TMA4212 - Project 1

Bjørnar Kaarevik Jo Andersson Stokke Kristian Holme

February 2022

1 Introduction

In this project, we will take a closer look on the one-dimensional linearized Black-Scholes equation

$$u_t - \frac{1}{2}\sigma^2 x^2 u_{xx} - rxu_x + cu = 0, x \in \mathbb{R}^+, t \in (0,T) \tag{1}$$

as well as the non-linear Black-Scholes equation

$$u_t - \frac{1}{2}x^2 \varphi\left(u_{xx}\right) u_{xx} = 0 \tag{2}$$

where the non-linear function φ is defined as

$$\varphi(x) = \sigma_1^2 + \frac{\sigma_2^2 - \sigma_1^2}{2}\left(1 + \frac{2}{\pi}\arctan x\right) \tag{3}$$

The linear case offers a linear operator $\mathcal{L}$ defined as

$$-\mathcal{L} = \partial_t - \frac{1}{2}\sigma^2 x^2 \partial_x^2 - rx\partial_x + c \tag{4}$$

We discretize and solve the linear problem using forward Euler (FE), backward Euler (BE) and Crank-Nicholson (CN). The discretized non-linear problem is solved with IMEX and BE.

The operators are discretized with central differences in space, and forward differences, backward differences and trapezoidal integration in time for the forward Euler, backward Euler and Crank-Nicholson schemes respectively. In the nonlinear case, the IMEX method is derived by making the nonlinear part of the BE method explicit (so it can be solved as a linear sytem).

When solving the discretized problems numerically, we solve on a grid of points $\{(x_m, t_n)\}_{m,n}$ where $x_m = m \cdot h, m = 0, 1, 2, \ldots, M+1$ and $t_n = n \cdot k, n = 0, 1, 2, \ldots, N$. The boundaries in the x- and t directions are denoted $x_{M+1} = (M+1)h = R$ and $t_N = N \cdot k = T$.

The equations require an initial condition (IC), we use the three following IC's:

$$\begin{aligned}
&u(x,0) = \max(K-x,0) =: (K-x)^+ && \text{(European put), (IC1)}\\
&u(x,0) = (x-K)^+ - 2(x-(K+H))^+ + (x-(K+2H))^+, && \text{(Butterfly spread), (IC2)}\\
&u(x,0) = \operatorname{sgn}^+(x-K), && \text{(Binary call), (IC3)}
\end{aligned}$$

2 Monotonicity and CFL conditions

Monotonicity and CFL conditions for the schemes hinge on derivations made from the coefficients of an expression on the form

$$\alpha_0 U_m^{n+1} = \sum_{i=1}^{s} \alpha_i U_{m_i}^n - \sum_{i=s+1}^{s+t} \alpha_i U_{m_t}^{n+1} \tag{5}$$

This expression defines a two-level finite difference scheme with s explicit contributions from the current timestep and t implicit contributions from the next timestep.

To make sure the scheme is monotonic the following conditions on the coefficients must be satisfied

$$\alpha_0 > 0, \quad \alpha_i \geq 0 \; \forall i \in \{1..s+t\}, \quad \alpha_0 \geq \sum_{i=1}^{s+t} \alpha_i$$

2.1 Forward Euler for the linear equations

The forward Euler scheme is characterized as

$$U_m^{n+1} = \frac{k}{2}\left(\sigma^2 m^2 + rm\right) U_{m+1}^n + \left(1 - k\sigma^2 m^2 - kc\right) U_m^n + \frac{k}{2}\left(\sigma^2 m^2 - rm\right) U_{m-1}^n$$

This scheme has the coefficients

$$\alpha_0 = 1, \quad \alpha_1 = \frac{k}{2}\left(\sigma^2 m^2 + rm\right), \quad \alpha_2 = \left(1 - k\sigma^2 m^2 - kc\right), \quad \alpha_3 = \frac{k}{2}\left(\sigma^2 m^2 - rm\right)$$

And the resulting conditions are

$$\sigma^2 \geq r \text{ and } \frac{1}{k} > \sigma^2\left(\frac{R}{h} - 1\right)^2 + c.$$

2.2 Backward Euler for the linear equations

The backward Euler scheme is characterized as

$$U_m^n = -\frac{k}{2}\left(\sigma^2 m^2 + rm\right) U_{m+1}^{n+1} + \left(1 + k\sigma^2 m^2 + kc\right) U_m^{n+1} - \frac{k}{2}\left(\sigma^2 m^2 - rm\right) U_{m-1}^{n+1}$$

The coefficients of the scheme are

$$\alpha_0 = \left(1 + k\sigma^2 m^2 + kc\right), \quad \alpha_1 = \frac{k}{2}\left(\sigma^2 m^2 + rm\right), \quad \alpha_2 = 1, \quad \alpha_3 = \frac{k}{2}\left(\sigma^2 m^2 - rm\right)$$

The backward Euler scheme is always diagonally dominant as $1 + kc \geq 1$, but from the coefficients above is still follows that $\sigma^2 \geq r$, otherwise α_3 would not be non-negative. There is no CFL-condition on the backward Euler scheme.

2.3 Crank-Nicholson for the linear equations

The Crank-Nicholson scheme is more involved and can be written as

$$\begin{aligned}
&-\frac{k}{4}\left(\sigma^2 m^2 + rm\right) U_{m+1}^{n+1} + \left(1 + \frac{k\sigma^2 m^2 + kc}{2}\right) U_m^{n+1} - \frac{k}{4}\left(\sigma^2 m^2 - rm\right) U_{m-1}^{n+1} \\
&= \frac{k}{4}\left(\sigma^2 m^2 + rm\right) U_{m+1}^n + \left(1 - \frac{k\sigma^2 m^2 + kc}{2}\right) U_m^n + \frac{k}{4}\left(\sigma^2 m^2 - rm\right) U_{m-1}^n
\end{aligned}$$

This yields a set of coefficients

$$\begin{aligned}
\alpha_0 &= \left(1 + \frac{k\sigma^2 m^2 + kc}{2}\right), \quad \alpha_1 = \alpha_4 = \frac{k}{4}\left(\sigma^2 m^2 + rm\right) \\
\alpha_2 &= \left(1 - \frac{k\sigma^2 m^2 + kc}{2}\right), \quad \alpha_3 = \alpha_5 = \frac{k}{4}\left(\sigma^2 m^2 - rm\right)
\end{aligned}$$

These coefficients result in the following constraints on the system

$$\sigma^2 \geq r \text{ and } \frac{2}{k} \geq \sigma^2\left(\frac{R}{h} - 1\right)^2 + c.$$

2.4 Monotonicity for IMEX and BE in the non-linear case

From appendix D we have that the scheme

$$\alpha_0 U_m^{n+1} \leq \beta_0 U_m^n + \beta_{-1} U_{m-1}^{n+1} + \beta_1 U_{m+1}^{n+1} \tag{6}$$

is monotone if $\alpha = \frac{1}{2} x_m^2 \max \varphi \frac{k}{h^2}$ and the following conditions hold

$$\alpha_0 = 1 + 2\alpha, \quad \beta_0 = 1, \quad \beta_1 = \beta_{-1} = \alpha$$

Due to the definition of α, this accounts for both IMEX and BE, thus both schemes are monotonic.

3 Consistency

3.1 Truncation error for backward Euler in the linear case

The truncation error is defined as the difference between the numerical scheme and the exact solution in one step. Hence we have

$$\tau := -\mathcal{L}_h u + \mathcal{L} u$$

.

For the backward Euler case, the discretized operator is

$$-\mathcal{L}_h = \nabla_{k,t} - \frac{1}{2}\sigma^2 x^2 \delta^2_{h,x} - rx\delta_{2h,x} + c$$

With the operators defined above and in equation 4, we can group terms by relevance and obtain

$$\tau = \left(\nabla_{k,t} - \partial t\right) u - \frac{1}{2}\sigma^2 x^2 \left(\delta^2_{h,x} - \partial^2_x\right) u - rx\left(\delta_{2h,x} - \partial_x\right) u + c\left(1-1\right) u \tag{7}$$

We may investigate each group of terms and decide their order of convergence without needing to take into account the coefficients and parameters before, seeing as it is only the result of the difference method and the exact differential operator which needs to be bounded for local truncation error. The terms in 7 presents the following differences

$$\begin{aligned}
\left(\nabla_{k,t} - \partial_t\right) u &= \frac{u(x,t+k) - u(x,t)}{k} - u_t \\
\left(\delta^2_{h,x} - \partial^2_x\right) u &= \frac{u(x-h,t) - 2u(x,t) + u(x+h,t)}{h^2} - u_{xx} \\
\left(\delta_{2h,x} - \partial_x\right) u &= \frac{u(x-h,t) - u(x+h,t)}{2h} - u_x
\end{aligned}$$

The remainder in these differences are

$$\begin{aligned}
\left(\nabla_{k,t} - \partial_t\right) u &= \frac{k}{2} u_{tt} + O(k^2) \\
\left(\delta^2_{h,x} - \partial^2_x\right) u &= \frac{h^2}{12}\partial^4_x u + O(h^3) \\
\left(\delta_{2h,x} - \partial_x\right) u &= \frac{h^2}{6}\partial^3_x u + O(h^3)
\end{aligned}$$

Thus we obtain an expression for the local truncation error

$$\begin{aligned}
|\tau^n_m| &\leq \frac{k}{2} u_{tt}(x_m,t_n) + \frac{1}{2}\sigma^2 x^2 \frac{h^2}{12}\partial^4_x u(x_m,t_n) + rx\frac{h^2}{6}\partial^3_x u(x_m,t_n) \\
&\leq k\left(\frac{1}{2}\max \partial^2_t u\right) + h^2\left(\frac{rR}{6}\max \partial^3_x u + \frac{\sigma^2 R^2}{24}\max \partial^4_x u\right) \\
&= O(k + h^2)
\end{aligned} \tag{8}$$

Since we assume the partial derivatives of u to be bounded on the domain, the local truncation error is dependent on k and h in such a way that τ tends to zero as the discretization steps tend to zero. Thus we've obtained consistency for the numerical scheme.

3.2 Truncation error in the cases of forward Euler and Crank-Nicholson

The truncation error for the forward Euler and Crank-Nicholson schemes follow much of the same derivation as the backward Euler case. The spatial discretizations are identical leading to the same contribution to τ from the partial derivatives of u wrt. x. Since the temporal discretization differs between the methods, this is where difference in convergence order will arise from.

A forward difference in time would result in the method being $O(k)$ and a trapezoidal integration (which is the case for Crank-Nicholson) would result in the method being $O(k^2)$, which means that the other methods are both consistent.

3.3 Truncation error in the non-linear case

From calculations in appendix D, we get that the truncation error in the IMEX and BE schemes for the non-linear Black-Scholes equation is

$$\|\tau^n\|_\infty \leq \frac{k}{2}\|u_{tt}^n\|_\infty + \frac{R^2}{2}\frac{h^2}{24}\|u_{x^4}^n\|_\infty\left(\frac{\sigma_2^2-\sigma_1^2}{\pi}\|u_{xx}^n\|_\infty + \sigma_2^2\right)$$

This expression is of order $O(k+h^2)$, and gives that the schemes are consistent.

4 Stability

4.1 Stability of the backward Euler scheme

The Black-Scholes equations as we have them (both continuous and discretized) can then be written as

$$\begin{aligned} -\mathcal{L}u = 0, u \in \Omega \quad u = g, u \in \partial\Omega \\ -\mathcal{L}_h U_p = 0, p \in \mathbb{G} \quad U_p = g_p, p \in \partial\mathbb{G} \end{aligned}$$

In order to test for stability with respect to the righthand side, we force the system to equal an arbitrary function on the righthand side. If the solution can be bounded by a constant with respect to the choice of the function, we've shown stability wrt. the righthand side. The constructed system will be as follows

$$\begin{aligned} -\mathcal{L}_h V_p = f_p, p \in \mathbb{G} \\ V_p = 0, p \in \partial\mathbb{G} \end{aligned}$$

If the scheme was explicit, as is the case for forward Euler, it would be enough to find a bound for for the subordinate matrix norm which can be done by spectral analysis on a tridiagonal matrix. Since backward Euler and Crank-Nicholson require matrix inversion for such spectral analysis, it seems easier to prove stability using the discrete max principle, seeing as both schemes are monotone. There is no need to consider any other particulars pertaining to the schemes, as monotonicity was shown for the schemes using Dirichlet boundary conditions.

If it is possible to construct a solution W_p to the discrete linear operator for the PDE with the property that $-\mathcal{L}_h W_p \leq 0$, we can apply the discrete max principle. Thus, we construct the solution W_p utilizing a supporting function φ, (not to be confused with the φ in the non-linear BS-equation).

$$\begin{aligned} W_p = V_p - \|\vec{f}\|_\infty \varphi_p \\ \varphi \geq 0 \text{ and } -\mathcal{L}_h\varphi_p = 1, p \in \mathbb{G}. \end{aligned}$$

.

From this construction it follows that

$$\begin{aligned} -\mathcal{L}_h W_p = -\mathcal{L}_h V_p - \|\vec{f}\|_\infty \cdot (-\mathcal{L}_h\varphi_p) \\ -\mathcal{L}_h W_p = f_p - \max_{p\in\mathbb{G}} |f_p| \cdot 1 \leq 0. \end{aligned}$$

Since this construction of W_p has the desired property described above, we may apply the discrete max principle. It is also worth noting that W_p is non-positive on the boundary of the grid, seeing as $V_p = 0$, $|f_p| \geq 0$ and $-\varphi_p \leq 0$. This gives that

$$W_p \leq \max_{p\in\partial\mathbb{G}} W_p \leq 0.$$

By definition of W_p we now know that

$$V_p - \|\vec{f}\|_\infty\varphi_p \leq 0 \iff V_p \leq \|\vec{f}\|_\infty\varphi_p.$$

Similarly, this bound is acheived for $-V_p$ and $-f_p$, as the vector norm is non-negative. At any rate, taking the maxima over all possible gridpoints p we obtain that

$$\max_{p\in\overline{\mathbb{G}}} |V_p| \leq \|\vec{f}\|_\infty \max_{p\in\overline{\mathbb{G}}} \varphi_p. \tag{9}$$

As long as $\max\varphi_p < \infty$, we can conclude that the scheme is stable with respect to the righthand side.

4.2 Stability of the IMEX and BE schemes for the non-linear case

Similarly, we want to choose a support function φ like before such that

$$W_p = V_p - \|\vec{f}\|_\infty \varphi_p$$
$$\varphi \geq 0 \text{ and } -\mathcal{L}_h \varphi_p = 1, p \in \mathbb{G}.$$

This time, we choose $\varphi = t$, which satisfies this criteria for $T > t > 0$ as

$$-\mathcal{L}_h \varphi_p = \partial_t \varphi_p = \partial_t t = 1$$

As the IMEX scheme and BE scheme has both been shown to be monotonic, this is enough to conclude by the same argument as for the linear case that

$$\max_{p \in \overline{\mathbb{G}}} |V_p| \leq \|\vec{f}\|_\infty \max_{p \in \overline{\mathbb{G}}} \varphi_p = \|\vec{f}\|_\infty T \tag{10}$$

As this is exactly the criteria of stability, it is thus proved that the given IMEX and BE schemes for the non-linear Black-Scholes equations are both stable with respect to a given right hand side which is bounded. Note: Together with the truncation error previously found, the two schemes are therefore also convergent in L^∞. (Given the zero-order polynomial interpolation of the points of the numerical solution on the grid.)

5 Error bounds

5.1 Error bound on the backward Euler method for the linear Black-Scholes equations

A shown in appendix B, the properties

$$\varphi \geq 0 \; \forall p \in \overline{\mathbb{G}} \quad -\mathcal{L}_h \varphi_p = 1 \; \forall p \in \overline{\mathbb{G}}$$

Together with constructing a $\varphi(x, t) = \varphi(t)$, result in a linear first order differential equation in time, to which the solution is

$$\varphi(t) = \frac{1}{c} + \kappa e^{-ct}$$

Here, c is the parameter from the description of the Black-Scholes system, while the κ is an undetermined coefficient. Determining the coefficient and finding the candidates for the maximum on all positive time gives

$$\max_{0 \leq t < \infty} \varphi(t) = \max \left\{ 0, \frac{1}{c} \right\} = \frac{1}{c} \tag{11}$$

5.2 Error bound

By utilizing stability and consistence together with the supporting function above, we have (by calculations in appendix C) the error bound for the backward Euler scheme as

$$\max_{p \in \overline{\mathbb{G}}} e_p \leq \frac{1}{c} \left(k \left(\max \frac{1}{2} \partial_t^2 u \right) + h^2 \left(\max \frac{1}{6} \partial_x^3 u + \max \frac{1}{12} \partial_x^4 u \right) \right)$$

6 Convergence

Since the Black-Scholes equation is a well-posed linear initial value problem, the Lax equivalence theorem states that consistency and stability is equivalent to convergence. From equations 8 and 10 we know that the backward Euler method applied to the linearized partial differential equation is consistent and convergent, thus the method is convergent. (the proof is the same as the proof for stability, with e instead of v, and τ instead of f).

7 Implementation and testing

When solving the discretized problems, we need Boundary conditions. At $x = 0$, the equations becomes ODE's that can easily be solved and forced for the numerical solution (linear case: $u(0,t) = u(0,0)e^{-ct}$, non-linear case: $u(0,t) = u(0,0)$). At $x = R$ we have used $u(R,t) = u(R,0)$ (called BC1) and $u_x(R,t) = u_x(R,0)$, (called BC2)

For the linear problem we have FE, BE and CN with all three IC's and the two BC's. In the non-linear case, we have implemented both IMEX and BE, with all three IC's and BC1.

The implementations are based on the matrix notation defined in appendix A. At the boundaries it is necessary to add a vector to the right hand side of the matrix equations, to make sure the boundary conditions are satisfied, completely analogous to what has been done in the lectures. For the second boundary condition we've introduced fictitious nodes to approximate the desired derivative with central differences and extend the matrix scheme to include the boundary points at $x = R$, as we have seen for Neumann boundary conditions in the lectures.

7.1 Solving the linear problem

See figure 1 for examples of solutions for the discretized linear problem.

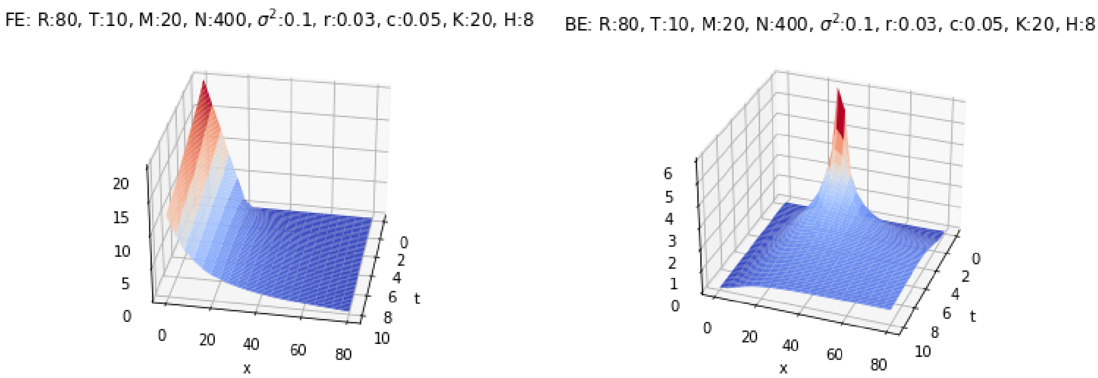


Figure 1: Solution to the linear problem. Left:FE, European IC, and boundary condition one. Right: BE, Butterfly IC, BC2.

7.2 Error and speed for the linear problem

7.2.1 CFL-conditions

To test error in the schemes we introduced a right hand side to the Black-Scholes equations, so we can design our own exact solution. We have chosen the test function $u(x,t) = 5\sin\left(\pi\frac{x-20}{20}\right)\cos\left(\pi\frac{t}{20}\right)$. This function was chosen because it works well when solved with Dirichlet boundary conditions, because when $T = 10, R = 80$, the function value is zero at all boundaries of $[0,R] \times [0,T]$. Figure 2 shows the solution to the test problem with the FE-method.

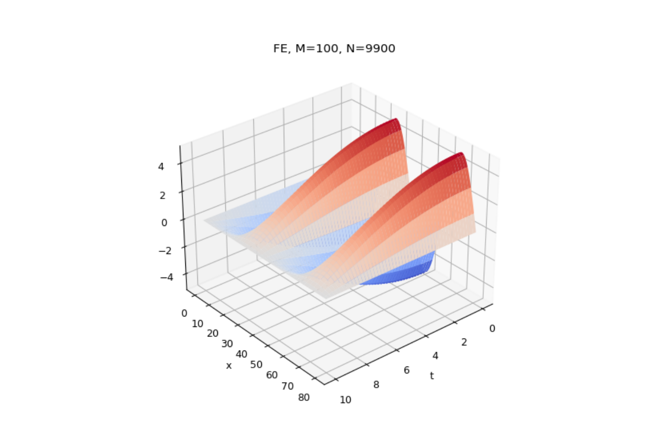


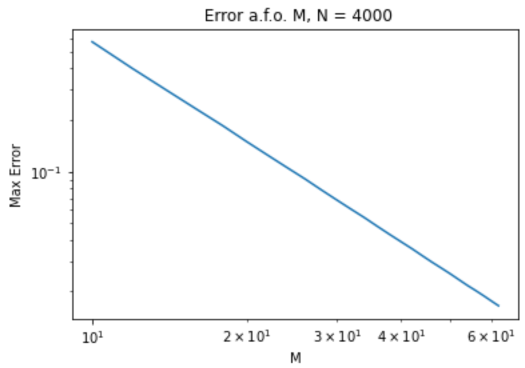
Figure 2: Solution with FE of the test problem.

The CFL-condition for the forward Euler method can be reformulated as $N/T > \sigma^2 M^2 + c$.

When $M = 400$, the FE-method resulted in reasonable error when N was 97% of the CFL-minimum, but the method broke down when N was lowered to 96% of the CFL-minimum. As M was increased further, the lower limit for stable N's approached the CFL-minimum, indicating that the CFL-condition from section 2.1 is correct. Doing the same experimentation with the Crank-Nicholson-method did not give similar results. The CN-method did not seem to care about the CFL-condition ($N/T > \frac{\sigma^2 M^2 + c}{2}$). Even when N was 0.05% of the CFL-minimum, the error was small.

7.2.2 Convergence/error for the linear problem

To check the convergence rates of the different methods, we kept either M (or N) constant, and recorded the error as N (or M) varied. Figure 3 shows a log-log plot of the max error with the FE-method when N is constant and M varies. We get a similar plot when M is held constant. The log-log plots of errors for the other methods are shown in the attached jupyter file. The slope on the left in 3 is approximately -2, indicating a quadratic error dependence on h. The slope of the error when N was varied was approximately 1, indicating a linear error dependence on k. The same results was achieved for backwards Euler. With the Crank-Nicholson method the error dependence was found to be quadratic in both h and k, as expected.



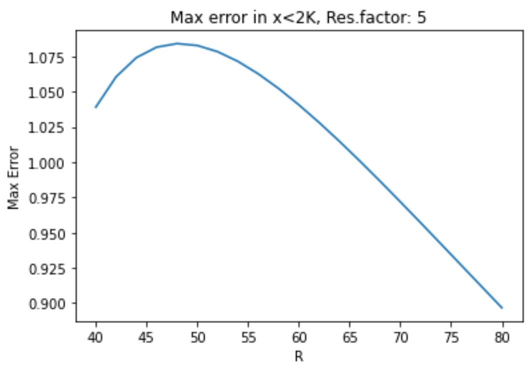


Figure 3: Left: FE Max error a.f.o. M as N is constant. Right: Max error with FE, Butterfly IC, as R increases

7.2.3 Computational speed

To establish which method is the fastest, we experimented with different values of M and N when solving the test problem. The error requirement was set to 0.008. The results are displayed in table 1

Method	M	N	Comp. time (s)
FE	90	8100	0.0457
BE	90	220	0.0255
CN	90	30	0.003

Table 1: Results of experimenting with M and N to achieve the desired error.

From this table we see that FE is slowest. BE is almost twice as fast as FE, and CN is about ten times faster than FE.

7.3 Error in the area of interest

Now we test the error with FE in the region $x < 2K$, as R is varied. We chose $K = 20$ and varied R from $2K = 40$ to $4K = 80$. When solving with FE, M is adjusted before each run to keep h constant. By experimentation with the exact solution, we found that the error of CN with M 100 times higher than for FE, the error for CN was 0.01% the error of FE. So here we treat a high resolution solution with CN as the exact solution to approximate the real error.

For the approximated exact solution with CN, we chose this resolution factor to be 5.

Generally we observe that the error in the region of interest decreases as R increases. The only slight anomaly is with Butterfly initial condition. Then we observe that as R increases from 40 to about 50, the error increases slightly, but then decreases as R is further increased, as we see in figure 3

Error plots for the different initial conditions with boundary condition one and two are included in the attached jupyter file.

7.4 IMEX for the non-linear problem

In figure 4 we see the solution to the non-linear BS-equation with Dirichlet type boundary condition and Butterfly type initial condition.

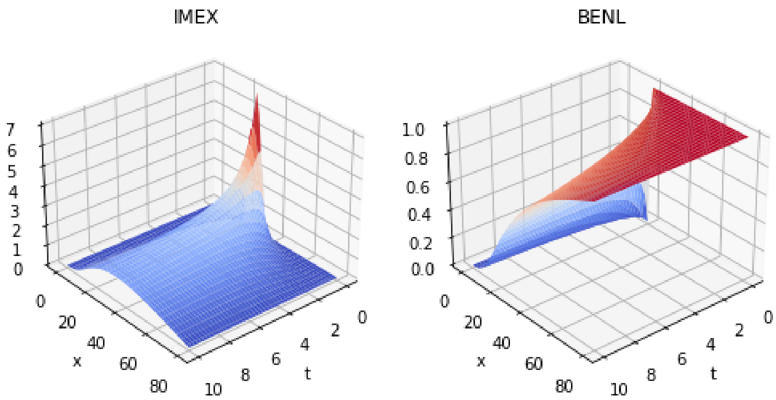


Figure 4: Solution of the non-linear BS-equation with BC 1. Left: IMEX and Butterfly IC. Right: BENL and Binary IC.

We have done the same convergence analysis for IMEX as we did for the methods in the linear case. Figure 4 shows the error decreasing as N increases. As with FE and BE in the linear case, the error was found be of order 2 in h and of order 1 in k.

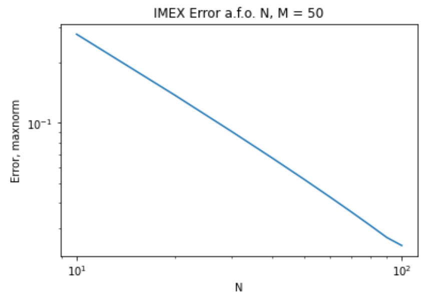


Figure 5: Log-log plot of max error with IMEX as N increases. Slope is approx. -1.

7.5 Non-linear Backwards Euler and speed comparison

The Non-linear Backwards Euler method (BENL) was implemented using SciPy's library for solving non-linear equations.

A solution with the Binary IC is shown in figure 4.

The convergence was found to be of same order as IMEX; linear in k and quadratic in h.

We did the same speed comparison between BENL and IMEX as we did for the linear solvers. For a required minimum error of 0.05, IMEX (0.015 s)was approximately twenty times faster than BENL (0.3 s). The details can be found in the jupyter file.

Statement on work division

Bjørnar and Jo was responsible for the theoretical derivations and proofs, while Kristian was responsible for the programming.

A Derivations of monotonicity and CFL conditions

1-D linear Black-Scholes Equation

$$u_t - \frac{1}{2}\sigma^2 x^2 u_{xx} - rxu_x + cu = 0$$

A.1 Forward Euler difference scheme

$$\begin{aligned}
\Delta_{t,k}U_m^n - \frac{1}{2}\sigma^2 x\delta_{x,h}^2 U_m^n - rx\delta_{x,2h}U_m^n + cU_m^n = 0\\
\frac{U_m^{n+1} - U_m^n}{k} - \frac{1}{2}\sigma^2 h^2 m^2 \frac{U_{m+1}^n - 2U_m^n + U_{m-1}^n}{h^2} - rmh\frac{U_{m+1}^n - U_{m-1}^n}{2h} + cU_m^n = 0\\
U_m^{n+1} = U_m^n + k\frac{1}{2}\sigma^2 m^2 (U_{m+1}^n - 2U_m^n + U_{m-1}^n) + k\frac{1}{2}rm(U_{m+1}^n - U_{m-1}^n) - kcU_m^n\\
U_m^{n+1} = \frac{k}{2}\left(\sigma^2 m^2 + rm\right)U_{m+1}^n + \left(1 - k\sigma^2 m^2 - kc\right)U_m^n + \frac{k}{2}\left(\sigma^2 m^2 - rm\right)U_{m-1}^n
\end{aligned}$$

This gives a linear system of equations ordered in a tridiagonal matrix as such

$$\vec{U^{n+1}} = A_{k,h}\vec{U^n}$$
$$A_{k,h} = \text{tridiag}\left\{\frac{k}{2}\left(\sigma^2 m^2 + rm\right), \left(1 - k\sigma^2 m^2 - kc\right), \frac{k}{2}\left(\sigma^2 m^2 - rm\right)\right\}$$

Notice that m is the row number in the matrix, so the diagonal elements are not constant.
This system results in the coefficients used to determine monotonicity and CFL conditions as

$$\begin{aligned}
\alpha_0 &= 1\\
\alpha_1 &= \frac{k}{2}\left(\sigma^2 m^2 + rm\right)\\
\alpha_2 &= \left(1 - k\sigma^2 m^2 - kc\right)\\
\alpha_3 &= \frac{k}{2}\left(\sigma^2 m^2 - rm\right)\\
\alpha_0 > 0, &\alpha_i \geq 0 \ \forall i \in \{1,2,3\}\\
\alpha_0 &\geq \alpha_1 + \alpha_2 + \alpha_3
\end{aligned}$$

Now for determining the conditions. The inequality $\alpha_0 > 0$ is trivially fulfilled as $\alpha_0 = 1$. Since $\sigma, r, m, k \geq 0$, it is also trivial that $\alpha_1 \geq 0$. The interesting conditions are those found by α_2 and α_3:

$$\begin{gathered}
\alpha_2 \geq 0 \Rightarrow 1 - k\sigma^2 m^2 - kc \geq 0\\
\frac{1}{k} \geq \sigma^2 m^2 + c\\
m = M, M+1 = \frac{R}{h} \Rightarrow \frac{1}{k} > \sigma^2(\frac{R}{h} - 1)^2 + c\\
\alpha_3 \geq 0 \Rightarrow \frac{k}{2}\left(\sigma^2 m^2 - rm\right) \geq 0\\
\sigma^2 m^2 \geq rm\\
m = 0 \Rightarrow 0 \geq 0\\
m \neq 0 \Rightarrow \sigma^2 m \geq r\\
m = 1 \Rightarrow \sigma^2 \geq r
\end{gathered}$$

Thus we have a CFL condition being $\frac{1}{k} > \sigma^2(\frac{R}{h} - 1)^2 + c$. In addition, there is a condition for monotonicity being $\sigma^2 \geq r$. Diagonal dominance with the sum gives

$$
\begin{aligned}
\alpha_0 &\geq \alpha_1 + \alpha_2 + \alpha_3 \\
1 &\geq \frac{k}{2}\left(\sigma^2 m^2 + rm\right) + 1 - k\sigma^2 m^2 - kc + \frac{k}{2}\left(\sigma^2 m^2 - rm\right) \\
0 &\geq \frac{k}{2}\left(\sigma^2 m^2\right) - k\sigma^2 m^2 - kc + \frac{k}{2}\left(\sigma^2 m^2\right) \\
0 &\geq -kc \\
c &\geq 0
\end{aligned}
$$

The diagonal dominance condition for monotonicity is fulfilled by c being defined by $c > 0$.
Thus the conditions for the forward euler scheme is

$$
\sigma^2 \geq r \text{ and } \frac{1}{k} > \sigma^2(\frac{R}{h} - 1)^2 + c.
$$

A.2 Backward Euler difference scheme

$$
\begin{aligned}
\nabla_{t,k} U_m^n - \frac{1}{2}\sigma^2 x \delta_{x,h}^2 U_m^{n+1} - rx\delta_{x,2h} U_m^{n+1} + cU_m^{n+1} &= 0 \\
\frac{U_m^{n+1} - U_m^n}{k} - \frac{1}{2}\sigma^2 h^2 m^2 \frac{U_{m+1}^{n+1} - 2U_m^n n + 1 + U_{m-1}^{n+1}}{h^2} - rmh\frac{U_{m+1}^{n+1} - U_{m-1}^{n+1}}{2h} + cU_m^{n+1} &= 0 \\
U_m^n = U_m^{n+1} - k\frac{1}{2}\sigma^2 m^2 (U_{m+1}^{n+1} - 2U_m^{n+1} + U_{m-1}^{n+1}) - k\frac{1}{2} rm(U_{m+1}^{n+1} - U_{m-1}^{n+1}) + kcU_m^{n+1} \\
U_m^n = -\frac{k}{2}\left(\sigma^2 m^2 + rm\right) U_{m+1}^{n+1} + \left(1 + k\sigma^2 m^2 + kc\right) U_m^{n+1} - \frac{k}{2}\left(\sigma^2 m^2 - rm\right) U_{m-1}^{n+1}
\end{aligned}
$$

This gives a system somewhat similar to the forward Euler, expcet for the fact that the indices for time are switched. Naturally, this leads to an implicit scheme requiring inversion of the matrix $A_{k,h}$, whereas the forward scheme required only matrix-vector products due to being explicit.

At any rate, the scheme can be summarized as:

$$
\begin{aligned}
B_{k,h} \vec{U^{n+1}} &= \vec{U^n} \\
B_{k,h} &= \text{tridiag}\left\{-\frac{k}{2}\left(\sigma^2 m^2 + rm\right), \left(1 + k\sigma^2 m^2 + kc\right), -\frac{k}{2}\left(\sigma^2 m^2 - rm\right)\right\} \\
\vec{U^{n+1}} &= B_{k,h}^{-1} \vec{U^n}
\end{aligned}
$$

A.3 Crank-Nicholson difference scheme

The scheme is nearly the average of the FE and BE shcemes, and hence, the equations should turn out as

$$
\begin{aligned}
&-\frac{k}{4}\left(\sigma^2 m^2 + rm\right) U_{m+1}^{n+1} + \left(1 + \frac{k\sigma^2 m^2 + kc}{2}\right) U_m^{n+1} - \frac{k}{4}\left(\sigma^2 m^2 - rm\right) U_{m-1}^{n+1} \\
&= \frac{k}{4}\left(\sigma^2 m^2 + rm\right) U_{m+1}^n + \left(1 - \frac{k\sigma^2 m^2 + kc}{2}\right) U_m^n + \frac{k}{4}\left(\sigma^2 m^2 - rm\right) U_{m-1}^n
\end{aligned}
$$

This gives the scheme in matrix for as

$$
\begin{aligned}
B\vec{U^{n+1}} &= A\vec{U^n} \\
B &= \frac{1}{2}\left(I + B_{k,h}\right) \\
A &= \frac{1}{2}\left(I + A_{k,h}\right) \\
\vec{U^{n+1}} &= B^{-1} A \vec{U^n}
\end{aligned}
$$

Luckily, we can ask the computer to just do it.
Using the equation of the scheme, at some point x_m, we have

$$
\begin{aligned}
&-\frac{k}{4}\left(\sigma^2 m^2 + rm\right) U_{m+1}^{n+1} + \left(1 + \frac{k\sigma^2 m^2 + kc}{2}\right) U_m^{n+1} - \frac{k}{4}\left(\sigma^2 m^2 - rm\right) U_{m-1}^{n+1} \\
&= \frac{k}{4}\left(\sigma^2 m^2 + rm\right) U_{m+1}^{n} + \left(1 - \frac{k\sigma^2 m^2 + kc}{2}\right) U_m^{n} + \frac{k}{4}\left(\sigma^2 m^2 - rm\right) U_{m-1}^{n}
\end{aligned}
$$

We label the coefficients as such

$$
\begin{aligned}
\alpha_0 &= \left(1 + \frac{k\sigma^2 m^2 + kc}{2}\right) \\
\alpha_1 &= \frac{k}{4}\left(\sigma^2 m^2 + rm\right) \\
\alpha_2 &= \left(1 - \frac{k\sigma^2 m^2 + kc}{2}\right) \\
\alpha_3 &= \frac{k}{4}\left(\sigma^2 m^2 - rm\right) \\
\alpha_4 &= -\frac{k}{4}\left(\sigma^2 m^2 + rm\right) = -\alpha_1 \\
\alpha_5 &= -\frac{k}{4}\left(\sigma^2 m^2 - rm\right) = -\alpha_3
\end{aligned}
$$

Yielding the system

$$
\begin{aligned}
\alpha_4 U_{m+1}^{n+1} + \alpha_0 U_m^{n+1} + \alpha_5 U_{m-1}^{n+1} &= \alpha_1 U_{m+1}^{n} + \alpha_2 U_m^{n} + \alpha_3 U_{m-1}^{n} \\
\alpha_0 U_m^{n+1} &= \alpha_1 U_{m+1}^{n} + \alpha_2 U_m^{n} + \alpha_3 U_{m-1}^{n} - \alpha_4 U_{m+1}^{n+1} - \alpha_5 U_{m-1}^{n+1} \\
\alpha_0 U_m^{n+1} &= \alpha_1\left(U_{m+1}^{n} + U_{m+1}^{n+1}\right) + \alpha_2 U_m^{n} + \alpha_3\left(U_{m-1}^{n} + U_{m-1}^{n+1}\right)
\end{aligned}
$$

And the conditions for monotonicity

$$
\alpha_0 > 0 \tag{12}
$$

$$
\alpha_i \geq 0 \forall i \in \{1, 2, 3\} \tag{13}
$$

$$
\alpha_0 \geq 2\alpha_1 + \alpha_2 + 2\alpha_3 \tag{14}
$$

These result in the following

$$
\begin{aligned}
\alpha_0 &> 0 \\
1 + \frac{k\sigma^2 m^2 + kc}{2} &> 0 \\
\frac{-2}{k} &< \sigma^2 m^2 + c \\
\sigma, c &> 0 \\
\therefore \frac{-2}{k} &< 0 < \sigma^2 m^2 + c \;\forall\; m \geq 0
\end{aligned}
$$

The first condition is always valid. To the conditions in the off-diagonals:

$$\alpha_1 \geq 0 \Rightarrow \frac{k}{4}\left(\sigma^2 m^2 + rm\right) \geq 0$$
$$\sigma^2 m^2 + rm \geq 0 \Rightarrow \text{ always valid}$$
$$\alpha_3 \geq 0 \Rightarrow \frac{k}{4}\left(\sigma^2 m^2 - rm\right) \geq 0$$
$$\sigma^2 m^2 \geq rm$$
$$m = 0 \Rightarrow 0 \geq 0 \Rightarrow \text{ always valid}$$
$$m \neq 0 \Rightarrow \sigma^2 m \geq r \ \forall m \geq 1$$
$$m_1 > m_2 \Rightarrow \sigma^2 m_1 > \sigma^2 m_2 \geq r$$
$$\therefore m = 1 \Rightarrow \sigma^2 \geq r$$

The condition for the coefficient of the current timestep on the diagonal is

$$\alpha_2 \geq 0 \Rightarrow 1 - \frac{k\sigma^2 m^2 + kc}{2} \geq 0$$
$$1 \geq \frac{k}{2}(\sigma^2 m^2 + c)$$
$$\frac{2}{k} \geq \sigma^2 m^2 + c$$
$$\frac{2}{k} \geq \sigma^2 (\frac{R}{h} - 1)^2 + c$$

The condition for the sum is

$$\alpha_0 \geq 2\alpha_1 + \alpha_2 + 2\alpha_3$$
$$1 + \frac{k\sigma^2 m^2 + kc}{2} \geq \frac{k}{2}\left(\sigma^2 m^2 + rm\right) + 1 - \frac{k\sigma^2 m^2 + kc}{2} + \frac{k}{2}\left(\sigma^2 m^2 - rm\right)$$
$$k\sigma^2 m^2 + kc \geq \frac{k}{2}\left(\sigma^2 m^2 + rm\right) + \frac{k}{2}\left(\sigma^2 m^2 - rm\right)$$
$$k\sigma^2 m^2 + kc \geq \frac{k}{2}\sigma^2 m^2 + \frac{k}{2}\sigma^2 m^2$$
$$c \geq 0$$

Thus we see that 12 does not put constraints on the solution of the system. This is also the case for 14. The conditions presented in 13 give the conditions

$$\sigma^2 \geq r \text{ and } \frac{2}{k} \geq \sigma^2 (\frac{R}{h} - 1)^2 + c.$$

B Error bound for the backward Euler scheme

The relevant properties of the supporting function φ are that

$$\varphi \geq 0 \ \forall p \in \overline{\mathbb{G}} \qquad -\mathcal{L}_h \varphi_p = 1 \ \forall p \in \overline{\mathbb{G}}$$

This is equivalent to finding a non-negative function solving the continuous problem $-\mathcal{L}\varphi = 1$ on the domain $(x, t) \in ([0, \infty) \times [0, \infty))$. One solution for this is to construct a function $\varphi(x, t) = \varphi(t)$, such that all dependency on x ceases, ignoring the terms in the linear operator pertaining to partial derivatives with respect to x. Thus we obtain

$$\varphi(x, t) = \varphi(t)$$
$$-\mathcal{L}\varphi = \left(\partial_t - \frac{1}{2}\sigma^2 x^2 \partial_x^2 - rx\partial_x + c\right)\varphi = 1$$
$$-\mathcal{L}\varphi = \varphi_t + c\varphi = 1$$

This is a linear first order differential equation in time and its solution is readily available as

$$\varphi(t) = \frac{1}{c} + \kappa e^{-ct}$$

where c is the parameter from the description of the Black-Scholes system, while the κ is an undetermined coefficient. Using that $\varphi \geq 0$, also for $t = 0$, we get that

$$\frac{1}{c} + \kappa e^{-ct} \geq 0 \iff \frac{1}{c} \geq -\kappa e^{-c \cdot 0} = -\kappa$$

The worst case for this is that $\kappa = -1/c$. Since $\exp\{-ct\}$ is a monotonically decreasing function, the maximal value of φ must occur either for $t = 0$ or $t \to \infty$.

$$\varphi(0) = \frac{1}{c} - \frac{1}{c} e^{-c \cdot 0} = \frac{1}{c} - \frac{1}{c} = 0$$
$$\lim_{t \to \infty} \varphi(t) = \frac{1}{c}$$

$$\max_{0 \leq t < \infty} \varphi(t) = \max \left\{ 0, \frac{1}{c} \right\} = \frac{1}{c} \tag{15}$$

C Error bound for backward Euler

We have obtained a bound for the solutions to the discretized linear operator for the Black-Scholes equation, and it is

$$\max_{p \in \overline{\mathbb{G}}} V_p \leq \frac{\|\vec{f}\|_\infty}{c}$$

Since, by assumption, both $\mathcal{L}_h U$ and $\mathcal{L}_h u$ solve the discretized linear operator and $e := u - U$, we have that

$$-\mathcal{L}_h e_p = -\mathcal{L}_h u_p + \mathcal{L}_h U_p = \tau + \mathcal{L}_h U - \mathcal{L}_h U = \tau$$

Thus we have a valid solution e with a righthand side. By stability we have

$$\max e_p \leq \frac{1}{c} \|\tau\|_\infty$$

Which has a known bound due to consistency. Hence the error bound is known as

$$\max_{p \in \overline{\mathbb{G}}} e_p \leq \frac{1}{c} \left(k \left(\max \frac{1}{2} \partial_t^2 u \right) + h^2 \left(\max \frac{1}{6} \partial_x^3 u + \max \frac{1}{12} \partial_x^4 u \right) \right)$$

D Stability and truncation error of 1D non-linear Black-Scholes

The 1D non-linear Black-Scholes equation is given by

$$u_t - \frac{1}{2} x^2 \varphi\left(u_{xx}\right) u_{xx} = 0, \quad x \in \mathbb{R}^+, t \in (0, T)$$

where φ is given by

$$\varphi(x) = \sigma_1^2 + \frac{\sigma_2^2 - \sigma_1^2}{2} \left(1 + \frac{2}{\pi} \arctan x \right)$$

On this equation we want to test the two schemes IMEX

$$\frac{1}{k}\nabla_t U_m^{n+1} = \frac{1}{2}x_m^2 \varphi\left(\frac{1}{h^2}\delta_x^2 U_m^n\right)\frac{1}{h^2}\delta_x^2 U_m^{n+1}$$

and BE

$$\frac{1}{k}\nabla_t U_m^{n+1} = \frac{1}{2}x_m^2 \varphi\left(\frac{1}{h^2}\delta_x^2 U_m^{n+1}\right)\frac{1}{h^2}\delta_x^2 U_m^{n+1}$$

Claim: Both methods are monotone Proof: If

$$\alpha = \frac{1}{2}x_m^2 max\{\varphi\}\frac{k}{h^2}$$

then

$$\begin{gathered}
\frac{1}{k}\nabla_t U_m^{n+1} = \frac{1}{2}x_m^2 \varphi\left(\frac{1}{h^2}\delta_x^2 U_m^n\right)\frac{1}{h^2}\delta_x^2 U_m^{n+1} \\
U_m^{n+1} - U_m^n \leq \alpha\delta_x^2 U_m^{n+1} \\
U_m^{n+1} - U_m^n \leq \alpha\left(U_{m-1}^{n+1} + U_{m+1}^{n+1} - 2U_m^{n+1}\right) \\
\alpha_0 U_m^{n+1} - \beta_0 U_m^{n+1} - \beta_{-1}U_{m-1}^{n+1} - \beta_1 U_{m+1}^{n+1} \leq 0
\end{gathered}$$

where

$$\begin{gathered}
\alpha_0 = 1 + 2\alpha, \\
\beta_0 = 1, \\
\beta_1 = \beta_{-1} = \alpha
\end{gathered}$$

As φ and x_m^2 are both strictly positive, the discrete maximum principle conditions

$$\begin{gathered}
\alpha_0 \geq \sum_i \beta \geq 0 \\
\alpha_0, \beta \geq 0
\end{gathered}$$

are fulfilled.

As the only difference between IMEX and BE is inside the φ, which we bound by its L^∞ norm anyway, this same result holds for BE as well as IMEX.

For the BE method the truncation error is

$$\begin{aligned}
\tau_m^n &= \mathcal{L}_{k,h}u_m^n - \mathcal{L}u_m^n \\
&= \left(\nabla_{k,t} - \partial_t\right)u_m^n - \frac{1}{2}x_m^2\left(\varphi(\delta_{h,x}^2 u_m^n)\delta_{h,x}^2 u_m^n - \varphi(\partial_x^2 u_m^n)\partial_x^2 u_m^n\right) \\
&= \frac{k}{2}u_{m,tt}^n + O(k^2) + \frac{x_m^2}{2}\left[\varphi(u_{mxx}^n)\partial_x^2 u_m^n + \frac{h^2}{12}u_{m,x^4}^n\varphi_x(u_{m,xx}^n)\partial_x^2 u_m^n + \frac{h^2}{12}u_{m,x^4}^n\varphi(u_{m,xx}^n) + O(h^4) - \varphi(u_{m,xx}^n)\partial_x^2 u_m^n\right] \\
&= \frac{k}{2}u_{m,tt}^n + \frac{x_m^2}{2}\frac{h^2}{12}u_{m,x^4}^n\left(\varphi_x(u_{m,xx}^n)u_{m,xx}^n + \varphi_x(u_{m,xx}^n)\right) + O(k^2 + h^4)
\end{aligned}$$

This means taking the norm of the truncation error yields

$$\begin{aligned}
\|\tau_m^n\|_\infty &\leq \frac{k}{2}\|u_{m,tt}^n\|_\infty + \frac{x_m^2}{2}\frac{h^2}{24}\|u_{m,x^4}^n\|_\infty\|\varphi_x(u_{m,xx}^n)u_{m,xx}^n + \varphi(u_{m,xx}^n)\|_\infty \\
&\leq \frac{k}{2}\|u_{m,tt}^n\|_\infty + \frac{x_m^2}{2}\frac{h^2}{24}\|u_{m,x^4}^n\|_\infty\left(\frac{\sigma_2^2 - \sigma_1^2}{\pi}\|u_{m,xx}^n\|_\infty + \sigma_2^2\right)
\end{aligned}$$

And by the relation $|x_m| \leq \|x\|_\infty$we get

$$\|\tau^n\|_\infty \leq \frac{k}{2}\|u_{tt}^n\|_\infty + \frac{R^2}{2}\frac{h^2}{24}\|u_{x^4}^n\|_\infty\left(\frac{\sigma_2^2 - \sigma_1^2}{\pi}\|u_{xx}^n\|_\infty + \sigma_2^2\right)$$

For the IMEX scheme, we get the same derivation with the minor difference that some of the $\|u^n\|_\infty$ are interchanged with $\|u^{n-1}\|_\infty$, which we could bound by just taking the max of the two and use that one in the truncation formula.