Convergence analysis

we analysed convergence of forward one-step schemes of the form

$$\begin{aligned} u^m &= Lu^{m-1}, \quad m > 1; && (6.38) \\ u^0 & \quad \text{given}; && (6.39) \\ \text{find } u^M & \quad \text{for some } M > 0, && (6.40) \end{aligned}$$

for forward PDEs. The fact that we are concerned with backward equations here, formally, merely changes the direction of travel,

$$\begin{aligned} V^{m-1} &= KV^m, \quad m > 1; && (6.41) \\ V^M & \quad \text{given for some } M > 0; && (6.42) \\ \text{find } V^0. & && (6.43) \end{aligned}$$

The backward PDE can be written as a forward PDE and vice versa by time reversal, $t \to T-t$, and a finite difference scheme for the forward equation thus turns into a finite difference scheme for the backward equation by counting timesteps backwards, $m \to M - m$.

The analysis in 4.1 and 4.2 reveals that convergence of finite difference schemes is a result of consistency and stability. To recap, the consistency analysis is based on the so-called truncation error of the scheme, which measures the closeness of the discretised and continuous equations. This translates immediately to the backward context. The same is true in principle for stability, the notion of boundedness of solutions, with the proviso that option pricing PDEs are routinely degenerate at boundaries which falls outside the previous framework. Maximum norm stability is analysed relatively easily because maximum principles only rely on a local comparison of the terms in the (discretised) PDE. This is more involved for l_2-stability, and we present a new technique, a so-called "energy method", and apply it to Black-Scholes-type PDEs.

6.2.1 Truncation error for Black-Scholes-type problems

Recall that the truncation error measures how well the solution to the PDE, evaluated at the grid points, satisfies the difference scheme. The notion of consistency expresses that in the limit for vanishing grid size and timestep the discrete equation approaches the continuous one. The consistency order measures the speed. We adapt the definition from the forward PDE for clarity.

Definition 6.2.1 (Truncation error for backward equations)**.** Let V be the solution to a PDE with time-coordinate t. The *truncation error* of a backward one-step difference scheme of the form (6.41) is defined as

$$T(.,t) = \frac{1}{\Delta t}\left(u(.,t-\Delta t) - Ku(.,t)\right).$$

Based on this definition of the truncation order, the notions of consistency and consistency order are as per points 2. and 3. in Definition 4.2.3.

Example 6.2.2. *The explicit finite difference scheme with central differences is consistent of order $p=2$, $q=1$. This can be verified by Taylor expansion of*

$$\begin{aligned} V(S_n, t_{m-1}) &= V - \Delta t\frac{\partial V}{\partial t} + \tfrac{1}{2}\Delta t^2\frac{\partial^2 V}{\partial t^2} + o(\Delta t^2) \\ V(S_{n\pm1}, t_m) &= V \pm \Delta S\frac{\partial V}{\partial S} + \tfrac{1}{2}\Delta S^2\frac{\partial^2 V}{\partial S^2} \pm \frac{1}{6}\Delta S^3\frac{\partial^3 V}{\partial S^3} + \frac{1}{24}\Delta S^4\frac{\partial^4 V}{\partial S^4} + o(\Delta S^4), \end{aligned}$$

where arguments (S_n, t_m) of V and its derivatives are omitted. From the definition of the truncation error,

$$\begin{aligned} T(S_n, t_m) &= \frac{V(S_n,t_{m-1}) - V(S_n,t_m)}{\Delta t} - \frac{1}{2}\sigma^2(S_n,t_m)\frac{V(S_{n+1},t_m) - 2V(S_n,t_m) + V(S_{n-1},t_m)}{\Delta S^2} \\ &\quad -\mu(S_n,t_m)\frac{V(S_{n+1},t_m) - V(S_{n-1},t_m)}{2\Delta S} + r(S_n,t_m)V(S_n,t_m) \\ &= -\frac{\partial V}{\partial t} + \tfrac{1}{2}\Delta t\frac{\partial^2 V}{\partial t^2} + o(\Delta t) - \frac{1}{2}\sigma^2(S_n,t_m)\frac{\partial^2 V}{\partial S^2} - \frac{1}{24}\sigma^2(S_n,t_m)\Delta S^2\frac{\partial^4 V}{\partial S^4} + o(\Delta S^2) \\ &\quad -\mu(S_n,t_m)\frac{\partial V}{\partial S} - \frac{1}{6}\Delta S^2\mu(S_n,t_m)\frac{\partial^3 V}{\partial S^3} + o(\Delta S^2) + rV \\ &= -\frac{1}{24}\sigma^2(S_n,t_m)\Delta S^2\frac{\partial^4 V}{\partial S^4} - \frac{1}{6}\Delta S^2\mu(S_n,t_m)\frac{\partial^3 V}{\partial S^3} + \tfrac{1}{2}\Delta t\frac{\partial^2 V}{\partial t^2} + o(\Delta S^2) + o(\Delta t). \end{aligned} \tag{6.44}$$

Similarly, by Taylor expansion exactly as for the forward PDEs, one shows the following for the truncation error of the θ-scheme.

Proposition 6.2.3. *The θ-scheme with central differences is consistent of order 2 in ΔS and of order 1 in Δt, unless $\theta = 1/2$ (Crank-Nicolson), for which the order in Δt is 2.*

For a put in the Black-Scholes model, as the analytical solution is known explicitly, one can actually compute the truncation error from 6.44 for illustration. The individual terms and the overall truncation error are plotted for continuous argument S and $t = 0$ in Fig. 6.4 for the explicit Euler scheme.

The truncation error is largest in absolute terms around the strike. By closer inspection, the location of the maximum error is somewhat below the strike. To understand this, note that the PDE (and the numerical scheme, respectively) does not only propagate the solution, but equally the error. The drift present in the PDE has moved the peak towards the left. This means that for most of the time the truncation error around the strike is positive, and explains why the numerical solution in Table 6.1 is smaller than the exact one.

It is also observed in Table 6.1 that reducing the timestep for fixed gridsize increases the error, if marginally, which may seem counter-intuitive at first. Looking at the individual terms of the truncation error, one sees that the peak in the term proportional to Δt, and the total truncation error, have opposite signs, which explains why the truncation error can become larger when the timestep is reduced. For $\Delta t \to 0$, and ΔS fixed, the numerical solution converges to a semi-discrete one given by a system of ODEs in time.

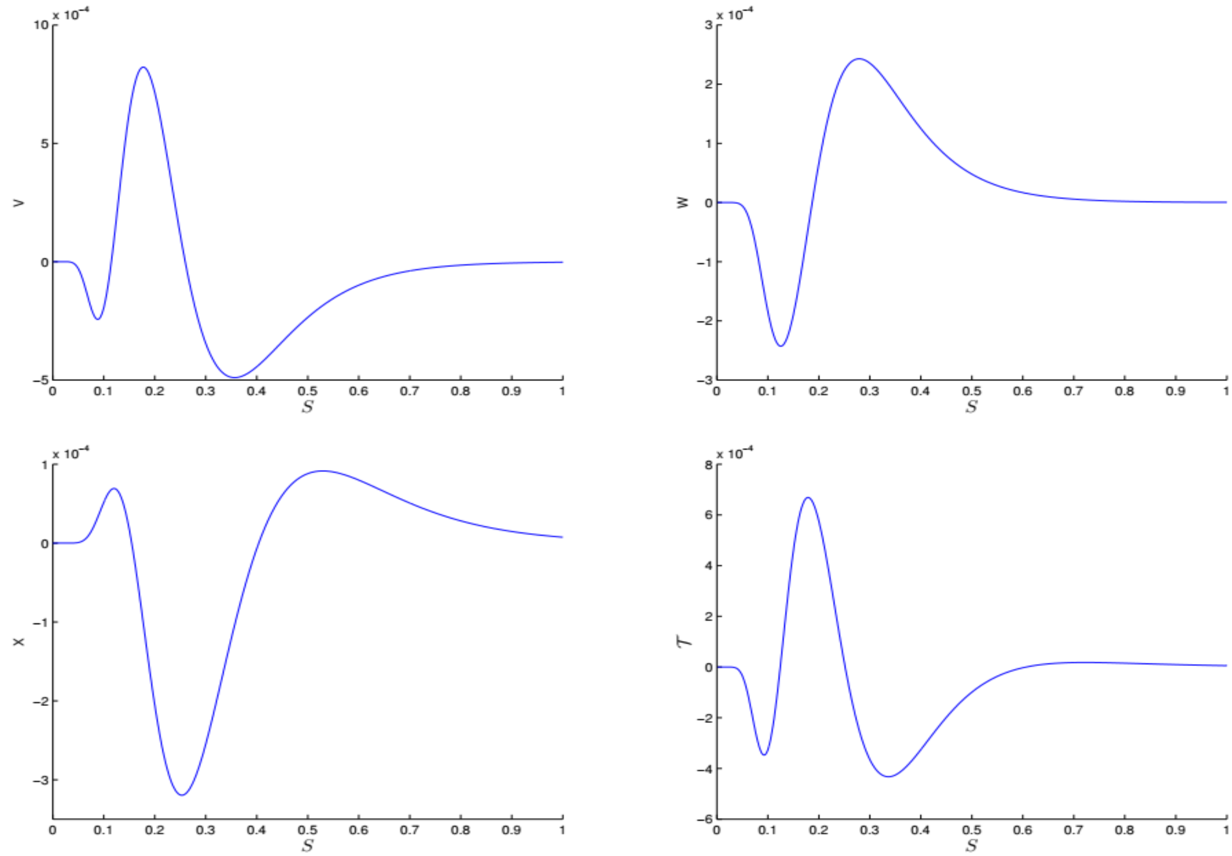

Figure 6.4: The truncation error for the explicit Euler central difference scheme for the Black-Scholes PDE, and its individual terms. Parameters are $\sigma = 0.4$, $r = 0.05$, $T = 1$, $t = 0$, $K = 0.25$.

6.2.2 Maximum norm stability for Black-Scholes-type problems

Recall the notion of stability from Definition 4.2.1, with the interpretation that perturbations of the solution stay within the order of magnitude of their initial value.

Example 6.2.4 (Explicit Euler with central differences). *By analogy with the forward equation (see Example 5.2.2 and Remark 5.2.3), a scheme of the form (6.20) satisfies a discrete maximum principle if*

$$A_n, B_n, C_n \geq 0, \tag{6.45}$$
$$A_n + B_n + C_n \leq 1. \tag{6.46}$$

This follows by induction from

$$|V_n^{m-1}| \leq \max(|V_{n-1}^m|, |V_n^m|, |V_{n+1}^m|),$$

and therefore $|V^0| \leq |V^1| \leq \dots |V^M|$, *where* $|\cdot|$ *the maximum norm.*

To check when (6.46) holds, first note that

$$A_n^m + B_n^m + C_n^m = 1 - r(S_n, t_m)\Delta t \leq 1$$

if $r \geq 0$, *which is generally the case. Now, for non-negativity of the coefficients, (6.45), it is necessary that*

$$B_n^m \geq 0 \quad \Leftrightarrow \quad \sigma^2(S_n, t_m)\Delta t/\Delta S^2 + r(S_n, t_m)\Delta t \leq 1 \tag{6.47}$$

for all $1 \leq n \leq N-1$, *and*

$$A_n^m, C_n^m \geq 0 \quad \Leftrightarrow \quad \Delta S|\mu(S_n, t_m)| \leq \sigma^2(S_n, t_m) \tag{6.48}$$

for $1 \leq n \leq N-1$. *The critical constraint is usually* $\sigma^2\Delta t/\Delta S^2 \leq 1$, *because* Δt *and* ΔS *and therefore the remaining terms in (6.47) and (6.48) are both small. But see the next example.*

Example 6.2.5 (Black-Scholes). *Continuing the previous example, specifically, applying Example 6.2.4 to the Black-Scholes PDE, with coefficients as in (6.20), one needs to satisfy the stability conditions*

$$B_n^m \geq 0 \quad \Leftrightarrow \quad n^2\sigma^2\Delta t + r\Delta t \leq 1$$

for all $0 \leq n \leq N-1$, *and*

$$A_n^m \geq 0 \quad \Leftrightarrow \quad r \leq \sigma^2 n \tag{6.49}$$

for $n \geq 1$. *The critical cases are for large and small n respectively. The explicit Euler scheme is therefore stable, if*

$$M = T/\Delta T \geq T\left(\sigma^2(N-1)^2 + r\right) \tag{6.50}$$

and

$$r \le \sigma^2. \tag{6.51}$$

Condition (6.50) is the expected timestep constraint for explicit schemes. For large N it essentially requires that $M \ge \sigma^2 T N^2$, which confirms precisely the empirical results from Table 6.1. Condition (6.51) is new and arises from the way in which the coefficients degenerate towards the zero boundary.

This can be addressed by upwinding as follows. For some k large enough, $r \le \sigma^2 n$ for $n \ge k$, so (6.49) is given for $n \ge k$. For $n < k$, upwinding can be used, resulting in slightly modified and always non-negative coefficients. As k is fixed and does not change when the total number N of grid points increases, the point S_k moves closer to 0 and this "local" upwinding has no adverse impact on the convergence order. This procedure is rarely necessary in practice though, even if (6.51) is violated.

Example 6.2.6 (Implicit Euler with central differences)**.** *Observe for the coefficients in (6.25)*

$$a_n^m + b_n^m + c_n^m = 1 + r(S_n, t_m)\Delta t \ge 1, \tag{6.52}$$

if $r \ge 0$ as usual. Assume now (6.48) holds again,

$$\sigma^2(S_n, t_m) \ge \Delta S |\mu(S_n, t_m)| \tag{6.53}$$

everywhere, as needed for stability of central differences in the explicit scheme in Example 6.2.4. Then this implies

$$a_n^m, c_n^m \le 0 \tag{6.54}$$

$$b_n^m \ge 1. \tag{6.55}$$

This guarantees convergence in the maximum norm, as seen for the forward equation in 5.2.2.

Recall, once again, that the θ-scheme is equivalent to a "fractional step" scheme with

1. an explicit Euler step of length $(1-\theta)\Delta t$,
2. followed by an implicit Euler step of length $\theta\Delta t$.

From this remark it follows that the θ-scheme will be stable if both substeps are stable, i.e. the coefficients A-C in (6.31)-(6.33) satisfy (6.45), and the coefficients a-c in (6.28)-(6.30) satisfy (6.52) and (6.54). This shows the following.

Proposition 6.2.7. *The θ-central difference scheme satisfies a discrete maximum principle, and is consequently stable in the maximum norm, if*

$$\sigma^2(S,t) \geq \Delta S\,|\mu(S,t)| \tag{6.56}$$

and

$$(1-\theta)\Delta t\left(\frac{\sigma^2(S,t)}{\Delta S^2}+r\right) \leq 1 \tag{6.57}$$

for all S and t.

The condition (6.57) can usually be satisfied by choosing ΔS small enough, provided $\sigma(S_n, t_m)$ is not degenerate. Indeed, the equation *is* degenerate for the Black-Scholes model, see Example 6.2.5, and the work-around explained there.

Similar to the heat equation, there appears to be a stability constraint on Δt if $\theta \neq 1$. Further analysis would reveal that such a condition is necessary for a maximum principle to apply, i.e. maximum norm stability with an amplification factor of $C = 1$ in Definition 4.2.1, but that the scheme is stable (however with $C > 1$) under weaker conditions, namely identical to those for l_2 stability. This is in line with numerical experiments, where all schemes with $\theta \geq 1/2$, including the Crank-Nicolson scheme (see Table 6.2), appear unconditionally stable. We refer for the analysis in the maximum norm to [Thomée, 1990] and sketch an analysis explaining this behaviour in the l_2 norm in 6.2.3.

6.2.3 Mean-square stability for Black-Scholes-type problems

Analysing l_2-stability for Black-Scholes problems becomes more involved than for the heat-equation. This is partly because of "non-normality", as discussed in 5.2.3 and especially Example 7 in 5.4, but more severely the degeneracy of the pricing equations at the boundary.

We follow [Achdou and Pironneau, 2005] here and consider in the following a model somewhat between (6.1) and (6.8) in generality, the so-called *local volatility* model

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{6.58}$$

We assume that σ stays within a certain range,

$$0 < \underline{\sigma} \leq \sigma(S,t) \leq \overline{\sigma} < \infty \tag{6.59}$$

for all S, t, so the model is in some sense close to Black-Scholes. This is especially crucial for large and small S, where (6.59) ensures that the diffusion coefficient σS is Lipschitz continuous.

We focus on the fully implicit Euler scheme of the form

$$K_1 V^{m-1} = (I + \Delta t A)V^{m-1} = V^m \tag{6.60}$$

with suitably defined matrices K_1 and A. The key to the analysis is an inequality of the form

$$V'AV \geq -C|V|^2 \quad \forall V \in \mathbb{R}^N, \tag{6.61}$$

where C is independent of ΔS and Δt. In [Achdou and Pironneau, 2005], the stronger inequality

$$V'AV \geq C_1|V|_w^2 - C_2|V|^2 \quad \forall V \in \mathbb{R}^N, \tag{6.62}$$

is derived, with a "weighted" norm

$$|V|_w^2 = \sum_{n=1}^{N}(S_n/\Delta S)^2(V_n - V_{n-1})^2 + (S_N/\Delta S)^2 V_N^2.$$

We refer to [Achdou and Pironneau, 2005] for the proof of this, and can then deduce from (6.61) for the fully implicit scheme (6.60) that

$$\begin{aligned} |V^m||V^{m-1}| &\geq V^{m-1\prime}V^m = V^{m-1\prime}(I + \Delta t A)V^{m-1} \geq |V^{m-1}|^2 - \Delta t C|V^{m-1}|^2 & (6.63)\\ &= (1 - \Delta t C)|V^{m-1}|^2. & (6.64)\end{aligned}$$

Assume now that $\Delta t < 1/C$, which is not very restrictive because C is assumed independent of ΔS. Dividing both sides by $|V^{m-1}|$, by induction,

$$|V^m| \leq (1 - C\Delta t)^{-(M-m)}|V^M| \leq \mathrm{e}^{CT}|V^M|.$$

Properties of the type (6.62) are called Gårding inequalities. They are weaker than *coercivity*,

$$V'AV \geq C|V|^2 \quad \forall V \in \mathbb{R}^N,$$

with $C > 0$, which holds e.g. for the heat equation, but not generally for problems with drift or variable coefficients. Equation (6.62) expresses that A is "not too negative". Loosely speaking, in (6.61), the negative eigenvalues are bounded below independent of N.

Remark 6.2.8. *The bad news is that the technique in [Achdou and Pironneau, 2005] to prove (6.62) appears limited to problems of the type (6.58), with well-behaved volatility as in (6.59). This excludes other relevant examples like the CIR process, where the diffusion coefficient is not linear at 0, but e.g. a square-root. A convergence analysis of difference schemes for some of these more general models is found in [Sun et al., 2003].*