The advection equation

June 13, 2020

1 Introduction

We consider the *advection equation*

$$\frac{\partial u}{\partial t}(x,t)+c\frac{\partial u}{\partial x}(x,t)=0,\ (x,t)\in\mathbb{R}\times[0,+\infty),$$

where $c\in\mathbb{R}\setminus\{0\}$, with the initial condition

$$u(x,0)=u_0(x),\ x\in\mathbb{R},$$

where $u_0:\mathbb{R}\to\mathbb{R}$ is a sufficientely smooth function with bounded derivatives on the whole $\mathbb{R}$.

The *characteristics* of the advection equation, i.e. the curves $x=\varphi(t)$ along which the solution is constant, are straight-lines

$$x=\varphi(t)=ct+q,\ t\geq 0,$$

where $q\in\mathbb{R}$. In fact, given a curve

$$x=\varphi(t),\ t\geq 0,$$

we have

$$\frac{d}{dt}u(\varphi(t),t)=\frac{\partial u}{\partial x}(\varphi(t),t)\,\varphi'(t)+\frac{\partial u}{\partial t}(\varphi(t),t)=\frac{\partial u}{\partial x}(\varphi(t),t)\,(\varphi'(t)-c)$$

and then $u(\varphi(t),t)$ is a constant function of $t\geq 0$ if

$$\varphi'(t)=c,\ t\geq 0.$$

So, for any $(x_1,t_1)\in\mathbb{R}\times[0,+\infty)$, the characteristic passing through (x_1,t_1) has

$$q=x_1-ct_1$$

and then we have

$$u(x_1,t_1)=u(\varphi(t_1),t_1)=u(\varphi(0),0)=u(q,0)=u_0(q)=u_0(x_1-ct_1).$$

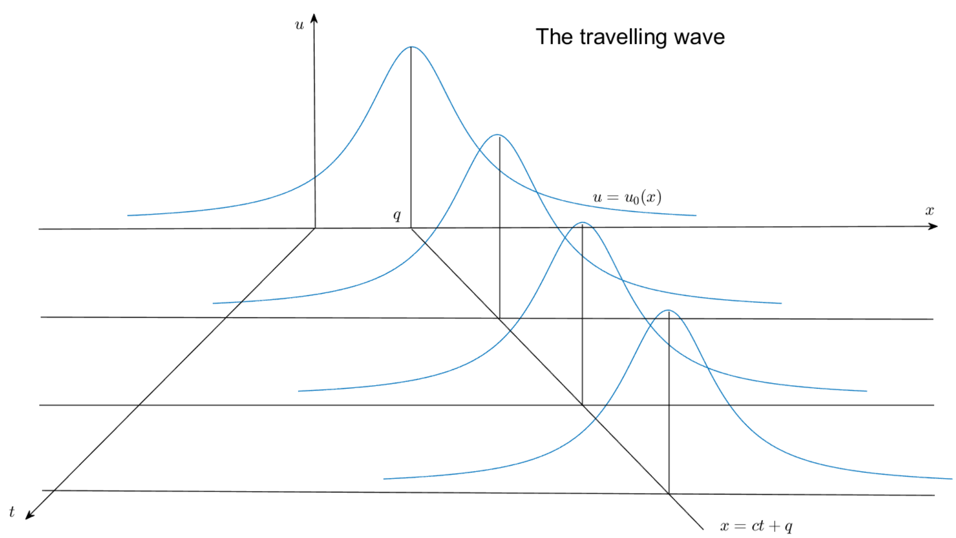
This shows that the solution is

$$u\left(x,t\right)=u_0\left(x-ct\right),\ \left(x,t\right)\in\mathbb{R}\times\left[0,+\infty\right).$$

This solution is called a *travelling wave*. Infact, the initial "profile" u_0 is travelling along the x-axis with speed c: for any $q\in\mathbb{R}$, the value

$$u_0(q)=u(q,0)=u(x,t),\ x=q+ct,$$

starting at time 0 from the space point q moves at the time t at the space point $x=q+ct$. The wave is travelling to the right if $c>0$ and to the left if $c<0$.



2 Numerical methods

Now, we introduce some simple numerical methods for the advection equation. We use a forward difference to discretize the time derivative and three different possibilities to discretize the space derivative: forward difference, backward difference and centered difference.

Since we are using a forward difference in time, all these three methods are explicit time integrations.

Let $h>0$ be the spatial stepsize and let $k>0$ be the time stepsize.

For $n\in\mathbb{Z}$ and $m\in\{0,1,2,\ldots\}$, let U_n^m and u_n^m be the discrete and continuous solutions, respectively, at (nh,mk).

The three methods are

$$\text{forward/forward:}\;\; \frac{U_n^{m+1}-U_n^m}{k}+c\frac{U_{n+1}^m-U_n^m}{h}=0$$

$$\text{backward/forward:}\;\; \frac{U_n^{m+1}-U_n^m}{k}+c\frac{U_n^m-U_{n-1}^m}{h}=0$$

$$\text{centered/forward:}\;\; \frac{U_n^{m+1}-U_n^m}{k}+c\frac{U_{n+1}^m-U_{n-1}^m}{2h}=0$$

$$n\in\mathbb{Z} \text{ and } m\in\{0,1,2,\ldots\}.$$

Equivalently, we can write:

$$\text{forward/forward:}\;\; U_n^{m+1}=(1+\lambda)\,U_n^m-\lambda U_{n+1}^m$$

$$\text{backward/forward:}\;\; U_n^{m+1}=\lambda U_{n-1}^m+(1-\lambda)\,U_n^m$$

$$\text{centered/forward:}\;\; U_n^{m+1}=\frac{\lambda}{2}U_{n-1}^m+U_n^m-\frac{\lambda}{2}U_{n+1}^m$$

$$n\in\mathbb{Z} \text{ and } m\in\{0,1,2,\ldots,\},$$

where

$$\lambda=\frac{ck}{h}.$$

The modulus of λ is known as the *Courant number*.

Exercise. Draw the stencils of forward/forward, backward/forward and centered/forward methods.

The consistency error

$$\varepsilon_n^{m+1},\;\; n\in\mathbb{Z} \text{ and } m\in\{0,1,2,\ldots\},$$

is defined in the usual way: left-hand side of the discrete equation minus the right-hand side (in this case of the advection equation the right-hand side is zero), with the discrete solution replaced by the continuous solution. For example, in case of forward/forward we have

$$\varepsilon_n^{m+1}:=\frac{u_n^{m+1}-u_n^m}{k}+c\frac{u_{n+1}^m-u_n^m}{h},\;\; n\in\mathbb{Z} \text{ and } m\in\{0,1,2,\ldots,\}.$$

Exercise. Prove that for the forward/forward and backward/forward methods, we have

$$\max_{\substack{n\in\mathbb{Z}\\ m\in\{1,2,3,\ldots\}}}|\varepsilon_n^m|\le Ch+Dk$$

for some constants $C, D \geq 0$ independent of h and k. Moreover, prove that for the centered/forward method we have

$$\max_{\substack{n\in\mathbb{Z} \\ m\in\{1,2,3,\ldots\}}} |\varepsilon_n^m| \leq Ch^2 + Dk$$

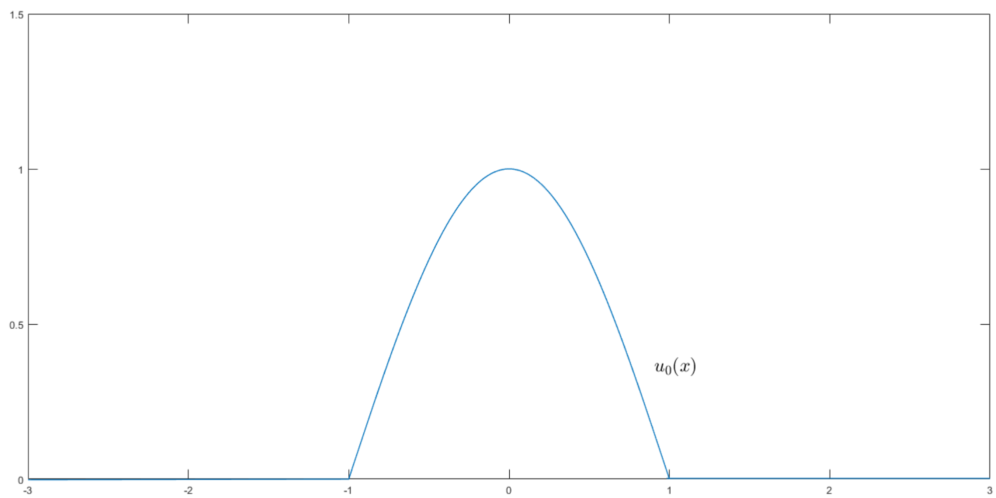
for some constants $C, D \geq 0$ independent of h and k.

2.1 A numerical test

As a numerical test, consider

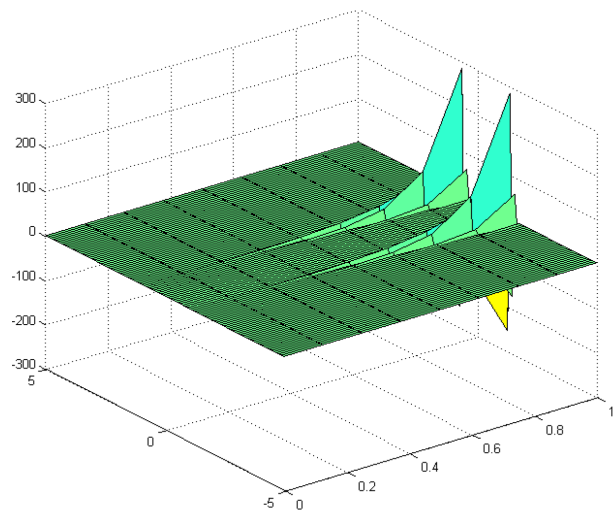
$$c = 1 \text{ and } u_0(x) = \begin{cases} \sin\left((1+x)\frac{\pi}{2}\right) \text{ if } |x| \leq 1 \\ 0 \text{ if } |x| > 1 \end{cases}, \ x \in \mathbb{R}.$$

The discretization is with $h = \frac{1}{10}$ and $k = \frac{1}{8}, \frac{1}{25}$.

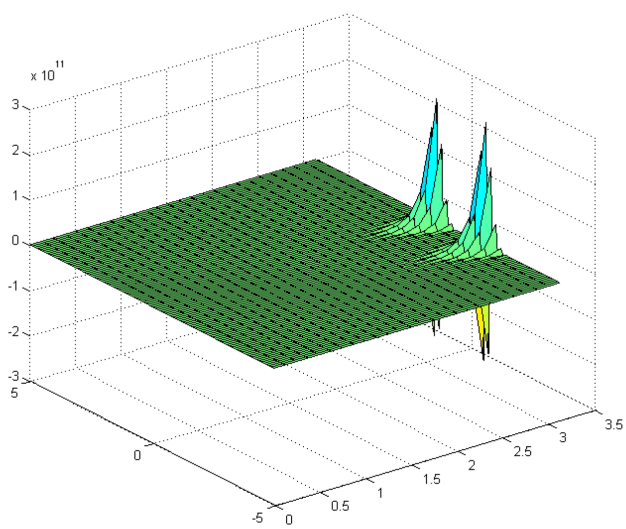


The solution is this initial "profile" moving in the positive direction of the x-axis.

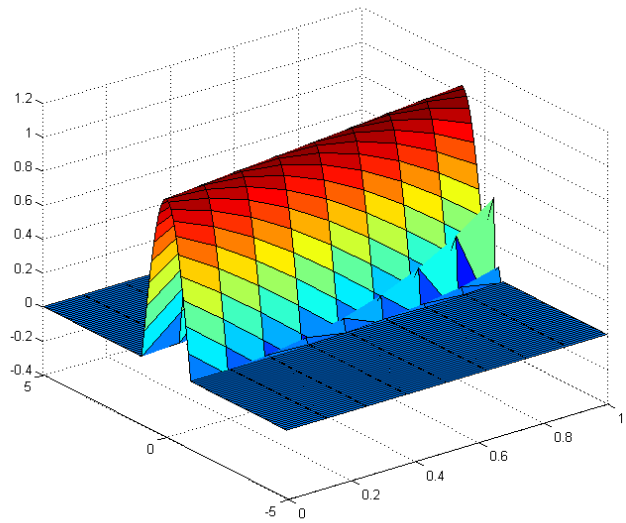
Forward/forward with $h = \frac{1}{10}$ and $k = \frac{1}{8}$ ($\lambda = 1.25$).



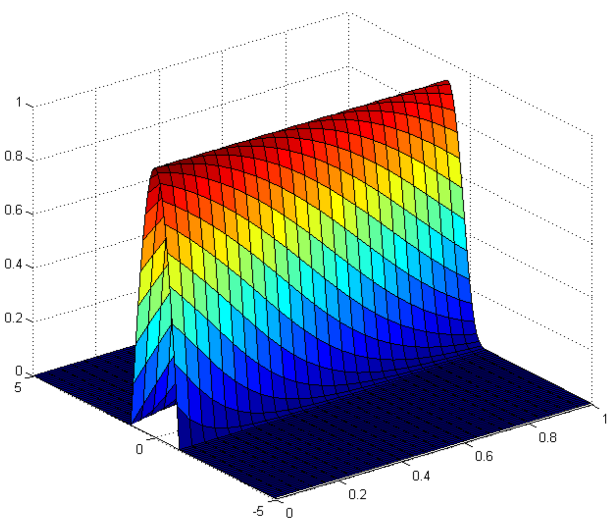
Forward/forward with $h = \frac{1}{10}$ and $k = \frac{1}{25}$ ($\lambda = 0.4$).



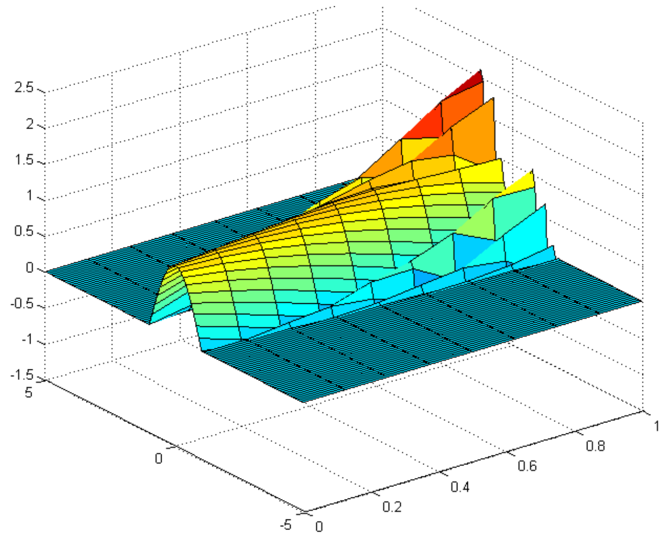
Backward/forward with $h = \frac{1}{10}$ and $k = \frac{1}{8}$ ($\lambda = 1.25$).



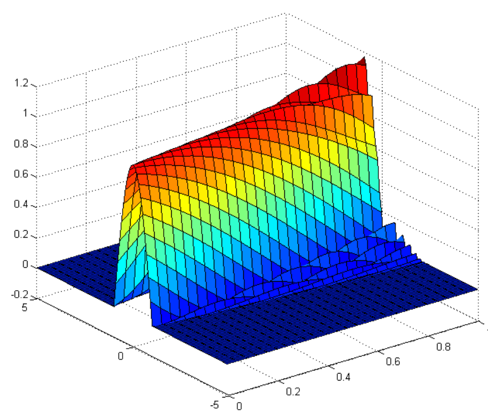
Backward/forward with $h = \frac{1}{10}$ and $k = \frac{1}{25}$ ($\lambda = 0.4$).



Centered/forward with $h = \frac{1}{10}$ and $k = \frac{1}{8}$ ($\lambda = 1.25$).



Centered/forward with $h = \frac{1}{10}$ and $k = \frac{1}{25}$ ($\lambda = 0.4$).



Numerical experiments on this equation by varying h and k show that:

- when $h, k \to 0$, the backward/forward method is convergent if and only if $\lambda = \frac{ck}{h} = \frac{k}{h} \leq 1$, i.e. the backward/forward method is conditionally convergent;
- when $h, k \to 0$, the forward/forward method and the centered/forward method are always not convergent.

Exercise. The function u_0 in this numerical test is not a smooth function. Explain why what we have stated up to now about the advection equation (in particular the order of convergence to zero of the consistency error) is still valid for such a function u_0.

3 The CFL condition

The CFL condition is used for establishing the non-convergence of a numerical method for the advection equation. It is based on the notions of domain of dependance of the continuous and discrete solutions.

The *domain of dependence of the continuous solution* at $(x,t) \in \mathbb{R} \times [0,+\infty)$ is the subset D of $\mathbb{R}$ such that $u(x,t)$ depends only on the values of u_0 on D. Since

$$u(x,t) = u_0(x - ct),$$

we have

$$D = \{x - ct\}.$$

Analogously, the *domain of dependence of the discrete (numerical) solution* at $(x,t) \in \mathbb{R} \times [0,+\infty)$ is the subset $D_{h,k}$ of $\mathbb{R}$ such that U_n^m, where $n = \frac{x}{h}$ and $m = \frac{t}{k}$, depends only on the values of u_0 on $D_{h,k}$.

A necessary condition for the convergence as $h, k \to 0$ of a numerical method for the advection equation is the so-called the *Courant-Friedrichs-Lewy (CFL) condition*:

- *for any $(x,t) \in \mathbb{R} \times [0,+\infty)$, the domain of dependence $x - ct$ of the continuous solution is contained in any closed interval I of $\mathbb{R}$ including the domain of dependence $D_{h,k}$ of the discrete solution, for all h and k sufficiently small.*

Equivalently, a sufficient condition for the non-convergence is the negation of the CFL condition:

- *there exist $(x,t) \in \mathbb{R} \times [0,+\infty)$ and a closed interval I of $\mathbb{R}$ such that $D_{h,k} \subseteq I$, for all h and k sufficiently small, and $x - ct \notin I$.*

In fact, suppose this last condition holds and consider a function u_0 such that

$$u_0(x - ct) = 1 \text{ and } u_0(y) = 0 \text{ for } y \in I.$$

We have

$$u(x,t) = u_0(x - ct) = 1$$

and

$$U_n^m = 0 \quad \text{for all } h \text{ and } k \text{ sufficiently small,}$$

since U_n^m depends linearly on the values of u_0 on $D_{h,k}$ (this linear dependence will be evident below, when we rewrite the numerical methods in a more suitable manner). Thus, at the point (x,t), the numerical solution does not converge to the continuous solution as $h, k \to 0$.

4 The CFL condition for the three methods

We see what is the domain of dependence of the numerical solution and whether or not the CFL condition holds, for the three methods we have introduced for the advection equation. First, we rewrite the methods in a more suitable manner.

Let

$$\mathbb{Z}h := \{nh : n \in \mathbb{Z}\}$$

and let $L(\mathbb{Z}h)$ be the space of the bounded functions $\mathbb{Z}h \to \mathbb{R}$.

Let $U \in L(\mathbb{Z}h)$. We denote the value $U(x)$ of U at $x = nh$, $n \in \mathbb{Z}$, by U_n. So

$$U = (U_n)_{n\in\mathbb{Z}}\,.$$

Let us introduce the *shift operator* $\tau_h : L(\mathbb{Z}h) \to L(\mathbb{Z}h)$ given by

$$(\tau_h U)(x) = U(x+h),\ x \in \mathbb{Z}h \ \text{ and } \ U \in L(\mathbb{Z}h)\,,$$

i.e.

$$(\tau_h U)_n = U_{n+1},\ n \in \mathbb{Z} \ \text{ and } \ U \in L(\mathbb{Z}h)\,.$$

Exercise. Prove that τ_h is a linear operator. Moreover, show that τ_h is invertible and describe τ_h^{-1}.

Let

$$U^m := (U_n^m)_{n\in\mathbb{Z}}\,,\ m \in \{0,1,2,\ldots\}.$$

By

$$\text{forward/forward: } U_n^{m+1} = (1+\lambda)\,U_n^m - \lambda U_{n+1}^m$$

$$\text{backward/forward: } U_n^{m+1} = \lambda U_{n-1}^m + (1-\lambda)\,U_n^m$$

$$\text{centered/forward: } U_n^{m+1} = \frac{\lambda}{2} U_{n-1}^m + U_n^m - \frac{\lambda}{2} U_{n+1}^m$$

$$n \in \mathbb{Z} \text{ and } m \in \{0,1,2,\ldots,\}\,,$$

we obtain

$$\text{forward/forward: } U^{m+1} = ((1+\lambda)\,I - \lambda\tau_h)\,U^m$$

$$\text{backward/forward: } U^{m+1} = ((1-\lambda)\,I + \lambda\tau_h^{-1})U^m$$

$$\text{centered/forward: } U^{m+1} = \left(I + \frac{\lambda}{2}\tau_h^{-1} - \frac{\lambda}{2}\tau_h\right)U^m$$

$$m \in \{0,1,2,\ldots\}.$$

Observe that

$$U^m \in L(\mathbb{Z}h),\ m \in \{0, 1, 2, \ldots\},$$

since

$$U^0 = (u_0\,(nh))_{n\in\mathbb{Z}} \in L(\mathbb{Z}h).$$

Moreover, we also see that U^m depends linearly on U^0 and then the discrete solution U_n^m at (x,t), where $n = \frac{x}{h}$ and $m = \frac{t}{k}$, depends linearly on the values of u_0 on the domain of dependance $D_{h,k}$, as previously announced.

Now, we assume *advection-to-the-right*, i.e. $c > 0$, i.e. $\lambda > 0$.

4.1 The CFL condition for the forward/forward method

As for the forward/forward method, we have, since τ_h and I commute,

$$\begin{aligned} U^m &= ((1+\lambda)\,I - \lambda\tau_h)^m\,U^0 \\ &= \sum_{i=0}^{m} \binom{m}{i} (1+\lambda)^{m-i}\,(-1)^i\,\tau_h^i U^0 \\ &\quad m \in \{0, 1, 2, \ldots\}\,. \end{aligned}$$

Now, we determine the domain of dependence of the discrete solution at $(x,t) \in \mathbb{R} \times [0, +\infty)$ with $t > 0$. The numerical solution

$$\begin{aligned} U_n^m &= \sum_{i=0}^{m} \binom{m}{i} (1+\lambda)^{m-i}\,(-1)^i \left(\tau_h^i U^0\right)_n \\ &= \sum_{i=0}^{m} \binom{m}{i} (1+\lambda)^{m-i}\,(-1)^i\,U_{n+i}^0 \end{aligned}$$

at (x,t), where $n = \frac{x}{h}$ and $m = \frac{t}{k}$, depends on

$$U_n^0 = u_0(nh), \ldots, U_{n+m}^0 = u_0((n+m)h).$$

Then, the domain of dependence of the numerical solution at (x,t) is

$$D_{h,k} = \{nh, \ldots, (n+m)h\}.$$

Now, we look at the CFL condition. We have

$$D_{h,k} \subseteq [nh, nh+mh] = \left[nh, nh + \frac{cmk}{\lambda}\right] = \left[x, x + \frac{ct}{\lambda}\right].$$

Thus, since $\lambda > 0$, we have

$$D_{h,k} \subseteq \left[x, x + \frac{ct}{\lambda}\right] \subseteq I = [x, +\infty)$$

and $x - ct \notin I$. The CFL condition does not hold.

Since the CFL condition does not hold, the forward/forward method is not convergent.

4.2 The CFL condition for the backward/forward method

As for the backward/forward method, we have, since τ_h^{-1} and I commute,

$$\begin{aligned} U^m &= ((1-\lambda)\,I + \lambda\tau_h^{-1})^m U^0 \\ &= \sum_{i=0}^{m} \binom{m}{i} (1-\lambda)^{m-i} (-1)^i (\tau_h^{-1})^i U^0 \\ & \quad m \in \{0,1,2,\dots\}. \end{aligned}$$

Now, we determine the domain of dependence of the discrete solution at $(x,t) \in \mathbb{R} \times [0,+\infty)$ with $t > 0$. The numerical solution

$$\begin{aligned} U_n^m &= \sum_{i=0}^{m} \binom{m}{i} (1-\lambda)^{m-i} (-1)^i \left((\tau_h^{-1})^i U^0\right)_n \\ &= \sum_{i=0}^{m} \binom{m}{i} (1-\lambda)^{m-i} (-1)^i\, U_{n-i}^0 \end{aligned}$$

at (x,t), where $n = \frac{x}{h}$ and $m = \frac{t}{k}$, depends on

$$U_n^0 = u_0(nh), \dots, U_{n-m}^0 = u_0((n-m)h).$$

Then, the domain of dependence of the numerical solution at (x,t) is

$$D_{h,k} = \{(n-m)h, \dots, nh\}.$$

Now, we look at the CFL condition. We have

$$D_{h,k} \subseteq [nh - mh, nh] = \left[nh - \frac{cmk}{\lambda}, nh\right] = \left[x - \frac{ct}{\lambda}, x\right].$$

Fix $\beta > 1$. If $\lambda \geq \beta$, we have

$$D_{h,k} \subseteq \left[x - \frac{ct}{\lambda}, x\right] \subseteq I = \left[x - \frac{ct}{\beta}, x\right].$$

and $x - ct \notin I$, since

$$x - ct < x - \frac{ct}{\beta}.$$

So, if $h,k \to 0$ with $\lambda \geq \beta$, for some $\beta > 1$, the CFL condition does not hold and then the backward/forward method is not convergent.

On the other hand, if $h,k \to 0$ with $\lambda \leq 1$, the CFL condition holds. In fact,

$$nh - mh = x - \frac{ct}{\lambda} \leq x - ct \leq x = nh$$

and so any closed interval including $D_{h,k} = \{(n-m)h, \dots, nh\}$ contains $x - ct$. We cannot conclude that the backward/forward method is convergent when $\lambda \leq 1$, since the CFL condition is a necessary condition for the convergence.

Indeed, as now we prove, the forward/backward method is convergent when $\lambda \leq 1$.

Suppose we want the integrate in time the advection equation on the interval $[0, T]$. Let M be a positive integer and let $k = \frac{T}{M}$.
We have

$$\begin{aligned}
&\frac{U_n^{m+1} - U_n^m}{k} + c\frac{U_n^m - U_{n-1}^m}{h} = 0 \\
&\frac{u_n^{m+1} - u_n^m}{k} + c\frac{u_n^m - u_{n-1}^m}{h} - \varepsilon_n^{m+1} = 0 \\
&n \in \mathbb{Z} \text{ and } m \in \{0, \dots, M-1\}
\end{aligned}$$

where ε_n^{m+1} is the consistency error. Then, by introducing the convergence error

$$e_n^m = U_n^m - u_n^m,\ m \in \mathbb{Z} \text{ and } m \in \{0, \dots, M\}$$

we obtain

$$\begin{aligned}
&\frac{e_n^{m+1} - e_n^m}{k} + c\frac{e_n^m - e_{n-1}^m}{h} + \varepsilon_n^{m+1} = 0, \\
&n \in \mathbb{Z} \text{ and } m \in \{0, \dots, M-1\}
\end{aligned}$$

i.e.

$$\begin{aligned}
&e_n^{m+1} = \lambda e_{n-1}^m + (1-\lambda)\, e_n^m - k\varepsilon_n^{m+1}, \\
&n \in \mathbb{Z} \text{ and } m \in \{0, \dots, M-1\}\,.
\end{aligned}$$

By using

$$e^m = \left(e_n^m\right)_{n \in \mathbb{Z}},\ m \in \{0, \dots, M\}\,,$$

and

$$\varepsilon^{m+1} = \left(\varepsilon_n^{m+1}\right)_{n \in \mathbb{Z}},\ m \in \{0, \dots, M-1\}\,,$$

we can write

$$e^{m+1} = ((1-\lambda)I + \lambda\tau_h^{-1})e^m + k(-\varepsilon^{m+1}),\ m \in \{0, \dots, M-1\}\,. \tag{1}$$

Next one is a stability result for the backward/forward method.

Theorem 1 *If $\lambda \leq 1$, then*

$$\begin{aligned}
&\max_{\substack{n \in \mathbb{Z} \\ m \in \{0, \dots, M\}}} |e_n^m| = \max_{m \in \{0, \dots, M\}} \|e^m\|_{L^\infty(\mathbb{Z}h)} \\
&\leq T \max_{i \in \{1, \dots, M\}} \|\varepsilon^i\|_{L^\infty(\mathbb{Z}h)} = T \max_{\substack{n \in \mathbb{Z} \\ i \in \{1, \dots, M\}}} |\varepsilon_n^i|.
\end{aligned}$$

Proof. By (1), we obtain

$$\|e^{m+1}\|_{L^\infty(\mathbb{Z}h)} \le \|\left(1-\lambda\right)I + \lambda\tau_h^{-1}\| \cdot \|e^m\|_{L^\infty(\mathbb{Z}h)} + k\|\varepsilon^{m+1}\|_{L^\infty(\mathbb{Z}h)}$$
$$m \in \{0,\dots,M-1\},$$

where $\left\|\left(1-\lambda\right)I + \lambda\tau_h^{-1}\right\|$ is the operator norm relevant to the norm L^∞ on $L\left(\mathbb{Z}h\right)$.

Now, we prove that

$$\left\|\left(1-\lambda\right)I + \lambda\tau_h^{-1}\right\| \le 1.$$

For $U \in L\left(\mathbb{Z}h\right)$, we have

$$\left\|\left(\left(1-\lambda\right)I + \lambda\tau_h^{-1}\right)U\right\|_{L^\infty(\mathbb{Z}h)} = \sup_{n\in\mathbb{Z}}\left|\left(\left(\left(1-\lambda\right)I + \lambda\tau_h^{-1}\right)U\right)_n\right|$$

and, for $n \in \mathbb{Z}$,

$$\begin{aligned}
\left|\left(\left(\left(1-\lambda\right)I + \lambda\tau_h^{-1}\right)U\right)_n\right| &= \left|\left(\left(1-\lambda\right)U + \lambda\tau_h^{-1}U\right)_n\right| \\
&= \left|\left(1-\lambda\right)U_n + \lambda U_{n-1}\right| \\
&\le \left(1-\lambda\right)\left|U_n\right| + \lambda\left|U_{n-1}\right| \quad \text{since } \lambda \le 1 \\
&\le \|U\|_{L^\infty(\mathbb{Z}h)},\ n \in \mathbb{Z}.
\end{aligned}$$

So

$$\left\|\left(\left(1-\lambda\right)I + \lambda\tau_h^{-1}\right)U\right\|_{L^\infty(\mathbb{Z}h)} \le \|U\|_{L^\infty(\mathbb{Z}h)}.$$

We conclude that

$$\left\|\left(1-\lambda\right)I + \lambda\tau_h^{-1}\right\| = \sup_{U\in L(\mathbb{Z}h)\setminus\{0\}} \frac{\left\|\left(\left(1-\lambda\right)I + \lambda\tau_h^{-1}\right)U\right\|_{L^\infty(\mathbb{Z}h)}}{\|U\|_{L^\infty(\mathbb{Z}h)}} \le 1.$$

As a consequence we have

$$\begin{aligned}
\|e^{m+1}\|_{L^\infty(\mathbb{Z}h)} &\le \|\left(1-\lambda\right)I + \lambda\tau_h^{-1}\| \cdot \|e^m\|_{L^\infty(\mathbb{Z}h)} + k\|\varepsilon^{m+1}\|_{L^\infty(\mathbb{Z}h)} \\
&\le \|e^m\|_{L^\infty(\mathbb{Z}h)} + k\|\varepsilon^{m+1}\|_{L^\infty(\mathbb{Z}h)},\ m \in \{0,\dots,M-1\},
\end{aligned}$$

and then

$$\|e^m\|_{L^\infty(\mathbb{Z}h)} \le k\sum_{i=1}^{m}\|\varepsilon^i\|_{L^\infty(\mathbb{Z}h)},\ m \in \{0,\dots,M\}.$$

Thus,

$$\max_{m\in\{0,\dots,M\}}\|e^m\|_{L^\infty(\mathbb{Z}h)} \le \underbrace{kM}_{=T}\max_{i\in\{1,\dots,M\}}\|\varepsilon^i\|_{L^\infty(\mathbb{Z}h)}.$$

■

We conclude that the backward/forward method is conditionally stable: it is stable if and only if $\lambda \le 1$. It is conditionally convergent: if $\lambda \le 1$, then

$$\max_{\substack{n\in\mathbb{Z}\\ m\in\{0,\dots,M\}}}|e_n^m| \le T\max_{\substack{n\in\mathbb{Z}\\ i\in\{1,\dots,M\}}}|\varepsilon_n^i| \le T(Ch + Dk).$$

Exercise. Explain why the previous argument for proving stability for the backward/forward method does not work for the other two methods.

4.3 The CFL condition for the centered/forward method

As for the centered/forward method, we have, since $\frac{\lambda}{2}\tau_h^{-1} - \frac{\lambda}{2}\tau_h$ and I commute as well as τ_h^{-1} and τ_h,

$$
\begin{aligned}
U^m &= \left(I + \frac{\lambda}{2}\tau_h^{-1} - \frac{\lambda}{2}\tau_h\right)^m U^0 \\
&= \sum_{i=0}^{m} \binom{m}{i} \left(\frac{\lambda}{2}\tau_h^{-1} - \frac{\lambda}{2}\tau_h\right)^i U^0 \\
&= \sum_{i=0}^{m} \binom{m}{i} \left(\frac{\lambda}{2}\right)^i \left(\tau_h^{-1} - \tau_h\right)^i U^0 \\
&= \sum_{i=0}^{m} \binom{m}{i} \left(\frac{\lambda}{2}\right)^i \sum_{j=0}^{i} \binom{i}{j} (-1)^j \left(\tau_h^{-1}\right)^j \tau_h^{i-j} U^0 \\
&= \sum_{i=0}^{m} \binom{m}{i} \left(\frac{\lambda}{2}\right)^i \sum_{j=0}^{i} \binom{i}{j} (-1)^l \, \tau_h^{i-2j} U^0 \\
&= \sum_{i=0}^{m} \binom{m}{i} \left(\frac{\lambda}{2}\right)^i \sum_{p=-i}^{i} \binom{i}{\frac{i-p}{2}} (-1)^{\frac{i-p}{2}} \tau_h^p U^0 \\
& \quad m \in \{0, 1, 2, \ldots\}.
\end{aligned}
$$

Now, we determine the domain of dependence of the discrete solution at $(x, t) \in \mathbb{R} \times [0, +\infty)$ with $t > 0$. The numerical solution

$$
\begin{aligned}
U_n^m &= \sum_{i=0}^{m} \binom{m}{i} \left(\frac{\lambda}{2}\right)^i \sum_{p=-i}^{i} \binom{i}{\frac{i-p}{2}} (-1)^{\frac{i-p}{2}} \left(\tau_h^p U^0\right)_n \\
&= \sum_{i=0}^{m} \binom{m}{i} \left(\frac{\lambda}{2}\right)^i \sum_{p=-i}^{i} \binom{i}{\frac{i-p}{2}} (-1)^{\frac{i-p}{2}} U_{n+p}^0
\end{aligned}
$$

at (x, t), where $n = \frac{x}{h}$ and $m = \frac{t}{k}$, depends on

$$U_{n-m}^0 = u_0((n-m)h), \ldots, U_n^0 = u_0(nh), \ldots, U_{n+m}^0 = u_0((n+m)h).$$

Then, the domain of dependence of the numerical solution at (x, t) is

$$D_{h,k} = \{(n-m)h, \ldots, nh, \ldots, (n+m)h\}.$$

Now, we look at the CFL condition. We have

$$D_{h,k} \subseteq [nh - mh, nh + mh] = \left[nh - \frac{cmk}{\lambda}, nh + \frac{cmk}{\lambda}\right] = \left[x - \frac{ct}{\lambda}, x + \frac{ct}{\lambda}\right].$$

Fix $\beta > 1$. If $\lambda \geq \beta$, we have

$$D_{h,k} \subseteq \left[x - \frac{ct}{\lambda}, x + \frac{ct}{\lambda}\right] \subseteq I = \left[x - \frac{ct}{\beta}, x + \frac{ct}{\beta}\right]$$

and $x - ct \notin I$ since

$$x - ct < x - \frac{ct}{\beta}.$$

We have the same situation of the backward/forward method: if $h, k \to 0$ with $\lambda \geq \beta$, for some $\beta > 1$, the CFL condition does not hold and then the centered/forward method is not convergent.

Moreover, as for the backward/forward method, the CFL condition holds if $h, k \to 0$ with $\lambda \leq 1$. However, unlike the the backward/forward method, the centered/forward method is not convergent. This will be proved below by using the Fourier analysis.

The non-convergence of the centered/forward method for $\lambda \leq 1$ shows that the CFL condition is in general not a sufficient condition for convergence.

5 Advection-to-the-left

Up to now, we have assumed advection-to-the-right, i.e. $c > 0$, i.e. $\lambda > 0$. Now, we assume *advection-to-the-left*, i.e. $c < 0$, i.e. $\lambda < 0$.

Associated to our advection-to-the-left equation

$$\frac{\partial v}{\partial t}(x,t) + c\frac{\partial v}{\partial x}(x,t) = 0,\ (x,t) \in \mathbb{R} \times [0, +\infty),$$

with the initial condition

$$v(x,0) = u_0(x),\ x \in \mathbb{R},$$

we consider the advection-to-the-right equation

$$\frac{\partial v}{\partial t}(x,t) + (-c)\frac{\partial v}{\partial x}(x,t) = 0,\ (x,t) \in \mathbb{R} \times [0, +\infty),$$

with the initial condition

$$v(x,0) = u_0(-x),\ x \in \mathbb{R}.$$

The solution of this associated advection-to-the-right equation is

$$v(x,t) = u(-x,t),\ (x,t) \in \mathbb{R} \times [0, +\infty),$$

In fact, for $(x,t) \in \mathbb{R} \times [0, +\infty)$,

$$v(x,t) = u_0(-(x - (-c)t)) = u_0(-x - ct) = u(-x,t).$$

Consider numerical solutions, for the same stepsizes h and k, for both advection equations. Let λ_r be the parameter λ for the associated advection-to-the-right equation. We have $\lambda_r = -\lambda$, where $\lambda = \frac{ck}{h} < 0$ is the parameter λ for our advection-to-the-left equation.

Let
$$V_n^m,\ n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\}.$$
be the numerical solution of the associated advection-to-the-right equation.

For the three methods, we have, with $\lambda = \frac{ck}{h} < 0$,

$$\begin{aligned}
&\text{forward/forward: } V^{m+1} = ((1+\lambda_r)\, I - \lambda_r \tau_h) V^m = ((1-\lambda)\, I + \lambda \tau_h) V^m \\
&\text{backward/forward: } V^{m+1} = \left((1-\lambda_r)\, I + \lambda_r \tau_h^{-1}\right) V^m = \left((1+\lambda)\, I - \lambda \tau_h^{-1}\right) V^m \\
&\text{centered/forward: } V^{m+1} = \left(I + \frac{\lambda_r}{2}\tau_h^{-1} - \frac{\lambda_r}{2}\tau_h\right) V^m = \left(I - \frac{\lambda}{2}\tau_h^{-1} + \frac{\lambda}{2}\tau_h\right) V^m \\
&m \in \{0, 1, 2, \ldots\}.
\end{aligned} \tag{2}$$

We introduce the operator $A : L(\mathbb{Z}h) \to L(\mathbb{Z}h)$ given by
$$(AU)_n = U_{-n},\ n \in \mathbb{Z} \ \text{ and } \ U \in L(\mathbb{Z}h).$$

Exercise. Prove that A is a linear operator and $A\tau_h = \tau_h^{-1} A$ and $A\tau_h^{-1} = \tau_h A$.

Let
$$U_n^m,\ n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\}.$$
be the numerical solution of our advection-to-the-left equation.

By applying A to both sides in (2), we obtain

$$\begin{aligned}
&\text{forward/forward: } AV^{m+1} = ((1-\lambda)\, I + \lambda \tau_h^{-1}) AV^m \\
&\text{backward/forward: } AV^{m+1} = ((1+\lambda)\, I - \lambda \tau_h)\, AV^m \\
&\text{centered/forward: } AV^{m+1} = \left(I - \frac{\lambda}{2}\tau_h + \frac{\lambda}{2}\tau_h^{-1}\right) AV^m. \\
&m \in \{0, 1, 2, \ldots\}
\end{aligned}$$

to be compared with the three methods for our advection-to-the-right equation:

backward/forward: $U^{m+1} = ((1-\lambda)\, I + \lambda \tau_h^{-1}) U^m$

forward/forward: $U^{m+1} = ((1+\lambda)\, I - \lambda \tau_h)\, U^m$

centered/forward: $U^{m+1} = \left(I + \frac{\lambda}{2} \tau_h^{-1} - \frac{\lambda}{2} \tau_h \right) U^m$

$m \in \{0, 1, 2, \ldots\}$.

Since

$$AV^0 = U^0,$$

we conclude that

$$AV^m = U^m,\ m \in \{0, 1, 2, \ldots\},$$

i.e.

$V_n^m = U_{-n}^m,\ n \in \mathbb{Z}$ and $m \in \{0, 1, 2, \ldots\}$,
(the discrete solutions V_n^m and U_n^m are the same
except for the inversion of the space index)

for

- V^m forward/forward and U^m backward/forward
- V^m backward/forward and U^m forward/forward
- V^m centered/forward and U^m centered/forward.

We draw the following conclusions.

- The forward/forward method for the advection-to-the-left equation becomes the backward/forward method for the advection-to-the-right equation.

  Thus, the forward/forward method for the advection-to-the-left equation is conditionally convergent: it is convergent if and only if $\lambda \geq -1$.

- The backward/forward method for the advection-to-the-left equation becomes the forward/forward method for the advection-to-the-right equation.

  Thus, the backward/forward method for the advection-to-the-left equation is not convergent.

- The centered/forward method for the advection-to-the-left equation becomes the centered/forward method for the advection-to-the-right equation.

  Thus, the centered/forward method for the advection-to-the-left equation is not convergent.

Up to now, we have not a numerical method for the advection equation which is convergent for both advection-to-the right ($c > 0$) and advection-to-the-left ($c < 0$). Such a method is now introduced.

6 The Lax-Friedrichs method

The *Lax-Friedrichs method* is a variant of the centered/forward method with the property to be conditionally stable for both advection to the right ($c > 0$) and to the left ($c < 0$).

The method is

$$\frac{U_n^{m+1} - \overbrace{\frac{U_{n+1}^m + U_{n-1}^m}{2}}^{U_n^m \text{ in centered/forward is replaced by}}}{k} + c\frac{U_{n+1}^m - U_{n-1}^m}{2h} = 0$$
$$n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\}$$

i.e.

$$U_n^{m+1} = \frac{1}{2}(1+\lambda) U_{n-1}^m + \frac{1}{2}(1-\lambda) U_{n+1}^m$$
$$n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\}.$$

i.e.

$$U^{m+1} = \left(\frac{1}{2}(1+\lambda)\tau_h^{-1} + \frac{1}{2}(1-\lambda)\tau_h\right) U^m$$
$$m \in \{0, 1, 2, \ldots\}$$

with $\lambda = \frac{ck}{h}$.

Observe that the Lax-Friedrichs method is an explicit time-integration.

Exercise. By writing the Lax-Friedrichs method as

$$\frac{U_n^{m+1} - U_n^m}{k} + c\frac{U_{n+1}^m - U_{n-1}^m}{2h} + \frac{U_n^m - \frac{U_{n+1}^m + U_{n-1}^m}{2}}{k} = 0$$
$$n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\},$$

show that, for the consistency error

$$\varepsilon_n^{m+1}, \; n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\},$$

(defined in the usual manner) we have

$$\max_{\substack{n \in \mathbb{Z} \\ m \in \{1,2,3,\ldots\}}} |\varepsilon_n^m| \leq Ch^2 + Dk + E\frac{h^2}{k}$$

for some constants $C, D, E \geq 0$ independent of h and k.

Exercise. Prove that the Lax-Friedrichs method does not satisfy the CFL condition for $|\lambda| > 1$.

Exercise. Consider the integration of the advection equation on the time interval $[0, T]$ with time stepsize $k = \frac{T}{M}$, M positive integer, by the Lax-Friedrichs method. Let

$$e_n^m = U_n^m - u_n^m, \ n \in \mathbb{Z} \text{ and } m \in \{0, \ldots, M\},$$

be the convergence error. Prove the stability result: if $|\lambda| \leq 1$, then

$$\begin{aligned} \max_{\substack{n\in\mathbb{Z} \\ m\in\{0,\ldots,M\}}} |e_n^m| &= \max_{m\in\{0,\ldots,M\}} \|e^m\|_{L^\infty(\mathbb{Z}h)} \\ &\leq T \max_{i\in\{1,\ldots,M\}} \|\varepsilon^i\|_{L^\infty(\mathbb{Z}h)} = T \max_{\substack{n\in\mathbb{Z} \\ i\in\{1,\ldots,M\}}} |\varepsilon_n^i|. \end{aligned}$$

By the two previous exercises, we conclude that the Lax-Friedrichs method is conditionally stable: it is stable if and only if $|\lambda| \leq 1$. Since the condition is on the modulus of λ, the method is conditionally stable for both advection to the right and to the left.

The method is conditionally convergent: if $|\lambda| \leq 1$, then

$$\begin{aligned} \max_{\substack{n\in\mathbb{Z} \\ m\in\{0,\ldots,M\}}} |e_n^m| &\leq T \max_{\substack{n\in\mathbb{Z} \\ i\in\{1,\ldots,M\}}} |\varepsilon_n^i| \\ &\leq T \left(Ch^2 + Dk + E\frac{h^2}{k} \right) \\ &= T \left(C \left(\frac{c}{\lambda}\right)^2 k^2 + Dk + E \left(\frac{c}{\lambda}\right)^2 k \right) \\ &= T \left(Ch^2 + D\frac{\lambda}{c}h + E\frac{c}{\lambda}h \right) \end{aligned}$$

We have

$$\max_{\substack{n\in\mathbb{Z} \\ m\in\{0,\ldots,M\}}} |e_n^m| = O(k) = O(h), \ M \to \infty,$$

when λ is held fixed.

Note that the Lax-Friedrichs method can be rewritten as

$$\begin{aligned} 0 &= \frac{U_n^{m+1} - U_n^m}{k} + c\frac{U_{n+1}^m - U_{n-1}^m}{2h} + \frac{U_n^m - \frac{U_{n+1}^m + U_{n-1}^m}{2}}{k} \\ &= \frac{U_n^{m+1} - U_n^m}{k} + c\frac{U_{n+1}^m - U_{n-1}^m}{2h} - \frac{h^2}{2k} \cdot \frac{U_{n-1}^m - 2U_n^m + U_{n+1}^m}{h^2} \\ & \quad n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\}. \end{aligned}$$

So the method is a centered/forward discretization of the equation

$$\frac{\partial u}{\partial t}(x,t) + c\frac{\partial u}{\partial x}(x,t) - \frac{h^2}{2k} \cdot \frac{\partial^2 u}{\partial x^2}(x,t) = 0, \ (x,t) \in \mathbb{R} \times [0, +\infty).$$

This is an advection/diffusion equation with a coefficient

$$\frac{h^2}{2k} = \frac{1}{2}\left(\frac{c}{\lambda}\right)^2 k = \frac{1}{2} \cdot \frac{c}{\lambda} h$$

multiplying the diffusion term $\frac{\partial^2 u}{\partial x^2}(x,t)$.

Thus the Lax-Friedrichs method can be viewed as a variant of the centered/forward method in which a small amount of artificial diffusion has been added to stabilize the numerical method. The amount is small because the coefficient $\frac{h^2}{2k}$ is $O(k) = O(h)$, $h, k \to 0$, when λ is held fixed.

7 The Fourier analysis

We can also use the Fourier analysis to study the stability of numerical methods for the advection equation. In this context, it is known as *von Neumann stability analysis*.

For simplicity, we consider initial functions u_0 which are periodic of period 1 and so also the solution u is periodic of period 1.

Exercise. Prove that if u_0 is periodic of period 1, also u is periodic of period 1.

Moreover, we consider complex-valued functions rather than real-valued functions.

Let $h = \frac{1}{N}$, where N is an integer. We introduce the space $L^{\text{per}}(\mathbb{Z}h)$ of the functions $U : \mathbb{Z}h \to \mathbb{C}$ of period 1, i.e. such that

$$U(x+1) = U(x),\ x \in \mathbb{Z}h,$$

or, equivalently,

$$U_{n+N} = U_n,\ n \in \mathbb{Z}.$$

The space $L^{\text{per}}(\mathbb{Z}h)$ is isomorphic to $\mathbb{C}^N$: each function $U \in L^{\text{per}}(\mathbb{Z}h)$ is determined by the values $U_0, \ldots, U_{N-1}$.

On the space $L^{\text{per}}(\mathbb{Z}h)$, we consider the scalar product

$$\langle U, V \rangle_h := h \sum_{n=0}^{N-1} U_n \overline{V_n},\ U, V \in L^{\text{per}}(\mathbb{Z}h),$$

and let $\|\cdot\|_h$ be the norm on $L^{\text{per}}(\mathbb{Z}h)$ derived by this scalar product. This norm is the discrete L^2 norm on $L^{\text{per}}(\mathbb{Z}h)$.

For any $p \in \{0, \ldots, N-1\}$, consider the function $\psi^{(p)} \in L^{\text{per}}(\mathbb{Z}h)$ given by

$$\psi_n^{(p)} = \mathrm{e}^{i2\pi pnh},\ n \in \mathbb{Z}.$$

Exercise. Prove that the function $\psi^{(p)} : \mathbb{Z}h \to \mathbb{C}$ just defined has period 1.

The functions $\psi^{(0)}, \ldots, \psi^{(N-1)}$ are orthonormal in the scalar product $\langle \cdot, \cdot \rangle_h$. In fact, for $p, q \in \{0, \ldots, N-1\}$, we have

$$\begin{aligned}
\left\langle \psi^{(p)}, \psi^{(q)} \right\rangle_h &= h \sum_{n=0}^{N-1} \mathrm{e}^{i2\pi pnh} \mathrm{e}^{-i2\pi qnh} \\
&= h \sum_{n=0}^{N-1} \mathrm{e}^{i2\pi(p-q)nh} \\
&= h \sum_{n=0}^{N-1} \left(\mathrm{e}^{i2\pi(p-q)h} \right)^n \\
&= \begin{cases} h \sum\limits_{n=0}^{N-1} 1 = 1 \text{ if } p = q \\ h \frac{\mathrm{e}^{i2\pi(p-q)Nh}-1}{\mathrm{e}^{i2\pi(p-q)h}-1} = \frac{\mathrm{e}^{i2\pi(p-q)}-1}{\mathrm{e}^{i2\pi(p-q)h}-1} = 0 \text{ if } p \neq q. \end{cases}
\end{aligned}$$

Therefore, the functions $\psi^{(0)}, \ldots, \psi^{(N-1)}$ constitute an orthonormal basis of $L^{\mathrm{per}}(\mathbb{Z}h)$ and then each $U \in L^{\mathrm{per}}(\mathbb{Z}h)$ has a discrete Fourier series

$$U = \sum_{p=0}^{N-1} c_p \psi^{(p)}$$

for unique coefficients $c_0, \ldots, c_{N-1}$.

Now, observe that the shift linear operator τ_h is an operator $L^{\mathrm{per}}(\mathbb{Z}h) \to L^{\mathrm{per}}(\mathbb{Z}h)$. In fact, for $U \in L^{\mathrm{per}}(\mathbb{Z}h)$, we have $\tau_h U \in L^{\mathrm{per}}(\mathbb{Z}h)$ since

$$(\tau_h U)_{n+N} = U_{n+N+1} = U_{n+1} = (\tau_h U)_n, \ n \in \mathbb{Z}.$$

Moreover, for $p \in \{0, \ldots, N-1\}$, we have

$$\left(\tau_h \psi^{(p)} \right)_n = \psi^{(p)}_{n+1} = \mathrm{e}^{i2\pi p(n+1)h} = \mathrm{e}^{i2\pi ph} \mathrm{e}^{i2\pi pnh} = \mathrm{e}^{i2\pi ph} \psi^{(p)}_n, \ n \in \mathbb{Z},$$

and then

$$\tau_h \psi^{(p)} = \mathrm{e}^{i2\pi ph} \psi^{(p)}.$$

So τ_h has the eigenvalues

$$\mu_{p,h} := \mathrm{e}^{i2\pi ph}, \ p \in \{0, \ldots, N-1\},$$

with relevant eigenvectors $\psi^{(p)}$, $p \in \{0, \ldots, N-1\}$.

Now, we can analyze the stability of the numerical methods for the advection equation by using, in the space discretization, the L^2 norm $\|\cdot\|_h$.

Let

$$U_n^m, u_n^m, e_n^m, \varepsilon_n^{m+1},\ n \in \mathbb{Z} \text{ and } m \in \{0,1,2,\ldots\},$$

be the discrete solution, the continuous solution, the convergence error and the consistency error, respectively, for the four methods previously seen for the advection equation. Moreover, let

$$U^m = (U_n^m)_{n\in\mathbb{Z}}, u^m = (u_n^m)_{n\in\mathbb{Z}}, e^m = (e_n^m)_{n\in\mathbb{Z}}, \varepsilon^{m+1} = \left(\varepsilon_n^{m+1}\right)_{n\in\mathbb{Z}}, m \in \{0,1,2,\ldots\}.$$

For the four methods, we have

$$U^{m+1} = R(\tau_h)U^m,\ m \in \{0,1,2\ldots\},$$

and

$$u^{m+1} = R(\tau_h)u^m + k\varepsilon^{m+1},\ m \in \{0,1,2\ldots\},$$

where

$$R(\tau_h) = (1+\lambda) I - \lambda\tau_h \text{ for forward/forward}$$

$$R(\tau_h) = (1-\lambda) I + \lambda\tau_h^{-1} \text{ for backward/forward}$$

$$R(\tau_h) = I + \frac{\lambda}{2}\tau_h^{-1} - \frac{\lambda}{2}\tau_h \text{ for centered/forward}$$

$$R(\tau_h) = \frac{1}{2}(1+\lambda)\tau_h^{-1} + \frac{1}{2}(1-\lambda)\tau_h \text{ for Lax-Friedrichs.}$$

with $\lambda = \frac{ck}{h}$. Thus

$$e^{m+1} = R(\tau_h)e^m + k(-\varepsilon^{m+1}),\ m \in \{0,1,2\ldots\}.$$

Exercise. Prove that

$$U^m, u^m, e^m, \varepsilon^{m+1} \in L^{\text{per}}(\mathbb{Z}h),\ m \in \{0,1,2,\ldots\},$$

under our assumption that u_0 is periodic of period 1.

So, by this exercise, we can analyze the convergence error by using the L^2 norm $\|\cdot\|_h$ on the space $L^{\text{per}}(\mathbb{Z}h)$.

Consider the integration of the advection equation on the time interval $[0,T]$ with time stepsize $k = \frac{T}{M}$, M positive integer. Exactly as in case of the Fourier analysis for the heat equation, we can prove the following stability and instability theorems. We recall that the spectral radius $\rho(A)$ of a (finite-dimensional) operator A is the maximum modulus of the eigenvalues of A.

Theorem 2 *If*

$$\rho\left(R\left(\tau_h\right)\right) \leq 1,$$

then

$$\max_{m \in \{0,\ldots,M\}} \|e^m\|_h \leq T \max_{i \in \{1,\ldots,M\}} \|\varepsilon^i\|_h.$$

Theorem 3 *Let $\beta > 1$. Consider discretizations with*

$$\rho\left(R\left(\tau_h\right)\right) \geq \beta$$

as $M, N \to \infty$. Then, there exists a consistency error $\left(\varepsilon^1, \ldots, \varepsilon^M\right)$ such that

$$\lim_{M,N\to\infty} \frac{\max\limits_{m \in \{0,\ldots,M\}} \|e^m\|_h}{\max\limits_{i \in \{1,\ldots,M\}} \|\varepsilon^i\|_h} = +\infty.$$

In particular, we have

$$\frac{\max\limits_{m \in \{0,\ldots,M\}} \|e^m\|_h}{\max\limits_{i \in \{1,\ldots,M\}} \|\varepsilon^i\|_h} \geq \frac{T}{M}\beta^M.$$

Thus, when the norm $\|\cdot\|_h$ is used, a method is stable if $\rho\left(R\left(\tau_h\right)\right) \leq 1$ for all positive integers M, N sufficiently large. Moreover, the method is unstable if $\rho\left(R\left(\tau_h\right)\right) \geq \beta$ for all positive integers M, N, for some $\beta > 1$.

7.1 Stability analysis

Now we analyze the stability of the four numerical methods under consideration by looking at $\rho\left(R\left(\tau_h\right)\right)$.

<u>Forward/forward</u>

In case of the forward/forward method, the eigenvalues of

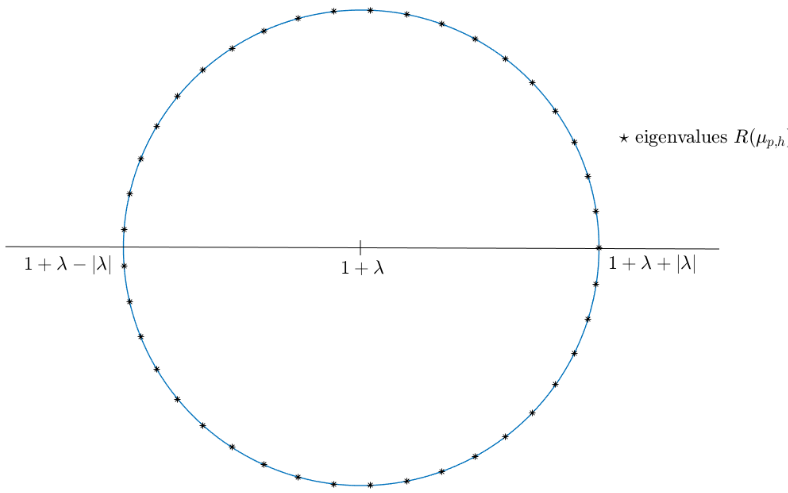
$$R(\tau_h) = (1+\lambda)\, I - \lambda\tau_h$$

are

$$R(\mu_{p,h}) = 1 + \lambda - \lambda\mu_{p,h},\ p \in \{0, \ldots, N-1\},$$

where $\mu_{p,h},\ p \in \{0, \ldots, N-1\}$ are the eigenvalues of τ_h.

In the complex plane, the eigenvalues of $R(\tau_h)$ lie on a circle of center $1+\lambda$ and radius $|\lambda|$.



Exercise. In this figure, is λ positive or negative?

Observe that the points of the circle at the maximum and minimum distance from the origin (the points of the circle of maximum and minimum modulus) are the extremal points $1+\lambda \pm |\lambda|$. One the extremal points is 1 and the other is $1+2\lambda$. So

$$\min\{1, |1+2\lambda|\} \leq |R(\mu_{p,h})| \leq \max\{1, |1+2\lambda|\}, \; p \in \{0, \ldots, N-1\}.$$

Finally, observe that $R(\mu_{0,h}) = 1$.

So for

$$|1+2\lambda| \leq 1,$$

i.e.

$$-1 \leq \lambda \leq 0,$$

the point of the circle at the maximum distance from the origin is $R(\mu_{0,h}) = 1$ and then

$$\rho(R(\tau_h)) = 1.$$

Exercise. Prove that if there exists $\gamma > 1$ such that

$$|1+2\lambda| \geq \gamma$$

for all positive integers M, N sufficiently large, then

$$\rho(R(\tau_h)) \geq \left| R\left(\mu_{\lceil \frac{N}{2} \rceil, h}\right) \right| \geq \beta = \frac{\gamma+1}{2} > 1$$

for all positive integers M, N sufficiently large.

Thus, the forward/forward method is stable if $-1 \leq \lambda \leq 0$ for all positive integers M, N sufficiently large. Moreover, it is unstable if $|1+2\lambda| \geq \gamma$ for all positive integers M, N, for some $\gamma > 1$, i.e. if λ remains at some positive distance from the interval $[-1, 0]$ for all positive integers M, N sufficiently large.

Backward/forward

In case of the backward/forward method, the eigenvalues of

$$R(\tau_h) = (1-\lambda)\, I + \lambda \tau_h^{-1}$$

are

$$R(\mu_{p,h}) = 1 - \lambda + \lambda \mu_{p,h}^{-1} = 1 - \lambda + \lambda \overline{\mu_{p,h}},\ p \in \{0, \ldots, N-1\},$$

and they lie on a circle of center $1-\lambda$ and radius $|\lambda|$. The extremal points $1 - \lambda \pm |\lambda|$ are 1 and $1-2\lambda$. So

$$\min\{1, |1-2\lambda|\} \leq |R(\mu_{p,h})| \leq \max\{1, |1-2\lambda|\},\ p \in \{0, \ldots, N-1\}.$$

We have $R(\mu_{0,h}) = 1$.

If

$$|1-2\lambda| \leq 1,$$

i.e.

$$0 \leq \lambda \leq 1,$$

the point of the circle at the maximum distance from the origin is $R(\mu_{0,h}) = 1$ and then

$$\rho(R(\tau_h)) = 1.$$

Moreover, similarly to the case of the forward/forward method, if there exists $\gamma > 1$ such that

$$|1-2\lambda| \geq \gamma$$

for all positive integers M, N sufficiently large, then

$$\rho(R(\tau_h)) \geq \left|R\left(\mu_{\lceil \frac{N}{2} \rceil, h}\right)\right| \geq \beta = \frac{\gamma+1}{2} > 1$$

for all positive integers M, N sufficiently large.

Thus, the backward/forward method is stable if $0 \leq \lambda \leq 1$ for all positive integers M, N sufficiently large. Moreover, it is unstable if $|1-2\lambda| \geq \gamma$ for all positive integers M, N, for some $\gamma > 1$, i.e. if λ remains at some positive distance from the interval $[0, 1]$ for all positive integers M, N sufficiently large.

Centered/forward

In case of the centered/forward method the eigenvalues of

$$R(\tau_h) = \frac{1}{2}(1+\lambda)\tau_h^{-1} + \frac{1}{2}(1-\lambda)\tau_h$$

are

$$\begin{aligned} R(\mu_{p,h}) &= 1+\frac{\lambda}{2}\mu_{p,h}^{-1}-\frac{\lambda}{2}\mu_{p,h} = 1+\frac{\lambda}{2}\left(\mu_{p,h}^{-1}-\mu_{p,h}\right) \\ &= 1+\frac{\lambda}{2}\left(\overline{\mu_{p,h}}-\mu_{p,h}\right) = 1-i\lambda\sin 2\pi ph \\ & \quad p\in\{0,\ldots,N-1\} \end{aligned}$$

with modulus

$$|R(\mu_{p,h})| = \sqrt{1+\lambda^2\sin^2 2\pi ph},\ p\in\{0,\ldots,N-1\}.$$

Exercise. Prove that if there exists $\gamma > 0$ such that

$$|\lambda| \geq \gamma$$

for all positive integers M, N sufficiently large, then

$$\rho\left(R\left(\tau_h\right)\right) \geq \left|R\left(\mu_{\lceil \frac{N}{4}\rceil,h}\right)\right| \geq \beta = \sqrt{1+\frac{\gamma^2}{2}} > 1$$

for all positive integers M, N sufficiently large.

Thus, the centered/forward method is unstable if $|\lambda| \geq \gamma$ for all positive integers M, N, for some $\gamma > 0$, i.e. if λ remains at some positive distance from 0 for all positive integers M, N sufficiently large.

<u>Lax-Friedrichs</u>

In the case of the Lax-Friedrichs method, the eigenvalues of

$$R(\tau_h) = \frac{1}{2}\left(1+\lambda\right)\tau_h^{-1}+\frac{1}{2}\left(1-\lambda\right)\tau_h$$

are

$$\begin{aligned} R(\mu_{p,h}) &= \frac{1}{2}\left(1+\lambda\right)\mu_{p,h}^{-1}+\frac{1}{2}\left(1-\lambda\right)\mu_{p,h} \\ &= \frac{1}{2}\left(\mu_{p,h}^{-1}+\mu_{p,h}\right)+\frac{\lambda}{2}\left(\mu_{p,h}^{-1}-\mu_{p,h}\right) \\ &= \frac{1}{2}\left(\overline{\mu_{p,h}}+\mu_{p,h}\right)+\frac{\lambda}{2}\left(\overline{\mu_{p,h}}-\mu_{p,h}\right) \\ &= \mathrm{Re}\left(\mu_{p,h}\right)-i\lambda\mathrm{Im}\left(\mu_{p,h}\right) \\ &= \cos 2\pi ph - i\lambda\sin 2\pi ph \\ & \quad p\in\{0,\ldots,N-1\} \end{aligned}$$

with modulus

$$|R(\mu_{p,h})| = \sqrt{\cos^2 2\pi ph+\lambda^2\sin^2 2\pi ph} = \sqrt{1+\left(\lambda^2-1\right)\sin^2 2\pi ph}$$
$$p\in\{0,1,\ldots,N-1\}.$$

Since

$$|R(\mu_{p,h})| \leq \sqrt{1 + (\lambda^2 - 1) \cdot 1} = |\lambda|,\ p \in \{0, \ldots, N-1\},$$

we see that

$$\rho\left(R\left(\tau_h\right)\right) \leq 1$$

if $|\lambda| \leq 1$.

Moreover, similarly to the case of the centered/forward method, if there exists $\gamma > 0$ such that

$$\left|\lambda^2 - 1\right| \geq \gamma$$

for all positive integers M, N sufficiently large, then

$$\rho\left(R\left(\tau_h\right)\right) \geq \left|R\left(\mu_{\lceil \frac{N}{4} \rceil, h}\right)\right| \geq \beta = \sqrt{1 + \frac{\gamma^2}{2}} > 1$$

for all positive integers M, N sufficiently large.

Thus, the Lax-Friedrichs method is stable if $|\lambda| \leq 1$ for all positive integers M, N sufficiently large. Moreover, it is unstable if $\left|\lambda^2 - 1\right| \geq \gamma$ for all positive integers M, N, for some $\gamma > 0$, i.e. if λ remains at some positive distance from the interval $[-1, 1]$ for all positive integers M, N sufficiently large.

8 The centered/backward method

None of the previous numerical methods for the advection equation is unconditionally stable. Now, we present a method which is unconditionally stable.

The *centered/backward method* is

$$\begin{gathered} \frac{U_n^{m+1} - U_n^m}{k} + c\frac{U_{n+1}^{m+1} - U_{n-1}^{m+1}}{2h} = 0 \\ n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\}. \end{gathered}$$

Here, we are approximating the time derivative $\frac{\partial u}{\partial t}(x,t)$ and spatial derivative $\frac{\partial u}{\partial x}(x,t)$ at $(x,t) = (nh, (m+1)k)$ rather than $(x,t) = (nh, mk)$ as in case of centered/forward.

Equivalently, the method can be written as

$$\begin{gathered} -\frac{\lambda}{2}U_{n-1}^{m+1} + U_n^{m+1} + \frac{\lambda}{2}U_{n+1}^{m+1} = U_n^m \\ n \in \mathbb{Z} \text{ and } m \in \{0, 1, 2, \ldots\} \end{gathered}$$

i.e.

$$\begin{gathered} \left(I - \frac{\lambda}{2}\tau_h^{-1} + \frac{\lambda}{2}\tau_h\right) U^{m+1} = U^m \\ m \in \{0, 1, 2, \ldots\} \end{gathered}$$

Exercise. Prove that for the consistency error

$$\varepsilon_n^{m+1},\ n \in \mathbb{Z} \text{ and } m \in \{0,1,2,\ldots\},$$

of the centered/backward method, we have

$$\max_{\substack{n\in\mathbb{Z}\\ m\in\{1,2,3,\ldots\}}} |\varepsilon_n^m| \leq Ch^2 + Dk$$

for some constants $C, D \geq 0$ independent of h and k.

With the usual notations, we have

$$\left(I - \frac{\lambda}{2}\tau_h^{-1} + \frac{\lambda}{2}\tau_h\right) U^{m+1} = U^m, \quad m \in \{0,1,2,\ldots\},$$

and

$$\left(I - \frac{\lambda}{2}\tau_h^{-1} + \frac{\lambda}{2}\tau_h\right) u^{m+1} = u^m + k\varepsilon^{m+1}, \quad m \in \{0,1,2,\ldots\}.$$

Then

$$\left(I - \frac{\lambda}{2}\tau_h^{-1} + \frac{\lambda}{2}\tau_h\right) e^{m+1} = e^m - k\varepsilon^{m+1}, \quad m \in \{0,1,2,\ldots\},$$

and so

$$e^{m+1} = R(\tau_h)e^m + kR(\tau_h)(-\varepsilon^{m+1}), \quad m \in \{0,1,2,\ldots\},$$

where

$$R(\tau_h) = \left(I - \frac{\lambda}{2}\tau_h^{-1} + \frac{\lambda}{2}\tau_h\right)^{-1}.$$

Exercise. Prove that $I - \frac{\lambda}{2}\tau_h^{-1} + \frac{\lambda}{2}\tau_h$ is invertible and that all the eigenvalues of $R(\tau_h)$ have modulus ≤ 1 and then $\rho\left(R\left(\tau_h\right)\right) \leq 1$. Then prove that the centered/backward method is unconditionally stable: by assuming to integrate in time on $[0,T]$ with stepsize $k = \frac{T}{M}$, M positive integer, the advection equation with the initial function u_0 periodic of period 1, we have

$$\max_{m\in\{0,\ldots,M\}} \|e^m\|_h \leq T \max_{i\in\{1,\ldots,M\}} \|\varepsilon^i\|_h$$

without any assumption of a relation between the stepsizes h and k.

On the other hand, the centered/backward method is an implicit method in the time integration. At each time step $m \in \{0,1,2,\ldots\}$, we need to solve the linear system on $L^{\mathrm{per}}(\mathbb{Z}h)$ (i.e. on $\mathbb{C}^N$)

$$\left(I - \frac{\lambda}{2}\tau_h^{-1} + \frac{\lambda}{2}\tau_h\right) U^{m+1} = U^m.$$