A framework for analysing difference schemes

The goal of this chapter is to develop general tools which allow us to assess discretisation schemes qualitatively and quantitatively. As an exact solution is usually unattainable, we look for strategies where the numerical solution can be made as accuurate as necessary. This gives rise to the concept of convergence. Convergence here is understood in the sense that by choosing the gridsize and timestep small enough, we can always ensure that the error measured against the exact solution of the underlying differential equation does not exceed a prescribed tolerance. This is not just of interest to the numerical analyst, but also decidedly important in practice, because it allows the financial engineer to disentangle the modelling error, i.e. the error incurred by using an inaccurate model, from the error of the numerical solution to the model equations. Even if we do not always have full control over the adequacy of a model, at least we can be sure that the numerical solution procedure does not add noticably to the error of the final result, and preserves relevant properties of the true solution adequately.

4.1 Time-stepping schemes and error propagation

For the vast majority of models in financial practice, the pricing equations are parabolic partial differential equations. They describe the evolution of a transition density or expectation forward or backward in time. The notation here is for a forward equation, but the same argument applies to backward equations, if we think of time as time-to-expiry, with expiry a given finite time-horizon.

Here, we assume an initial value at $t = 0$ to be known and are ultimately interested in the solution u at a time T. Numerical methods for these equations naturally have the structure of time-stepping schemes. In order to find an approximation to u, M discrete time steps $t_0 = 0, t_1 = \Delta t, t_2 = 2\Delta t, \ldots, t_M = M\Delta t = T$ are applied. As previously, we introduce u^m an approximation to u at time t_m. The idea is that we can approximate the equation by a more tractable one over a small time interval with sufficient accuracy.

A finite difference scheme of the form

$$u^m = Lu^{m-1}, \quad 1 \leq m \leq M,$$

with a linear operator (i.e. matrix) L takes us from an approximation at t_{m-1} to one at t_m.

The schemes introduced earlier can all be written in this form: For the implict scheme (2.28), for instance, define $L = K^{-1}$. Starting at an initial value u^0, the solution at t_M is then

$$u^M = L^{M-m}u^m = L^M u^0, \quad 0 \le m \le M.$$

The lowest curve on Fig. 4.1 gives an illustration of this for $M = 4$, the top curve being the true solution. We keep in the back of our mind that u usually has a continuous spatial dimension x as well, and that u^m is a vector of grid values approximating u at spatially distributed grid points. We suppress this in our notation. In this sense, we can think of Fig. 4.1 as a cross-section through the solution for fixed x, or some other derived quantity.

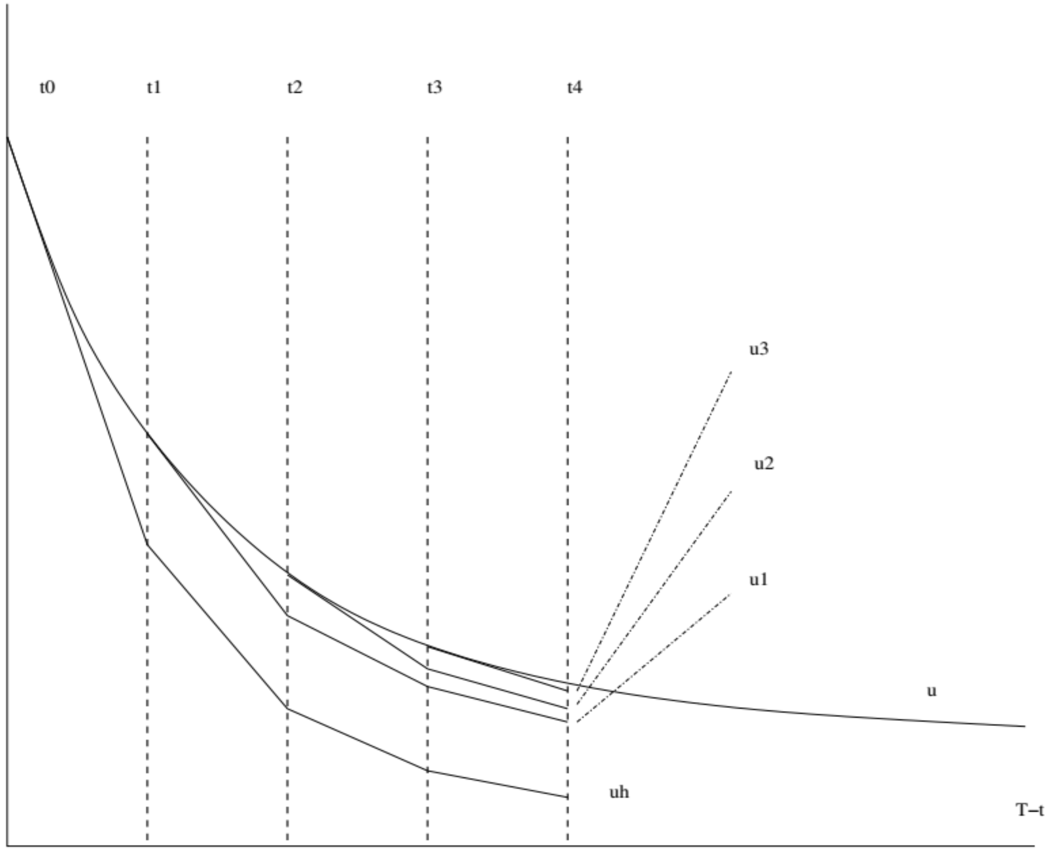

Figure 4.1: Propagation of discretisation errors for $M = 4$ timesteps.

At $t = t_0$, the solution is given by the known initial value $u(t_0)$. Stepping from $t = t_0$ to $t = t_1$, the difference scheme gives an approximation $u^1 = Lu^0 = Lu(t_0)$ to $u(t_1)$.[1] If we apply a second finite difference step at t_1, the error found at t_2 has two contributions, the error from approximating the model by a finite difference scheme in $[t_1, t_2]$, plus the error from departing at t_1 from the approximate solution u^1 rather than the exact one $u(t_1)$:

$$u^2 - u(t_2) = L\left(Lu(t_0) - u(t_1)\right) + \left(Lu(t_1) - u(t_2)\right).$$

---

[1] Note that $u(t_m)$ is a continuous function in the spatial direction, whereas u^m is a vector and L is a matrix acting on u^m. We understand $Lu(t_m)$ as first restricting $u(t_m)$ to the grid to obtain a grid vector, and then applying L.

Here $Lu(t_1)$ is the solution we would have obtained had we used as initial condition at t_1 the exact solution $u(t_1)$. This is somewhat hypothetical since we do not know $u(t)$ for $t > 0$, but useful for the analysis. Compare again to Fig. 4.1 for illustration.

More generally, $L^{m-j}u(t_j)$ for $M \geq m \geq j \geq 0$ is the hypothetical solution of a finite difference scheme starting from the exact solution $u(t_j)$ at t_j. Special boundary cases are

$$\begin{aligned} L^m u(t_0) &= u^m, \\ L^0 u(t_m) &= u(t_m). \end{aligned}$$

With this in mind,

$$u^M - u(T) = L^M(t_0) - L^0 u(t_M) = \sum_{m=1}^{M} L^{M-m} \left(Lu(t_{m-1}) - u(t_m) \right). \tag{4.1}$$

The error at time T is made up from the propagation of all local errors introduced before T. In each term (of the sum), the second factor measures how well the finite difference scheme tracks the continuous evolution over a single timestep. The first factor controls how these individual errors evolve over time. The art of designing a good numerical scheme lies in optimising the local approximation for small timesteps $\Delta t = T/M$, while keeping the error propagation under control when the total number of timesteps $M = T/\Delta t$ increases accordingly.

An added complexity is that a grid size Δx in the spatial direction has to be chosen suitably for every Δt, such that both go to zero following a functional relation $\Delta x = g(\Delta t)$. A good choice will balance the discretisation errors from the time- and space-discretisations, under the constraint that the scheme is stable. Take such a dependence as given as basis for the present analysis. (We will analyse this question for various finite difference schemes at length later.) We also make the dependence of L on Δt explicit now by writing $L(\Delta t)$.

From the above analysis, summarised in (4.1), we can deduce two key desiderata for a good numerical scheme:

1. *Consistency*: The error introduced in a single time step by replacing the original model with its finite difference approximation has to be small. Specifically, it has to go to zero faster than the stepsize,

   $$\frac{1}{\Delta t} \left(L(\Delta t) u(t_{m-1}) - u(t_m) \right) \to 0 \text{ for } \Delta t \to 0. \tag{4.2}$$

   This is because the number of error sources is $M = T/\Delta t$, so for the accumulated error to vanish for $\Delta t \to 0$, it is necessary that each individual term goes to zero faster than Δt. Comparing with Fig. 4.1, this says intuitively that in the limit the approximation is tangential to the exact solution. Such a scheme is called *consistent* with the differential equation. The term in (4.2) is called *truncation error*.

2. *Stability*: Once a numerical error is introduced, which invariably happens, it is necessary that it does not blow up through iteration with the numerical scheme, if the number of time steps increases. This requires that

   $$L(\Delta t)^M \text{ is bounded for } M \to \infty,\ \Delta t = T/M. \tag{4.3}$$

   Such a scheme is called *stable*.

From (4.1), we could potentially afford for L^M to increase as $\Delta t \downarrow 0$ if only the truncation error goes to zero fast enough. This may be true in an ideal world, but in practice other effects including boundary conditions, rounding errors etc introduce additional errors and (4.3) ensures that their influence on the final solution stays within bounds.

To make all these statements precise, we have to decide on a suitable measure for the error. The notions of convergence, boundedness etc are then to be interpreted in this sense. We analyse this a bit more formally in the next sections.

4.2 Convergence analysis of difference schemes

The pragmatic goal of a convergence analysis is to understand if and how a sufficiently "accurate" solution can be found by an appropriate choice of the numerical parameters. What exactly "accurate" means, depends on the computational question at hand.

If the desired result is a transition density u of a process X_t, there are various choices to measure the distance between u and a numerically computed density $\hat{u}$, for instance the Hellinger distance

$$d(u, \hat{u}) = \sqrt{\int_{\mathbb{R}} (\sqrt{u} - \sqrt{\hat{u}})^2 \, \mathrm{d}x}.$$

This will be difficult to come by through a numerical analysis of the PDE, and the probability density is usually used to compute some derived quantity.

We might ultimately be interested in the expectation of a function of the process at some time T, say $g(X_T)$,

$$\mathbb{E}[g(X_T)] = \int_{\mathbb{R}} g(x) p(x, T) \, \mathrm{d}x.$$

Think of X_t modelling a stock and g the payoff of an option. The error of the computation is

$$\text{error} = \int_{\mathbb{R}} g(x) \left(u(x,T) - \hat{u}(x,T) \right) \mathrm{d}x \leq \sqrt{\int_{\mathbb{R}} g(x)^2 \, \mathrm{d}x} \; \sqrt{\int_{\mathbb{R}} \left(u(x,T) - \hat{u}(x,T) \right)^2 \, \mathrm{d}x}.$$

Assuming the first factor is finite, e.g. for an option with bounded pay-off, a mean-square error norm would tell us how accurate the final result is.

Most commonly, the relevant result of the computation is an option price directly, by solving a Black-Scholes-type equation for instance. Typically we are most interested in the value function evaluated at the current time and the current spot price of the underlying asset. A pointwise error measure would then seem appropriate. It is however difficult to analyse the error at a single point in isolation, and we may want to reuse the solution to find the value of the solution at a later time, for a different value of the underlying process. A (usually pessimistic) upper bound which accounts for this the maximum norm.

We will develop a generic framework for the analysis first. Recall a one-step difference scheme in the form

$$u^m = L(\Delta t) u^{m-1}. \tag{4.4}$$

The igredients for the analysis are a vector norm $|\cdot|$, for the numerical solution u^m, and the associated operator/matrix norm $\|\cdot\|$ via

$$\|L\| = \max_{u, |u|=1} |Lu|,$$

for the discretisation matrix.

4.2.1 Stability

Definition 4.2.1. A scheme (4.4) is called *stable* in a norm $\| \cdot \|$ if

$$\|L(\Delta t)^m\| \leq C$$

for all $m \leq M$, $\Delta t = T/M$, and $C, T \geq 0$ are fixed.

Example 4.2.2 (Maximum stability of the θ-scheme)**.** *We have seen that discrete minimum/maximum principles,*

$$\min u^0 \leq u^m = L^m u^0 \leq \max u^0,$$

hold for the implicit scheme, and for the explicit scheme if $\Delta t \leq \Delta x^2$. *Here the minimum/maximum are taken over the range of grid points,* $\min u^m = \min_{n \in \mathbb{Z}} u_n^m$ *etc. From*

$$u^m \leq \max u^0 \leq \max |u^0|,$$
$$-u^m \leq -\min u^0 = \max -u^0 \leq \max |u^0|,$$

follows

$$|L^m u^0| = |u^m| \leq \max |u^0|,$$

for all u^0 *and therefore* $\|L^m\| \leq 1$.

The explicit Euler scheme is conditionally stable *in the maximum norm, i.e. stable (only) if* $\Delta t \leq \Delta x^2$. *The implicit scheme is* unconditionally stable.

Now note that the θ*-scheme can be seen as an explicit step of size* $(1-\theta)\Delta t$*, followed by an implicit step of size* $\theta \Delta t$*. The implict step is always stable, the explicit one if*

$$(1-\theta)\Delta t \leq \Delta x^2. \tag{4.5}$$

There is a timestep constraint (4.5) for maximum norm stability of the θ*-scheme for* $\theta < 1$.

4.2.2 Truncation error and consistency

We cannot easily measure how well the finite difference solution satisfies the PDE, since it is only defined at discrete points. What we can measure is how well the solution to the PDE, evaluated at the grid points, satisfies the difference scheme. Notation is as in 4.1.

Definition 4.2.3 (Truncation error and consistency)**.** Let u be the solution to a PDE with time-coordinate t.

1. The *truncation error* of a one-step difference scheme of the form (4.4) is defined as

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

2. The scheme is called *consistent* with the PDE if

$$T \to 0$$

for $\Delta x, \Delta t \to 0$, where Δx is a spatial grid size related to Δt.

3. It is called *consistent of order* (p,q), if for sufficiently smooth u

$$T = O(\Delta x^p, \Delta t^q)$$

Consistency expresses that in the limit for vanishing grid size and timestep, the discrete equation approaches in a certain sense the continuous one. The consistency order measures the speed of this.

Example 4.2.4 (Consistency of the explicit and implicit scheme). *The truncation error is easily calculated by Taylor expansion of*

$$\begin{aligned} u(x,t+\Delta t) &= u + \Delta t \frac{\partial u}{\partial t} + \tfrac{1}{2}\Delta t^2 \frac{\partial^2 u}{\partial t^2} + o(\Delta t^2), \\ u(x \pm \Delta x, t) &= u \pm \Delta x \frac{\partial u}{\partial x} + \tfrac{1}{2}\Delta x^2 \frac{\partial^2 u}{\partial x^2} \pm \frac{1}{6}\Delta x^3 \frac{\partial^3 u}{\partial x^3} + \frac{1}{24}\Delta x^4 \frac{\partial^4 u}{\partial x^4} + o(\Delta x^4), \end{aligned}$$

where arguments (x,t) of u and its derivatives are omitted on the right-hand side. In the case of the explicit scheme, directly from the definition of the truncation error,

$$\begin{aligned} T(x,t+\Delta t) &= \frac{u(x,t+\Delta t) - u(x,t)}{\Delta t} - \frac{1}{2}\frac{u(x+\Delta x,t) - 2u(x,t) + u(x-\Delta x,t)}{\Delta x^2} \\ &= \frac{\partial u}{\partial t} - \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \frac{1}{2}\Delta t \frac{\partial^2 u}{\partial t^2} - \frac{1}{24}\Delta x^2 \frac{\partial^4 u}{\partial x^4} + o(\Delta t) + o(\Delta x^2) \qquad (4.6) \\ &= O(\Delta t) + O(\Delta x^2), \qquad (4.7) \end{aligned}$$

using in (4.6) that u is a solution to the heat equation. For the implicit scheme, is is helpful to write

$$\begin{aligned} L^{-1}T(x,t) &= \frac{u(x,t) - u(x,t-\Delta t)}{\Delta t} - \frac{1}{2}\frac{u(x+\Delta x,t) - 2u(x,t) + u(x-\Delta x,t)}{\Delta x^2} \\ &= \frac{\partial u}{\partial t} - \frac{1}{2}\frac{\partial^2 u}{\partial x^2} - \frac{1}{2}\Delta t \frac{\partial^2 u}{\partial t^2} - \frac{1}{24}\Delta x^2 \frac{\partial^4 u}{\partial x^4} + o(\Delta t) + o(\Delta x^2) \\ &= O(\Delta t) + O(\Delta x^2), \qquad (4.8) \end{aligned}$$

and now use the stability of L to deduce $T = O(\Delta t) + O(\Delta x^2)$. The explicit and implicit finite difference scheme with central differences are both consistent with the heat equation of order $p = 2$ in x and $q = 1$ in t.

Remark 4.2.5. *As concerns technique, we note from the implicit case that if a scheme is more naturally written as*

$$K_1 u^m = K_2 u^{m-1},$$

as is the case for the implicit scheme above – with $K_2 = I$, $L = K_1^{-1}$ – or more generally for the θ-scheme, it is sufficient to use Taylor expansion on the scheme in its simpler form, and then appeal to stability of the implicit part.

It is left as an exercise to show that the θ-scheme is of second order accurate in Δt if and only if $\theta = 1/2$, i.e. for the Crank-Nicolson scheme.

Another intersting observation is to be made from the analysis of the explicit scheme. Differentiating the heat equation,

$$\frac{\partial^2 u}{\partial t^2} = \frac{1}{2}\frac{\partial}{\partial t}\frac{\partial^2 u}{\partial x^2} = \frac{1}{4}\frac{\partial^4 u}{\partial x^4},$$

so by inspection of (4.6) one sees that the terms proportional to Δt and Δx^2 can be made to cancel by choosing $\Delta t/\Delta x^2 = 1/3$. The scheme is then of order $p = 2$, $q = 4$, at no extra cost. Note this choice is in the stability range $\Delta t \le \Delta x^2$.

4.2.3 Consistency + stability = convergence

Theorem 4.2.6 (Lax Equivalence Theorem)**.** *A consistent scheme is convergent if and only if it is stable.*

Proof. We only show the direction implying convergence. For a consistent and stable scheme, it follows from (4.3) and (4.2) that

$$\begin{aligned} |u(T) - u^M| &\le \sum_{m=1}^{M} \|L^{M-m}\| |Lu(t_{m-1}) - u(t_m)| \\ &\le M \max_{1\le m\le M} \|L^{M-m}\| \max_{1\le m\le M} |Lu(t_{m-1}) - u(t_m)| \\ &= T \max_{1\le m\le M} \|L^m\| \max_{1\le m\le M} \frac{1}{\Delta t} |Lu(t_{m-1}) - u(t_m)| \qquad (4.9) \\ &\to 0 \quad \text{for } \Delta t \to 0, \qquad (4.10) \end{aligned}$$

i.e. the scheme is convergent.

□

Looking at the proof, one also sees that a a stable scheme that is consistent (of order (p,q)) is convergent (of order (p,q)).

A direct consequence is:

Corollary 4.2.7. *The explicit Euler schemes is conditionally convergent of order* $(2,1)$*, i.e. provided* $\Delta t \le \Delta x^2$*. The implicit Euler scheme is unconditionally convergent of order* $(2,1)$.

The error can be made as small as required by reducing Δx and Δt, i.e. by increasing N and M. For stability of the explicit method, one has to choose $M \sim N^2$, whereas there is no such restriction for the fully implict scheme. This implies that if we keep doubling the number of grid points, we have to increase the number of timesteps by a factor of four each time. The error of the explicit and implicit schemes is $O(M^{-1}) + O(N^{-2})$. For both error terms to be of the same order of magnitude, it is also necessary that $M \sim N^2$. So even if there was no constraint on the stability of the method, it would be optimal to keep this relation between M and N. This puts severe constraints on the viability of the method. If we double N, we must take four times as many timesteps, and the overall computation time will be eight times as long. This is just to reduce the error by a factor of four.

For the Crank-Nicolson scheme, the error is $O(M^{-2}) + O(N^{-2})$, which would suggest and optimal choice of $M \sim N$, however (4.5) dictates $M \sim N^2$ for maximum norm stability. Experiments suggest that under certain circumstances, notably smooth initial data, we can get away without (4.5), hence improving the efficiency of the Crank-Nicolson method. We need a more refined analysis to explain this.

4.3 Spectral analysis and mean-square convergence

The previous analysis went some distance towards explaining the behaviour of finite difference schemes, however there are still at least two main open questions. The first concerns the unexplained empirical convergence of the Crank-Nicolson scheme for large timesteps in the case of smooth solutions. The second concerns the properties of solutions for non-smooth data, which are not covered by the analysis at all. While explicit and implicit schemes may have given optimism that the same criteria might hold as for the smooth case, this was not the case for the Crank-Nicolson scheme.

4.3.1 Von Neumann stability analysis

It was observed in numerical experiments earlier that instabilities resulted in the explosion of highly oscillatory modes. A (historically) very popular tack of analysis uses a more or less explicit representation of the finite difference solution, akin the classical Fourier series solution of the heat equation. This allows to investigate the dependence of amplification on the wave number. For the continuous as for the discretised form, the strategy it relies on linear constant coefficient equations to achieve a separation of variables with exponential/trigonometric simple solutions, and a superposition principle to construct more complex solutions which match initial data.

Recall the θ-scheme as

$$\frac{u_n^m - u_n^{m-1}}{\Delta t} = \frac{\theta}{2}\frac{u_{n+1}^m - 2u_n^m + u_{n-1}^m}{\Delta x^2} + \frac{(1-\theta)}{2}\frac{u_{n+1}^{m-1} - 2u_n^{m-1} + u_{n-1}^{m-1}}{\Delta x^2}.$$

For simplicity look at the explicit scheme, $\theta = 0$, first, then

$$u_n^m = u_n^{m-1} + \frac{\lambda}{2}\left(u_{n-1}^{m-1} - 2u_n^{m-1} + u_{n+1}^{m-1}\right)$$

with $\lambda = \Delta t/\Delta x^2$. We ignore boundary conditions for now and consider the problem for $n \in \mathbb{Z}$.

We could run through a similar derivation as for the heat equation itself, but to cut to the chase guess solutions with a spatial variation of the form $\cos(kn)$, which we extend to

$$u_n^m = R_0^m \mathrm{e}^{ikn}$$

with $i = \sqrt{-1}$ to simplify the algebra with trigonometric functions. To get the original solution back simply take the real part.

From

$$\frac{1}{\Delta x^2}\left(u_{n+1}^m - 2u_n^m + u_{n-1}^m\right) = \frac{1}{\Delta x^2}\mathrm{e}^{ikn}\left(\mathrm{e}^{ik/2} - \mathrm{e}^{-ik/2}\right)^2 = -\frac{4}{\Delta x^2}\sin^2(k/2)\, u_n^m$$

one obtains

$$u_n^m = R_0(\Delta x, \Delta t; k) u_n^{m-1}$$

with

$$R_0(\Delta x, \Delta t; k) = 1 - 2\Delta t/\Delta x^2 \sin^2(k/2)$$

R is called the *symbol* of the method.

Similar to the continuous case, differencing has turned into a multiplication.

Indeed, for the θ-scheme one gets by similar reasoning (Exercise 4, Section 4.4)

$$u_n^m = R_\theta(\Delta x, \Delta t; k) u_n^{m-1}$$

with

$$R_\theta(\Delta x, \Delta t; k) = \frac{1 - 2(1-\theta)\Delta t/\Delta x^2 \sin^2(k/2)}{1 + 2\theta \Delta t/\Delta x^2 \sin^2(k/2)}. \tag{4.11}$$

If we carry on, this gives

$$u_n^M = R_\theta(\Delta x, \Delta t; k)^M u_n^0$$

We expect the method can generally only be stable, if

$$|R_\theta(\Delta x, \Delta t; k)|^M \leq C \quad \forall k \tag{4.12}$$

where C is independent of M, $\Delta t = T/M$, and Δx chosen as a function of Δt. A sufficient criterion is clearly $|R_\theta| \leq 1$. It is left as an exercise to show that (4.12) is true if and only if there is $c > 0$ such that

$$|R_\theta(\Delta x, \Delta t; k)| \leq 1 + c\Delta t.$$

Observe that always $R_\theta \leq 1$, and the difficulty comes from "overshooting" beyond -1 for large timesteps. Fig. 4.2 illustrates this for different values of θ and for increasing mesh rations $\lambda = \Delta t/\Delta x^2$.

Simple algebra shows that the explicit scheme is stable in this sense if

$$\Delta t \leq \Delta x^2$$

More generally, a sufficient criterion for the θ-scheme is

$$(1 - 2\theta)\Delta t \leq \Delta x^2. \tag{4.13}$$

This is true for all values of Δt, if and only if

$$\theta \geq \frac{1}{2},$$

then the method is unconditionally stable. The Crank-Nicolson method is the stable limiting case. For $\theta < 1/2$, there is a stability constraint on Δt. Note in how far (4.13) is an improvement to (4.5).

4.3.2 Mean-square stability

To see that this gives stability for "general" initial conditions, we can use a dicrete-continuous Fourier decomposition

$$u_n^0 = \frac{1}{2\pi\Delta x} \int_{-\pi}^{\pi} \widehat{u}^0(k) \mathrm{e}^{ink} \, \mathrm{d}k, \tag{4.14}$$

where

$$\widehat{u}^0(k) = \Delta x \sum_{n=-\infty}^{\infty} u_n^0 \mathrm{e}^{-ink}. \tag{4.15}$$

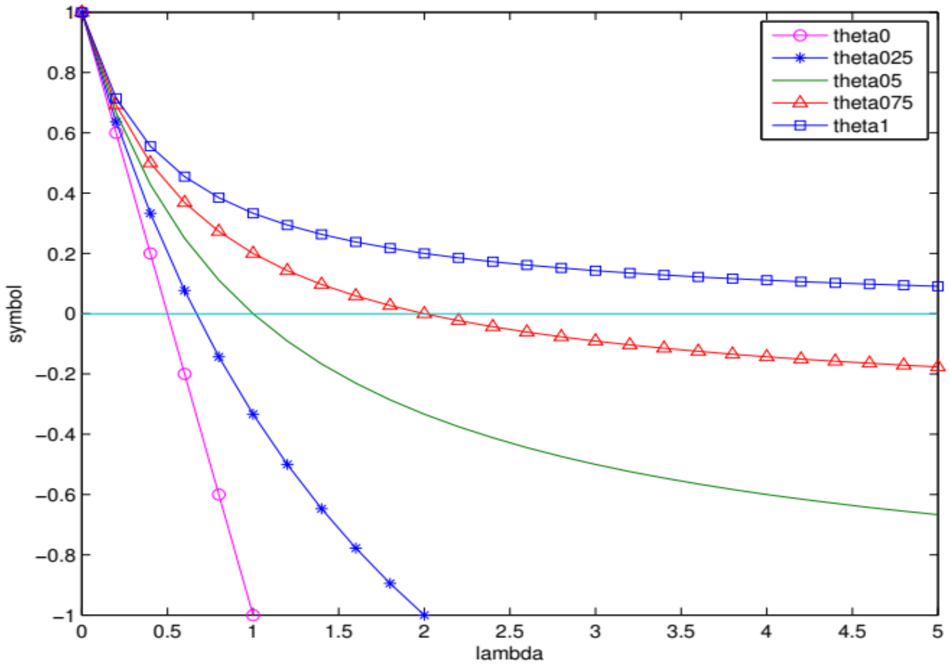

Figure 4.2: The amplification factor (4.11) for increasing mesh ratio $\lambda = \Delta t/\Delta x^2$ and different values of θ, $k = \pi$ fixed.

Then by linearity of the equation,

$$u_n^m = \int_{-\pi}^{\pi} R(\Delta x, \Delta t; k)^m \widehat{u}^0(k) \mathrm{e}^{ink}\, \mathrm{d}k. \tag{4.16}$$

The appropriate measure is a mean-square norm. For a vector $u = (u_n)_{-\infty<n<\infty}$, define

$$|u|_2 = \left(\Delta x \sum_{n=-\infty}^{\infty} u_n^2 \right)^{1/2}.$$

Then *Parseval's equality*

$$|u|_2^2 = \frac{1}{2\pi\Delta x} \int_{-\pi}^{\pi} |\widehat{u}(k)|^2 \, \mathrm{d}k$$

holds, which allows us to relate the norm of the solution to its Fourier modes,

$$\begin{aligned} |u^m|_2^2 &= \frac{1}{2\pi\Delta x} \int_{-\infty}^{\infty} |\widehat{u}^m(k)|^2 \, \mathrm{d}k \\ &= \frac{1}{2\pi\Delta x} \int_{-\infty}^{\infty} |R(\Delta x, \Delta t; k) \widehat{u}^{m-1}(k)|^2 \, \mathrm{d}k \\ &\leq \sup_k |R(\Delta x, \Delta t; k)|^2 |u^{m-1}|_2^2. \end{aligned}$$

So if $|R| \leq 1$,

$$|u^m|_2 \leq |u^0|_2$$

for $m \geq 0$.

Remark 4.3.1 (Analogy to the continuous problem)**.** *The Fourier (von Neumann) analysis of the discretisation is a semi-discrete version of the solution of the continuous equation via Fourier transform.*

$$u(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{u}(k,t)\mathrm{e}^{ikx}\,\mathrm{d}k$$

where

$$\widehat{u}(k,t) = \mathrm{e}^{-\frac{1}{2}k^2 t}\widehat{u}(k,0)$$

Note that for the continuous problem with infinite range the spectrum is continuous (Fourier transform), whereas for the discrete infinite problem the spectrum is continuous but of finite range.

Note that this analysis requires that the initial condition is square integrable. In particular, it breaks down for Dirac initial data. We analyse this separately in the next chapter.