KTH Computer Science
and Communication

Lecture Notes 4

Convergence Theory for Linear Methods

Let q_j^n be the numerical approximation of the exact cell average,

$$q_j^n \approx u_j^n := \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} u(t_n, x)dx, \qquad t_n = n\Delta t. \tag{1}$$

We want to check

- Convergence $q_j^n \to u_j^n$ as $\Delta x, \Delta t \to 0$,
- Accuracy and convergence rate

$$q_j^n = u_j^n + O(\Delta x^p + \Delta t^r),$$

  for some $p, r \geq 1$.

Notation

We consider two cases of the numerical approximation: with and without boundaries. When there are boundaries, we let $\boldsymbol{q}^n$ be the finite length vector

$$\boldsymbol{q}^n = (q_0^n, \ldots, q_N^n)^T.$$

When there are no boundaries we let $\boldsymbol{q}^n$ denote the infinite sequence

$$\boldsymbol{q}^n = (\ldots, q_{-1}^n, q_0^n, q_1^n \ldots),$$

and similarly for the exact solution $\boldsymbol{u}^n$. We write the numerical scheme compactly as an operator $\mathcal{N}$ acting on $\boldsymbol{q}^n$,

$$\boldsymbol{q}^{n+1} = \mathcal{N}(\boldsymbol{q}^n, \Delta t, \Delta x).$$

When $\mathcal{N}$ only depends on the CFL number $\lambda = \Delta t/\Delta x$ we simply write $\mathcal{N}(\boldsymbol{q}^n, \lambda)$ or just $\mathcal{N}(\boldsymbol{q}^n)$ when there is no risk for confusion.

Linear methods

We assume that $\mathcal{N}$ is a linear method, i.e. if α, β are scalars,

$$\mathcal{N}(\alpha\boldsymbol{q} + \beta\boldsymbol{u}) = \alpha\mathcal{N}(\boldsymbol{q}) + \beta\mathcal{N}(\boldsymbol{u}).$$

A linear method can always be represented by sequences of numbers, $\{b_{j,\ell}\}$, that depend on the time and spatial step size,

$$q_j^{n+1} = \sum_{\ell=-m}^{M} b_{j,\ell}(\Delta t, \Delta x) q_{j+\ell}^n, \tag{2}$$

for some finite m and M, which determines the width of the spatial stencils used. (All methods we considered so far are of this type. When the equation itself is nonlinear the scheme would in general not be linear, however.)

Example 1 *When $\mathcal{N}$ is the upwind scheme applied to*

$$u_t + a(x)u_x = 0, \qquad a(x) > 0,$$

then

$$q_j^{n+1} = q_j^n - \frac{\Delta t}{\Delta x} a(x_j)(q_j^n - q_{j-1}^n),$$

so that

$$b_{j,0} = 1 - a(x_j)\frac{\Delta t}{\Delta x}, \qquad b_{j,-1} = a(x_j)\frac{\Delta t}{\Delta x},$$

and all other $b_{j,\ell}$ are zero, i.e. $m = 1$ and $M = 0$.

Norms

To measure errors we need norms. We use the discrete version of the L^2 norm,

$$||\boldsymbol{q}||_{2,\Delta x}^2 = \sum_{j=0}^{N} |q_j|^2 \Delta x,$$

for the case with boundaries. Note that by the scaling with $N\Delta x$ =constant, the size of the norm should not increase as we refine the grid. In the case of no boundaries we similarly use

$$||\boldsymbol{q}||_{2,\Delta x}^2 = \sum_{j=-\infty}^{\infty} |q_j|^2 \Delta x,$$

which mimics a trapezoidal rule approximation of the continuous L^2 norm when q_j approximates a smooth function – again the norm should then be bounded as $\Delta x \to 0$.

One can also use the discrete L^1 norm,

$$||\boldsymbol{q}||_{1,\Delta x} = \sum_{j=0}^{N} |q_j| \Delta x, \qquad ||\boldsymbol{q}||_{1,\Delta x} = \sum_{j=-\infty}^{\infty} |q_j| \Delta x.$$

Sometimes we just write $|| \cdot ||_{\Delta x}$ when the precise norm type is not important.

1 Convergence theory

Convergence is usually established using the Lax equivalence theorem which states that a scheme is convergent if and only if it is *consistent* and *stable*. To check convergence for our scheme $\mathcal{N}$ we must thus verify these two properties and then apply the theorem. We go through the three steps here.

1.1 Consistency

A scheme is consistent if the exact solution fits the scheme well. More precisely, we define the *local truncation error* $\boldsymbol{\tau}^n$ as

$$\boldsymbol{u}^{n+1} = \mathcal{N}(\boldsymbol{u}^n) + \Delta t \boldsymbol{\tau}^n.$$

The local truncation error is thus the residual when the exact solution $\boldsymbol{u}^n$ (instead of $\boldsymbol{q}^n$) is entered into the scheme, scaled by Δt. One can also think of it as the error performed in one time step, scaled by Δt.

For convergence we need a small $\boldsymbol{\tau}^n$. We say that the method is *consistent* if

$$\max_{0\leq n\Delta t\leq T} ||\boldsymbol{\tau}^n||_{\Delta x} \to 0$$

as $\Delta t, \Delta x \to 0$, for a fixed T. Moreover, if there is a number C independent of Δt and Δx such that

$$\max_{0\leq n\Delta t\leq T} ||\boldsymbol{\tau}^n||_{\Delta x} \leq C(\Delta x^p + \Delta t^r)$$

we say that the method is of order p in space and r in time. If we use a constant $\lambda = \Delta t/\Delta x$, or more generally if $\lambda = \mathcal{O}(1)$, then $||\boldsymbol{\tau}^n||_{\Delta x} = \mathcal{O}((\Delta x)^p + (\Delta x)^r) = \mathcal{O}((\Delta x)^q)$ where $q = \min(p, r)$ and we simply say that the method is of order q.

Consistency and order can usually be checked by simple Taylor expansion of the exact solution and using the fact that it satisfies the PDE.

Example 2 *Consider the upwind method for $u_t + au_x = 0$. The local truncation error τ_j^n is defined by*

$$u_j^{n+1} = u_j^n - a\frac{\Delta t}{\Delta x}(u_j^n - u_{j-1}^n) + \Delta t\tau_j^n,$$

where u_j^n is the exact local average defined in (1). We can rewrite this as

$$\begin{aligned}\tau_j^n &= \frac{u_j^{n+1} - u_j^n}{\Delta t} + a\frac{u_j^n - u_{j-1}^n}{\Delta x}\\ &= \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} \frac{u(t_{n+1}, x) - u(t_n, x)}{\Delta t} + a\frac{u(t_n, x) - u(t_n, x - \Delta x)}{\Delta x}dx.\end{aligned}$$

Now we can Taylor expand the expressions inside the integral

$$\frac{u(t_{n+1}, x) - u(t_n, x)}{\Delta t} = u_t(t_n, x) + \frac{1}{2}\Delta t u_{tt}(t_n, x) + O(\Delta t^2),$$

and

$$a\frac{u(t_n, x) - u(t_n, x - \Delta x)}{\Delta x} = au_x(t_n, x) - \frac{a}{2}\Delta x u_{xx}(t_n, x) + O(\Delta x^2).$$

Then, since $u_t + au_x = 0$,

$$\begin{aligned}\tau_j^n &= \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} u_t(t_n, x) + \frac{1}{2}\Delta t u_{tt}(t_n, x) + au_x(t_n, x) - \frac{a}{2}\Delta x u_{xx}(t_n, x) + O(\Delta x^2 + \Delta t^2)dx\\ &= \frac{1}{2\Delta x}\int_{x_j}^{x_{j+1}} \Delta t u_{tt}(t_n, x) - a\Delta x u_{xx}(t_n, x)dx + O(\Delta x^2 + \Delta t^2)\\ &= \frac{\Delta t}{2}\cdot\frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} u_{tt}(t_n, x)dx - \frac{a\Delta x}{2}\cdot\frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} u_{xx}(t_n, x)dx + O(\Delta x^2 + \Delta t^2).\end{aligned}$$

Since the integrals are both bounded as $\Delta t, \Delta x \to 0$ (they are local averages of u_{tt} and u_{xx}) we conclude that

$$\tau_j^n = O(\Delta x + \Delta t),$$

showing that upwind is a consistent method that is first order in time and space.

We can make a more precise characterization of the local truncation error by differentiating the equation once in time and space to get

$$u_{tt} + au_{xt} = 0, \qquad u_{tx} + au_{xx} = 0.$$

Together this shows that $u_{tt} = a^2 u_{xx}$. *Therefore*

$$\tau_j^n = \frac{a(a\Delta t - \Delta x)}{2} \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} u_{xx}(t_n, x)dx + O(\Delta x^2 + \Delta t^2).$$

This kind of characterization is useful when one derives modified equations (see Leveque 8.6). It also shows that if one chooses the "magic time step" $\Delta t = \Delta x/a$ *the method is more accurate. In fact, then the numerical scheme is exact and* $\tau_j^n \equiv 0$. *This is, however, very special to the constant coefficient advection equation, and does not happen in general.*

1.2 Stability

The scheme is called (Lax-Richtmyer) stable if

$$||\mathcal{N}(\boldsymbol{q})||_{\Delta x} \leq (1 + \alpha \Delta t)||\boldsymbol{q}||_{\Delta x}$$

for all $\boldsymbol{q}$ and with α independent of $\boldsymbol{q}$, Δt and Δx. We will get back later to how this can be shown for a scheme.

Note that when $\mathcal{N}$ is nonlinear then we need instead an "almost contraction" property,

$$||\mathcal{N}(\boldsymbol{q}) - \mathcal{N}(\boldsymbol{q}')||_{\Delta x} \leq (1 + \alpha \Delta t)||\boldsymbol{q} - \boldsymbol{q}'||_{\Delta x}$$

for all $\boldsymbol{q}, \boldsymbol{q}'$ and with α independent of $\boldsymbol{q}$, $\boldsymbol{q}'$, Δt and Δx. See Leveque 8.3.

1.3 Convergence

By the Lax equivalence theorem

"stability + consistency $\Leftrightarrow$ convergence".

More precisely, the right implication $\Rightarrow$ gives both convergence and an error estimate — if the method is stable and consistent with order p in space and r in time (in the way defined above) we get

$$\max_{0 \leq n\Delta t \leq T} ||\boldsymbol{q}^n - \boldsymbol{u}^n||_{\Delta x} \leq C(\Delta x^p + \Delta t^r),$$

where C is independent of Δx and Δt, but in general depends on T and the exact solution $u(t, x)$. This is quite straightforward to prove, as follows:

We assume

- Stability

$$||\mathcal{N}(\boldsymbol{q})||_{\Delta x} \leq (1 + \alpha \Delta t)||\boldsymbol{q}||_{\Delta x}, \qquad \forall \boldsymbol{q},$$

- Consistency such that

$$\tau := \max_{0 \leq n\Delta t \leq T} ||\boldsymbol{\tau}^n||_{\Delta x} \leq C(\Delta x^p + \Delta t^r),$$

- Exact initial data,

$$q_j^0 = u_j^0, \qquad ||\boldsymbol{q}^0 - \boldsymbol{u}^0||_{\Delta x} = 0.$$

Let us first define the error $e_j^n = u_j^n - q_j^n$ and in vector form $\boldsymbol{e}^n = \boldsymbol{u}^n - \boldsymbol{q}^n$. Then

$$\boldsymbol{e}^{n+1} = \boldsymbol{u}^{n+1} - \boldsymbol{q}^{n+1} = \mathcal{N}(\boldsymbol{u}^n) + \Delta t \boldsymbol{\tau}^n - \mathcal{N}(\boldsymbol{q}^n) = \mathcal{N}(\boldsymbol{e}^n) + \Delta t \boldsymbol{\tau}^n,$$

where we used the definition of the truncation error and the linearity of $\mathcal{N}$. Then

$$\begin{aligned}
||\boldsymbol{e}^{n+1}||_{\Delta x} &\leq ||\mathcal{N}(\boldsymbol{e}^n)||_{\Delta x} + \Delta t ||\boldsymbol{\tau}^n||_{\Delta x} \leq \{\text{stability and def. of } \tau\} \\
&\leq (1+\alpha\Delta t)||\boldsymbol{e}^n||_{\Delta x} + \Delta t \tau \leq \{\text{applying same estimate to } \boldsymbol{e}^n\} \\
&\leq (1+\alpha\Delta t)^2||\boldsymbol{e}^{n-1}||_{\Delta x} + (1+\alpha\Delta t)\Delta t\tau + \Delta t\tau \leq \{\text{induction}\} \\
&\leq (1+\alpha\Delta t)^{n+1}||\boldsymbol{e}^0||_{\Delta x} + \sum_{j=0}^{n}(1+\alpha\Delta t)^j \Delta t \tau = \{\text{exact initial data}\} \\
&= \Delta t \tau \sum_{j=0}^{n} (1+\alpha\Delta t)^j.
\end{aligned}$$

The sum is a geometric series,

$$\sum_{j=0}^{n}(1+\alpha\Delta t)^j = \frac{(1+\alpha\Delta t)^{n+1}-1}{(1+\alpha\Delta t)-1} = \frac{(1+\alpha\Delta t)^{n+1}-1}{\alpha\Delta t}.$$

Hence,

$$||\boldsymbol{e}^n||_{\Delta x} \leq \tau \frac{(1+\alpha\Delta t)^n - 1}{\alpha}.$$

Now, using the fact that $1 + x \leq e^x$ for all x, we get

$$\max_{0\leq n\Delta t\leq T} ||\boldsymbol{e}^n||_{\Delta x} \leq \max_{0\leq n\Delta t\leq T} \tau \frac{e^{\alpha n \Delta t}-1}{\alpha} \leq \tau \frac{e^{\alpha T}-1}{\alpha}.$$

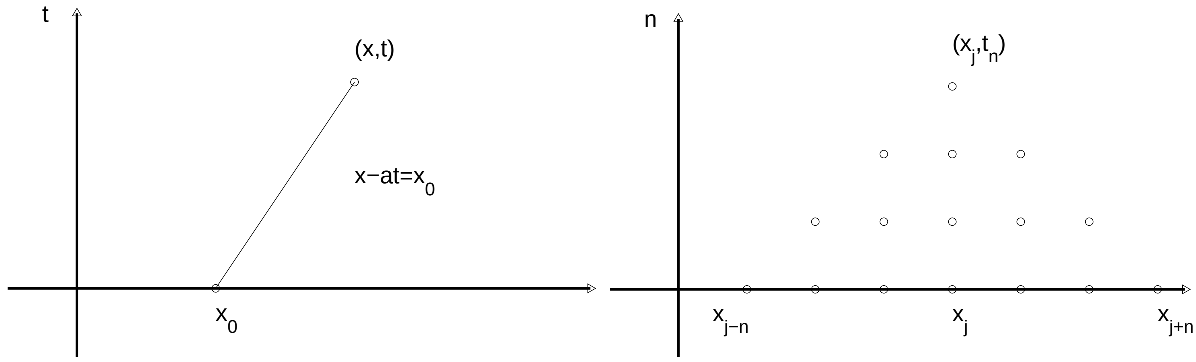
Hence,

$$\max_{0\leq n\Delta t\leq T} ||\boldsymbol{u}^n - \boldsymbol{q}^n||_{\Delta x} \leq C'(\Delta x^p + \Delta t^r),$$

where the number C' is the number C in the consistency assumption multiplied by $(e^{\alpha T} - 1)/\alpha$. This shows the convergence and error estimate.

Remark 1 *In general the boundary conditions can have a significant effect on stability, accuracy and convergence. The above analysis is not always sharp. For instance, the local truncation error can, sometimes, be allowed to have lower order at the boundaries without ruining the overall convergence rate.*

Also note that for higher order approximations wider spatial stencils are needed, which means that more ghost cells are needed. Then also more boundary conditions for these cells are needed. However, the PDE itself has a fixed number of boundary conditions. Hence, the number of numerical boundary conditions is often larger than the number of PDE boundary conditions. Chosing these extra conditions can be a delicate issue.



(a) Continuous problem (b) Numerical approximation

Figure 1. The CFL condition.

2 Checking stability

Checking stability of a scheme is usually the most difficult part when proving convergence. There are severel different approaches, for instance

1. CFL condition (necessary condition)
2. von Neumann analysis (sufficient condition, constant coefficients)
3. Energy method (sufficient condition, variable coefficients)

We will discuss the first two. The L^1 version of the energy method is briefly explained in Leveque 8.3.4.

Note that von Neumann analysis can only handle periodic boundary conditions or no boundaries. With the energy method more general boundary conditions can be handled.

2.1 CFL condition

Consider the constant coefficient advection equation

$$\begin{aligned} u_t + au_x &= 0, \quad x \in \mathbb{R},\ t > 0, \\ u(0,x) &= g(x). \end{aligned}$$

and an explicit "3-point method", i.e.

$$u_j^{n+1} = c_{-1}u_{j-1}^n + c_0 u_j^n + c_1 u_{j+1}^n,$$

for some coefficients c_j. In its most simple form, the CFL condition says that if this scheme is consistent then

$$|a|\frac{\Delta t}{\Delta x} \leq 1, \tag{3}$$

is a necessary condition for stability. Note that all methods we have considered so far have been consistent 3-point methods.

An intuitive explanation of the CFL condition can be made as follows. Consider the continuous case first (Figure 1a). The exact solution in the point (x,t) then only depends on the initial data $g(x)$ evaluated in the point $x_0 = x - at$. (The line $x - at = x_0$ is the characteristic passing through (x, t)). We say that the "domain of dependence" of (x, t) is $D(x, t) = \{x_0\} = \{x - at\}$. In the numerical case (Figure 1b), each value at time level n depends on the three surrounding values at time level $n-1$. Therefore, by induction, the numerical solution u_j^n at (x_j, t_n) depends on the initial data $g(x)$ evaluated at the points $x_{j-n}, \ldots, x_{j+n}$. In the limit as $\Delta t, \Delta x \to 0$, it will depend on all points between x_{j-n} and x_{j+n} and we set $D_{\text{num}}(x_j, t_n) = [x_{j-n}, x_{j+n}]$. The general CFL condition then says that:

> A consistent method can only be stable if the continuous domain of dependence $D(x_j, t_n)$ is a subset of the numerical domain of dependence $D_{\text{num}}(x_j, t_n)$.

This is a quite natural condition, since if $D(x_j, t_n) \not\subset D_{\text{num}}(x_j, t_n)$ then the numerical method cannot "know" what the exact solution in (x_j, t_n) should be, and there is no hope of getting a convergent method.

We should emphasize again that the CFL condition is only a necssary condition. The scheme may still be unstable of the condition is satisfied.

In our case, $D \subset D_{\text{num}}$ is equivalent to

$$x_{j-n} = x_j - n\Delta x \leq x_j - at_n \leq x_j + n\Delta x = x_{j+n},$$

which, implies

$$|a|t_n \leq n\Delta x,$$

which implies the CFL condition (3) since $t_n = n\Delta t$.

For a system of p equations, $u_t + Au_x = 0$, with $u \in \mathbb{R}^p$ and $A \in \mathbb{R}^{p\times p}$, the domain of dependence is

$$D(x, t) = \bigcup_{k=1}^{p} \{x - \lambda_k t\},$$

where λ_k are the eigenvalues of A. The same arguments as above then leads to the CFL condition

$$|\lambda_k| \frac{\Delta t}{\Delta x} \leq 1, \qquad k = 1, \ldots, p.$$

2.2 von Neumann analysis

In order to derive a sufficient condition for stability one can use von Neumann analysis. This is based on Fourier analysis and stability is shown in a way similar to how well-posedness is shown for the continuous problem. As in that case, von Neumann analysis requires that the problem has constant coefficients. More precisely, the scheme should have the same form at all grid points. This means that the (2) simplifies to

$$q_j^{n+1} = \sum_{\ell=-m}^{M} b_\ell(\Delta t, \Delta x) q_{j+\ell}^n, \tag{4}$$

i.e. b_ℓ does not depend on j. Moreover, there should either be no boundaries or periodic boundary conditions.

Example 3 *The upwind scheme in Example 1 for the constant coefficient case $u_t + au_x = 0$ reduces to $m = 1$, $M = 0$ and*

$$b_0 = 1 - a\lambda, \quad b_{-1} = a\lambda, \quad \lambda = \frac{\Delta t}{\Delta x}.$$

Remark 2 *von Neumann analysis works for any equation, not just hyperbolic equations. The natural relation between Δt and Δx may then differ, however. (E.g. $\Delta t/\Delta x^2 = O(1)$ for explicit methods for parabolic problems.)*

2.2.1 Periodic boundary conditions

We first assume that we have periodic boundary conditions. The space discretization is

$$x_j = j\Delta x, \qquad \Delta x = \frac{2\pi}{N},$$

and the approximation q_j^n satisfies

$$q_j^n = q_{j+N}^n, \qquad \forall j, \quad \forall n \geq 0.$$

Hence, we only compute the q_j^n values for $j = 0, \ldots, N-1$, but then define q_j^n for all j by periodicity. We assume also that N is even. Let $\hat{q}_k^n$ be the discrete Fourier transform of q_j^n, so that

$$q_j^n = \sum_{k=-N/2}^{N/2-1} \hat{q}_k^n e^{ikx_j}.$$

The Fourier coefficients can be obtained by the transform

$$\hat{q}_k^n = \frac{1}{N} \sum_{j=0}^{N-1} q_j^n e^{-ikx_j}.$$

By using the scheme $\mathcal{N}$ in (4) we can derive an expression for $\hat{q}_k^{n+1}$ in terms if $\hat{q}_k^n$ as follows.

$$\begin{aligned}
\hat{q}_k^{n+1} &= \frac{1}{N} \sum_{j=0}^{N-1} q_j^{n+1} e^{-ikx_j} \\
&= \frac{1}{N} \sum_{\ell=-m}^{M} \sum_{j=0}^{N-1} b_\ell q_{j+\ell}^n e^{-ikx_j} = \{\text{by periodicity}\} \\
&= \frac{1}{N} \sum_{\ell=-m}^{M} \sum_{j=0}^{N-1} b_\ell q_j^n e^{-ikx_{j-\ell}} = \{x_{j-\ell} = x_j - \ell\Delta x\} \\
&= \frac{1}{N} \sum_{\ell=-m}^{M} \sum_{j=0}^{N-1} b_\ell e^{ik\ell\Delta x} q_j^n e^{-ikx_j} \\
&= \hat{q}_k^n \sum_{\ell=-m}^{M} b_\ell e^{ik\ell\Delta x}.
\end{aligned}$$

We can thus write

$$\hat{q}_k^{n+1} = g_k(\Delta t, \Delta x)\hat{q}_k^n, \qquad g_k(\Delta t, \Delta x) = \sum_{\ell=-m}^{M} b_\ell(\Delta t, \Delta x) e^{ik\ell\Delta x}.$$

The factor g_k is called the amplification factor since it shows how the different frequencies in the solution are amplified in each time steps.

Example 4 *The amplification factor for upwind is*

$$g_k(\Delta t, \Delta x) = b_0 + b_{-1} e^{-i\ell\Delta x} = 1 - a\frac{\Delta t}{\Delta x} + a\frac{\Delta t}{\Delta x} e^{-ik\Delta x}.$$

We now use Parseval's theorem, which says that

$$||\boldsymbol{q}^n||_{2,\Delta x}^2 = \sum_{j=0}^{N-1} |q_j^n|^2 \Delta x = \sum_{k=-N/2}^{N/2-1} |\hat{q}_k^n|^2.$$

We get

$$\begin{aligned} ||\mathcal{N}(\boldsymbol{q}^n)||_{2,\Delta x}^2 = ||\boldsymbol{q}^{n+1}||_{2,\Delta x}^2 &= \sum_{k=-N/2}^{N/2-1} |\hat{q}_k^{n+1}|^2 = \sum_{k=-N/2}^{N/2-1} |g_k \hat{q}_k^n|^2 \\ &\leq \max_{-N/2\leq k\leq N/2-1} |g_k|^2 \sum_{k=-N/2}^{N/2-1} |\hat{q}_k^n|^2 = \max_{-N/2\leq k\leq N/2-1} |g_k|^2 ||\boldsymbol{q}^n||_{2,\Delta x}^2. \end{aligned}$$

Hence,

$$||\mathcal{N}(\boldsymbol{q}^n)||_{2,\Delta x} \leq \max_{-N/2\leq k\leq N/2-1} |g_k| ||\boldsymbol{q}^n||_{2,\Delta x},$$

and we see that a sufficient condition for stability is

$$\max_{-N/2\leq k\leq N/2-1} |g_k| \leq 1 + \alpha \Delta t. \tag{5}$$

In most cases when the exact solution does not grow exponentially we can actually show the stronger version

$$\max_{-N/2\leq k\leq N/2-1} |g_k| \leq 1.$$

Example 5 *When the CFL condition $|a|\Delta t/\Delta x \leq 1$ holds for the upwind scheme and $a > 0$ we have for the amplification factor*

$$|g_k(\Delta t, \Delta x)| \leq \left|1 - a\frac{\Delta t}{\Delta x}\right| + \left|a\frac{\Delta t}{\Delta x}\right| = 1 - a\frac{\Delta t}{\Delta x} + a\frac{\Delta t}{\Delta x} = 1.$$

The CFL condition is hence both necessary and sufficient for the upwind scheme.

Example 6 *Consider forward Euler + central differences for the heat equation. This can be written as*

$$u_j^{n+1} = u_j^n + \mu(u_{j+1}^n - 2u_j^n + u_{j-1}^n), \qquad \mu = \frac{\Delta t}{(\Delta x)^2}.$$

Here $m = M = 1$ and

$$b_{-1} = \mu, \quad b_0 = 1 - 2\mu, \quad b_1 = \mu.$$

Then

$$g_k(\Delta t, \Delta x) = \mu e^{-ik\Delta x} + 1 - 2\mu + \mu e^{ik\Delta x} = 1 + 2\mu(\cos(k\Delta x) - 1) = 1 - 4\mu \sin^2\left(\frac{k\Delta x}{2}\right).$$

Since

$$\max_{|k|\leq N/2} \sin^2\left(\frac{k\Delta x}{2}\right) = \max_{|k|\leq N/2} \sin^2\left(\frac{k\pi}{N}\right) = \sin^2\left(\frac{\pi}{2}\right) = 1,$$

g_k takes values in the interval $[1 - 4\mu, 1]$ and the method is stable if $\mu \leq 1/2$.

Example 7 *Consider forward Euler + central differences for the advection equation. This can be written as*

$$u_j^{n+1} = u_j^n + \frac{1}{2}\lambda(u_{j+1}^n - u_{j-1}^n), \qquad \lambda = \frac{\Delta t}{\Delta x}.$$

Again, here $m = M = 1$ but

$$b_{-1} = -\frac{\lambda}{2}, \quad b_0 = 1, \quad b_1 = \frac{\lambda}{2}.$$

Then

$$g_k(\Delta t, \Delta x) = -\frac{\lambda}{2}e^{-ik\Delta x} + 1 + \frac{\lambda}{2}e^{ik\Delta x} = 1 + \lambda i \sin(k\Delta x).$$

Hence,

$$\max_k |g_k(\Delta t, \Delta x)| = \max_k \sqrt{1 + \lambda^2 \sin(k\Delta x)^2} > 1,$$

and the method is unstable for all fixed λ. However, if one uses the "parabolic" CFL condition $\Delta t \sim \Delta x^2$ then $\lambda \sim \Delta x \sim \sqrt{\Delta t}$ and

$$\sqrt{1 + \lambda^2 \sin(k\Delta x)^2} \sim \sqrt{1 + \Delta t \sin(k\Delta x)^2} \leq 1 + \alpha \Delta t,$$

for some α and small enough Δt. This choice makes the otherwise unstable method stable.

2.2.2 No boundaries

The case of no boundaries is simlar to the periodic case, but instead of the discrete Fourier transform we now use Fourier series. The infinite sequence $\boldsymbol{q}^n$ defines a 2π-periodic funtion $\hat{q}^n(\xi) \in L^2([0, 2\pi])$ via

$$\hat{q}^n(\xi) = \sum_{j=-\infty}^{\infty} q_j^n e^{ij\xi}, \qquad q_j^n = \frac{1}{2\pi}\int_0^{2\pi} \hat{q}^n(\xi) e^{-ij\xi} d\xi.$$

Like before we derive an expression for $\hat{q}^{n+1}(\xi)$ in terms of $\hat{q}^n(\xi)$:

$$\begin{aligned}
\hat{q}^{n+1}(\xi) &= \sum_{j=-\infty}^{\infty} q_j^{n+1} e^{ij\xi} \\
&= \sum_{\ell=-m}^{M} \sum_{j=-\infty}^{\infty} b_\ell q_{j+\ell}^n e^{ij\xi} \\
&= \sum_{\ell=-m}^{M} \sum_{j=-\infty}^{\infty} b_\ell q_j^n e^{i(j-\ell)\xi} \\
&= \sum_{\ell=-m}^{M} \sum_{j=-\infty}^{\infty} b_\ell e^{-i\ell\xi} q_j^n e^{ij\xi} \\
&= \hat{q}^n(\xi) \sum_{\ell=-m}^{M} b_\ell e^{-i\ell\xi}.
\end{aligned}$$

We thus have

$$\hat{q}^{n+1}(\xi) = \bar{g}(\xi, \Delta t, \Delta x)\hat{q}^n(\xi), \qquad \bar{g}(\xi, \Delta t, \Delta x) = \sum_{\ell=-m}^{M} b_\ell(\Delta t, \Delta x) e^{-i\ell\xi}.$$

For this setting Parseval's theorem says

$$\frac{1}{2\pi}\int_0^{2\pi} |\hat{q}^n(\xi)|^2 dx = \sum_{j=-\infty}^{\infty} |q_j^n|^2 = \frac{1}{\Delta x}||\boldsymbol{q}^n||_{2,\Delta x}^2$$

Therefore,

$$\begin{aligned}||\mathcal{N}(\boldsymbol{q}^n)||_{2,\Delta x}^2 = ||\boldsymbol{q}^{n+1}||_{2,\Delta x}^2 &= \frac{\Delta x}{2\pi}\int_0^{2\pi} |\hat{q}^{n+1}(\xi)|^2 dx = \frac{\Delta x}{2\pi}\int_0^{2\pi} |\bar{g}(\xi)\hat{q}^n(\xi)|^2 dx \\ &\leq \sup_{\xi\in[0,2\pi]} |\bar{g}(\xi)|^2 \frac{\Delta x}{2\pi}\int_0^{2\pi} |\hat{q}^n(\xi)|^2 dx = \sup_{\xi\in[0,2\pi]} |\bar{g}(\xi)|^2 ||\boldsymbol{q}^n||_{2,\Delta x}^2.\end{aligned}$$

Hence,

$$||\mathcal{N}(\boldsymbol{q}^n)||_{2,\Delta x} \leq \sup_{\xi\in[0,2\pi]} |\bar{g}(\xi)| ||\boldsymbol{q}^n||_{2,\Delta x},$$

and we see that a sufficient condition for stability is

$$\sup_{\xi\in[0,2\pi]} |\bar{g}(\xi)| \leq 1 + \alpha\Delta t. \tag{6}$$

We note here the relationship

$$g_k(\Delta t, \Delta x) = \bar{g}(-k\Delta x, \Delta t, \Delta x).$$

Since $\bar{g}$ is 2π-periodic in ξ this shows that as $\Delta x \to 0$ the two stability conditions (5) and (6) are identical.

Remark 3 *For problems with variable coefficients, one can apply von Neumann analysis to the scheme for a* fixed *value of the coefficient. Stability for each such* frozen coefficient *problem is a necessary condition for stability of the whole scheme. It is often also sufficient. For instance, the von Neumann analysis above shows that for the upwind scheme we should have* $a\Delta t/\Delta x \leq 1$. *In the variable coefficient case we would then require* $a(x)\Delta t/\Delta x \leq 1$