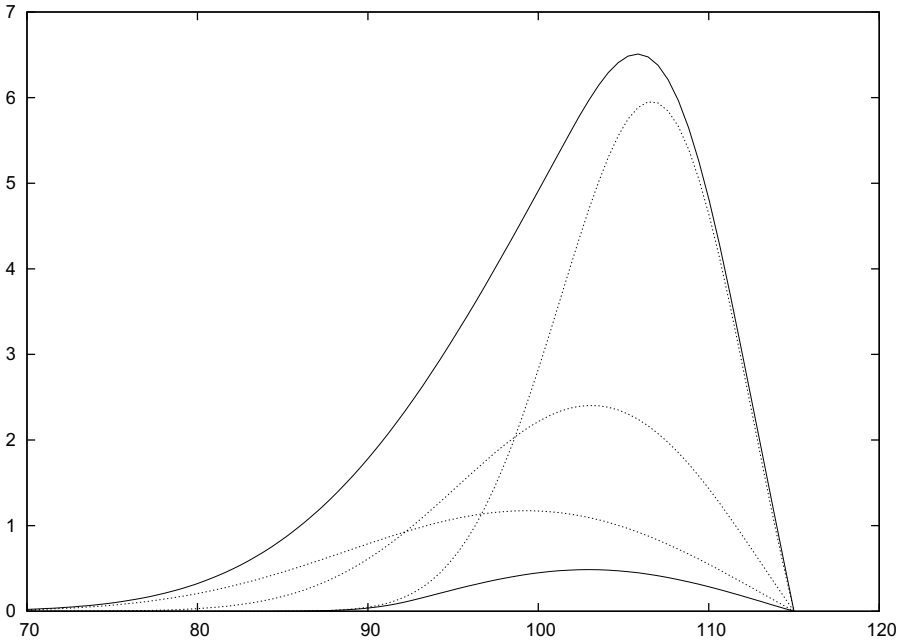

Fig. 7.3. $V(S,0)$ of a European up-and-out barrier call, uncertain-volatility model of Avellaneda at al., Section 7.1.3; with barrier $B = 115$, and $K = 100$, $r = 0.1$, $\sigma_{\min} = 0.1$, $\sigma_{\max} = 0.3$, $\delta = 0$, $T = 0.2$. In addition to the two curves V^+ and V^- (solid lines) three V curves are shown (with dotted lines) of the standard Black–Scholes model with constant volatilities $\sigma = 0.1$ (the steepest) and $\sigma = 0.2$, 0.3.

The above has discussed convergence towards the viscosity solution. An application of the uncertain-volatility model to a butterfly is shown in Figure 7.2. Another illustration is the barrier option in Figure 7.3. — When in case of an American-style option a penalty approach is applied, further assumptions are needed to assert convergence to the solution for $\hat{p} \to \infty$, even though one keeps $\hat{p}$ fixed.

7.3 Option Valuation Under Jump Processes

In this section, we sketch some instruments of Lévy processes as background to the application of partial integro-differential equations. The focus is on one important example, namely Merton's jump diffusion, and on strategies for a numerical valuation of options under such processes. This is no introduction to Lévy processes; for expositions on Lévy processes consult, for instance, [Sato99], [Shi99], [ConT94].

For a Lévy process X_t, all increments $X_{t+\Delta t} - X_t$ are stochastically independent. Further, they are stationary, which means that all increments have the distribution of X_t. Instead of requiring continuity, Lévy processes must be "càdlàg" (French for "continu à droite avec limites à gauche"): For all t, the process X_t is right-continuous ($X_t = X_{t^+}$), and the left limit X_{t^-} exists. Important examples of Lévy processes are the Wiener process (Section 1.6.1), and the Poisson process (Section 1.9).

7.3.1 Characteristic Functions

A classification of Lévy processes X_t is based on the Fourier transformation[2]

$$\phi_{X_t}(\zeta) := \mathsf{E}(\exp(\mathrm{i}\zeta X_t)) \,. \tag{7.21}$$

The function ϕ_{X_t} singles out characteristic properties of a random variable X_t. ϕ_{X_t} is called *characteristic function* of X_t, and $\psi_{X_t}(\zeta)$ [shorter: $\psi(\zeta)$] defined by $\exp(t\psi(\zeta)) = \phi_{X_t}(\zeta)$ is the *characteristic exponent*. It suffices to take $t = 1$, since the distribution of X_1 characterizes the process. The characteristic exponent $\psi(\zeta)$ satisfies the **Lévy-Khinchin representation**

$$\psi(\zeta) = \mathrm{i}\gamma\zeta - \frac{1}{2}\sigma^2\zeta^2 + \int_{-\infty}^{\infty} \left(\exp(\mathrm{i}\zeta x) - 1 - \mathrm{i}\zeta x\, \mathbf{1}_{\{|x|\le 1\}}\right) \nu(\mathrm{d}x) \,. \tag{7.22}$$

The three terms in this representation characterize different aspects of X_t. $\gamma \in \mathbb{R}$ corresponds to a deterministic trend, σ^2 to the variance of a diffusion (Brownian-motion) part of X_t, and ν is a measure on $\mathbb{R}$ characterizing the activity of jumps $\Delta X_t := X_t - X_{t^-}$,

$$\nu(A) := \mathsf{E}\left[\#\{t \in [0,1] \mid \Delta X_t \neq 0,\ \Delta X_t \in A\}\right].$$

The Lévy measure $\nu(A)$ counts the (expected) number of jumps of "size" within A per unit time [ConT04]. $\nu(A)$ is not a probability measure. For the Lévy measure ν, require $\int_{\mathbb{R}} \min(x^2, 1)\, \nu(\mathrm{d}x) < \infty$ and $\nu(\{0\}) = 0$. In the integrand of (7.22), the subtracted term $\mathrm{i}\zeta x\, \mathbf{1}_{\{|x|\le 1\}}$ causes the integrand to be of the order $O(|x|^2)$ for $x \to 0$. This compensation along with the constraints on ν implies existence of the integral. For many important Lévy processes, $\nu(\mathrm{d}x)$ has a convenient representation

$$\nu(\mathrm{d}x) = f_{\mathrm{L}}(x)\, \mathrm{d}x \tag{7.23}$$

with a Lévy density f_{L}. The three items γ, σ^2, ν ("characteristic triplet") characterize a Lévy process in a unique way.

[2] For the Fourier transform, see Section 7.4.

Example 7.4 (compound Poisson process)
For a Poisson process J_t with jump intensity λ, a compound Poisson process is

$$X_t := \sum_{j=1}^{J_t} \Delta X_{\tau_j} ,$$

where the jump sizes ΔX_{τ_j} are assumed i.i.d. with distribution density f, and independent of the Poisson process J. The characteristic function $\phi_{X_t}(\zeta)$ of the compound Poisson process is

$$\begin{aligned} \mathsf{E}(\exp[\mathrm{i}\zeta X_t)) &= \exp[\lambda t\,(\phi_{\Delta X}(\zeta) - 1)] \\ &= \exp\left[t \int_{\mathbb{R}} (\mathrm{e}^{\mathrm{i}\zeta x} - 1)\nu(\mathrm{d}x)\right] \end{aligned} \tag{7.24}$$

with Lévy measure $\nu(\mathrm{d}x) = \lambda f(x)\,\mathrm{d}x$. The first of the equations in (7.24) uses rules of the conditional expectation [ConT04], whereas the second just applies (7.21) with the definition (B1.4) of the expectation, including $\int_{\mathbb{R}} \nu(\mathrm{d}x) = \lambda$. The characteristic exponent ψ_{cP} is the integral in (7.24), $\gamma = \sigma = 0$. □

As in (1.49), financial models typically arise in exponential form. For such exponential Lévy processes there is a useful criterion for the martingale property, and hence for risk-neutral valuation:

Lemma 7.5 (martingale criterion)
Let X_t be a Lévy process. e^{X_t} is a martingale if and only if $\psi_X(-\mathrm{i}) = 0$ and $\mathsf{E}(\mathrm{e}^{X_t}) < \infty$.

Proof: We extend ζ to complex numbers, and note that

$$\mathsf{E}(\mathrm{e}^{X_t}) = \mathsf{E}(\mathrm{e}^{-\mathrm{i}\mathrm{i}X_t}) = \phi_{X_t}(-\mathrm{i}) = \mathrm{e}^{t\psi(-\mathrm{i})} .$$

Then by independence and stationarity,

$$\mathsf{E}(\mathrm{e}^{X_t} \,|\, \mathcal{F}_s) - \mathrm{e}^{X_s} = \mathsf{E}(\mathrm{e}^{X_{t-s}}) - \mathrm{e}^{X_0} = \mathrm{e}^{(t-s)\psi(-\mathrm{i})} - 1 .$$

(⟶ Exercise 7.6)

In finance applications, with an asset price S_t for $t \geq 0$, the absence of arbitrage implies that the discounted $\mathrm{e}^{-rt}S_t$ is a martingale with respect to a risk-neutral measure. This suggests to represent S_t in the form $S_t = S_0 \exp(rt + X_t)$. Then the discounted S_t is the situation to which the Lemma 7.5 applies.

Example 7.6 (Brownian motion with drift)
A Lévy process X_t is Brownian motion if and only if $\nu \equiv 0$ (no jump). For ease of comparison with (1.54) and (1.59) we take the drift γ in the form $\gamma = \mu - \frac{1}{2}\sigma^2$. For the Brownian motion with drift (Bwd) $X_t := \gamma t + \sigma W_t$ we use a result from probability[3] and conclude for the characteristic exponent

[3] $\mathsf{E}(\mathrm{e}^{\mathrm{i}\zeta X}) = \exp(\mathrm{i}\zeta\gamma - \zeta^2\sigma^2/2)$ holds for $X \sim \mathcal{N}(\gamma, \sigma^2)$, see [JaP03] p. 108.

$$\psi_{\mathrm{Bwd}}(\zeta) = \mathrm{i}(\mu - \frac{1}{2}\sigma^2)\zeta - \frac{1}{2}\sigma^2\zeta^2 \,.$$

Clearly, $\psi_{\mathrm{Bwd}}(-\mathrm{i}) = \mu$. Hence by Lemma 7.5 e_t^X is martingale for $\mu = 0$. Hence the discounted

$$S_0\mathrm{e}^{-rt}\exp(rt + X_t) = S_0\mathrm{e}^{-rt}\exp[(r - \tfrac{1}{2}\sigma^2)t + \sigma W_t]$$

is martingale. This recovers the well-known riskless drift rate r for a numerical simulation of GBM in the Black-Scholes model.

Example 7.7 (Merton' s jump diffusion)

We now combine Examples 7.4 and 7.6. As a special case of Example 7.4 we choose as in Section 1.9 the jump sizes ΔY in the log process $Y_t := \log S_t$ to be normally distributed, $\Delta Y \sim \mathcal{N}(\mu_{\mathrm{J}}, \sigma_{\mathrm{J}}^2)$. ($\log q$ in Section 1.9) Furnished with a drifted Brownian motion, this is Merton's jump-diffusion model (1.57) with jump intensity λ and $\gamma = \mu - \frac{1}{2}\sigma^2$. The Lévy density of the compound Poisson process (cP) is λ times the density of the normal distribution,

$$f_{\mathrm{L}}(x) = f_{\mathrm{cP}}(x) := \lambda \frac{1}{\sigma_{\mathrm{J}}\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_{\mathrm{J}})^2}{2\sigma_{\mathrm{J}}^2}\right] . \tag{7.25}$$

Since the two processes are independent, and by the exponential structure in (7.21), the two characteristic exponents add:

$$\begin{aligned}\psi(\zeta) &= \psi_{\mathrm{Bwd}}(\zeta) + \psi_{\mathrm{cP}}(\zeta) \\ &= \mathrm{i}\gamma\zeta - \frac{1}{2}\sigma^2\zeta^2 + \int_{\mathbb{R}} (\mathrm{e}^{\mathrm{i}\zeta x} - 1)\nu(\mathrm{d}x)\end{aligned}$$

and

$$\psi(-\mathrm{i}) = \gamma + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^x - 1)\nu(\mathrm{d}x) \,.$$

Similar as in Exercise 1.12 we calculate the integral

$$\int_{-\infty}^{\infty} (\mathrm{e}^x - 1) f_{\mathrm{cP}}(x)\,\mathrm{d}x = \lambda\left(\exp\left[\mathrm{i}\mu_{\mathrm{J}}\zeta - \frac{1}{2}\sigma_{\mathrm{J}}^2\zeta^2\right] - 1\right) .$$

Hence to see whether $S_t = \exp(Y_t)$ is a martingale, we check $\psi(-\mathrm{i}) = \gamma + \frac{1}{2}\sigma^2 + \lambda(\exp[\mu_{\mathrm{J}} + \frac{1}{2}\sigma_{\mathrm{J}}^2] - 1)$. By Lemma 7.5, a martingale can be obtained by choosing a drift with

$$\gamma = -\frac{\sigma^2}{2} - \lambda\left(\exp\left[\mu_{\mathrm{J}} + \frac{1}{2}\sigma_{\mathrm{J}}^2\right] - 1\right) .$$

This makes $S_0\mathrm{e}^{-rt}\exp(rt + \gamma t + \sigma W_t + \sum_{j=1}^{J_t} \log q_j)$ a martingale. When applied to simulation of SDEs under the risk-neutral measure for Monte Carlo, this risk-neutral valuation amounts to the drift rate in Example 1.21. That is, the SDE is

$$\frac{\mathrm{d}S}{S} = (r - \lambda(\exp[\mu_\mathrm{J} + \tfrac{1}{2}\sigma_\mathrm{J}^2] - 1))\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\,.$$

In case of a dividend yield with rate δ, the term $\delta \mathrm{d}t$ is subtracted on the right-hand side, similar as in Section 3.5. □

For other models, a risk-neutral growth rate can be obtained in an analogous way. A table of risk-neutral drift rates is given in [Sch03], p.80. For a jump diffusion, jumps are comparably "rare," there is only a finite number of them in any time interval. Apart from Merton's model another jump-diffusion model is Kou's model, which works with an asymmetric double exponential distribution of jump sizes [Kou02].

There are Lévy processes of infinite activity: Then in every time interval an infinite number of jumps occurs. Examples include the VG-process (Variance Gamma) [MaS90], the NIG-process (Normal Inverse Gaussian), the hyperbolic process [EbK95] and the CGMY process [CaGMY03]. Specifically for VG and NIG, see also [Gla04]. Time deformation plays an important role for constructing Lévy processes. For example, with a Wiener process W_t and a Gamma process G_t as *subordinator* replacing time, VG can be represented as

$$S_t = S_0 \mathrm{e}^{rt + X_t} \text{ with } X_t = \theta G_t + \sigma W_{G_t}\,.$$

This includes GBM with the standard time $G_t = t$ and parameter $\theta = -\sigma^2/2$. Such a subordinating process G_t can be regarded as "business time," which runs faster than the calendar time when the trading volume is high, and slower otherwise. Then, for a Wiener process W_t, a class of Lévy processes is defined by W_{G_t}. With a t-grid as in Algorithm 1.8, a time-changed process can be generated as $W_j = W_{j-1} + Z\sqrt{G_{j\Delta t} - G_{(j-1)\Delta t}}$ (⟶ Exercise 2.17).

7.3.2 Option Valuation with PIDEs

Assume European options based on a price process $S_t = S_0 \exp(rt + X_t)$, where X_t is a Lévy process such that e^{X_t} is a martingale, with Lévy measure ν, and the integral $\int_{|y|\geq 1} \mathrm{e}^{2y}\nu(\mathrm{d}y)$ exists. Then the value function $V(S,t)$ satisfies

$$\begin{aligned} &\frac{\partial V(S,t)}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV \\ &+ \int_{\mathbb{R}} \left[V(S\mathrm{e}^y, t) - V(S,t) - (\mathrm{e}^y - 1)S\frac{\partial V(S,t)}{\partial S} \right] \nu(\mathrm{d}y) = 0 \end{aligned} \tag{7.26}$$

A proof can be found in [ConT04] p. 385-387.

Definition 7.8 (PIDE)
An equation of the above type (7.26) is called *partial integro-differential equation (PIDE).*

The integral term in (7.26) complicates the numerical solution since it is a nonlocal term accumulating information on all $-\infty < y < \infty$, in contrast to the local character of the partial derivatives. For general Lévy processes, the three terms under the integral can not be separated, otherwise the integral may fail to converge. It can be separated in the case of Merton's jump-diffusion model, because this process is of finite activity, $\lambda = \nu(\mathbb{R}) < \infty$.

In what follows, we discuss Merton's jump-diffusion process, with lognormal distribution for $q = \mathrm{e}^y$. The integral in (7.26) can be split into three terms with three integrals

$$\int_{\mathbb{R}} V(S\mathrm{e}^y, t)\nu(\mathrm{d}y) - V(S,t)\int_{\mathbb{R}} \nu(\mathrm{d}y) - S\frac{\partial V(S,t)}{\partial S}\int_{\mathbb{R}} (\mathrm{e}^y - 1)\nu(\mathrm{d}y)\,.$$

In view of $\nu(\mathrm{d}y) = \lambda f(y)\mathrm{d}y$, factors λ show up. f is the standard normal density, and the integrals become expectations. Then the first integral can be written $\lambda\mathsf{E}(V(S\mathrm{e}^y, t))$, and the second integral is λ. The integral $c := \mathsf{E}(\mathrm{e}^y - 1)$ does not depend on V and can be calculated beforehand since the distribution for $q = \mathrm{e}^y$ is stipulated.[4] The lognormal density for q is

$$f_q(x) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{J}} \cdot x}\exp\left\{-\frac{(\log x - \mu_{\mathrm{J}})^2}{2\sigma_{\mathrm{J}}^2}\right\}\mathbf{1}_{\{x>0\}}$$

and we recover the constant of Example 7.7:

$$\begin{aligned} c :&= \int_0^\infty (x-1) f_q(x)\,\mathrm{d}x \\ &= \int_{-\infty}^\infty (\mathrm{e}^y - 1) f(y)\,\mathrm{d}y = \exp[\mu_{\mathrm{J}} + \frac{1}{2}\sigma_{\mathrm{J}}^2] - 1\,. \end{aligned}$$

With the precalculated number c, the resulting equation can be ordered into

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + (r - \lambda c)S\frac{\partial V}{\partial S} - (\lambda + r)V + \lambda\mathsf{E}(V(qS,t)) = 0\,. \quad (7.27)$$

The last term is an integral taken over the unknown solution function $V(S,t)$. So the resulting equation is a PIDE, a special case of (7.26). Note that the product λc is the drift compensation in Example 7.7. The standard Black–Scholes PDE is included for $\lambda = 0$. A simplified derivation of (7.27) can be found in Appendix A4. For further discussions, see for example [Mer76], [Wil98], [Tsay02], [ConT04].

[4] The parameters are not the same as those in (1.48).

7.3.3 Transformation of the PIDE

We approach the PIDE (7.27) with the transformation

$$\tau := T - t\,, \quad x := \log S\,, \quad u(x,\tau) := V(\mathrm{e}^x, T-\tau)\,, \tag{7.28}$$

which appears moderate as compared to (4.3). Substituting accordingly

$$u_x = \frac{\partial V}{\partial S} S\,, \quad u_{xx} = u_x + S^2 \frac{\partial^2 V}{\partial S^2}$$

into (7.27) leads to

$$-u_\tau + \tfrac{1}{2}\sigma^2(u_{xx} - u_x) + (r - \lambda c)u_x - (\lambda + r)u + \lambda \mathsf{E}(V(q\mathrm{e}^x, T-\tau)) = 0\,,$$

which is organized into

$$u_\tau - \tfrac{1}{2}\sigma^2 u_{xx} - (r - \lambda c - \tfrac{1}{2}\sigma^2)u_x + (\lambda + \tau)u - \lambda \mathsf{E}(V(q\mathrm{e}^x, T-\tau)) = 0\,.$$

After the above transformation $S = \mathrm{e}^x$ we next transform the jump-size variable $q = \mathrm{e}^y$. Ignoring the factor λ, the integral term changes to

$$\begin{aligned} &\mathsf{E}(V(q\mathrm{e}^x, T-\tau)) = \mathsf{E}(V(\mathrm{e}^{x+y}, T-\tau)) = \mathsf{E}(u(x+y,\tau)) \\ &= \int_{\mathbb{R}} u(x+y,\tau) f(y)\,\mathrm{d}y = \int_{\mathbb{R}} u(z,\tau) f(z-x)\,\mathrm{d}z\,, \end{aligned} \tag{7.29}$$

where we have applied the substitution $z := x+y$. The function f for Merton's jump-diffusion model is the density of $y = \log q \sim \mathcal{N}(\mu_\mathrm{J}, \sigma_\mathrm{J}^2)$. In summary, the PIDE of Merton's jump-diffusion model is

Problem 7.9 (Merton's jump-diffusion PIDE)

$$\begin{aligned} &u_\tau - \tfrac{1}{2}\sigma^2 u_{xx} - (r - \lambda c - \tfrac{1}{2}\sigma^2)u_x + (\lambda + r)u \\ &\quad - \lambda \int_{\mathbb{R}} u(z,\tau) f(z-x)\,\mathrm{d}z = 0\,, \\ &\text{with} \quad f(y) = \frac{1}{\sqrt{2\pi}\sigma_\mathrm{J}} \exp\left[-\frac{(y-\mu_\mathrm{J})^2}{2\sigma_\mathrm{J}^2}\right] \\ &\text{and} \quad c = \exp[\mu_\mathrm{J} + \tfrac{1}{2}\sigma_\mathrm{J}^2] - 1\,. \end{aligned} \tag{7.30}$$

This is the problem to be solved numerically.

7.3.4 Numerical Approximation

For an approximation of the integral (7.29) we truncate the domain to a finite interval $x_{\min} \le x \le x_{\max}$. In view of the meaning of the integral, this truncation amounts to disregard large jumps. This might be seen as a weakness of the approach, but jumps that large are highly improbable. The simplest discretization approach is to use an equidistant x-grid with

$$\Delta x := \frac{x_{\max} - x_{\min}}{m}, \quad x_i := x_{\min} + i\Delta x, \quad i = 0, \dots, m,$$

for a suitable integer m. As in Chapter 4, the time-stepping nodes are τ_ν, and the approximations of $u(x_i, \tau_\nu)$ are denoted by $w_{i,\nu}$. The integral in (7.30) is evaluated at each node $(x, \tau) = (x_i, \tau_\nu)$. That is, for each i, ν, the numbers

$$\int_{\mathbb{R}} u(z, \tau_\nu) f(z - x_i)\, \mathrm{d}z \approx \int_{x_{\min}}^{x_{\max}} u(z, \tau_\nu) f(z - x_i)\, \mathrm{d}z$$

are to be approximated. Applying the composite trapezoidal rule (C1.2) with

$$f_{i,l} := f(x_l - x_i) = f((l - i)\Delta x),$$

the approximation of the integral for each i, ν is

$$\Delta x \left[\frac{w_{0,\nu} f_{i,0}}{2} + \sum_{l=1}^{m-1} w_{l,\nu} f_{i,l} + \frac{w_{m,\nu} f_{i,m}}{2} \right]. \tag{7.31}$$

The numbers $f_{i,l}$ are elements of a Toeplitz matrix.[5] That is, the entries take only $2m + 1$ different numbers. Due to the exponential structure of f, the elements in the northeast and southwest corners of the $f_{i,l}$-matrix go to zero. In this sense, this Toeplitz matrix has a "banded" structure. In summary, for each i, ν the integral is approximated by a scalar product of the row vector

$$\Delta x \left(\frac{f_{i,0}}{2}, f_{i,1}, \dots, f_{i,m-1}, \frac{f_{i,m}}{2} \right)$$

times the vector $w^{(\nu)}$. In (7.31) the first term $w_{0,\nu}$ and the last term $w_{m,\nu}$ (where boundary conditions enter) must be treated separately in case we deal with the short vector $(w_1, \dots, w_{m-1})$ as in Section 4.2.3. Now assemble all the rows into an $(m + 1)^2$-matrix C. Then for all i within time level ν, the integrals are represented by the product

$$C w^{(\nu)}.$$

Neglecting the fact that many of its elements are close to zero, the matrix C is dense, which reflects the nonlocal character of the integral. This is in contrast to the local character of standard finite differences with its tridiagonal

[5] The entries of a Toeplitz matrix are constant along each diagonal.

matrices. The transformation (7.28) is different from (4.3), but tridiagonal matrices can be derived from (7.30) in a similar way as done in Chapter 4. The dense matrix C adds to the tridiagonal matrices, which makes the solution of linear systems with full matrices in each time step $\nu \to \nu + 1$ more expensive. In an attempt to save costs, *splitting* has been suggested. This means to evaluate the integral at the previous line (ν). In this way, the multiplication Cw only shows up in the right-hand side of the known terms. The tridiagonality of the left-hand side matrices is maintained, and the method still converges. Up to boundary conditions, this splitting can be represented by an Euler-type implicit scheme

$$\frac{w^{(\nu+1)} - w^{(\nu)}}{\Delta\tau} = Gw^{(\nu+1)} + \lambda C w^{(\nu)} \,,$$

where the matrix G represents the local information of the differentials. Neither G nor C are symmetric. We leave it to the reader to set up the system of equations (⟶ Exercise 7.7). The matrices G and C are used for the analysis, no matrix is needed for the algorithm. — For an illustration how a larger intensity λ increases the value of an option see Figure 7.4.

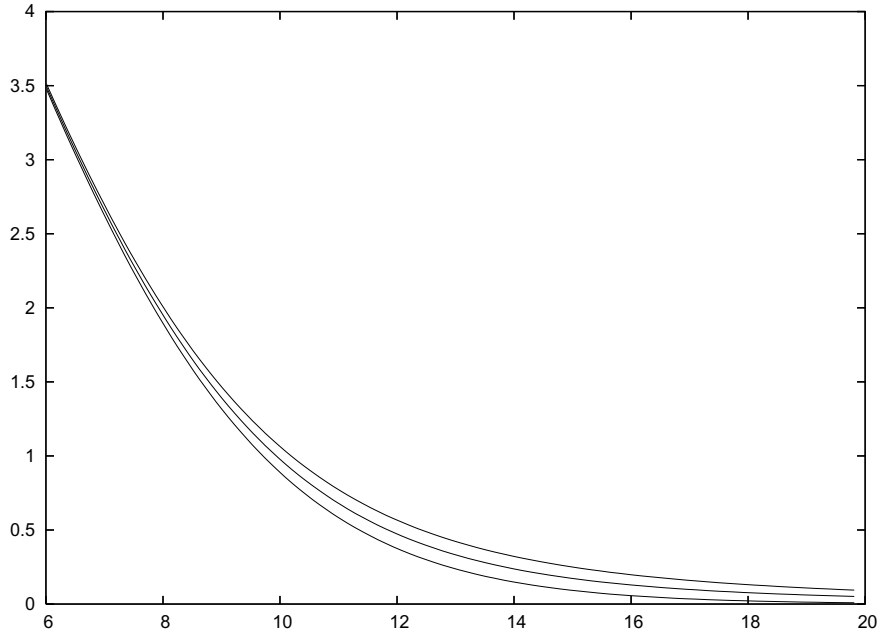

Fig. 7.4. $V(S,0)$ of a European put option, solution of Problem 7.9; parameters as in Example 1.21: $K = 10$, $r = 0.06$, $\sigma = 0.3$, $T = 1$, with Merton's jump diffusion, $\mu_J = -0.3$, $\sigma_J = 0.4$, and three values of jump intensity λ: 0 (lower curve, no jump), 0.1, and 0.2 (top curve); $x_{\min} = -3$, $x_{\max} = \log(K) + 1.6 = 3.9$. The chosen value of $\mu_J = -0.3$ corresponds to $q = \exp(\mu_J) = 0.74$, or a 26% fall in the asset price.

Since the splitting can deteriorate the accuracy, a fixed point iteration has been suggested [dHaFV05]. The integral term $\mathsf{E}(V)$ with its truncation and discretization challenges the control of the involved errors. For example, [CoV05] give an estimate of the error induced by truncating the integral, as well as a convergence proof for finite differences applied to general Lévy models. Codes for American options based on a penalty formulation or on an LCP formulation can be easily modified and extended by an integral term. The techniques of Chapter 4 or Chapter 5 can be applied. Application of FFT increases the efficiency [dHaFV05]. Typically, each Lévy process calls for a separate algorithm. A Monte Carlo approach is [MeA02]. For Merton's model and European options, an analytic solution is given [Mer76], which allows to test corresponding algorithms.

7.4 Application of the Fourier Transform

The Fourier transform $\mathcal{F}$ of a real function f is defined by[6]

$$\mathcal{F}[f(u)] := \int_{y=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}uy} f(y)\,\mathrm{d}y\,. \tag{7.32}$$

This requires integrability of f. The inverse Fourier transformation is

$$\mathcal{F}^{-1}[g(x)] = \frac{1}{2\pi}\int_{u=-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}xu} g(u)\,\mathrm{d}u\,, \tag{7.33}$$

A sufficiently well-behaved f is recovered by the inversion,

$$f = \mathcal{F}^{-1}\mathcal{F} f\,.$$

We perform this process of transform and inverse transform for a function $c(k)$ to be defined below. The application of the Fourier transform in our context and the outline of three steps of the subsequent analysis is symbolized as follows:

$$\begin{array}{ccc} & \overset{(1)}{\longrightarrow} & \\ c(k) & \circ \quad \bullet & g(u) = \text{integral} \\ & & \downarrow (2) \\ c(k) & \circ \quad \bullet & g(u) = \text{formula} \\ & \underset{(3)}{\longleftarrow} & \end{array}$$

Step (1) is the forward Fourier transform (7.32) of a function $c(k)$. The result is an integral expression $g(u)$. In our context this integral can be solved

[6] There are different conventions for the Fourier transform; for background, see special literature, for example [Vre03]. To get used to it try Exercise 7.8.

analytically (step (2)), which produces a formula for $g(u)$. The inverse transformation (7.33) in step (3) is approximated numerically by the Fast Fourier Transformation (FFT), based on (C1.8). The detour (1)–(3) is worth the effort, because the FFT calculation of $c(k)$ is faster to evaluate than the original $c(k)$.

Recall the characteristic function (7.21) ϕ of a Lévy process X_t. These functions are the Fourier transform of the density function of X,

$$\phi_{X_t}(u) := \mathsf{E}(\exp(\mathrm{i}uX_t)) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}uX} f_{\text{density}X}\,\mathrm{d}x = \mathcal{F}[f_{\text{density}X}]\,. \tag{7.34}$$

The characteristic functions ϕ of many processes X are known and available as analytical expressions, for example, in [Sch03], [ConT04], [KwLW12].

In the following, we investigate a European call with vanilla payoff $\Psi(S) = (S-K)^+$ with an arbitrary underlying Lévy process S_t. The integral representation of the call's value under the risk-neutral measure Q is

$$\begin{aligned} V(S_t,t;\,K) &= \mathrm{e}^{-r(T-t)}\mathsf{E}_{\mathsf{Q}}[\Psi(S_T)\,|\,S_t] \\ &= \mathrm{e}^{-r(T-t)}\int_{S_T=K}^{\infty}(S_T-K)\,f_{\text{density}}(S_T)\,\mathrm{d}S_T \end{aligned}$$

where f is the density of S_T of the Lévy process starting at t with the value S_t. Transform

$$S_T = \mathrm{e}^s,\ K = \mathrm{e}^k,\ \mathrm{d}S_T = \mathrm{e}^s\mathrm{d}s\,; \tag{7.35}$$

note that $k \in \mathbb{R}$. Then

$$V(S_t,t;\,K) = \mathrm{e}^{-r(T-t)}\int_k^{\infty}(\mathrm{e}^s-\mathrm{e}^k)\hat{f}(s)\,\mathrm{d}s$$

where $\hat{f}(s) = \mathrm{e}^s f(\mathrm{e}^s)$ is the density of $\log S$, similar as in Section 1.8.2. Following [CaM99], in order to make the function integrable, we scale the integral with a factor $\exp(\alpha k)$ (a constant):

$$c(k) := \mathrm{e}^{\alpha k}\mathrm{e}^{-r(T-t)}\int_k^{\infty}(\mathrm{e}^s-\mathrm{e}^k)\hat{f}(s)\,\mathrm{d}s = \mathrm{e}^{\alpha k}V(S_t,t;\,K) \tag{7.36}$$

and denote $\mathcal{F}[c(u)]$ its Fourier transform. We leave the choice of α open until later.

As outlined above, when $\mathcal{F}[c]$ is calculated, then the call's value $V(S,t)$ is recovered from the inverse Fourier transformation,

$$V(S_t,t;\mathrm{e}^k) = \left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}ux}\mathcal{F}[c(u)]\,\mathrm{d}u\right)\cdot\mathrm{e}^{-\alpha k}\,,$$

which can approximated efficiently by the Fast Fourier Transform (FFT). This outlines the program of the three steps (1),(2),(3), and now we turn to its realization.

The Fourier transform of $c(k)$ is

$$\begin{aligned}\mathcal{F}[c(u)] &= \int_{k=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}uk} c(k)\,\mathrm{d}k \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}uk}\mathrm{e}^{\alpha k}\mathrm{e}^{-r(T-t)} \int_{s=k}^{\infty} (\mathrm{e}^{s} - \mathrm{e}^{k})\hat{f}(s)\,\mathrm{d}s\,\mathrm{d}k \\ &= \mathrm{e}^{-r(T-t)} \int_{k=-\infty}^{\infty}\int_{s=k}^{\infty} \mathrm{e}^{(\mathrm{i}u+\alpha)k}(\mathrm{e}^{s} - \mathrm{e}^{k})\hat{f}(s)\,\mathrm{d}s\,\mathrm{d}k \\ &= \mathrm{e}^{-r(T-t)} \int_{s=-\infty}^{\infty}\int_{k=-\infty}^{s} \mathrm{e}^{(\mathrm{i}u+\alpha)k}(\mathrm{e}^{s} - \mathrm{e}^{k})\hat{f}(s)\,\mathrm{d}k\,\mathrm{d}s\end{aligned}$$

where the last equation holds since

$$\{k \le s < \infty \,|\, -\infty < k < \infty\} = \{-\infty < k \le s \,|\, -\infty < s < \infty\}\,.$$

This leads to

$$\begin{aligned}\mathcal{F}[c(u)] &= \mathrm{e}^{-r(T-t)} \int_{-\infty}^{\infty} \hat{f}(s) \int_{-\infty}^{s} [\mathrm{e}^{(\mathrm{i}u+\alpha)k+s} - \mathrm{e}^{(\mathrm{i}u+\alpha+1)k}]\,\mathrm{d}k\,\mathrm{d}s \\ &= \mathrm{e}^{-r(T-t)} \int_{-\infty}^{\infty} \hat{f}(s) \left[\frac{\mathrm{e}^{s}\mathrm{e}^{(\mathrm{i}u+\alpha)k}}{\mathrm{i}u+\alpha} - \frac{\mathrm{e}^{(\mathrm{i}u+\alpha+1)k}}{\mathrm{i}u+\alpha+1}\right]_{k=-\infty}^{s} \mathrm{d}s\,.\end{aligned} \tag{7.37}$$

To have the integral exist, we require the factor $\mathrm{e}^{\alpha k}$ to vanish for $k \to -\infty$, which leads to choose $\alpha > 0$. That is, the factor $\exp(\alpha k)$ amounts to a damping of the integral. The bracketed term in (7.37) is

$$\frac{(\mathrm{i}u+\alpha+1)\mathrm{e}^{s(\mathrm{i}u+\alpha+1)} - (\mathrm{i}u+\alpha)\mathrm{e}^{s(\mathrm{i}u+\alpha+1)}}{\mathrm{i}u(2\alpha+1) + \alpha(\alpha+1) - u^2}\,,$$

and we come up with

$$\mathcal{F}[c(u)] = \frac{\mathrm{e}^{-r(T-t)}}{\mathrm{i}u(2\alpha+1) + \alpha(\alpha+1) - u^2} \int_{-\infty}^{\infty} \hat{f}(s)\mathrm{e}^{\mathrm{i}s(u-(\alpha+1)\mathrm{i})}\,\mathrm{d}s\,.$$

We denote the integral therein $\phi(u-(\alpha+1)\mathrm{i})$, because it is the characteristic function of the density $\hat{f}$. For ϕ an analytic expression is known. Hence

$$\mathcal{F}[c(u)] = \frac{\mathrm{e}^{-r(T-t)}\,\phi(u-(\alpha+1)\mathrm{i})}{\alpha^2+\alpha-u^2+\mathrm{i}u(2\alpha+1)} =: g(u) \tag{7.38}$$

can be considered to be a known function g, and step (2) is completed. For the final choice of the parameter $\alpha > 0$ we further request $g(u) = \mathcal{F}[c(u)]$ to be integrable as well. Since the integration is along real values of u one has to take care that the denominator has only imaginary roots in u. The choice of α is discussed in the literature [CaM99], [KwLW12]. Usually $\alpha = 3$ works well.

The inverse Fourier transformation evaluates

$$\mathrm{e}^{-\alpha k}\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}ku}g(u)\,\mathrm{d}u\,.$$

The integral is real, and hence its integrand is real too. Think of g from (7.38) being split into real part and imaginary part, $g(u) = g_1(u) + \mathrm{i}g_2(u)$. Then $\mathrm{i}(\cos(ku)g_2(u) - \sin(ku)g_1(u)) = 0$, and we conclude that $g_1(u)$ is an even function, and $g_2(u)$ is an odd function. Hence the integrand

$$\cos(ku)g_1(u) + \sin(ku)g_2(u)$$

is even, and the value of the call is

$$V(S_t, t; \mathrm{e}^k) = \frac{\mathrm{e}^{-\alpha k}}{\pi}\int_0^{\infty}\mathrm{e}^{-\mathrm{i}ku}g(u)\,\mathrm{d}u\,. \tag{7.39}$$

Next, the semi-infinite integration interval is truncated to finite length A. Thereby, for most Lévy models the truncation error can be made arbitrarily small because the characteristic function ϕ decays exponentially fast at infinity.[7] With the restriction to the integration interval $0 \le u \le A$ and $M-1$ subintervals with equal length Δu, the discrete grid points are

$$u_j := j\Delta u = j\frac{A}{M-1}\,,\quad j = 0, \ldots, M-1\,.$$

Choosing the trapezoidal sum (C1.2) for the quadrature, the approximation is

$$\int_0^{\infty}\mathrm{e}^{-\mathrm{i}ku}g(u)\,\mathrm{d}u \;\approx\; \frac{A}{M-1}\sum_{j=0}^{M-1}\beta_j\, g(u_j)\,\mathrm{e}^{-\mathrm{i}ku_j} \tag{7.40}$$

with weights $\beta_0 = \beta_{M-1} = \frac{1}{2}$ and $\beta_j = 1$ for $1 \le j \le M-2$. The trapezoidal sum goes along with a sampling error of the order $O(\Delta u^2)$.

So far, the log-strike $k = \log K$ is not specified. The aim is to exploit the potential of FFT, which calculates sums of the type

$$\sum_{j=0}^{M-1} a_j\,\mathrm{e}^{-\mathrm{i}\nu j\frac{2\pi}{M}} \tag{7.41}$$

for complex numbers $a_0, \ldots, a_{M-1}$, one sum for each ν. This amounts to calculate a *vector* of M such sums, for $\nu = 0, \ldots, M-1$. Applying FFT we gain the possibility to calculate for M strikes simultaneously. Let us calculate the call values for the log-strike values

$$k_\nu := -b + \Delta k \cdot \nu\,,\quad \nu = 0, \ldots, M-1\,, \tag{7.42}$$

[7] This does not hold for the VG process, see [ConT04], [KwLW12].

for suitable values of b and Δk, which define the k-range and the strike spacing of interest. Substituting these values k_ν into the above sum (7.40) produces

$$\frac{A}{M-1}\sum_{j=0}^{M-1}\beta_j\, g(u_j)\exp[-\mathrm{i}(-b+\Delta k\,\nu)j\frac{A}{M-1}]\,.$$

The argument of the exponential function is

$$\mathrm{i}bj\frac{A}{M-1}-\mathrm{i}\nu j\Delta k\frac{A}{M-1}\,.$$

To apply FFT aiming at (7.41), steps Δk and $\Delta u = \frac{A}{M-1}$ must be chosen such that

$$\Delta k\frac{A}{M-1}=\Delta k\;\Delta u=\frac{2\pi}{M}\,. \tag{7.43}$$

Then the sum in (7.40) is

$$\frac{A}{M-1}\sum_{j=0}^{M-1}\beta_j g(u_j)\exp[\mathrm{i}bj\frac{A}{M-1}]\,\mathrm{e}^{-\mathrm{i}\nu j\frac{2\pi}{M}}\,,$$

which is the standard FFT applied to (7.41) for the complex numbers

$$a_j := A\beta_j g(u_j)\exp[\mathrm{i}bj\frac{A}{M-1}]\,,\quad i=0,\ldots,M-1\,. \tag{7.44}$$

This completes the calculation of a bunch of European call values: The integral in (7.39) is approximated by the FFT sum (7.41) with coefficients (7.44). For the highly efficient calculation of the FFT sums (7.41) consult standard literature on numerical analysis (such as [PrTVF92]), and related software packages.

The above method amounts to a fast algorithm in case option prices are to be calculated on a grid of many strikes, all options with the same maturity T. The log-strike grid of the values k_ν is defined by (7.42) with the parameters b and Δk, which in turn are based on A, M. By (7.43),

$$\Delta k=\frac{2\pi}{A}\frac{M-1}{M}\,.$$

And to cover log strikes in the at-the-moment range around $k=0$, one aims at

$$b=\frac{(M-1)\Delta k}{2}\,.$$

Efficiency of FFT is maximal for M a power of 2. The equation (7.43) is a limitation that requests a careful design of parameters M and A.

In this section, we have explained the classical FFT approach of Carr and Madan [CaM99]. The Fast Fourier Transform can be applied also for early-exercise options [LoFBO08]. A novel transform is based on Fourier-cosine

expansions [FaO08], which is also applied to barrier options [FaO09]. The resulting algorithms converge exponentially fast. In summary, FFT-based methods have shown a rich potential, in particular for option pricing under Lévy models.

Notes and Comments

on Section 7.1:

For a critical account of Leland's approach see [ZhZ07]. The nonlinear version (7.4) – (7.6) is due to [HoWW94]. A piecewise linear treatment is suggested in [ChHK04]. The paper [AvP96] discusses equation (7.5), suggesting a modification for the case $\gamma \geq 1$, where $\hat{\sigma}^2$ would be negative for $\Gamma < 0$. For bounds on V in case of "misspecified" volatility, see [ElKJS98]. For related work, consult also [Gra01], [Ehr08], [GlDN10].

Apart from the one-factor case, ranges for parameters play a role also in multiasset cases. For example, consider two assets with prices S_1, S_2, and assume a correlation in the range $-1 \leq \rho_{\min} \leq \rho \leq \rho_{\max} \leq 1$. In the Black–Scholes equation (6.2), the term

$$\rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 V}{\partial S_1 \partial S_2}$$

occurs. Depending on the sign of the cross derivative $\frac{\partial^2 V}{\partial S_1 \partial S_2}$, ρ is chosen either as $\rho_{\min}$ or $\rho_{\max}$ in order to characterize a "worst-case," see [Top05].

To complete the introduction into more general models we have outlined the Dupire equation in Appendix A6.

on Section 7.2:

For reference and examples consult [Hei10], [HeS10], [FoV12]. The assumption of a constant c_+ in Theorem 7.2 is not always easily satisfied. For example, in the Barles and Soner model of Section 7.1.2 and a payoff with jump discontinuity (as digital option), $c_+ = c_+(\Delta x) = O(\Delta x^2)$, which affects the assumptions of Theorem 7.2, and has strong implications on stability. Apart from nonsmooth payoffs, also the PDE itself is typically not smooth. For American options, the penalty term in (7.12) causes a lack of smoothness. Also the volatility function $\tilde{\sigma}$ may be nonsmooth. This happens, for example, in Leland's model when V_{SS} changes sign. Newton method then works with a generalized derivative. The higher the degree of "non-smoothness," the worse the convergence rate of CN. The BDF method (7.20) is highly recommended. An a priori check of convergence criteria is advisable.

on Section 7.3:

The definition of Lévy processes includes stochastic continuity. A table of Lévy densities f_{L} is in [Sch03] p.154. The Lévy-Khinchin representation (7.22) is a scalar setting; [CaW04] develops analytic expressions for the characteristic function of time-changed Lévy process in a general vector setting. In this framework, Heston's stochastic-volatility model can be represented as time-changed Brownian motion.

For time-changed Lévy processes, consult [AnéG00], [CaGMY03], [ConT04], [CaW04]. Time-changed Lévy processes have been successfully applied to match empirical data. For processes with density function (Merton, VG, NIG), Algorithm 1.18 can be applied [Que07]. Lévy-process models have been extended by incorporating stochastic volatilities [CaGMY03], [Kal06]. A subordinator $\tau(t)$ can be constructed as integral of a square-root process.

[Pha97] investigates properties of American options. Heston presents the characteristic function for his model in [Hes93]. His model extended by jump diffusion [Bat96] can be cast into the above framework: in this case a two-dimensional PDE is considered. For computational approaches see [AnA00], [MaPS02], [BrLN04], [AlO05], [dHaFV05], [dHaFL05], [CoV05], [AlO06].

on Section 7.4:

Choosing the weights w_j of Simpson's sums instead of trapezoidal sums, the integrations get more accurate. An application to VG is found in [CaM99]. Modifications and extensions of the above basic approach are described and reviewed in [KwLW12]. For references on transform methods in option pricing, see [FaO09].

Exercises

Exercise 7.1

Let ΔW be the increment of a Wiener process, see Section 1.6.1. Show

$$\mathsf{E}(|\Delta W|) = \sqrt{\Delta t}\,\sqrt{\frac{2}{\pi}}\,.$$

Exercise 7.2 Barles–Soner Model

The differential equation of Barles and Soner is:

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{f(x)+1}{2\sqrt{xf(x)}-x} \quad \text{with } f(0)=0\,.$$

a) By numerical computations, analyze the solution for $-2 \le x \le 2$.
b) Construct an approximating function $\hat{f}(x)$ in a piecewise fashion.

Exercise 7.3 Transformation of Nonlinear Black–Scholes Models

According Section 7.2, consider the following nonlinear PDE

$$V_t + \frac{1}{2}\sigma^2(t, S, V_{SS})S^2 V_{SS} + (r - \delta)SV_S - rV + \hat{p}\max(\Psi - V, 0) = 0\,,$$

where $\sigma^2(t, S, V_{SS})$ depends on the particular model; r is the risk-free interest rate and δ is the continuous dividend yield. Apply the transformation (7.13)

$$x = \log(S/K),\ \tau = \sigma_0^2(T - t)/2,\ u(x, \tau) = \mathrm{e}^{-x}V(S, t)/K,$$

with $K > 0$ and a model-dependent parameter σ_0, and derive a PDE for u.

Exercise 7.4 Payoffs of Spreads

We consider portfolios of two or more options of the same type with the same underlying stock. K_1, K_2, K are strikes with $K_1 < K_2$.

a) A *butterfly spread* is a portfolio with
- one long call with strike K_1,
- one long call with strike K_2,
- two short calls with strike $K = \frac{K_2 - K_1}{2}$.

The payoff is

$$\Psi(S) = \begin{cases} 0 & \text{for } S \le K_1 \\ S - K_1 & \text{for } K_1 < S \le K \\ K_2 - S & \text{for } K < S \le K_2 \\ 0 & \text{for } K_2 \le S \end{cases}$$

b) A *bull spread* is a portfolio with
- one long call with strike K_1,
- one short call with strike K_2,

The payoff is

$$\Psi(S) = \begin{cases} 0 & \text{for } S \le K_1 \\ S - K_1 & K_1 < S \le K_2 \\ K_2 - K_1 & K_2 < S \end{cases}$$

For both spreads explain and sketch the payoff. Apply the transformation (7.13) (Exercise 7.3) to derive the transformed payoff $u^*(x)$. For b), apply the transformation with K_2.

Exercise 7.5 Convergence of the Fully Implicit Method

Two out of the three criteria for monotony in Theorem 7.2 are (i) and (ii). For

a) Leland's model of transaction costs, with parameter γ, and

b) the model of uncertain volatility with $\sigma_{\min} \le \sigma \le \sigma_{\max}$,

show that (i) and (ii) are satisfied. What are the constants c_+? For b), σ^- of (7.8b) suffices.

Exercise 7.6

For a Lévy process X_t adapted to a filtration $\mathcal{F}_t$ show

$$\mathsf{E}(\mathrm{e}^{X_t} \,|\, \mathcal{F}_s) - \mathrm{e}^{X_s} = \mathsf{E}(\mathrm{e}^{X_{t-s}}) - \mathrm{e}^{X_0} \,.$$

Exercise 7.7 Project: Implementing a PIDE

Set up a computer program to solve Merton's jump diffusion (7.30) numerically. To this end, concentrate on European-style vanilla options. Set up boundary conditions using (4.18), and use a BDF implicit scheme. Think of how to choose $x_{\min}$, $x_{\max}$ in relation to the strike K.
Hint: For testing the core part of the program, set the jump intensity $\lambda = 0$ and compare to the Black–Scholes value.

Exercise 7.8 Fourier Transform

Consider the Fourier transform

$$\mathcal{F}[f(u)] := \int_{-\infty}^{\infty} e^{\mathrm{i}uy} f(y)\,\mathrm{d}y \,.$$

For the example $f(y) := e^{-a|y|}$ and complex a show that

$$\int_{-A}^{A} e^{\mathrm{i}uy} f(y)\,\mathrm{d}y$$

converges for $A \to \infty$ and $\mathrm{Re}(a) > 0$.