Numerical Solution of Partial Differential Equations

Endre Süli

Mathematical Institute
University of Oxford
2022

Lecture 15

First-order hyperbolic equations: initial-boundary-value problem and energy estimate

Let Ω be a bounded open set in $\mathbb{R}^n$, $n \geq 1$, with boundary $\Gamma = \partial\Omega$, and let $T > 0$. In $Q = \Omega \times (0, T]$, we consider the initial boundary-value problem

$$\frac{\partial u}{\partial t} + \sum_{i=1}^{n} b_i(x) \frac{\partial u}{\partial x_i} + c(x, t)u = f(x, t), \quad x \in \Omega, \;\; t \in (0, T], \tag{1}$$

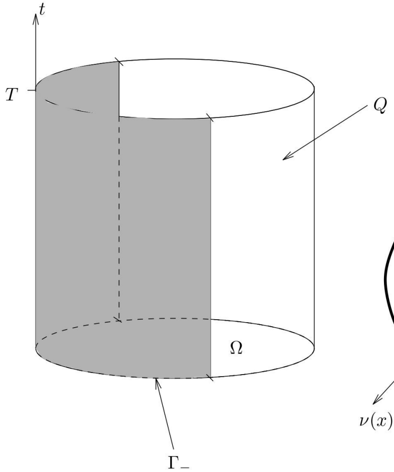
$$u(x, t) = 0, \quad x \in \Gamma_-, \;\; t \in [0, T], \tag{2}$$

$$u(x, 0) = u_0(x) \quad x \in \overline{\Omega}, \tag{3}$$

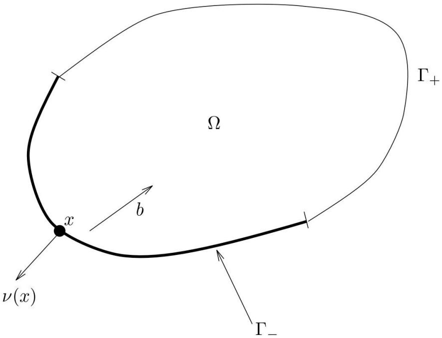
where

$$\Gamma_- = \{x \in \Gamma : b(x) \cdot \nu(x) < 0\},$$

$b = (b_1, \ldots, b_n)$ and $\nu(x)$ denotes the unit outward normal to Γ at $x \in \Gamma$. Γ_- will be called the *inflow boundary*. Its complement, $\Gamma_+ = \Gamma \setminus \Gamma_-$, will be referred to as the *outflow boundary*.



t
T
Q
Ω
Γ_-



Γ_+
Ω
b
x
$\nu(x)$
Γ_-

Continuous dependence of the solution on the data

We shall assume that

$$b_i \in C^1(\overline{\Omega}), \quad i = 1, \ldots, n, \tag{4}$$

$$c \in C(\overline{Q}), \quad f \in L_2(Q), \tag{5}$$

$$u_0 \in L^2(\Omega). \tag{6}$$

In order to ensure consistency between the initial and the boundary condition, we shall suppose that $u_0(x) = 0$, $x \in \Gamma_-$.

We make the additional hypothesis:

$$c(x,t) - \frac{1}{2}\sum_{i=1}^{n} \frac{\partial b_i}{\partial x_i}(x) \geq 0, \qquad x \in \overline{\Omega}, \quad t \in [0, T]. \tag{7}$$

By taking the inner product in $L_2(\Omega)$ of the equation (1) with $u(\cdot, t)$, performing partial integration and noting the boundary condition (2):

$$\left(\frac{\partial u}{\partial t}(\cdot, t), u(\cdot, t)\right) + \left(c(\cdot, t) - \frac{1}{2}\sum_{i=1}^{n} \frac{\partial b_i}{\partial x_i}(\cdot), u^2(\cdot, t)\right) + \frac{1}{2}\int_{\Gamma_+} \left[\sum_{i=1}^{n} b_i(x)\nu_i(x)\right] u^2(x, t)\,\mathrm{d}s(x) = (f(\cdot, t), u(\cdot, t)), \tag{8}$$

where $\nu(x) = (\nu_1(x), \ldots, \nu_n(x))$ is the unit outward normal vector to Γ at $x \in \Gamma$.

By virtue of (7) and noting that

$$\begin{aligned}\left(\frac{\partial u}{\partial t}, u\right) &= \int_\Omega \frac{\partial u}{\partial t}(x,t)\, u(x,t)\,\mathrm{d}x \\ &= \int_\Omega \frac{1}{2}\frac{\partial}{\partial t} u^2(x,t)\,\mathrm{d}x = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u^2(x,t)\,\mathrm{d}x \\ &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(\cdot,t)\|^2,\end{aligned}$$

it follows from (8) that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(\cdot,t)\|^2 + \frac{1}{2}\int_{\Gamma_+}\left[\sum_{i=1}^n b_i(x)\nu_i(x)\right] u^2(x,t)\,\mathrm{d}s(x) \le (f,u).$$

By the Cauchy–Schwarz inequality,

$$\begin{aligned}(f,u) &\le \|f(\cdot,t)\|\,\|u(\cdot,t)\| \\ &\le \frac{1}{2}\|f(\cdot,t)\|^2 + \frac{1}{2}\|u(\cdot,t)\|^2,\end{aligned}$$

and therefore, for any $t \in [0,T]$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(\cdot,t)\|^2 + \int_{\Gamma_+}\left[\sum_{i=1}^n b_i(x)\nu_i(x)\right] u^2(x,t)\,\mathrm{d}s(x) - \|u(\cdot,t)\|^2 \le \|f(\cdot,t)\|^2.$$

Multiplying both sides by e^{-t}, this inequality can be rewritten as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-t}\|u(\cdot,t)\|^2\right) + \mathrm{e}^{-t}\int_{\Gamma_+}\left[\sum_{i=1}^{n} b_i(x)\nu_i(x)\right] u^2(x,t)\,\mathrm{d}s \le \mathrm{e}^{-t}\|f(\cdot,t)\|^2,$$

Integrating this inequality w.r.t. t and noting the initial condition (3),

$$\mathrm{e}^{-t}\|u(\cdot,t)\|^2 + \int_0^t \mathrm{e}^{-\tau}\int_{\Gamma_+}\left[\sum_{i=1}^{n} b_i(x)\nu_i(x)\right] u^2(x,\tau)\,\mathrm{d}s(x)\,\mathrm{d}\tau$$
$$\le \|u_0\|^2 + \int_0^t \mathrm{e}^{-\tau}\|f(\cdot,\tau)\|^2\,\mathrm{d}\tau, \quad t \in [0,T].$$

It therefore follows that

$$\|u(\cdot,t)\|^2 + \int_0^t \mathrm{e}^{t-\tau}\int_{\Gamma_+}\left[\sum_{i=1}^{n} b_i(x)\nu_i(x)\right] u^2(x,\tau)\,\mathrm{d}s(x)\,\mathrm{d}\tau$$
$$\le \mathrm{e}^{t}\|u_0\|^2 + \int_0^t \mathrm{e}^{t-\tau}\|f(\cdot,\tau)\|^2\,\mathrm{d}\tau, \quad t \in [0,T]. \tag{9}$$

This 'energy inequality' expresses continuous dependence of the solution to (1)–(3) on the data. It also implies uniqueness of the solution.

Let us consider a particularly important case when

$$c \equiv 0, \quad f \equiv 0, \quad \text{and} \quad \text{div } b = \sum_{i=1}^{n} \frac{\partial b_i}{\partial x_i} \equiv 0,$$

where $b(x) = (b_1(x), \ldots, b_n(x))$. Then, thanks to the identity (8),

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|u(\cdot, t)\|^2 + \frac{1}{2} \int_{\Gamma_+} [b(x) \cdot \nu(x)]\, u^2(x, t)\, \mathrm{d}s(x) = 0,$$

and therefore,

$$\|u(\cdot, t)\|^2 + \int_0^t \int_{\Gamma_+} [b(x) \cdot \nu(x)]\, u^2(x, \tau)\, \mathrm{d}s(x)\, \mathrm{d}\tau = \|u_0\|^2,$$

which can be viewed as an identity expressing 'conservation of energy' for the initial-boundary-value problem (1)–(3).

Explicit finite difference approximation

We focus on a special case of the problem: the constant- coefficient hyperbolic equation in one space dimension

$$\frac{\partial u}{\partial t} + b\frac{\partial u}{\partial x} = f(x,t), \quad x \in (0,1), \;\; t \in (0,T], \tag{10}$$

subject to the boundary and initial conditions

$$u(x,t) = 0, \qquad x \in \Gamma_-, \;\; t \in [0,T], \tag{11}$$

$$u(x,0) = u_0(x), \quad x \in [0,1]. \tag{12}$$

If $b > 0$ then $\Gamma_- = \{0\}$, and if $b < 0$ then $\Gamma_- = \{1\}$. Let us assume, for example, that $b > 0$. Then the appropriate boundary condition is

$$u(0,t) = 0, \quad t \in [0,T]. \tag{13}$$

To construct a finite difference approximation of (10)–(13) let $\Delta x := 1/J$ be the mesh-size in the x-direction and $\Delta t := T/M$ the mesh-size in the time-direction, t. Let us also define

$$x_j := j\Delta x, \;\; j = 0, \dots, J, \qquad t_m := m\,\Delta t, \;\; m = 0, \dots, M.$$

At the mesh-point (x_j, t_m), (10) is approximated by the explicit finite difference scheme

$$\frac{U_j^{m+1} - U_j^m}{\Delta t} + b\, D_x^- U_j^m = f(x_j, t_m), \qquad j = 1, \dots, J, \quad m = 0, \dots, M-1, \tag{14}$$

subject to the boundary and initial condition, respectively:

$$U_0^m = 0, \qquad m = 0, \dots, M, \tag{15}$$

$$U_j^0 = u_0(x_j), \qquad j = 0, \dots, J. \tag{16}$$

Equivalently, this can be written as follows:

$$U_j^{m+1} = (1-\mu)U_j^m + \mu U_{j-1}^m + \Delta t\, f(x_j, t_m), \qquad \begin{cases} j = 1, \dots, J, \\ m = 0, \dots, M-1, \end{cases}$$

in conjunction with

$$\begin{aligned} U_0^m &= 0, & m &= 0, \dots, M, \\ U_j^0 &= u_0(x_j), & j &= 0, \dots, J, \end{aligned}$$

where

$$\mu = \frac{b\Delta t}{\Delta x};$$

μ is called the CFL (or Courant–Friedrichs–Lewy) number.

The explicit finite difference scheme (14) is frequently called the *first-order upwind scheme*.

We shall explore the stability of this scheme in the discrete maximum norm. Suppose that $0 \leq \mu \leq 1$; then

$$\begin{aligned}
\left|U_j^{m+1}\right| &\leq (1-\mu)\left|U_j^m\right| + \mu\left|U_{j-1}^m\right| + \Delta t\left|f(x_j,t_m)\right| \\
&\leq (1-\mu)\max_{0\leq j\leq J}\left|U_j^m\right| + \mu\max_{1\leq j\leq J+1}\left|U_{j-1}^m\right| + \Delta t\max_{0\leq j\leq J}\left|f(x_j,t_m)\right| \\
&= \max_{0\leq j\leq J}\left|U_j^m\right| + \Delta t\max_{0\leq j\leq J}\left|f(x_j,t_m)\right|.
\end{aligned}$$

Thus we have that

$$\max_{0\leq j\leq J}\left|U_j^{m+1}\right| \leq \max_{0\leq j\leq J}\left|U_j^m\right| + \Delta t\max_{0\leq j\leq J}\left|f(x_j,t_m)\right|.$$

Let us define the mesh-dependent norm

$$\|U\|_\infty := \max_{0\leq j\leq J}\left|U_j\right|;$$

then

$$\|U^{m+1}\|_\infty \leq \|U^m\|_\infty + \Delta t\|f(\cdot,t_m)\|_\infty, \quad m = 0,\ldots,M-1.$$

Summing through m, we get

$$\max_{1\leq k\leq M} \|U^k\|_\infty \leq \|U^0\|_\infty + \sum_{m=0}^{M-1} \Delta t \|f(\cdot, t_m)\|_\infty, \tag{17}$$

which expresses the stability of the finite difference scheme (14)–(16) under the condition

$$0 \leq \mu = \frac{b\Delta t}{\Delta x} \leq 1. \tag{18}$$

Thus we have proved that the finite difference scheme (14)–(16) is conditionally stable, the condition being that the CFL number $\mu \in [0, 1]$.

It is possible to show that the scheme (14)–(16) is also stable in the mesh-dependent L_2-norm, $\|\cdot]|$, defined by

$$\|V]|^2 = \sum_{i=1}^{J} \Delta x \, V_i^2.$$

The associated inner product is

$$(V, W] := \sum_{i=1}^{J} \Delta x \, V_i W_i.$$

Since

$$U_j^m = \frac{U_j^m + U_{j-1}^m}{2} + \frac{U_j^m - U_{j-1}^m}{2},$$

and $U_0^m = 0$, it follows that

$$\begin{aligned}
(D_x^- U^m, U^m] &= \sum_{j=1}^{J} \Delta x \, \frac{U_j^m - U_{j-1}^m}{\Delta x} \, U_j^m \\
&= \frac{1}{2} \sum_{j=1}^{J} \{(U_j^m)^2 - (U_{j-1}^m)^2\} + \frac{\Delta x}{2} \sum_{j=1}^{J} \Delta x \left(\frac{U_j^m - U_{j-1}^m}{\Delta x} \right)^2 \qquad (19) \\
&= \frac{1}{2} (U_J^m)^2 + \frac{\Delta x}{2} \| D_x^- U^m]|^2.
\end{aligned}$$

In addition, since

$$U_j^m = \frac{U_j^{m+1} + U_j^m}{2} - \frac{U_j^{m+1} - U_j^m}{2}$$

for $m = 0, \ldots, M-1$, we have for such m that

$$\left(\frac{U^{m+1} - U^m}{\Delta t}, U^m \right] = \frac{1}{2\Delta t} \left(\| U^{m+1}]|^2 - \| U^m]|^2 \right) - \frac{\Delta t}{2} \left\| \frac{U^{m+1} - U^m}{\Delta t} \right]\right|^2 . \quad (20)$$

By taking the $(\cdot, \cdot]$-inner product of (14) with U^m and using (19) and (20):

$$\begin{aligned} &\| U^{m+1}]|^2 + \Delta t \, b (U_J^m)^2 + b \, \Delta x \, \Delta t \, \| D_x^- U^m]|^2 - \| U^m]|^2 \\ &\qquad - \Delta t^2 \left\| \frac{U^{m+1} - U^m}{\Delta t} \right]\right|^2 = 2\Delta t \, (f^m, U^m], \qquad m = 0, \ldots, M-1. \quad (21) \end{aligned}$$

First suppose that $f \equiv 0$; then,

$$\frac{U^{m+1} - U^m}{\Delta t} = -b\, D_x^- U^m,$$

and by substituting this into the last term on the left-hand side of the equality (21) we have that, for $m = 0, \ldots, M-1$,

$$\| U^{m+1}]|^2 + \Delta t\, b\, |U_J^m|^2 + b\, \Delta x\, \Delta t\, (1-\mu) \| D_x^- U^m]|^2 = \| U^m]|^2.$$

Summing through m,

$$\| U^k]|^2 + \sum_{m=0}^{k-1} \Delta t\, b\, |U_J^m|^2 + b\, \Delta x\, (1-\mu) \sum_{m=0}^{k-1} \Delta t\, \| D_x^- U^m]|^2 = \| U^0]|^2 \qquad (22)$$

for $k = 1, \ldots, M$, which proves the stability of the scheme in the case when $f \equiv 0$ whenever

$$0 \le \mu = \frac{b\, \Delta t}{\Delta x} \le 1.$$

In particular, if $\mu = 1$, we have from (22) that

$$\|U^k]|^2 + \sum_{m=0}^{k-1} \Delta t\, b\, |U_J^m|^2 = \|U^0]|^2, \quad k = 1, \ldots, M,$$

which is the discrete version of the identity (9), and expresses 'conservation of energy' in the discrete sense.

More generally, for $0 \leq \mu \leq 1$, (22) implies

$$\|U^k]|^2 + \sum_{m=0}^{k-1} \Delta t\, b\, |U_J^m|^2 \leq \|U^0]|^2, \quad k = 1, \ldots, M.$$

Now consider the question of stability in the $\|\cdot]|$-norm for $f \not\equiv 0$. Since

$$\left\| \frac{U^{m+1} - U^m}{\Delta t} \right]\right|^2 = \|f^m - bD_x^- U^m]|^2 \le \{\|f^m]| + b\|D_x^- U^m]|\}^2$$
$$\le \left(1 + \frac{1}{\epsilon'}\right) \|f^m]|^2 + (1+\epsilon')b^2\|D_x^- U^m]|^2, \quad \epsilon' > 0,$$

and

$$(f^m, U^m] \le \|f^m]|\,\|U^m]| \le \tfrac{1}{2}\|f^m]|^2 + \tfrac{1}{2}\|U^m]|^2,$$

it follows from the equality (21) that

$$\|U^{m+1}]|^2 + \Delta t\, b\,|U_n^m|^2 + b\,\Delta x\,\Delta t \left[1 - (1+\epsilon')\frac{b\Delta t}{\Delta x}\right] \|D_x^- U^m]|^2$$
$$\le \Delta t \left[\left(1 + \frac{1}{\epsilon'}\right)\Delta t + 1\right] \|f^m]|^2 + (1+\Delta t)\|U^m]|^2.$$

Letting $\epsilon = 1 - 1/(1+\epsilon') \in (0,1)$ and assuming that

$$0 \le \mu = \frac{b\,\Delta t}{\Delta x} \le 1 - \epsilon,$$

we have, for $m = 0, \ldots, M-1$, that

$$\|U^{m+1}]|^2 + \Delta t\, b\,|U^m_J|^2 \le \|U^m]|^2 + \Delta t\left(1 + \frac{\Delta t}{\epsilon}\right)\|f^m]|^2 + \Delta t\|U^m]|^2.$$

Upon summation of this inequality over $m = 0, \ldots, k-1$, we deduce that

$$\begin{aligned} \|U^k]|^2 + \left(\sum_{m=0}^{k-1} \Delta t\, b\, |U^m_J|^2\right) \le \|U^0]|^2 + \left(1 + \frac{\Delta t}{\epsilon}\right)\sum_{m=0}^{k-1} \Delta t\,\|f^m]|^2 \\ + \sum_{m=0}^{k-1} \Delta t\,\|U^m]|^2 \end{aligned} \tag{23}$$

for $k = 1, \ldots, M$.

To complete the proof of stability of the finite difference scheme we require the next lemma, which is easily proved by induction.

Lemma (Discrete Gronwall lemma)

Let (a_k), (b_k), (c_k) and (d_k) be four sequences of non-negative numbers such that the sequence (c_k) is non-decreasing and

$$a_k + b_k \leq c_k + \sum_{m=0}^{k-1} d_m a_m, \quad k \geq 1; \;\; a_0 + b_0 \leq c_0.$$

Then

$$a_k + b_k \leq c_k \exp\left(\sum_{m=0}^{k-1} d_m\right), \quad k \geq 1.$$

By applying this lemma to the inequality (23) with

$$
\begin{aligned}
a_k &:= \|U^k]|^2, \quad k \geq 0, \\
b_k &:= \sum_{m=0}^{k-1} \Delta t\, b\, \left|U_J^m\right|^2, \quad k \geq 1; \;\; b_0 = 0, \\
c_k &:= \|U^0]|^2 + \left(1 + \frac{\Delta t}{\epsilon}\right) \sum_{m=0}^{k-1} \Delta t\, \|f^m]|^2, \quad k \geq 1; \;\; c_0 = \|U^0]|^2, \\
d_k &:= \Delta t, \quad\; k = 1, 2, \ldots, M,
\end{aligned}
$$

we obtain for $k = 1, \ldots, M$:

$$
\|U^k]|^2 + \sum_{m=0}^{k-1} \Delta t\, b\, |U_J^m|^2 \leq \mathrm{e}^{t_k} \left(\|U^0]|^2 + \left(1 + \frac{\Delta t}{\epsilon}\right) \sum_{m=0}^{k-1} \Delta t \|f^m]|^2 \right),
$$

where $t_k = k\Delta t$. Hence we deduce stability of the scheme, in the sense that

$$
\max_{1 \leq k \leq M} \left(\|U^k]|^2 + \sum_{m=0}^{k-1} \Delta t\, b\, \left|U_J^m\right|^2 \right) \leq \mathrm{e}^{T} \left(\|U^0]|^2 + \left(1 + \frac{\Delta t}{\epsilon}\right) \sum_{m=0}^{M-1} \Delta t \|f^m]|^2 \right).
$$