An Introduction to Numerical Partial Differential Equations

CHUN-HAO TENG

Department of Applied Mathematics
National Chiao Tung University
Hsin-Chu, Taiwan

Outline

An Overview

1. What is numerical partial differential equations?
   - Many natural phenomena are formulated by partial differential equations (PDEs) and by solving the equations one gains a better understanding on the mechanism of these phenomena.
   - The topics of numerical PDEs is about constructing computational algorithms to analyze problems.
2. Does there exist a general methodology for constructing numerical schemes?
   - The answer is "Yes" to certain types of problems. General speaking we need to construct a well-posed analysis on the problem wherever possible. This analysis is our guide line for constructing numerical schemes.

Well-posedness of Problems Described by Partial Differential Equations

Example (Model Wave Problem)

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x \in [0,1], \quad t \geq 0,$$

$$u(x,t) = f(x), \quad x \in [0,1], \quad t = 0,$$

$$u(0,t) = g(t), \quad x = 0, \quad t \geq 0.$$

Smoothness condition : $f(0) = g(0)$

We say the problem is well-posed if

1. The solution exists.
2. The solution is unique.
3. The solution is stable.

Well-posedness II

- It would be meaningful if there exists a solution to the problem.
- If the problem does have a solution we need to ask whether the problem has other solutions.
- In physics we consider that a physical quantity is a finite number and can be measured by certain methods or devices. Moreover, we wish that the system of a physical problem is stable, in the sense that when a small perturbation is introduced into the system, the solution does not deviate away from the unperturbed one.

Well-posedness of the Initial Value Problem

Consider the following example.

Example (2π-periodic scalar wave equation)

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x \in [0, 2\pi], \quad t \geq 0, \tag{1}$$

$$u(x,t) = f(x), \quad t = 0, \tag{2}$$

$$u^{(p)}(0,t) = u^{(p)}(2\pi, t), \quad u^{(p)} = \frac{\partial^{(p)} u}{\partial x^{(p)}}, \quad p = 0, 1, 2, \ldots \tag{3}$$

Does the solution exist ?

Well-posedness of the Initial Value Problem II

Assume

$$u(x,t) = \hat{u}_k(t)\, e^{ikx}, \quad k \in \mathbb{Z} \tag{4}$$

If (4) is a solution to (1) then

$$\begin{aligned}
& \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \\
\Rightarrow \quad & \frac{d\hat{u}_k(t)}{dt} \cdot e^{ikx} = (-ik)\, \hat{u}_k(t) \cdot e^{ikx} \\
\Rightarrow \quad & \hat{u}_k(t) = \hat{u}_k(0) \cdot e^{-ikt} \\
\Rightarrow \quad & u(x,t) = \hat{u}_k(0) \cdot e^{-ikt} \cdot e^{ikx} = \hat{u}_k(0) \cdot e^{ik(x-t)}
\end{aligned}$$

Well-posedness of the Initial Value Problem III

Invoking linear superposition we have

$$u(x,t) = \sum_{k=-\infty}^{\infty} \hat{u}_k(0) e^{ik(x-t)}.$$

Take $t = 0$

$$u(x,0) = f(x) \quad \Rightarrow \sum_{k=-\infty}^{\infty} \hat{u}_k(0) e^{ikx} = f(x)$$

where $\hat{u}_k(0)$ is the fourier coefficients of the function $f(x)$, provided that f has a fourier series representation.
We have a solution to the problem.

Well-posedness of the Initial Value Problem IV

Uniqueness: Is this the only one ?

Assume that $v \neq u$ is also a solution to the problem, i.e.,

$$\frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} = 0, \quad v(x,0) = f(x), \quad v^{(p)}(0,t) = v^{(p)}(2\pi,t).$$

Let $w = u - v$ then

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} = 0, \quad w(x,0) = 0, \quad w^{(p)}(0,t) = w^{(p)}(2\pi,t).$$

Hence

$$w(x,t) = \sum_{k=-\infty}^{\infty} \hat{w}_k(0)\, e^{ik(x-t)} \quad \text{and} \quad w(x,0) = 0 = \sum_{k=-\infty}^{\infty} \hat{w}_k(0)\, e^{ikx}.$$

This implies

$$\hat{w}_k(0) = 0 \quad \Rightarrow \quad w(x,t) = u(x,t) - v(x,t) = 0$$

$u(x,t)$ and $v(x,t)$ are identical.

Well-posedness of the Initial Value Problem V

How do we know a solution is stable?

- In addition to the issues concerning the existence and uniqueness of the solution, we also need to know whether the solution is stable.

Consider the problem:

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0,$$
$$u(x,0) = f(x),$$
$$u^{(p)}(0,t) = u^{(p)}(2\pi,t).$$

Define the energy of the system as

Definition

$$E(t) = \int_0^{2\pi} u^2(x,t)\,dx$$

Energy of Physical Systems

- We observe similar definitions of energy for various types of physical systems, for example
  1. in electromagnetism

$$\text{Energy} = \int_{\Omega} \frac{1}{2}\epsilon E^2 + \frac{1}{2}\mu H^2 \, dx$$

  2. in fluid dynamics

$$\text{Energy} = \int_{\Omega} \rho(V \cdot V) \, d\Omega$$

Energy Estimate for $\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0$

Multiplying u to the equation and integrating over the entire domain, we have the energy rate equation

$$\frac{dE(t)}{dt} = \frac{d}{dt}\int_0^{2\pi} u^2(x,t)\,dx = \int_0^{2\pi} 2u\frac{\partial u(x,t)}{\partial t}\,dx$$

$$= \int_0^{2\pi} 2u\left(-\frac{\partial u}{\partial x}\right)\,dx = -\int_0^{2\pi}\frac{\partial u^2}{\partial x}\,dx = -u^2(x,t)\Big|_0^{2\pi} = 0$$

Hence,

$$E(t) = E(0) \Rightarrow \int_0^{2\pi} u^2(x,t)\,dx = \int_0^{2\pi} f^2(x)\,dx$$

Well-posedness of the Initial Value Problem VI

Energy of a Perturbed System

Let us now consider the following problems.

Unperturbed Problem:

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0$$
$$u(x,0) = f(x)$$
$$u^{(p)}(0,t) = u^{(p)}(2\pi,t)$$

Perturbed Problem:

$$\frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} = 0$$
$$v(x,0) = f(x) + \epsilon(x), \ |\epsilon(x)| << 1$$
$$v^{(p)}(0,t) = v^{(p)}(2\pi,t)$$

Well-posedness of the Initial Value Problem VII

How do we know a solution is stable?

Let $w = v - u$

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} = 0$$
$$w(x,0) = \epsilon(x)$$
$$w^{(p)}(0,t) = w^{(p)}(2\pi,t)$$

Then

$$\int_0^{2\pi} w^2(x,t)\,dx = \int_0^{2\pi} (v(x,t) - u(x,t))^2\,dx = \int_0^{2\pi} \epsilon^2(x)\,dx$$

At any given time the difference between u and v measured in the sense of energy is bounded by the energy of the initial perturbation.

Consistency

Define the grid points:

$$x_j = j \cdot h = j \cdot \frac{2\pi}{N+1}, \quad j = 0, 1, 2, \ldots, N$$

Let

$$u(x_j, t) = u_j(t)$$

Recall that $\frac{\partial u(x,t)}{\partial x}$ can be approximated by

- forward difference: $\frac{u_{j+1} - u_j}{h} + \mathcal{O}(h)$
- backward difference: $\frac{u_j - u_{j-1}}{h} + \mathcal{O}(h)$
- central difference: $\frac{u_{j+1} - u_{j-1}}{2h} + \mathcal{O}(h^2)$

Upwind Scheme: $\frac{dv_i}{dt} = -\frac{v_i - v_{i-1}}{h}$ I

Define the numerical solution as

$$v_i(t), i = 0, 1, 2, ..., N$$

satisfying the semi-discrete scheme

$$\frac{dv_i}{dt} + \frac{v_i - v_{i-1}}{h} = 0, \quad i = 0, 1, ..., N$$
$$v_i(t) = f(x_i) = f_i$$
$$v_{-1} = v_N,$$

Upwind Scheme: $\frac{dv_i}{dt} = -\frac{v_i - v_{i-1}}{h}$ II

Consistency

Substituting the exact solution $u_i(t) = u(x_i, t)$ to the scheme

$$\frac{dv_i}{dt} + \frac{v_i - v_{i-1}}{h} = 0$$

we get the truncation error (TE)

$$TE = \frac{du_i}{dt} + \frac{u_i - u_{i-1}}{h} = \frac{\partial u(x_i,t)}{\partial t} + \frac{\partial u(x_i,t)}{\partial x} + \mathcal{O}(h) = \mathcal{O}(h).$$

Observe that $TE \to 0$ as $h \to 0$. The scheme is consistent.

Upwind Scheme: $\frac{dv_i}{dt} = -\frac{v_i - v_{i-1}}{h}$ III

Energy Estimate

$$
\begin{aligned}
\sum_{i=0}^{N} v_i \frac{dv_i}{dt} h &= -\sum_{i=0}^{N} v_i v_i + \sum_{i=0}^{N} v_i v_{i-1} \\
&\leq -\sum_{i=0}^{N} v_i^2 + \frac{1}{2}\sum_{i=0}^{N} v_i^2 + \frac{1}{2}\sum_{i=0}^{N} v_{i-1}^2 \\
&= -\frac{1}{2}\sum_{i=0}^{N} v_i^2 + \frac{1}{2}\sum_{i=0}^{N} v_{i-1}^2 \quad (\text{periodicity } v_{-1} = v_N) \\
&= 0
\end{aligned}
$$

Upwind Scheme: $\frac{dv_i}{dt} = -\frac{v_i - v_{i-1}}{h}$ IV

Energy Estimate

Hence

$$\frac{1}{2}\frac{d}{dt}\sum_{i=0}^{N} v_j^2 h \leq 0 \quad \Rightarrow \quad \sum_{i=0}^{N} v_i^2 h \leq \sum_{i=0}^{N} f_i^2 h$$

The scheme has a bounded energy estimate for a given terminal time.

Recall the energy estimate of the continuous system:

$$\int_0^{2\pi} u^2(x,t)\,dx = \int_0^{2\pi} f^2(x)\,dx.$$

Down Wind Scheme: $\frac{dv_i}{dt} = -\frac{v_{i+1} - v_i}{h}$

Consider the scheme:

$$\frac{dv_i}{dt} = -\frac{v_{i+1} - v_i}{h} \quad i = 0, 1, 2, \ldots, N$$
$$v_i(0) = f(x_i) = f_i$$
$$v_{N+1} = v_0$$

Consistency Check: Substituting the exact solution $u_i(t)$ to the scheme we get the truncation error (TE)

$$TE = \frac{du_i}{dt} + \frac{u_{i+1} - u_i}{h} = \frac{\partial u(x_i, t)}{\partial t} + \frac{\partial u(x_i, t)}{\partial x} + \mathcal{O}(h) = \mathcal{O}(h).$$

Observe that $TE \to 0$ as $h \to 0$. The scheme is consistent

Down Wind Scheme: $\frac{dv_i}{dt} = -\frac{v_{i+1} - v_i}{h}$ II

Energy Estimate

$$\begin{aligned}\frac{dE(t)}{dt} &= \sum_{i=0}^{N} v_i \frac{dv_i}{dt} h = -\sum_{i=0}^{N} v_{i+1} v_i + \sum_{i=0}^{N} v_i^2 \\ &= \frac{1}{2}\sum_{i=0}^{N} v_i^2 + \frac{1}{2}\sum_{i=0}^{N} v_i^2 - \sum_{i=0}^{N} v_{i+1} v_i + \frac{1}{2}\sum_{i=0}^{N} v_{i+1}^2 - \frac{1}{2}\sum_{i=0}^{N} v_{i+1}^2 \\ &= \frac{1}{2}\sum_{i=0}^{N} (v_i - v_{i+1})^2 \geq 0 \quad (= 0 \quad \text{when constant})\end{aligned}$$

Down Wind Scheme: $\frac{dv_i}{dt} = -\frac{v_{i+1} - v_i}{h}$ III

Energy Estimate

The fact that

$$\frac{dE(t)}{dt} = \sum_{i=0}^{N} v_i \frac{dv_i}{dt} h \geq 0$$

leads to

$$\sum_{i=0}^{N} v_i^2 h \geq \sum_{i=0}^{N} f_i^2 h.$$

We have a scheme with an energy estimate that can not be bounded by the prescribed data.

What Does the Bounded Energy Estimate Imply?

Energy of the Error Equation

Consider vectors

$$\boldsymbol{a} = [a_0, a_1, ..., a_N]^T, \quad \boldsymbol{b} = [b_0, b_1, ..., b_N]^T$$

we define the vector inner product and norm as

$$(\boldsymbol{a}, \boldsymbol{b})_h = \boldsymbol{a}^T \boldsymbol{b} h = \sum_{i=0}^{N} a_i b_i h, \quad ||\boldsymbol{a}||^2 = (\boldsymbol{a}, \boldsymbol{a})_h$$

Useful inequality: For real numbers a and b, we have

$$2ab \leq a^2 + b^2$$

For vectors $\boldsymbol{a}$ and $\boldsymbol{b}$

$$(\boldsymbol{a}, \boldsymbol{b})_h^2 \leq ||\boldsymbol{a}|| \cdot ||\boldsymbol{b}||$$

What Does the Bounded Energy Estimate Imply?

Energy of the Error Equation (Upwind Scheme)

Recall the scheme

$$\frac{dv_i}{dt} + \frac{v_i - v_{i-1}}{h} = 0$$
$$v_i(0) = f(x_i) = f_i$$
$$v_{-1} = v_N,$$

Let $u_i(t) = u(x_i, t)$. Then

$$\frac{du_i}{dt} + \frac{u_i - u_{i-1}}{h} = te_i, \quad te_i = \mathcal{O}(h)$$
$$u_i(0) = f(x_i) = f_i$$
$$u_{-1} = u_N,$$

Define $e_i = u_i - v_i$. Then e_i satisfies

$$\frac{de_i}{dt} + \frac{e_i - e_{i-1}}{h} = te_i,$$
$$e_i(0) = 0$$
$$e_{-1} = e_N,$$

We now examine the energy of the discrete system.

What Does the Bounded Energy Estimate Imply?

Energy of the Error Equation (Upwind Scheme)

Denote

$$\boldsymbol{e} = [e_0, e_1, ..., e_N]^T, \quad \boldsymbol{te} = [te_0, te_1, ..., te_N]^T$$

$$\begin{aligned}
\frac{1}{2}\frac{d||\boldsymbol{e}||^2}{dt} &= \sum_{i=0}^{N} e_i \frac{de_i}{dt} h = \sum_{i=0}^{N} (e_i - e_{i-1}) e_i + \sum_{i=0}^{N} e_i te_i h \\
&\leq -\frac{1}{2}\sum_{i=0}^{N}(e_i - e_{i-1})^2 + \frac{1}{2}\sum_{i=0}^{N} e_i^2 h + \frac{1}{2}\sum_{i=0}^{N} te_i^2 h \\
&\leq \frac{1}{2}||\boldsymbol{e}||^2 + \frac{1}{2}||\boldsymbol{te}||^2
\end{aligned}$$

What Does the Bounded Energy Estimate Imply?

Energy of the Error Equation (Upwind Scheme)

$$
\begin{aligned}
& e^{-t}\frac{d||\boldsymbol{e}||^2}{dt} - e^{-t}||\boldsymbol{e}||^2 \leq e^{-t}||\boldsymbol{te}||^2 \\
\Rightarrow & \frac{d}{dt}\left(e^{-t}||\boldsymbol{e}||^2\right) \leq e^{-t}||\boldsymbol{te}||^2 \\
\Rightarrow & e^{-t}||\boldsymbol{e}(t)||^2 - e^{0}||\boldsymbol{e}(0)||^2 \leq \int_0^t e^{-\xi}||\boldsymbol{te}(\xi)||^2\, d\xi \\
\Rightarrow & ||\boldsymbol{e}(t)||^2 \leq e^{t}\int_0^t e^{-\xi}||\boldsymbol{te}(\xi)||^2\, d\xi \\
\Rightarrow & ||\boldsymbol{e}(t)||^2 \leq e^{t}\left(\max_{\xi\in[0,t]}||\boldsymbol{te}(\xi)||^2\right)\int_0^t e^{-\xi}\, d\xi \\
\Rightarrow & ||\boldsymbol{e}(t)||^2 \leq \left(\max_{\xi\in[0,t]}||\boldsymbol{te}(\xi)||^2\right)\left(e^{t}-1\right)
\end{aligned}
$$

What Does the Bounded Energy Estimate Imply?

Energy of the Error Equation (Upwind Scheme)

From

$$||\boldsymbol{e}(t)||^2 \leq \left(\max_{\xi \in [0,t]} ||\boldsymbol{te}(\xi)||^2 \right) (e^t - 1)$$

we have

1. As time evolves the error may grow exponentially.
2. For constant t, $e^t - 1$ is fixed. As $h \to 0$, $\max_{\xi \in [0,t]} ||\boldsymbol{te}(\xi)||^2 \to 0$. Hence

$$||\boldsymbol{e}(t)||^2 = ||\boldsymbol{u}(t) - \boldsymbol{v}(t)||^2 \to 0$$

implying convergence.

What Does the Bounded Energy Estimate Imply?

Energy of the Error Equation (Down Wind Scheme)

For down wind scheme one can follow a similar and yield

$$||\boldsymbol{e}|| \geq \min_{\xi \in [0,t]} c(t) \frac{h^2}{\lambda_0} (e^{\frac{\lambda_0 t}{h}} - 1)$$

where $c(t) = \mathcal{O}(1)$ and $\lambda_0 > 0$ is a constant. From the result we observe that

1. as $h \to 0$ $h^2 (e^{\frac{\lambda_0 t}{h}} - 1)$

Classical Theory on Convergence

Theorem (Lax-Richtmyer Equivalence Theorem)

A consistent approximation to a linear well-posed partial differential equation is convergent if and only if it is stable.

Remark:

- We need a consistent scheme which has an energy estimate bounded by the prescribed data.
- The upwind scheme is consistent and stable. Thus, the numerical solution converges to the exact solution.
- The downwind scheme is consistent and unstable. Thus, the numerical solution does not converge to the exact one.

Phase Error Analysis I

Consider the linear wave problem

$$\frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x}, \quad 0 \le x \le 2\pi, \qquad (5)$$
$$u(x,0) = e^{ikx}.$$

The solution to equation (5) is a travelling wave

$$u(x,t) = e^{ik(x-ct)}$$

with phase speed c

Phase Error Analysis II

We use the equidistant grid

$$x_j = j\,\Delta x = \frac{2\pi j}{N+1}, \quad j \in [0,\dots,N].$$

The $2m$-order approximation of the derivative of a function $f(x)$ is

$$\left.\frac{df}{dx}\right|_{x_j} = \sum_{n=1}^{m} \alpha_n^m \mathcal{D}_n f(x_j)$$

where

$$\mathcal{D}_n f(x_j) = \frac{f(x_j + n\,\Delta x) - f(x_j - n\,\Delta x)}{2n\,\Delta x} = \frac{f_{j+n} - f_{j-n}}{2n\,\Delta x}$$

$$\alpha_n^m = -2(-1)^n \frac{(m!)^2}{(m-n)!\,(m+n)!}$$

Phase Error Analysis III

In the semi-discrete version of Equation (5) we seek a vector $\boldsymbol{v} = (v_0(t),\ v_1(t),\ \ldots,\ v_N(t))$ which satisfies

$$\frac{dv_j}{dt} = -c\sum_{n=1}^{m} \alpha_n^m \mathcal{D}_n v_j, \quad v_j(0) = e^{ikx_j}. \tag{6}$$

We may interpret the grid vector, $\boldsymbol{v}$, as a vector of grid point values of a trigonometric polynomial, $v(x,t)$, with $v(x_j,t) = \boldsymbol{v}_j(t)$, such that

$$\frac{\partial v(x,t)}{\partial t} = -c\sum_{n=1}^{m} \alpha_n^m \mathcal{D}_n v(x,t), \quad v(x,0) = e^{ikx} \tag{7}$$

If $v(x,t)$ satisfies Equation (7), the solution to Equation (6) is given by $v(x_j,t)$.

Phase Error Analysis IV

The solution to Equation (7) is

$$v(x,t) = e^{ik(x-c_m(k)t)},$$

where $c_m(k)$ is the numerical wave speed. The dependence of c_m on the wave number k is known as the dispersion relation.

The phase error $e_m(k)$, is defined as the leading term in the relative error between the actual solution $u(x,t)$ and the approximate solution $v(x,t)$:

$$\left|\frac{u(x,t)-v(x,t)}{u(x,t)}\right| = \left|1-e^{ik(c-c_m(k))t}\right| \simeq |k(c-c_m(k))t| = e_m(k).$$

Phase Error Analysis V

Applying phase error analysis to the second-order finite difference scheme

$$\frac{\partial v(x,t)}{\partial t} = -c\frac{v(x+\Delta x,t) - v(x-\Delta x,t)}{2\Delta x},$$
$$v(x,0) = e^{ikx},$$

we obtain the numerical phase speed

$$c_1(k) = c\,\frac{\sin(k\Delta x)}{k\Delta x}.$$

For $\Delta x \ll 1$,

$$c_1(k) = c\left(1 - \frac{(k\,\Delta x)^2}{6} + \mathcal{O}\left((k\,\Delta x)^4\right)\right),$$

confirming the second-order accuracy of the scheme.

Phase Error Analysis VI

For the fourth-order scheme

$$\frac{\partial v(x,t)}{\partial t} = -\frac{c}{12\,\Delta x}\Big(v(x-2\,\Delta x,t) - 8v(x-\Delta x,t) + 8v(x+\Delta x,t) - v(x+2\,\Delta x,t)\Big),$$

we obtain

$$c_2(k) = c\left(\frac{8\sin(k\Delta x) - \sin(2k\Delta x)}{6k\Delta x}\right).$$

For $\Delta x \ll 1$ we recover the approximation

$$c_2(k) = c\left(1 - \frac{(k\Delta x)^4}{30} + \mathcal{O}\big((k\Delta x)^6\big)\right).$$

illustrating the expected fourth-order accuracy.

Phase Error Analysis VII

Denoting $e_1(k,t)$ as the phase error of the second-order scheme and $e_2(k,t)$ as the phase error of the fourth-order scheme, with the corresponding numerical wave speeds $c_1(k)$ and $c_2(k)$, we obtain

$$e_1(k,t) = kct\left|1 - \frac{\sin(k\Delta x)}{k\Delta x}\right|, \quad (8)$$

$$e_2(k,t) = kct\left|1 - \frac{8\sin(k\Delta x) - \sin(2k\Delta x)}{6k\Delta x}\right|. \quad (9)$$

Phase Error Analysis VIII

Introduce

$$p = \frac{N+1}{k} = \frac{2\pi}{k\Delta x} \quad \text{(the number of points per wavelength)}$$
$$\nu = \frac{kct}{2\pi} \quad \text{(the number of periods in time)}$$

Rewriting the phase error in term of p and ν yields

$$e_1(p,v) = 2\pi v \left| 1 - \frac{\sin(2\pi p^{-1})}{2\pi p^{-1}} \right|, \tag{10}$$
$$e_2(p,v) = 2\pi v \left| 1 - \frac{8\sin(2\pi p^{-1}) - \sin(4\pi p^{-1})}{12\pi p^{-1}} \right|. \tag{11}$$

Phase Error Analysis IX

The leading order approximation to (10) is

$$e_1(p,v) \simeq \frac{\pi v}{3}\left(\frac{2\pi}{p}\right)^2, \quad (12)$$

$$e_2(p,v) \simeq \frac{\pi v}{15}\left(\frac{2\pi}{p}\right)^4 \quad (13)$$

from which we immediately observe that the phase error is directly proportional to the number of periods ν i.e., the error grows linearly in time.

Phase Error Analysis X

We arrive at a more straightforward measure of the error of the scheme by introducing $p_m(\epsilon_p, \nu)$ as a measure of the number of points per wavelength required to guarantee a phase error, $e_p \leq \epsilon_p$, after ν periodic for a $2m$-order scheme. Indeed, from (12) we directly obtain the lower bounds

$$p_1(\epsilon, v) \geq 2\pi \sqrt{\frac{v\pi}{3\epsilon_p}} \tag{14}$$

$$p_2(\epsilon, v) \geq 2\pi \sqrt[4]{\frac{\pi v}{15\epsilon_p}} \tag{15}$$

Phase Error Analysis XI

Example

$\epsilon_p = 0.1$ Consider the case in which the desired phase error is $\leq 10\%$. For this relatively large error,

$$p_1 \geq 20\sqrt{v}, \quad p_2 \geq 7\sqrt[4]{v}.$$

$\epsilon = 0.01$ When the desired phase error is within 1%, we have

$$p_1 \geq 64\sqrt{v}, \quad p_2 \geq 13\sqrt[4]{v}.$$

$\epsilon = 10^{-5}$

$$p_1 \geq 643\sqrt{v}, \quad p_2 \geq 43\sqrt[4]{v}$$

Phase Error Analysis XII

Sixth-order method

As an illustration of the general trend in the behavior of the phase error, we give the bound on $p_3(\epsilon_p, \nu)$ for the sixth-order centered-difference scheme as

$$p_3(\epsilon_p, \nu) \geq 2\pi \sqrt[6]{\frac{\pi\nu}{70\epsilon_p}},$$

for which the above special cases become

$$p_3(0.1, \nu) = 5\sqrt[6]{\nu}, \;\; p_3(0.01, \nu) = 8\sqrt[6]{\nu}, \;\; p_3(10^{-5}, \nu) = 26\sqrt[6]{\nu},$$

confirming that when high accuracy is required, a high-order is the optimal choice.

Energy Estimate of IBVP

Consider the initial boundary value problem:

$$
\begin{aligned}
\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \quad & x \in [0,1] \quad t \geq 0 \\
u(x,0) = f(x) \quad & x \in [0,1] \quad t = 0 \\
u(-1,t) = g(t) \quad & x = 0 \quad t \geq 0.
\end{aligned}
$$

Energy Estimate of IBVP

Multiplying u to the partial differential equation, integrating over the domain and applying the boundary condition, we have

$$\int_0^1 u\frac{\partial u}{\partial t}\,dx = -\int_0^1 u\frac{\partial u}{\partial x}\,dx$$

$$\Rightarrow \frac{1}{2}\frac{dE}{dt} = -\left.\frac{1}{2}u^2\right|_0^1 = \frac{1}{2}u^2(0,t) - \frac{1}{2}u^2(1,t), \quad E(t) = \int_0^1 u^2(x,t)\,dx$$

$$\Rightarrow \frac{dE}{dt} = g^2(t) - u^2(1,t) \le g^2(t)$$

Integrating the energy rate equation with respect to time and invoking the initial condition, we have

$$E(t) \le E(0) + \int_0^t g^2(\xi)\,d\xi \le E(0) + t\cdot G, \quad G = \max_{\xi\in[0,t]} g^2(\xi)$$

or $$\int_0^1 u^2(x,t)dx \le \int_0^1 f^2(x)dx + t\cdot G$$

Schemes for the Model Wave Equation

We now construct finite difference schemes for the model wave problem

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} &= 0, \quad x \in [0,1], \quad t \geq 0, \\ u(x,t) &= f(x), \quad x \in [0,1], \quad t = 0, \\ u(0,t) &= g(t), \quad x = 0, \quad t \geq 0. \end{aligned}$$

We define the grid points as

$$x_i = ih, \quad h = 1/N, \quad i = 0, 1, 2, ..., N$$

Denote $v_i(t)$ the approximation of $u(x_i, t)$ at x_i.

Upwind Scheme I

Strongly Enforced Boundary Condition

Consider the scheme

$$\frac{dv_i}{dt} + \frac{v_i - v_{i-1}}{h} = 0, \quad i = 1, 2, \ldots, N$$

$$v_i(0) = f(x_i)$$

$$v_0(t) = g(t)$$

Accuracy: first order in space

Stability

$$\sum_{i=1}^{N} v_i \frac{dv_i}{dt} h = -\sum_{i=1}^{N} v_i (v_i - v_{i-1})$$

Upwind Scheme II

Strongly Enforced Boundary Condition

$$\frac{1}{2}\frac{d}{dt}E_D(t) = \frac{1}{2}\frac{d}{dt}\sum_{i=1}^{N} h\, v_i^2(t) = -\sum_{i=1}^{N} v_i^2 + \sum_{i=1}^{N} v_i v_{i-1}$$

$$= -\frac{1}{2}\sum_{i=1}^{N} v_i^2 - \frac{1}{2}\sum_{i=1}^{N} v_i^2 + \sum_{i=1}^{N} v_i v_{i-1} - \frac{1}{2}\sum_{i=1}^{N} v_{i-1}^2 + \frac{1}{2}\sum_{i=1}^{N} v_{i-1}^2$$

$$= -\frac{1}{2}\sum_{i=1}^{N} v_i^2 - \frac{1}{2}\sum_{i=1}^{N} (v_i - v_{i-1})^2 + \frac{1}{2}\sum_{i=0}^{N-1} v_i^2$$

$$= \frac{1}{2}v_0^2 - \frac{1}{2}v_N^2 - \frac{1}{2}\sum_{i=1}^{N} (v_i - v_{i-1})^2$$

$$= \frac{1}{2}g^2(t) - \frac{1}{2}v_N^2 - \frac{1}{2}\sum_{i=1}^{N} (v_i - v_{i-1})^2 \leq \frac{1}{2}g^2(t)$$

Upwind Scheme III

Strongly Enforced Boundary Condition

Then

$$\frac{dE_D(t)}{dt} \leq g^2(t) \quad \Rightarrow \quad E_D(t) \leq E_D(0) + \int_0^t g^2(\xi)\, d\xi,$$

or explicitly

$$\sum_{i=1}^{N} h\, v_i^2(t) \leq \sum_{i=1}^{N} h\, f_i^2 + \int_0^t g^2(\xi)\, d\xi$$

Recall that for the continuous system we have

$$\int_0^1 u^2(x,t)\, dx \leq \int_0^1 f^2(x)\, dx + \int_0^t g^2(\xi)\, d\xi$$

Upwind Scheme IV

Weakly Enforced Boundary Condition

Consider the scheme

$$
\begin{aligned}
\frac{dv_i}{dt} + \frac{v_i - v_{i-1}}{h} &= 0 \quad i = 1, 2, \ldots, N \\
\frac{dv_0}{dt} + \frac{v_1 - v_0}{h} &= -\tau(v_0 - g(t)) \\
v_i(0) = f(x_i) & \quad i = 0, 1, 2, \ldots, N
\end{aligned}
$$

τ: free parameter.

(1) $\tau \to 0$ (the scheme behaves like the PDE)

$$\frac{dv_0}{dt} + \frac{v_1 - v_0}{h} = -\tau(v_0 - g(t)) \to 0$$

(2) $\tau \to \infty$ (the scheme behaves like the boundary condition)

$$v_0 - g(t) = \frac{1}{\tau}\left(\frac{dv_0}{dt} + \frac{v_1 - v_0}{h}\right) \to 0$$

Upwind Scheme V

Weakly Enforced Boundary Condition

Accuracy: 1st order in space
Stability: We need to check whether the discrete energy

$$E_D(t) = \sum_{i=0}^{N} v_i^2(t)h$$

has an estimate bounded by the prescribed data and τ.

Upwind Scheme VI

Weakly Enforced Boundary Condition

Multiplying $v_i h$ to the scheme and summing up the resultant equations, we have

$$\sum_{i=0}^{N} v_i \frac{dv_i}{dt} h = -\sum_{i=1}^{N} v_i(v_i - v_{i-1}) - v_0(v_1 - v_0) - \tau h v_0(v_0 - g(t))$$

Recall

$$-\sum_{i=1}^{N} v_i(v_i - v_{i-1}) = \frac{1}{2}v_0^2 - \frac{1}{2}v_N^2 - \frac{1}{2}\sum_{i=1}^{N}(v_i - v_{i-1})^2$$

$$\begin{aligned}\frac{1}{2}\frac{dE_D}{dt} &= \frac{1}{2}v_0^2 - \frac{1}{2}v_N^2 - \frac{1}{2}\sum_{i=2}^{N}(v_i - v_{i-1})^2 - \frac{1}{2}(v_1^2 - 2v_1v_0 + v_0^2) \\ &\quad - v_0v_1 + v_0^2 - \tau h v_0(v_0 - g(t)) \\ &= -\frac{1}{2}v_N^2 - \frac{1}{2}\sum_{i=2}^{N}(v_i - v_{i-1})^2 - \frac{1}{2}v_1^2 + v_0^2(1 - \tau h) + \tau h v_0 g(t)\end{aligned}$$

Upwind Scheme VII

Weakly Enforced Boundary Condition

$$\frac{dE_D}{dt} = -v_N^2 - v_1^2 - \sum_{i=2}^{N}(v_i - v_{i-1})^2 + 2(1-\tau h)\left(v_0 + \frac{\tau h g(t)}{2(1-\tau h)}\right)^2 - \frac{\tau^2 h^2 g^2(t)}{2(1-\tau h)}$$

Take $1 - \tau h < 0 \;\Rightarrow\; \tau > \frac{1}{h}$

$$\frac{dE_D}{dt} \leq \frac{\tau^2 h^2 g^2(t)}{2(\tau h - 1)}$$

If $\tau h = 2$

$$\frac{dE_D}{dt} \leq 2g^2(t) \Rightarrow E_D(t) \leq E_D(0) + 2\int_0^t g^2(\xi)\, d\xi$$

Upwind Scheme VII

Weakly Enforced Boundary Condition

Remarks:

1. By properly choosing the value of the parameter τ the scheme has a bounded energy estimate, implying stability.
2. Since $\tau > 1/h$ as $h \to 0$, $\tau \to \infty$, the equation at $x_0 = 0$,

$$v_0 - g(t) = \frac{1}{\tau}\left(\frac{dv_0}{dt} + \frac{v_1 - v_0}{h}\right)$$

converges to the boundary condition as $h \to 0$.

Central Difference Scheme I

Strongly Enforced Boundary Condition

Consider the scheme

$$
\begin{aligned}
\frac{dv_i}{dt} + \frac{v_{i+1} - v_{i-1}}{2h} &= 0 \quad i = 1, 2, \ldots, N \\
\frac{dv_N}{dt} + \frac{v_N - v_{N-1}}{h} &= 0 \\
v_0(t) &= g(t) \\
v_i(0) &= f(x_i)
\end{aligned}
$$

Accuracy: 2nd order at interior points
1st order at boundary points
globally second order[1]

[1] Gustafsson (1975) *The Convergence Rate for Difference Approximations to Mixed Initial Boundary Value Problems*

Central Difference Scheme II

Strongly Enforced Boundary Condition

Let $c_N = \frac{1}{2}$, $c_i = 1$ for $i \neq N$. Multiplying $c_i v_i h$ to the scheme we have

$$\sum_{i=1}^{N} c_i v_i \frac{dv_i}{dt} h = -\sum_{i=1}^{N-1} v_i \left(\frac{v_{i+1} - v_{i-1}}{2} \right) - v_N \left(\frac{v_N - v_{N-1}}{2} \right)$$

$$\Rightarrow \frac{d}{dt} \sum_{i=1}^{N} c_i v_i^2 h = -\sum_{i=1}^{N-1} v_i (v_{i+1} - v_{i-1}) - v_N (v_N - v_{N-1})$$

$$= -\sum_{i=1}^{N-1} v_i v_{i+1} + \sum_{i=1}^{N-1} v_i v_{i-1} - v_N^2 + v_N v_{N-1}$$

Central Difference Scheme III

Strongly Enforced Boundary Condition

$$\begin{aligned}
\frac{d}{dt}\sum_{i=1}^{N} c_i v_i^2 h &= -\sum_{i=1}^{N-1} v_i v_{i+1} + \sum_{i=1}^{N} v_i v_{i-1} - v_N^2 \\
&= -\sum_{i=1}^{N-1} v_i v_{i+1} + \sum_{i=0}^{N-1} v_{i+1} v_i - v_N^2 \\
&= v_1 v_0 - v_N^2
\end{aligned}$$

We have no idea on bounding the discrete energy rate.

Central Difference Scheme IV

Weakly Enforced Boundary Condition

Consider the scheme

$$\begin{aligned}\frac{dv_i}{dt} + \frac{v_{i+1} - v_{i-1}}{2h} &= 0 \quad i = 1, 2, \ldots, N-1\\ \frac{dv_N}{dt} + \frac{v_N - v_{N-1}}{h} &= 0\\ \frac{dv_0}{dt} + \frac{v_1 - v_0}{h} &= -\tau(v_0 - g(t))\end{aligned}$$

$$\frac{1}{2}v_0\frac{dv_0}{dt}h + \sum_{i=1}^{N-1} v_i\frac{dv_i}{dt}h + \frac{1}{2}v_N\frac{dv_N}{dt}h$$

$$= -\sum_{i=1}^{N-1}\frac{v_i(v_{i+1} - v_{i-1})}{2} - \frac{v_N(v_N - v_{N-1})}{2} - \frac{(v_1 - v_0)v_0}{2} - \frac{\tau h v_0}{2}(v_0 - g(t))$$

Central Difference Scheme V

Weakly Enforced Boundary Condition

$$\frac{d}{dt}\sum_{i=0}^{N} c_i v_i^2 h = v_1 v_0 - v_N^2 - v_1 v_0 + v_0^2 - \tau h v_0^2 + \tau h v_0 g(t)$$

$$= -v_N^2 + (1-\tau h) v_0^2 + \tau h v_0 g(t)$$

$$= -v_N^2 + (1-\tau h)\left[v_0^2 + \frac{\tau h}{1-\tau h} v_0 g(t) + \left(\frac{\tau h}{2(1-\tau h)}\right)^2 g^2(t)\right]$$

$$- (1-\tau h)\left(\frac{\tau h}{2(1-\tau h)}\right)^2 g^2(t)$$

If $\tau h = 2$ $$\frac{d}{dt}\sum_{i=0}^{N} c_i v_i^2 h = -v_N^2 - (v_0 - g(t))^2 + g^2(t) \le g^2(t)$$

$$\Rightarrow \sum_{i=0}^{N} c_i v_i^2(t) h \le \sum_{i=0}^{N} c_i f_i^2 h + \int_0^t g^2(\xi)\, d\xi$$

Central Difference Scheme VI

Weakly Enforced Boundary Condition

Remarks:

1. By properly choosing the value of τ the scheme has a discrete energy estimate that is bounded by the prescribed data f_i and $g(t)$ and it is independent of N. This implies stability.
2. A scheme is stable at the semi-discrete level does not ensure the stability of the scheme at the fully discrete level. This is because the stability condition at the semidiscrete level is only a necessary condition.
3. The advantage of using semidiscrete analysis is that we can check whether the boundary closure will cause instability, and possibly fix the problem.

Central Difference Scheme VI

Matrix Vector Representation

Let $\boldsymbol{v} = \left[v_0(t) \; v_1(t) \; \cdots \; v_N(t) \right]$ The scheme has a matrix-vector representation as

$$\frac{d}{dt}\begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ \vdots \\ v_{N-1} \\ v_N \end{bmatrix} + \frac{1}{2h}\begin{bmatrix} -2 & 2 & 0 & & & \\ -1 & 0 & 1 & \ddots & & \\ 0 & -1 & 0 & 1 & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & \\ & & & -1 & 0 & 1 \\ & & & 0 & -2 & 2 \end{bmatrix}\begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ \vdots \\ v_{N-1} \\ v_N \end{bmatrix} = \begin{bmatrix} -\tau(v_0 - g(t)) \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

or

$$\frac{d}{dt}\boldsymbol{v}(t) + \boldsymbol{D}\boldsymbol{v}(t) = -\tau(v_0 - g(t))\boldsymbol{e}_0, \quad \boldsymbol{e}_0 = \left[1\ 0\ 0\ \cdots\ 0\ 0\right]^T$$

Define $\boldsymbol{H} = \text{diag}(\frac{1}{2},\ 1,\ 1,\ \cdots,\ 1,\ \frac{1}{2})$

$$\boldsymbol{v}^T\boldsymbol{H}\frac{d\boldsymbol{v}}{dt}h + \boldsymbol{v}^T\,\boldsymbol{H}\,\boldsymbol{D}\,h\,\boldsymbol{v}(t) = -\tau(v_0 - g(t))\boldsymbol{v}^T\,\boldsymbol{H}\,\boldsymbol{e}_0\,h$$

Central Difference Scheme VII

Interesting Property: Summation-by-Parts Rule

Observe

$$
\boldsymbol{H}\,\boldsymbol{D}\,h = \frac{1}{2}\begin{bmatrix} 1/2 & 0 & & & \\ 0 & 1 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 1 & 0 \\ & & & 0 & 1/2 \end{bmatrix}\begin{bmatrix} -2 & 2 & 0 & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -2 & 2 \end{bmatrix}
$$

$$
= \frac{1}{2}\begin{bmatrix} -1 & 1 & 0 & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 & & \\ -1 & 0 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ & & & -1 & 0 \end{bmatrix}
$$

$$
= \boldsymbol{Q}_S + \boldsymbol{Q}_A
$$

So

$$
\boldsymbol{v}^T \boldsymbol{H}\boldsymbol{D}h\boldsymbol{v} = \boldsymbol{v}^T \boldsymbol{Q}_S \boldsymbol{v} + \boldsymbol{v}^T \boldsymbol{Q}_A \boldsymbol{v} = \frac{1}{2}(-v_0^2 + v_N^2)
$$

Central Difference Scheme VII

Interesting Property: Summation-by-Parts Rule

Hence

$$\boldsymbol{v}^T \boldsymbol{H} \frac{d\boldsymbol{v}}{dt} h = -\boldsymbol{v}^T \boldsymbol{H}\, \boldsymbol{D}\, h\, \boldsymbol{v} - \tau (v_0 - g(t)) \boldsymbol{v}^T \boldsymbol{H}\, \boldsymbol{e_0}\, h$$

$$\frac{1}{2}\frac{d}{dt}\left[\boldsymbol{v}^T \boldsymbol{H} \boldsymbol{v} h\right] = \frac{1}{2} v_0^2 - \frac{1}{2} v_N^2 - \frac{\tau h}{2} v_0 (v_0 - g(t))$$

The rule

$$\boldsymbol{v}^T \boldsymbol{H}\boldsymbol{D} h \boldsymbol{v} = \boldsymbol{v}^T \boldsymbol{Q}_S \boldsymbol{v} + \boldsymbol{v}^T \boldsymbol{Q}_A \boldsymbol{v} = \frac{1}{2}(-v_0^2 + v_N^2)$$

in fact mimic the action of

$$\int_0^1 u(x) \frac{\partial u(x)}{\partial x}\, dx = \frac{1}{2} u(1) - \frac{1}{2} u(0)$$

From Low-Order to High-Order Methods

- To construct a high-order scheme we basically seek a differentiation matrix D (resulting from central and one-sided difference scheme) and a positive definite matrix H such that a rule similar to

$$\mathbf{v}^T \boldsymbol{H}\boldsymbol{D}h\mathbf{v} = \mathbf{v}^T \boldsymbol{Q}_S \mathbf{v} + \mathbf{v}^T \boldsymbol{Q}_A \mathbf{v} = \frac{1}{2}(-v_0^2 + v_N^2)$$

  exists.
- Notice that a summation-by-parts rule is only for estimating the energy of the system. To stably impose boundary conditions we still use the penalty methodology.

Basic Concepts and Notations I

Let $\mathsf{I} = [0,1]$. Consider two functions $f(x)$ and $g(x)$ defined on I. We define the continuous L_2 inner product and norm for functions over I as

$$(f,g) = \int_{\mathsf{I}} f\,g\,dx, \quad ||f||_{\mathsf{I}}^2 = (f,f)$$

Consider $\mathsf{I}^2 = [0,1] \times [0,1]$. The continuous L_2 inner product and norm for functions over I^2 are defined as

$$(f,g) = \int_{\mathsf{I}^2} f\,g\,dx\,dy, \quad ||f||_{\mathsf{I}^2}^2 = (f,f).$$

Likewise, for functions defined on $\mathsf{I}^3 = [0,1] \times [0,1] \times [0,1]$ we denote the continuous L_2 inner product and norm for functions over I^3 as

$$(f,g) = \int_{\mathsf{I}^3} f\,g\,dx\,dy\,dz, \quad ||f||_{\mathsf{I}^3}^2 = (f,f).$$

Basic Concepts and Notations II

We introduce a set of uniformly spaced grid points:

$$x_i = ih, \quad i = 0, 1, 2, ..., L, \quad h = 1/L,$$

where h is the grid distance. Consider two vectors, $\boldsymbol{f}, \boldsymbol{g} \in \mathrm{V}_{L+1}$, explicitly given by

$$\boldsymbol{f} = [f_0,\, f_1,\, ...,\, f_L]^T, \quad \boldsymbol{g} = [g_0,\, g_1,\, ...,\, g_L]^T,$$

We define a weighted discrete L_2 inner product and norm, with respect to the step size h and the matrix $\boldsymbol{M}$, for vectors as

$$(\boldsymbol{f}, \boldsymbol{g})_{h,M} = h\boldsymbol{f}^T\boldsymbol{M}\boldsymbol{g}, \quad ||f||^2_{h,M} = (\boldsymbol{f}, \boldsymbol{f})_{h,M}.$$

If $\boldsymbol{M}$ is an identity matrix then

$$(\boldsymbol{f}, \boldsymbol{g})_h = h\boldsymbol{f}^T\boldsymbol{g}, \quad ||f||^2_h = (\boldsymbol{f}, \boldsymbol{f})_h.$$

Basic Concepts and Notations III

To numerically approximate a function u and its derivative du/dx, we consider the difference approximation of the form

$$\boldsymbol{P}\boldsymbol{v}_x = h^{-1}\boldsymbol{Q}\boldsymbol{v}, \quad \text{or} \quad \boldsymbol{v}_x = \boldsymbol{D}\boldsymbol{v} = h^{-1}\boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{v}, \tag{16}$$

where

$$\boldsymbol{v} = [v_0,\, v_1,\, \ldots,\, v_L]^T, \quad \boldsymbol{v}_x = [v_{x0},\, v_{x1},\, \ldots,\, v_{xL}]^T,$$

denote the numerical approximations of u and u' evaluated at the grid points, and $\boldsymbol{D}$, $\boldsymbol{P}$, $\boldsymbol{Q} \in \mathrm{M}_{L+1}$.

Basic Concepts and Notations IV

Let $\boldsymbol{u}$ and $\boldsymbol{u}_x$ denote vectors with components being, respectively, the collocated values of the functions u and du/dx at the grid points, i.e.,

$$\boldsymbol{u} = [u(x_0),\, u(x_1),\, \ldots,\, u(x_L)]^T, \quad \boldsymbol{u}_x = \left[\frac{du(x_0)}{dx}, \frac{du(x_1)}{dx}, \ldots, \frac{du(x_L)}{dx}\right]^T.$$

The truncation error $\boldsymbol{t}_e$ of the scheme Eq.(16) is defined by

$$\boldsymbol{P}\boldsymbol{t}_e = \boldsymbol{P}\boldsymbol{u}_x - h^{-1}\boldsymbol{Q}\boldsymbol{u},$$

and $|\boldsymbol{t}_e| = O(h_x^{\alpha}, h_x^{\beta})$ where α and β are the convergence rates of the approximation at interior and boundary grid points, respectively.

Basic Concepts and Notations V

We devise implicit difference methods for approximating the differential operator d/dx by constructing a special class of $\boldsymbol{P}$ and $\boldsymbol{Q}$ satisfying the following properties;

SBP1: The matrix $\boldsymbol{P}$ is symmetric positive definite.

SBP2: The matrix $\boldsymbol{Q}$ is nearly skew-symmetric and satisfies the constraint

$$\frac{\boldsymbol{Q}+\boldsymbol{Q}^T}{2} = \mathrm{diag}(q_{00}, 0, \ldots, 0, q_{LL}), \quad q_{00} < 0, \quad q_{LL} = -q_{00} > 0.$$

where q_{00} and q_{LL} are the upper most and lower most diagonal elements of $\boldsymbol{Q}$.

Summation-by-Parts Rule in 1D Space I

Lemma (Summation-by-Parts)

Consider the difference operator $\boldsymbol{D} = h^{-1}\boldsymbol{P}^{-1}\boldsymbol{Q}$ where $\boldsymbol{P}$ and $\boldsymbol{Q}$ satisfy **SBP1** and **SBP2**, respectively. We have

$$(\boldsymbol{v}, \boldsymbol{D}\boldsymbol{v})_{h,P} = (\boldsymbol{v}, h^{-1}\boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{v})_{h,P} = q_{00}v_0^2 + q_{LL}v_L^2,$$

for $\boldsymbol{v} \in \mathrm{V}_{L+1}$.

Summation-by-Parts Rule in 1D Space II

Proof.
First we rewrite the inner product as

$$(\boldsymbol{v}, \boldsymbol{D}\boldsymbol{v})_{h,P} = \left(\boldsymbol{v}, h^{-1}\boldsymbol{Q}\boldsymbol{v}\right)_h = \boldsymbol{v}^T\boldsymbol{Q}^S\boldsymbol{v} + \boldsymbol{v}^T\boldsymbol{Q}^A\boldsymbol{v}$$

where $\boldsymbol{Q}^S = (\boldsymbol{Q} + \boldsymbol{Q}^T)/2$ and $\boldsymbol{Q}^A = (\boldsymbol{Q} - \boldsymbol{Q}^T)/2$ are, respectively, the symmetric and anti-symmetric parts of the matrix $\boldsymbol{Q}$. Notice that $\boldsymbol{v}^T\boldsymbol{Q}^A\boldsymbol{v} = 0$ since $\boldsymbol{Q}^A$ is antisymmetric. Thus, we have

$$(\boldsymbol{v}, \boldsymbol{D}\boldsymbol{v})_{h,P} = \boldsymbol{v}^T\boldsymbol{Q}^{\boldsymbol{S}}\boldsymbol{v} = q_{00}v_0^2 + q_{LL}v_L^2,$$

where the last equality is due to **SPB2**. This completes the proof. □

One Dimensional Advection Equation I

Consider the advection equation

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x \in \mathsf{I}, \quad t \geq 0, \tag{17}$$

with the initial condition

$$u(x,0) = f(x), \quad x \in \mathsf{I}, \tag{18}$$

and the boundary condition

$$u(0,t) = g(t), \quad t \geq 0. \tag{19}$$

One Dimensional Advection Equation II

Equation (17) leads to an energy rate

$$\frac{d}{dt}||u||_I^2 = g^2(t) - u^2(1,t),$$

For well-posed analysis it is sufficient to consider $g = 0$, and we obtain an energy estimate

$$||u(x,t)||_I^2 \leq ||u(x,0)||_I^2 = ||f(x)||_I^2. \tag{20}$$

One Dimensional Advection Equation III

Consider a equally spaced partition:

$$x_i = ih, \quad h = 1/L.$$

With v_i denoting the approximation of $u(x_i)$, we seek a numerical solution $\boldsymbol{v}$ of the form

$$\boldsymbol{v}(t) = [v_0(t),\, v_1(t),\, ...,\, v_L(t)]^T\,,$$

which satisfies the semidiscrete scheme

$$\frac{d\boldsymbol{v}}{dt} + h^{-1}\boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{v} = h^{-1}\tau q_{00}(v_0 - g(t))\boldsymbol{P}^{-1}\boldsymbol{e}_0, \tag{21a}$$

$$\boldsymbol{v}(0) = \boldsymbol{f} = [f(x_0), f(x_1), ..., f(x_L)]^T, \tag{21b}$$

where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are defined by Eq.(16), and $\boldsymbol{e}_0 = [1,\, 0, ...,\, 0]^T$.

One Dimensional Advection Equation IV

Theorem

Assume that there exists a smooth solution to the one dimensional wave problem described by Eqs.(17-19). Then Eq.(21a) is stable at the semi-discrete level provided that

$$\tau \geq 1.$$

Moreover, $\boldsymbol{v}(t)$ satisfies the estimate

$$||\boldsymbol{v}(t)||_{h,P}^2 \leq ||\boldsymbol{f}||_{h,P}^2.$$

One Dimensional Advection Equation V

Proof.
Multiplying $h\boldsymbol{v}^T\boldsymbol{P}$ to the scheme and invoking Lemma 11 we obtain

$$\frac{1}{2}\frac{d}{dt}||\boldsymbol{v}||^2_{h,P} = -(q_{00}v_0^2 + q_{LL}v_L^2) + \tau v_0 q_{00}(v_0 - g(t)).$$

For the stability analysis, it is sufficient enough to consider the scheme subject to $g(t) = 0$. Hence,

$$\frac{1}{2}\frac{d}{dt}||\boldsymbol{v}||^2_{h,P} = q_{00}(\tau - 1)v_0^2 - q_{LL}v_L^2 \leq q_{00}(\tau - 1)v_0^2,$$

where the last inequality results from $q_{LL} > 0$ demanded by **SBP2**. Recall that $q_{00} < 0$.

One Dimensional Advection Equation VI

So, taking $\tau \geq 1$ immediately yields a non-increasing energy rate

$$\frac{1}{2}\frac{d}{dt}||\boldsymbol{v}||_{h,P}^2 \leq 0,$$

which leads to the estimate

$$||\boldsymbol{v}(t)||_{h,P}^2 \leq ||\boldsymbol{v}(0)||_{h,P}^2 = ||\boldsymbol{f}||_{h,P}^2.$$

Thus, the scheme is stable. This completes the proof. □

Important Reference: Bo Strand (1994)

Difference method

$$\boldsymbol{v}_x = \frac{1}{h}\boldsymbol{Q}\,\boldsymbol{v},$$

Summation-By-Parts

$$\boldsymbol{v}\boldsymbol{H}\boldsymbol{v}_x = \frac{1}{h}\boldsymbol{v}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{v} = \frac{1}{2}(v_N^2 - v_0^2)$$

Fourth-Order (Class 1: Accuracy $\alpha = 4, \beta = 3$)

$$H = \begin{bmatrix} \begin{bmatrix} H_U \end{bmatrix} & & \\ & \ddots & \\ & & \begin{bmatrix} H_L \end{bmatrix} \end{bmatrix}$$

$$hQ = \begin{bmatrix} \otimes & \times & \times & \times & & & & & & \\ \times & \otimes & \times & \times & \times & \times & & & & \\ \times & \times & \otimes & \times & \times & \times & & & & \\ \times & \times & \times & \otimes & \times & \times & \times & & & \\ \times & \times & \times & \times & \otimes & \times & \times & & & \\ & & & \times & \times & \otimes & \times & \times & & \\ & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Bo Strand (Class 2: Accuracy $\alpha = 6, \beta = 3$)

$$H = \text{diag}\left[h_{00},\ h_{11},\ h_{22},\ \ldots\ldots\right]$$

$$hQ = \begin{bmatrix}
\otimes & \times & \times & \times & \times & & & & & & & \\
\times & \otimes & \times & \times & \times & \times & & & & & & \\
\times & \times & \otimes & \times & \times & \times & & & & & & \\
\times & \times & \times & \otimes & \times & \times & \times & & & & & \\
\times & \times & \times & \times & \otimes & \times & \times & \times & & & & \\
 & \times & \times & \times & \times & \otimes & \times & \times & \times & & & \\
 & & & \times & \times & \times & \otimes & \times & \times & \times & & \\
 & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots
\end{bmatrix}$$

Other References

- Bo Strand (1998) *Numerical studies of hyperbolic IBVP with high-order finite difference operators satisfying a summation by parts rule*
- Jan Nordström, Mark H. Carpenter (2001) *High-Order Finite Difference Methods, Multidimensional Linear Problems, and Curvilinear Coordinates*
- Ken Mattsson, Jan Nordström (2004), *Summation by parts operators for finite difference approximations of second derivatives*
- Magnus Saärd, Jan Nordström (2006), *On the order of accuracy for difference approximations of initial- boundary value problems*

Carpenter, Gottlieb, Abarbanel (1994)

Difference Method:

$$\boldsymbol{P}\,\boldsymbol{v}_x = \boldsymbol{Q}\boldsymbol{v}$$

Summation-by-parts rule:

$$\boldsymbol{v}^T\boldsymbol{H}\boldsymbol{P}\boldsymbol{v}_x = \boldsymbol{v}^T\boldsymbol{H}\boldsymbol{Q}\boldsymbol{v} = \frac{1}{2}\left(v_N^2 - v_0^2\right)$$

$\boldsymbol{P}$ is tridiagonal (implicit method) and $\boldsymbol{HP}$ is symmetric positive definite.

Class 1: Accuracy $\alpha = 4,\ \beta = 3$

Class 2: Accuracy $\alpha = 6,\ \beta = 5$ ($\boldsymbol{H}$ is a identity matrix)