Chapter 3

Basic Theoretical Concepts for Time-Dependent Problems

3.1 Consistency, convergence, and stability

Three fundamental properties that every finite difference approximation of a partial differential equation should possess are consistency, convergence, and stability. Roughly speaking, consistency implies that the finite difference equation is a good approximation of the partial differential equation, convergence implies that the solution of the difference equation approaches the solution of the partial differential equation as the computational mesh is refined, and stability implies that the solution of the difference equation is not too sensitive to small perturbations in, say, the initial data. These properties are often difficult to verify for realistic problems, but they can be explained and illustrated quite easily using difference schemes for some simple model problems, such as (2.2.2b) for the kinematic wave equation (2.2.1) and (2.3.3) for the one-dimensional heat conduction equation (2.3.1).

**Definition 3.1.1.** Consider a differential equation $\mathcal{P}u = 0$ and a finite difference approximation of it $\mathcal{P}_\Delta U^n_j = 0$. Let $v(x, t)$ be any smooth function, then the local discretization or local truncation error is

$$
\tau^n_j = \mathcal{P}v(j \Delta x, n \Delta t) - \mathcal{P}_\Delta v(j \Delta x, n \Delta t).
$$

**Remark 1.** $\mathcal{P}$ is a differential operator, e.g.,

$$
\mathcal{P}v = v_t - \sigma v_{xx}
$$
for the heat conduction equation (2.3.1a). Similarly, $\mathcal{P}_\Delta$ is a finite difference operator; thus, the forward time-centered space difference operator (2.3.2b) for the heat conduction operator is

$$\mathcal{P}_\Delta v = \frac{v_j^{n+1} - v_j^n}{\Delta t} - \sigma \frac{v_{j+1}^{n+1} - 2v_j^n + v_{j-1}^n}{\Delta x^2}.$$ 

Remark 2. The function $v$ is often regarded as the exact solution $u$ of the partial differential equation. This is convenient but not necessary; thus, $v$ may be any smooth function. This gives us a way of defining the local discretization error even when the solution of the partial differential equation has singularities.

Example 3.1.1. The local discretization error of the forward time-centered space difference approximation (2.3.2b) for the heat conduction equation (2.3.1a) is

$$\tau_j^n = (v_t - \sigma v_{xx})_j^n - \left(\frac{v_j^{n+1} - v_j^n}{\Delta t} - \sigma \frac{v_{j+1}^{n+1} - 2v_j^n + v_{j-1}^n}{\Delta x^2}\right)$$

or

$$\tau_j^n = (v_t)_j^n - \frac{v_j^{n+1} - v_j^n}{\Delta t} = \sigma \left[(v_{xx})_j^n - \frac{v_{j+1}^{n+1} - 2v_j^n + v_{j-1}^n}{\Delta x^2}\right]. \quad (3.1.2a)$$

Using the Taylor’s series (2.1.8, 2.1.9)

$$\tau_j^n = -\frac{\Delta t}{2} (v_{tt})_j^{n+\theta} + \frac{\sigma \Delta x^2}{12} (v_{xxxx})_j^{n+\xi}. \quad (3.1.2b)$$

If $v$ is the solution $u$ of the partial differential equation

$$\tau_j^n = -\frac{\Delta t}{2} (u_{tt})_j^{n+\theta} + \frac{\sigma \Delta x^2}{12} (u_{xxxx})_j^{n+\xi}. \quad (3.1.2c)$$

The local error is often used in place of the local discretization error.

Definition 3.1.2. The local error is the difference $u_j^{n+1} - U_j^{n+1}$ between solutions of the partial differential equation and its finite difference approximation assuming that no errors were committed prior to time level $n+1$.

Example 3.1.2. The forward time-centered space difference approximation (2.3.2b) of the heat conduction equation (2.3.1a) is

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} - \sigma \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{\Delta x^2} = 0.$$
According to Definition 3.1.2, we commit no errors prior to time level \( n+1 \); thus, \( U^n_j = u^n_j \), \( j = 0, 1, \ldots, J \), and
\[
\frac{U^{n+1}_j - u^n_j}{\Delta t} - \sigma \frac{u^n_{j+1} - 2u^n_j + u^n_{j-1}}{\Delta x^2} = 0.
\]
Using (3.1.2a) with \( v \) replaced by \( u \) and the heat conduction equation (2.3.1a) reveals
\[
\tau^n_j = -\frac{u^{n+1}_j - u^n_j}{\Delta t} + \sigma \frac{u^n_{j+1} - 2u^n_j + u^n_{j-1}}{\Delta x^2}.
\]
Combining the previous two relations, we obtain
\[
u^{n+1}_j - U^{n+1}_j = -\Delta t \tau^n_j.
\]
Thus, the local error is the negative of the product of the local discretization error and the time step. This relationship between local and local discretization errors generally holds for explicit finite difference schemes. Explicit difference schemes are the type that we have been discussing. The solution at time level \( n+1 \) is explicitly obtained from known solution values at previous time levels. The alternative, an implicit finite difference scheme, couples other unknowns at time level \( n+1 \) to \( U^n_j \). Such schemes will be defined in Chapter 4.

**Definition 3.1.3.** A finite difference scheme \( \mathcal{P}_n U^n_j = 0 \) is consistent with a partial differential equation \( \mathcal{P}u = 0 \) if the local discretization error tends to zero as \( \Delta x \to 0 \) and \( \Delta t \to 0 \).

**Remark 3.** An equivalent definition expressed in terms of the local error would be that a difference scheme is consistent if the local error tends to zero at least to \( O(\Delta t) \) as \( \Delta x \to 0 \) and \( \Delta t \to 0 \).

**Example 3.1.3.** Using (3.1.2b), we see that the forward time-centered space scheme for the heat conduction equation (2.3.1a) is consistent.

Let’s emphasize the time dependence of the partial differential equation by switching notation from \( \mathcal{P}u = 0 \) to
\[
u_t = \mathcal{L}u,
\]
where \( \mathcal{L} \) is a spatial differential operator, e.g., \( \mathcal{L}u \equiv \sigma u_{xx} \) for the heat conduction equation (2.3.1a) and \( \mathcal{L}u \equiv -au_x \) for the kinematic wave equation (2.2.1a).
The solution \( u \) of the partial differential equation can be regarded as being an element of a function space. We’ll not be too precise about the nature of such function spaces at the moment, but simply note that there is a positive scalar, called a norm, that is associated with many function spaces and which is used to measure the “size” of member functions and the “distances” between them.

**Definition 3.1.4.** The norm \( \|u\| \) of a function \( u \) is a scalar that satisfies:

1. \( \|u\| \geq 0 \) and \( \|u\| = 0 \) if and only if \( u = 0 \),
2. \( \|\alpha u\| = |\alpha|\|u\| \) for any constant \( \alpha \), and
3. \( \|u + v\| \leq \|u\| + \|v\| \).

**Remark 4.** Condition 2 is called the condition of homogeneity and Condition 3 is the triangular inequality.

The two norms that are most suited to our purposes are the *maximum norm*

\[
\|u\|_\infty \equiv \max_{x \in \Omega} |u(x, t)|, \tag{3.1.5a}
\]

and the *Euclidean or \( L^2 \) norm*

\[
\|u\|_2 \equiv \left( \int_{\Omega} u(x, t)^2 \, dx \right)^{1/2}. \tag{3.1.5b}
\]

In both instances, \( \Omega \) is the spatial domain of the partial differential equation. As noted in Section 1.2, norms often give us useful estimates of the growth or decay of solutions with time.

While solutions \( u \) of the partial differential equation are regarded as elements of a function space, we may think of finite difference solutions as elements of a linear vector space. To this end, we’ll collect all of the unknowns at a given time level \( n \) into a vector \( U^n \). For example, the unknowns at the time level \( n \) for the forward time-centered space scheme (2.3.2b) for the heat-conduction problem (2.3.1) are \( U^n_1, U^n_2, \ldots, U^n_{J-1} \); hence, we may define the vector

\[
U^n \equiv [U^n_1, U^n_2, \ldots, U^n_{J-1}]^T. \tag{3.1.6a}
\]
The superscript $T$ denotes transposition. Similarly, the unknowns at a time level $n$ for a problem that is periodic in $x$ on an interval $(0, X)$ might be $U^n_0, U^n_1, \ldots, U^n_{J-1}$ and we would introduce the vector

$$ U^n \equiv [U^n_0, U^n_1, \ldots, U^n_{J-1}]^T. \quad (3.1.6b) $$

Using this notation, we’ll write the finite difference approximation as a matrix equation of the form

$$ U^{n+1} = L_\Delta U^n. \quad (3.1.7a) $$

An inhomogeneous difference equation would have the form

$$ U^{n+1} = L_\Delta U^n + f^n. \quad (3.1.7b) $$

where the vector $f^n$ is independent of $U^n$. We’ll ignore this complication at present.

*Example 3.1.4.* The matrix form of the forward time-centered space difference scheme (2.3.3) for the heat-conduction problem (2.3.1) is

$$
\begin{bmatrix}
U^{n+1}_1 \\
U^{n+1}_2 \\
\vdots \\
U^{n+1}_{J-1}
\end{bmatrix} =
\begin{bmatrix}
1 - 2r & r & & & \\
r & 1 - 2r & r & & \\
& r & 1 - 2r & r & \\
& & & \ddots & \\
& & & r & 1 - 2r
\end{bmatrix}
\begin{bmatrix}
U^n_1 \\
U^n_2 \\
\vdots \\
U^n_{J-1}
\end{bmatrix}. \quad (3.1.8)
$$

Elements not shown in the above matrix are zero; thus, $L_\Delta$ is a $J - 1 \times J - 1$ tridiagonal matrix.

You may recall the definition of a linear vector space [4].

**Definition 3.1.5.** $V$ is a linear vector space if

1. $U, V \in V$ then $U + V \in V$ and

2. $\alpha$ is a scalar constant and $U \in V$ then $\alpha U \in V$.

The properties of a norm on a vector space are identical to those for a function space, but we repeat them for completeness.

**Definition 3.1.6.** The norm $\|U\|$ of a vector $U$ is a scalar that satisfies:
1. \( \|U\| \geq 0 \) and \( \|U\| = 0 \) if and only if \( U = 0 \),

2. \( \|\alpha U\| = |\alpha| \|U\| \) for any constant \( \alpha \), and

3. \( \|U + V\| \leq \|U\| + \|V\| \).

Again, the two norms that most suit our purposes are the maximum norm

\[
\|U^n\|_\infty \equiv \max_j |U^n_j|, \tag{3.1.9a}
\]

and the Euclidean or \( \mathcal{L}^2 \) norm

\[
\|U^n\|_2 \equiv \left( \sum_j (U^n_j)^2 \right)^{1/2}. \tag{3.1.9b}
\]

The range on \( j \) is over all components of the vector \( U^n \).

Matrix norms provide estimates of the “size” of the discrete operator \( L_\Delta \) and the growth or decay of \( U^n \) with \( n \). They are typically defined by their effect on vector norms.

**Definition 3.1.7.** The norm \( \|A\| \) of a matrix \( A \) induced by a vector norm is the scalar

\[
\|A\| = \max_{\|z\| \neq 0} \frac{\|Az\|}{\|z\|} = \max_{\|z\| = 1} \|Az\|. \tag{3.1.10}
\]

The matrix norm satisfies the properties of a vector norm as well as

\[
\|AB\| \leq \|A\| \|B\|, \tag{3.1.11a}
\]

for any matrices \( A \) and \( B \).

Using (3.1.10), we immediately see that

\[
\|Az\| \leq \|A\| \|z\|. \tag{3.1.11b}
\]

This inequality also follows from (3.1.11a) with \( B \) replaced by the vector \( z \).

The maximum and Euclidean norms of a matrix are

\[
\|A\|_\infty = \max_i \sum_j |a_{ij}|, \quad \|A\|_2 = \max_j |\lambda_j(A^T A)|^{1/2}, \tag{3.1.12}
\]
3.1. Consistency, Convergence, Stability

where $\lambda_j$ is an eigenvalue of the matrix $A^TA$. Additional matrix norms and properties of norms appear in Strikwerda [5], Appendix A.

**Example 3.1.5.** The maximum norm of the matrix $L_{\Delta}$ of (3.1.8) is $\|L_{\Delta}\|_\infty = |1 - 2r| + 2r$. If $r \leq 1/2$ then $(1 - 2r) > 0$ and $\|L_{\Delta}\|_\infty = 1$.

With these preliminary considerations established, we define the concepts of convergence and stability for finite difference schemes.

**Definition 3.1.8.** A finite difference approximation converges to the solution of a partial differential equation on $0 < t \leq T$ in a particular vector norm if

$$\|u^n - U^n\| \to 0, \quad n \to \infty, \quad \Delta x \to 0, \quad \Delta t \to 0, \quad n\Delta t \leq T. \quad (3.1.13)$$

**Remark 5.** The vector $u^n$ is the restriction of the continuous solution $u(x, t_n)$ to the mesh.

**Remark 6.** When applying Definition 3.1.8, visualize a sequence of computations performed on $0 < t \leq T$ using finer-and-finer meshes. Convergence implies that the discrete and continuous solutions approach each other for $t \in (0, T]$ in a particular vector norm as the mesh spacing decreases.

**Definition 3.1.9.** Let $U^n$ and $V^n$ satisfy homogeneous finite-difference initial value problems with different initial conditions, i.e.,

$$U^{n+1} = L_{\Delta} U^n, \quad U^0 = \phi,$$
$$V^{n+1} = L_{\Delta} V^n, \quad V^0 = \psi.$$

A finite difference scheme is stable if there exists a positive constant $C$, independent of the mesh spacing and initial data, such that

$$\|U^n - V^n\| \leq C\|U^0 - V^0\|, \quad n \to \infty, \quad \Delta x \to 0, \quad \Delta t \to 0, \quad n\Delta t \leq T. \quad (3.1.14a)$$

**Remark 7.** The requirement that $C$ be independent of $\Delta x$, $\Delta t$, $U^0$, and $V^0$ implies that the bound expressed in (3.1.14a) is uniform.

**Remark 8.** Definition 3.1.9, like Definition 3.1.8, can be thought of in relation to a sequence of calculations performed on finer-and-finer meshes.
Remark 9. This concept of stability implies that initial bounded differences between two solutions remain bounded for finite times when the mesh spacing is sufficiently small. Nothing in Definition 3.1.9 implies that $C \leq 1$; hence, contrary to notions of stability that arise in physics and engineering, some growth of the difference between solutions is permitted.

Our most successful stability analyses will occur when $L_\Delta$ is linear. In this case, the matrix $L_\Delta$ is independent of $U^n$ and $V^n$ and we have

$$U^{n+1} - V^{n+1} = L_\Delta(U^n - V^n), \quad U^0 - V^0 = \phi - \psi.$$ 

Replacing $U^n - V^n$ by $U^n$, we see that stability can be defined without introducing a perturbation. This alternate definition of stability is used throughout numerical analysis (cf., e.g., Strikwerda [5], Section 1.5) and we repeat it here for completeness.

Definition 3.1.10. A finite difference scheme (3.1.7) for a homogeneous initial value problem is stable if there exists a positive constant $C$, independent of the mesh spacing and initial data, such that

$$\|U^n\| \leq C\|U^0\|, \quad n \to \infty, \quad \Delta x \to 0, \quad \Delta t \to 0, \quad n\Delta t \leq T. \quad \text{(3.1.14b)}$$

Remark 10. Definitions 3.1.9 and 3.1.10 are equivalent for linear homogeneous initial value problems; however, both forms are used for nonlinear problems where they are not equivalent.

Remark 11. The notion of stability expressed by Definition 3.1.10 implies a bound on the growth of the solution and, as such, is similar to the concept of a “well posed” partial differential equation.

Let us use these definitions to establish the convergence or stability of some of the finite difference schemes that were introduced in Chapter 2.

Example 3.1.6. We show that the solution of the forward time-centered space scheme (2.3.3) for the heat-conduction problem (2.3.1) converges in the maximum norm when $r \leq 1/2$.

Using (2.3.3a) and (3.1.2a) with $v$ replaced by $u$, the finite difference and partial differential equation solutions satisfy

$$U_j^{n+1} = r U_{j-1}^n + (1 - 2r)U_j^n + r U_{j+1}^n.$$
3.1. Consistency, Convergence, Stability

\[ u_j^{n+1} = ru_j^n + (1 - 2r)u_j^n + ru_{j+1}^n + \Delta t\tau_j^n, \quad j = 1, 2, \ldots, J - 1. \]

Letting \( e_j^n = u_j^n - U_j^n \) and subtracting

\[ e_j^{n+1} = re_j^n + (1 - 2r)e_j^n + \Delta t\tau_j^n, \quad j = 1, 2, \ldots, J - 1. \]

Taking the absolute value and using the triangular inequality

\[ |e_j^{n+1}| \leq |r||e_j^n| + |1 - 2r||e_j^n| + |r||e_j^{n+1}| + \Delta t|\tau_j^n|, \quad j = 1, 2, \ldots, J - 1. \]

Replacing the error terms on the right side of the above expression by their maximum values \( \|e^n\|_\infty \) and \( \|\tau^n\|_\infty \), we obtain

\[ |e_j^{n+1}| \leq (|r| + |1 - 2r| + |r|)\|e^n\|_\infty + \Delta t\|\tau^n\|_\infty, \quad j = 1, 2, \ldots, J - 1. \]

We know that \( r > 0 \). If, in addition, \( r \leq 1/2 \), then \( 1 - 2r \geq 0 \) and the absolute value signs can be removed from the terms involving \( r \) to obtain

\[ |e_j^{n+1}| \leq \|e^n\|_\infty + \Delta t\|\tau^n\|_\infty, \quad j = 1, 2, \ldots, J - 1. \]

Since this result is valid for \( j \in [1, J - 1] \), it holds for the particular value of \( j \) where the left side attains its maximum value. Thus, we can replace the left side absolute value by its maximum norm to obtain

\[ \|e^{n+1}\|_\infty \leq \|e^n\|_\infty + \Delta t\|\tau^n\|_\infty. \]

Iterating the inequality over \( n \)

\[ \|e^{n+1}\|_\infty \leq \|e^{n-1}\|_\infty + \Delta t(\|\tau^n\|_\infty + \|\tau^{n-1}\|_\infty) \leq \ldots. \]

Thus,

\[ \|e^n\|_\infty \leq \|e^0\|_\infty + n\Delta t\tau, \]

where

\[ \tau = \max_{0 \leq k < n-1} \|\tau^k\|_\infty. \]

Neglecting round off errors, solutions of the partial differential and difference equations both satisfy the same initial conditions, so \( \|e^0\|_\infty = 0 \). Additionally, if \( T \) is the total time of interest, \( n\Delta t \leq T \) for all combinations of \( n \) and \( \Delta t \). Thus,

\[ \|e^n\|_\infty \leq T\tau. \]
Using (3.1.2c) we bound $\tau$ as

$$\tau \leq \frac{\Delta t}{2} K + \frac{\sigma \Delta x^2}{12} M,$$

where

$$K = \max_{0 \leq t \leq T, 0 \leq x \leq 1} |u_t|, \quad M = \max_{0 \leq t \leq T, 0 \leq x \leq 1} |u_{xxxx}|.$$

Thus,

$$\|e^n\|_\infty \leq T \left( \frac{\Delta t}{2} K + \frac{\sigma \Delta x^2}{12} M \right).$$

Letting $\Delta x$ and $\Delta t$ approach zero, we see that $\|e^n\|_\infty \to 0$; hence, the forward time-centered space scheme (2.3.3) converges to the solution of the heat-conduction problem (2.3.1) in the maximum norm when $r \leq 1/2$.

*Example 3.1.7.* We show that the forward time-backward space scheme (2.2.4)

$$U_j^{n+1} = (1 - \alpha_j^n)U_j^n + \alpha_j^n U_{j-1}^n,$$

for the kinematic wave equation (2.2.1) is stable in the maximum norm when the Courant number $\alpha_j^n \equiv a(U_j^n)\Delta t/\Delta x \in (0, 1]$. (Recall that this condition must be imposed to satisfy the Courant, Friedrichs, Lewy Theorem 2.2.1.)

We’ll use the stability definition (3.1.10) and begin by taking the absolute value of the difference scheme and using the triangular inequality to obtain

$$|U_j^{n+1}| \leq |1 - \alpha_j^n||U_j^n| + |\alpha_j^n||U_{j-1}^n|.$$

Since $0 < \alpha_j^n \leq 1$ by the Courant, Friedrichs, Lewy Theorem, we can remove the absolute value signs from the terms involving $\alpha_j^n$ to get

$$|U_j^{n+1}| \leq (1 - \alpha_j^n)|U_j^n| + \alpha_j^n|U_{j-1}^n|.$$

Replacing the solution values on the right side by their maximum values

$$|U_j^n| \leq \|U^n\|_\infty, \quad \forall j.$$

As in Example 3.1.6, the inequality must hold for the value of $j$ that maximizes its left side; thus,

$$\|U^{n+1}\|_\infty \leq \|U^n\|_\infty.$$
This inequality holds for all $n$ and may be iterated to yield
\[ \|U^n\|_\infty \leq \|U^0\|_\infty, \]
which establishes stability in the sense of (3.1.14b).

In the two previous examples, we analyzed the convergence and stability of finite difference schemes using very similar arguments. These arguments can be generalized further to obtain a “maximum principle.”

**Theorem 3.1.1.** A sufficient condition for stability of the one-level finite difference scheme
\[ U_{j+1} = \sum_{|s| \leq S} c_s U_{j+s} \]
in the maximum norm is that all coefficients $c_s$, $|s| \leq S$, be positive and add to unity.

**Proof.** The arguments follow those used in Examples 3.1.6 and 3.1.7. \qed

**Remark 1.2.** A one-level difference scheme is one that only involves solutions at time levels $n$ and $n+1$. A multilevel difference scheme might involve time levels prior to $n$ as well.

Typically, we not only want to know that a given scheme converges but the rate at which the numerical and exact solutions approach each other. This prompts the notion of order of accuracy.

**Definition 3.1.11.** A consistent finite-difference approximation of a partial differential equation is of order $p$ in time and $q$ in space if
\[ \tau_j^n = O(\Delta t^p) + O(\Delta x^q). \] (3.1.15)

**Problems**

1. Using (3.1.2b), we easily see that the forward time-centered space scheme (3.2.3a) for the heat conduction equation is first order in time and second order in space. Show that this scheme has an $O(\Delta x^2)$ local discretization error when $r = \sigma \Delta t / \Delta x^2 = 1/6$. 
2. Show that the Euclidean norm of a $J \times J$ matrix $A$ given by $(3.1.12)$ follows from from the definition of a matrix norm $(3.1.10)$ upon use of the Euclidean vector norm $(3.1.9b)$.

3. In Example 3.1.7, we used the maximum principle (Theorem 3.1.1) to establish stability of the forward time-backward space scheme $(2.2.4)$. It can likewise be used to establish stability in the maximum norm of the forward time-forward space scheme $(2.2.2b)$ when $-1 \leq a^n_j \leq 0$. Can it be applied to the forward time-centered space scheme $(2.2.5)$?

### 3.2 Stability Analysis Using a Discrete Fourier Series

A discrete Fourier series can be used to analyze the stability of constant coefficient finite-difference problems with periodic initial data in much the same way that an infinite Fourier series can be used to solve constant coefficient partial differential equations. This method was introduced by John von Neumann and is called *von Neumann stability analysis*.

Thus, suppose that the solution of a finite difference problem is periodic in $j$ with period $J$ and express its solution as the discrete complex Fourier series

\[
U^n_j = \sum_{k=0}^{J-1} A^n_k \omega^k_j, \quad j = 0, 1, \ldots, J - 1, \tag{3.2.1a}
\]

where

\[
\omega_j \equiv e^{2\pi i j / J}. \tag{3.2.1b}
\]

The complex form of the Fourier series is much more convenient for our present purposes than the more common representation in terms of sines and cosines. In this form, the solution $U^n_j, j = 0, 1, \ldots, J - 1,$ is an approximation of the solution of a partial differential equation that is periodic in $x$ with period $2\pi$. The mesh spacing that is inferred by $(3.2.1a)$ is $\Delta x = 2\pi / J$. 

The discrete Fourier series has many properties in common with the infinite Fourier series or the continuous Fourier transform (cf. Gottlieb and Orszag [1] or Strikwerda [5], Section 2.1). For example, the discrete Fourier series satisfies the orthogonality relation

\[ \sum_{j=0}^{J-1} \omega_j^k \bar{\omega}_l^j = \begin{cases} J, & \text{if } k \equiv l \mod J, \\ 0, & \text{otherwise} \end{cases}, \]  

(3.2.2)

where \( k \equiv l \mod J \) means that \( k \) and \( l \) differ by an integral multiple of \( J \) and a superimposed bar denotes a complex conjugate, e.g., \( \bar{\omega}_j = e^{-2\pi ij/J} \). The relationship (3.2.2) is easily established using properties of the complex roots of unity.

Given the solution \( U^n_j, j = 0, 1, \ldots, J - 1 \), we can find the Fourier coefficients \( A^n_k, k = 0, 1, \ldots, J - 1 \), by inverting the discrete Fourier series using the orthogonality relation (3.2.2). Thus, multiplying (3.2.1a) by \( \omega_j^l \) and summing over \( j \) yields

\[ \sum_{j=0}^{J-1} U^n_j \omega_j^l = \sum_{j=0}^{J-1} \omega_j^l \sum_{k=0}^{J-1} A^n_k \omega_j^k. \]

Interchanging the order of the summations on the right side and using (3.2.2) gives

\[ \sum_{j=0}^{J-1} U^n_j \omega_j^l = \sum_{k=0}^{J-1} A^n_k \sum_{j=0}^{J-1} \omega_j^k \bar{\omega}_j^l = JA^n_l. \]

Thus,

\[ A^n_l = \frac{1}{J} \sum_{j=0}^{J-1} U^n_j \omega_j^l. \]  

(3.2.3)

The inverse (3.2.3) of the discrete Fourier series (3.2.1a) is called the discrete Fourier transform.

Another property of discrete Fourier series and transforms that follows directly from the orthogonality condition (3.2.2) is the discrete form of Parseval’s relation (cf. Problem 1 at the end of this section)

\[ \|U^n\|_2^2 = J \|A^n\|_2^2, \]  

(3.2.4a)

where, for complex quantities,

\[ \|U^n\|_2^2 = \sum_{j=0}^{J-1} U^n_j U^n_j, \quad \|A^n\|_2^2 = \sum_{k=0}^{J-1} A^n_k A^n_k. \]  

(3.2.4b)
Given the link between physical and Fourier coefficients expressed by Parseval’s relation (3.2.4a), it is natural to use von Neumann’s stability analysis in the Euclidean ($L^2$) norm. Let us illustrate the technique with some examples.

**Example 3.2.1.** We use von Neumann’s approach to show that the forward time-backward space finite difference scheme (2.2.4) is stable in $L^2$ when $\alpha_n^j$ is a positive constant, say, $\alpha \in (0, 1]$ and periodic initial data is applied. In this case, (2.2.4) becomes

$$U_{j}^{n+1} = (1 - \alpha)U_{j}^{n} + \alpha U_{j-1}^{n}.$$  

Substitute (3.2.1a) into the above expression to obtain

$$\sum_{k=0}^{J-1} A_{k}^{n+1} \omega_{j}^{k} = \sum_{k=0}^{J-1} [(1 - \alpha) A_{k}^{n} \omega_{j}^{k} + \alpha A_{k}^{n} \omega_{j-1}^{k}].$$

According to (3.2.1b), $\omega_{j-1} = \omega_{j} e^{-2\pi i/J}$; thus,

$$\sum_{k=0}^{J-1} [A_{k}^{n+1} - (1 - \alpha) A_{k}^{n} - \alpha A_{k}^{n} e^{-2\pi i k / J}] \omega_{j}^{k} = 0.$$

Using the orthogonality relation (3.2.2) and the discrete Fourier transform (3.2.3), we infer that the (bracketed) coefficient of $\omega_{j}^{k}$ must vanish for each $k$, i.e.,

$$A_{k}^{n+1} - [(1 - \alpha) + \alpha e^{-2\pi i k / J}] A_{k}^{n} = 0.$$

Solving for $A_{k}^{n+1}$

$$A_{k}^{n+1} = M_{k} A_{k}^{n}, \quad M_{k} = (1 - \alpha) + \alpha e^{-2\pi i k / J}.$$

The amplification factor $M_{k}$ gives the growth or decay of the $k$th Fourier mode in one time step.

The above recurrence can be iterated to obtain

$$A_{k}^{n} = (M_{k})^{n} A_{k}^{0},$$

where the Fourier coefficient $A_{k}^{0}$, $k = 0, 1, \ldots, J - 1$, can be determined from the prescribed initial data upon use of (3.2.3). Explicit determination of $A_{k}^{0}$ is rarely necessary since we seek to establish stability independently of the initial data.
We can use Euler's identity
\[ e^{i\phi} = \cos \phi + i \sin \phi \]
to determine the magnitude of the amplification factor as
\[ |M_k|^2 = M_k \overline{M_k} = (1 - \alpha + \alpha \cos \frac{2\pi k}{J})^2 + (\alpha \sin \frac{2\pi k}{J})^2 \]
or
\[ |M_k|^2 = 1 - 2\alpha (1 - \alpha)(1 - \cos \frac{2\pi k}{J}). \]
The half-angle formula
\[ \sin^2 \frac{\phi}{2} = \frac{1 - \cos \phi}{2} \]
can be used to simplify the above relation to
\[ |M_k|^2 = 1 - 4\alpha (1 - \alpha) \sin^2 \frac{\pi k}{J}. \]

We now see that the initial Fourier mode \( A_0^k \) will grow or decay depending on whether \( |M_k| \) is greater than or less than unity. In this example, no Fourier mode grows since \( 0 \leq 4\alpha (1 - \alpha) \leq 1 \) when \( 0 < \alpha \leq 1 \).

Using (3.2.1a), we see that the finite difference solution satisfies
\[ U_j^n = \sum_{k=0}^{J-1} (M_k)^n A_0^k \omega_j^k. \]
Stability in the Euclidean norm follows by applying Parseval's relation (3.2.4a) to the above equation to obtain
\[ \| U^n \|_2^2 = J \sum_{k=0}^{J-1} |M_k|^{2n} |A_0^k|^2. \]
Since \( |M_k| \leq 1 \) for all \( k \) when \( 0 < \alpha \leq 1 \), we can use Parseval's relation once more to find
\[ \| U^n \|_2^2 \leq J \sum_{k=0}^{J-1} |A_0^k|^2 = \| U^0 \|_2^2. \]
This establishes the \( L^2 \) stability of (2.2.4) when the Courant number is a constant \( \alpha \in (0,1] \).

**Example 3.2.2.** Let us, at long last, use von Neumann’s method to examine the \( L^2 \) stability of the forward time-centered space scheme (2.2.5) for the kinematic wave
equation (2.2.1). Necessarily, we assume that $\alpha_j^n$ is a constant $\alpha$ and the initial data is periodic in $j$ with period $J$ so that (2.2.5) becomes

$$U_{j}^{n+1} = U_{j}^{n} - \frac{\alpha}{2}(U_{j+1}^{n} - U_{j-1}^{n}).$$

Expanding $U_{j}^{n}$ in a discrete Fourier series (3.2.1a), we find

$$\sum_{k=0}^{J-1} [A_{k}^{n+1} - A_{k}^{n} + \frac{\alpha}{2}(e^{2\pi ik/J} - e^{-2\pi ik/J})A_{k}^{n} \omega_{j}^{k}] = 0.$$ 

Once again, the orthogonality relation (3.2.2) may be used to infer that

$$A_{k}^{n} = (M_{k})^{n} A_{k}^{0},$$

where

$$M_{k} = 1 - \frac{\alpha}{2}(e^{2\pi ik/J} - e^{-2\pi ik/J}) = 1 - i\alpha \sin \frac{2\pi k}{J}.$$ 

The magnitude of the amplification factor is

$$|M_{k}|^2 = 1 + \alpha^2 \sin^2 \frac{2\pi k}{J}.$$ 

There is no possibility of restricting $\alpha$ so that $|M_{k}| \leq 1$ for all $k$; thus, most Fourier modes will grow. As bad as this appears, (2.2.5) is still stable according to either (3.1.14a) or (3.1.14b). To demonstrate this, observe that

$$|M_{k}|^2 \leq 1 + \alpha^2.$$ 

The following inequality is useful in many situations, so we'll record it as a lemma.

**Lemma 3.2.1.** For all real $z$

$$1 + z \leq e^z. \quad (3.2.5)$$

**Proof.** Using Taylor’s formula for the exponential function

$$e^z = 1 + z + \frac{1}{2} e^\xi z^2 \geq 1 + z.$$ 

□
In the present case, (3.2.5) can be used to bound the amplification factor as

$$|M_k|^2 \leq e^{\alpha^2}.$$ 

Taking the $n$th power and using the definition of the Courant number (2.2.3)

$$|M_k|^{2n} \leq e^{n \alpha^2 \Delta t^2 / \Delta x^2} = e^{(n \Delta t \alpha^2 / \Delta x^2)}.$$ 

As usual, introduce $T$ so that $n \Delta t \leq T$. Also, let us agree to select meshes having temporal and spatial steps satisfying $\alpha^2 \Delta t / \Delta x^2 = \beta$, where $\beta$ is a constant. Then

$$|M_k|^{2n} \leq e^{T \beta} = C^2.$$ 

Using Parseval’s relation (3.2.4a) and following the steps used in Example 3.2.1, we find

$$\|U^n\|^2_2 = J \sum_{k=0}^{J-1} |M_k|^{2n} |A_k^0|^2 \leq C^2 \|U^0\|^2_2.$$ 

We have found a constant $C$ satisfying

$$\|U^n\|^2_2 \leq C \|U^0\|^2_2;$$

hence, (2.2.5) is stable in $L^2$. The constant $C$ is greater than unity, so that most Fourier modes will grow. In fact, if $T$ and $\beta$ are not small enough, initial disturbances will grow to such an extent that they dominate the computation. Additionally, we are forced to restrict the time step $\Delta t \leq \beta \Delta x^2 / \alpha^2$, while the directional methods (2.2.2b, 4) only require $\Delta t \leq \Delta x / \alpha$. This leads us to conclude that, while technically stable, the forward time-centered space scheme (2.2.5) is not a practical method to apply to the kinematic wave equation (2.2.1).

When applying von Neumann’s method we will always find a solution of the form

$$U^n_j = \sum_{k=0}^{J-1} (M_k)^n A_k^0 \omega_j^k.$$ 

In Example 3.2.1, we found that $|M_k| \leq 1$ so that initial perturbations did not grow and the forward time-backward space scheme was stable in $L^2$. In Example 3.2.2, $|M_k| > 1$ and perturbations grew, but were bounded for finite times $T$. While the (forward time-centered space) scheme was stable, we saw that it was impractical and, thus, may ask
whether or not we should keep $|M_k| \leq 1$ as a practical matter. The answer to this question depends on the behavior of the problem under investigation. If the solution of the partial differential equation is unstable and, hence, growing in time, then some growth of initial perturbations can and should be tolerated. However, if the solution of the partial differential equation does not grow, then either $|M_k|$ must be bounded by unity or we must be willing to compute for sufficiently small times so that the bound $C$ on the growth of $\|U^n\|$ remains small. There are only a few instances where the latter option will be desirable; thus, we should typically maintain $|M_k| \leq 1$ whenever solutions of the partial differential equation under study do not grow.

In those cases where $|M_k|$ may exceed unity, we may ask the amount by which it may do so. This leads to the von Neumann condition.

**Definition 3.2.1.** A finite difference scheme satisfies the von Neumann condition if there exists a positive constant $c$ that is independent of $\Delta t$, $\Delta x$, and $k$ which satisfies

$$|M_k| \leq 1 + c\Delta t, \quad \forall \Delta t \leq \Delta t^*, \; \Delta x \leq \Delta x^*. \quad (3.2.6)$$

Finite difference schemes that satisfy the von Neumann condition are stable and conversely as expressed by the following theorem.

**Theorem 3.2.1.** A constant coefficient scalar one-level difference scheme is stable in the Euclidean norm if and only if it satisfies the von Neumann condition.

**Proof.** Suppose the von Neumann condition is satisfied. We follow the arguments used in Example 3.2.2 and use Parseval’s relation (3.2.4a) to obtain

$$\|U^n\|_2^2 = J \sum_{k=0}^{J-1} |M_k|^{2n} |A_k^0|^2 \leq (1 + c\Delta t)^{2n}\|U^0\|_2^2.$$ 

Using (3.2.5), we have

$$\|U^n\|_2^2 \leq e^{2cn\Delta t}\|U^0\|_2^2 \leq e^{2cT}\|U^0\|_2^2,$$

where, $n\Delta t \leq T$, with $T$ being the total time of interest. Choosing $C^2 = e^{cT}$ establishes stability according to (3.1.14b).
To prove the converse, we suppose that there is a value of \( k = k^* \) such that
\[
|M_k*| > (1 + c\Delta t), \quad \forall c.
\]
Further suppose that initial data is selected so that \( A^0_k \neq 0 \) and \( A^0_k = 0, k \neq k^* \). Thus, the initial condition is
\[
U_j^0 = A^0_{k^*}\omega_j^{k^*}.
\]
The solution after \( n \) time steps is
\[
U_j^n = (M_{k^*})^n A^0_k\omega_j^{k^*} = (M_{k^*})^n U_j^0.
\]
Taking the Euclidean norm
\[
\|U^n\|_2^2 = |M_{k^*}|^{2n}\|U^0\|_2^2 > (1 + c\Delta t)^{2n}\|U^0\|_2^2, \quad \forall c.
\]
Since this must hold for any and all values of \( c, \|U^n\|_2 \) must be unbounded and, hence, the finite difference method is unstable.

The von Neumann condition is a necessary condition for stability in much more general situations than stated by Theorem 3.2.1. Sufficiency has also been verified under less restrictive conditions [2, 3].

**Example 3.2.3.** It may seem like we are restricted to using the directional derivatives when solving the kinematic wave equation (2.2.1); however, the Lax-Friedrichs scheme uses centered space difference with a forward-averaged time difference to obtain the approximation
\[
\frac{U_{j+1}^n - (U_{j+1}^n + U_{j-1}^n)/2}{\Delta t} + a_j^n \frac{U_{j+1}^n - U_{j-1}^n}{2\Delta x} = 0
\]

or
\[
U_{j+1}^n = \frac{1 + a_j^n}{2} U_{j-1}^n + \frac{1 - a_j^n}{2} U_{j+1}^n.
\]

The computational stencil for this scheme is shown in Figure 3.2.1. This spatially-centered difference scheme is clearly stable in the maximum norm by the Maximum Principle as long as the Courant, Friedrichs, Lewy Theorem is satisfied, *i.e.*, as long as \( |a_j^n| \leq 1 \). It is also stable in \( L^2 \) under similar conditions (*cf.* Problems 2 and 3 at the end of this section).

**Problems**
1. Let $U_j^n$, $j = 0, 1, \ldots, J - 1$, and $A_k^n$, $k = 0, 1, \ldots, J - 1$, be discrete Fourier pairs according to (3.2.1a). Derive the discrete form of Parseval’s relation (3.2.4a).

2. Consider the Lax-Friedrichs difference scheme (3.2.7) for the constant-coefficient kinematic wave equation

$$u_t + au_x = 0.$$  

2.1. Calculate the leading terms in the local discretization error for this scheme. Is the Lax-Friedrichs scheme consistent for all choices of mesh spacings? Explain.

2.2. Use a von Neumann stability analysis to show that the Lax-Friedrichs scheme is stable in $L^2$ whenever the Courant, Friedrichs, Lewy Theorem is satisfied

3. Consider finite-difference schemes for the constant-coefficient kinematic wave equation of Problem 2 that have the form

$$U_j^{n+1} = U_j^n - \frac{\alpha}{2} (U_{j+1}^n - U_{j-1}^n) + \frac{\beta}{2} (U_{j+1}^n - 2U_j^n + U_{j-1}^n)$$  

(3.2.8a)

with

$$\alpha = a \frac{\Delta t}{\Delta x}, \quad \beta = \alpha z. \quad (3.2.8b)$$

The parameter $\alpha$ is the Courant number and $\beta$ and $z$ are dissipation factors. Some common schemes having this form for specific choices of $z$ and $\beta$ appear in Table 3.2.1. The function $\text{sgn}(x) = x/|x|$ and the Lax-Wendroff scheme is described in Section 3.3. Find the region in the $(\alpha, \beta)$-plane where the amplification factor associated with the scheme (3.2.8) does not exceed unity in magnitude. You may,
### Table 3.2.1: Values of the dissipation factors for schemes of the form (3.2.8).

<table>
<thead>
<tr>
<th>Method</th>
<th>$z$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centered</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lax-Wendroff</td>
<td>$\alpha$</td>
<td>$\alpha^2$</td>
</tr>
<tr>
<td>Upwind</td>
<td>$\text{sgn}\alpha$</td>
<td>$\text{sgn}\alpha$</td>
</tr>
<tr>
<td>Lax-Friedrichs</td>
<td>$1/\alpha$</td>
<td>1</td>
</tr>
</tbody>
</table>

if you want, confine your analysis to positive values of $\alpha$. This region of parameter space is often called a “region of absolutely stability.” Present a graph of the absolute stability region. The specific methods presented in the table correspond to curves in the $(\alpha, \beta)$-plane. Superimpose these curves on your stability diagram. Comment on the stability properties of the various methods.

### 3.3 Matrix Stability Analysis

The tools that are currently available to us for performing stability analyses have severe limitations. The maximum principle (Theorem 3.1.1) requires finite difference schemes to have positive coefficients that sum to unity and the von Neumann method requires constant coefficient linear problems with periodic initial data. Matrix methods can be used to analyze more general linear problems. With the goal of describing these methods, consider a linear finite difference scheme in the form of (3.1.7). For simplicity, assume that the difference operator $L_\Delta$ is independent of the time level $n$ and iterate this (3.1.7) to obtain

$$ U^n = (L_\Delta)^n U^0. $$

(If $L_\Delta$ depended on $n$, we would get a product of the operators at each time level.) Taking a norm

$$ \| U^n \| = \|(L_\Delta)^n U^0\| \leq \|(L_\Delta)^n\| \| U^0 \|. $$

The finite difference scheme (3.1.7) is stable according to (3.1.14b) if and only if there exists a constant $C$ such that

$$ \|(L_\Delta)^n\| \leq C, \quad n \to \infty, \quad \Delta x, \Delta t \to 0, \quad n\Delta t \leq T. \quad (3.3.1) $$

Using matrix stability methods, we try to limit the growth of $\|L_\Delta\|$. Before beginning, we review some material about matrix eigenvalue problems. If $x^i$ is an eigenvector of $L_\Delta$ corresponding to the eigenvalue $\lambda_i$, then

$$L_\Delta x^i = \lambda_i x^i, \quad i = 1, 2, \ldots, N,$$

where $N$ is the dimension of $L_\Delta$.

**Definition 3.3.1.** The spectral radius $\rho(L_\Delta)$ of $L_\Delta$ is the modulus of its largest eigenvalue, i.e.,

$$\rho(L_\Delta) \equiv \max_{1 \leq i \leq N} |\lambda_i(L_\Delta)|. \quad (3.3.3)$$

The spectral radius provides a lower bound to any matrix norm.

**Lemma 3.3.1.** Let $\rho(L_\Delta)$ and $\|L_\Delta\|$ be the spectral radius and any vector-induced matrix norm of $L_\Delta$, then

$$\rho(L_\Delta) \leq \|L_\Delta\|. \quad (3.3.4)$$

**Proof.** Take a norm of (3.3.2)

$$\|L_\Delta x^i\| = |\lambda_i||x^i|, \quad i = 1, 2, \ldots, N,$$

and use (3.1.10, 3.1.11)

$$|\lambda_i| = \frac{\|L_\Delta x^i\|}{\|x^i\|} \leq \|L_\Delta\|, \quad i = 1, 2, \ldots, N.$$

The result (3.3.4) follows since this relation holds for the value of $i$ corresponding to the maximum eigenvalue. \qed

Since the eigenvalues of $(L_\Delta)^n$ are $(\lambda_i)^n$, we can use (3.3.4) and (3.1.11) to obtain

$$\rho^n(L_\Delta) \leq \|(L_\Delta)^n\| \leq \|L_\Delta\|^n. \quad (3.3.5)$$

The left and right sides of (3.3.4) lead, respectively, to necessary and sufficient stability conditions.
Theorem 3.3.1. (The von Neumann necessary stability condition.) A necessary condition for the stability of the linear homogeneous finite difference scheme (3.1.7) is that there exist a constant $c$, independent of $\Delta x$ and $\Delta t$, satisfying

$$\rho(L_\Delta) \leq 1 + c\Delta t. \quad (3.3.6)$$

Proof. Using (3.3.5), we see that $\rho^n(L_\Delta)$ must be bounded otherwise $\|(L_\Delta)^n\|$ couldn’t be. Thus, a necessary condition for the stability of (3.1.7) is that there exist a constant $C$ such that $\rho^n(L_\Delta) \leq C$ for all $n\Delta t \leq T$ as $\Delta x, \Delta t \to 0$. Taking an $nth$ root and using the maximal value of $n$ gives

$$\rho(L_\Delta) \leq C^{1/n} \leq C^{\Delta t/T}. \quad (3.3.7)$$

When $\Delta t$ is sufficiently small, Taylor’s formula can be used to find a value of $c$ so that $C^{\Delta t/T}$ is bounded by $1 + c\Delta t$. \qed

Theorem 3.3.2. A sufficient condition for the stability of the linear finite difference scheme (3.1.7) is that there exist a constant $c$, independent of $\Delta x$ and $\Delta t$, such that

$$\|L_\Delta\| \leq 1 + c\Delta t. \quad (3.3.7)$$

Proof. The proof follows the arguments used in Theorem 3.2.1. \qed

Remark 1. The von Neumann condition (3.3.6) is also sufficient for stability in many cases. For example, it is sufficient for stability when $L_\Delta$ is either symmetric or similar to a symmetric matrix [2].

Example 3.3.1. In Example 3.1.4 we showed that $L_\Delta$ for the forward time-centered space scheme (2.3.3) for the heat conduction problem (2.3.1) satisfies

$$L_\Delta = \begin{bmatrix}
1 - 2r & r & & & \\
r & 1 - 2r & r & & \\
r & r & 1 - 2r & r & \\
& r & r & 1 - 2r & \\
& & \ddots & \ddots & \ddots \\
& & r & 1 - 2r & 
\end{bmatrix}.$$  

If $r \leq 1/2$, then $\|L_\Delta\|_\infty = 1$; therefore, from Theorem 3.3.2, this scheme is stable in the maximum norm.
Example 3.3.2. Consider the finite difference scheme

\[ U^{n+1}_j = c_{-1} U^n_{j-1} + c_1 U^n_{j+1} \]

with periodic initial data in \( j \) with period \( J \). (This problem is similar to Exercise 1.5 of Strikwerda [5].) The Lax-Friedrichs scheme (3.2.7) for the kinematic wave equation (2.2.1) has this form with

\[ c_{-1} = \frac{1 + \alpha}{2}, \quad c_1 = \frac{1 - \alpha}{2}, \]

where, for simplicity, we have assumed the Courant number \( \alpha \) to be constant.

The difference scheme has the form of (3.1.7) with \( L_\Delta \)

\[ L_\Delta = \begin{bmatrix} 0 & c_1 & c_{-1} \\ c_{-1} & 0 & c_1 \\ c_1 & \cdots & 0 \end{bmatrix}, \quad U^n = \begin{bmatrix} U^n_0 \\ U^n_1 \\ \vdots \\ U^n_{J-1} \end{bmatrix}. \]

If \( |c_{-1}| + |c_1| \leq 1 \) then \( \|L_\Delta\|_\infty \leq 1 \) and the scheme is stable in the maximum norm. This is the case for the Lax-Friedrichs scheme (even for nonlinear problems) when \( |\alpha| \leq 1 \).

We can analyze the stability of finite difference schemes using the basic definitions (3.1.14a, 3.1.14b). Nonlinear problems can be analyzed in this manner; thus, in some sense, it is the most powerful analytical technique at our disposal, but it is also the most difficult one to apply. Let us try an example.

Example 3.3.3. The Lax-Wendroff scheme for the kinematic wave equation

\[ u_t + au_x = 0 \]

is obtained from the Taylor’s series in time

\[ u^{n+1}_j = u^n_j + \Delta t (u^n_t)_j + \frac{\Delta t^2}{2} (u^n_{tt})_j + \ldots. \]

Time derivatives are replaced by spatial derivatives that are obtained from the partial differential equation. For simplicity, let us suppose that \( a \) is a constant, then

\[ u_t = -au_x, \quad u_{tt} = -au_{xt} = -a(u_t)_x = a^2 u_{xx}, \ldots \]
(In fact, the Lax-Wendroff scheme is rather difficult to implement when $\alpha$ is not constant. We’ll discuss an alternative in Chapter 6 that is superior to this approach in this case.)

The Taylor’s series is truncated and the spatial derivatives are approximated by centered differences (2.1.7, 2.1.9). Retaining $O(\Delta t^2)$ terms leads to the second-order Lax-Wendroff scheme

\[ U_j^{n+1} = U_j^n - \frac{\alpha}{2} (U_{j+1}^n - U_{j-1}^n) + \frac{\alpha^2}{2} (U_{j+1}^n - 2U_j^n + U_{j-1}^n) \]

or

\[ U_j^{n+1} = \frac{\alpha(\alpha + 1)}{2} U_{j-1}^n + (1 - \alpha^2)U_j^n + \frac{\alpha(\alpha - 1)}{2} U_{j+1}^n. \quad (3.3.8) \]

where, as usual, $\alpha$ is the Courant number. Taylor’s series arguments immediately show that the local error is $O(\Delta t^2) + O(\Delta x^2)$.

The coefficients of this Lax-Wendroff scheme are not all positive. For example, if $0 < \alpha < 1$ then the coefficient of $U_{j+1}^n$ will be negative. Likewise, the coefficient of $U_{j-1}^n$ is negative when $\alpha$ is negative. Thus, Theorem 3.1.1 cannot be used to establish stability in the maximum norm. As expected, the Lax-Wendroff scheme gives the exact solution of the kinematic wave equation when $|\alpha| = 1$.

We’ll establish the stability of a periodic initial value problem in the Euclidean norm when $|\alpha| \leq 1$ using definition (3.1.14b). The von Neumann method could also be used in this case (cf. Problem 3.2.3). To begin, we square (3.3.8) to obtain

\[ (U_j^{n+1})^2 = \left[ \frac{\alpha(\alpha + 1)}{2} U_{j-1}^n + (1 - \alpha^2)U_j^n + \frac{\alpha(\alpha - 1)}{2} U_{j+1}^n \right]^2. \]

If $|\alpha| \leq 1$ then

\[ \frac{\alpha^2(1 - \alpha^2)}{4} [U_{j-1}^n - 2U_j^n + U_{j+1}^n]^2 \geq 0. \]

Add this term to the right side of the previous expression and sum over a period to get

\[ \sum_{j=0}^{J-1} (U_j^{n+1})^2 \leq \sum_{j=0}^{J-1} \left[ \frac{\alpha^2(1 + \alpha)}{2} (U_{j-1}^n)^2 + (1 - \alpha^2)(U_j^n)^2 + \frac{\alpha^2(1 - \alpha)}{2} (U_{j+1}^n)^2 \right. \\
\left. - \alpha(1 - \alpha^2)(U_{j+1}^n U_j^n - U_{j+1}^n U_{j+1}^n) \right]. \]

This expression can be simplified by reindexing the summations, e.g.,

\[ \sum_{j=0}^{J-1} (U_{j-1}^n)^2 = \sum_{k=-1}^{J-2} (U_k^n)^2 = \sum_{k=0}^{J-1} (U_k^n)^2 + (U_{J-1}^n)^2 - (U_{J-1}^n)^2. \]
The solution is periodic so \( U_{j-1}^n = U_j^{n-1} \); hence,

\[
\sum_{j=0}^{J-1} (U_{j-1}^n)^2 = \sum_{k=0}^{J-1} (U_k^n)^2 .
\]

Similar reindexing of other terms yields

\[
\sum_{j=0}^{J-1} (U_{j+1}^n)^2 = \sum_{k=0}^{J-1} (U_k^n)^2 , \quad \sum_{j=0}^{J-1} U_j^n U_{j-1}^n = \sum_{k=0}^{J-1} U_{k+1}^n U_k^n .
\]

Thus, the summation simplifies to

\[
\sum_{j=0}^{J-1} (U_j^{n+1})^2 \leq \sum_{j=0}^{J-1} \left[ \frac{\alpha^2 (1 + \alpha)}{2} + (1 - \alpha^2) + \frac{\alpha^2 (1 - \alpha)}{2} \right] (U_j^n)^2 = \sum_{j=0}^{J-1} (U_j^n)^2 .
\]

Therefore, \( \|U^{n+1}\|_2 \leq \|U^n\|_2 \) and the Lax-Wendroff scheme is stable when the Courant, Friedrichs, Lewy Theorem (\( |\alpha| \leq 1 \)) is satisfied.

Centered schemes like the Lax-Wendroff (3.3.8) and Lax-Friedrichs (3.2.7) methods may require artificial boundary conditions for initial-boundary value problems. For instance, suppose the kinematic wave equation with a positive wave speed \( a \) is to be solved on \( 0 < x < 1 \). This problem only requires a boundary condition at \( x = 0 \) (Section 1.3). A numerical solution at the right-most point \( j = J \), however, cannot be computed by either the Lax-Wendroff or Lax-Friedrichs schemes. Another method is needed to compute \( U_J^n \).

As we’ll learn in Chapter 6, a poor choice could affect the stability of the method.

**Problems**

1. Write a computer program for the difference scheme (3.2.8). Implement it with \( \beta = \alpha^2 \), which corresponds to the Lax-Wendroff scheme. Assume that the initial data

\[
u(x, 0) = \phi(x) \]

is periodic in \( x \) with period 2. The problem should be solved on \(-1 \leq x \leq 1\), \( 0 < t \leq T \). Use \( J(= 2/\Delta x) \), \( \alpha \), and \( N(= T/\Delta t) \) as input parameters.

1.1. Execute your program when \( a = 1 \) and \( \phi(x) \) has the form

\[
\phi(x) = \begin{cases} 
\Delta x + x, & \text{if } -\Delta x \leq x < 0 \\
\Delta x - x, & \text{if } 0 \leq x < \Delta x \\
0, & \text{elsewhere for } x \in (-1, 1)
\end{cases}
\]
1.2. This data is an attempt to simulate the effects of a small error introduced by, say, round off. Execute your program for about 20 time steps with $J = 10$ and $\alpha = 0.5, 0.999, 1.1$ (more if you like). Plot the numerical and exact solutions as functions of $x$ for a few times. Comment on the solutions for each value of $\alpha$. Which choice of $\alpha$ would be preferred if the initial data actually did correspond to a rounding error?

1.3. Solve a problem with $a = 1$ and

$$
\phi(x) = \begin{cases} 
-2(1 + x), & \text{if } -1 \leq x < -1/2 \\
2x, & \text{if } -1/2 \leq x < 1/2 \\
2(1 - x), & \text{if } 1/2 \leq x < 1 
\end{cases}
$$

Compute solutions on $0 < t \leq 2$ using $J = 10, 20, 40$ and $\alpha = 1/2, 1$. (Values of $N$ follow from the Courant number.) Compare the accuracy of the solutions at $t = 2$. Tabulate errors at $t = 2$ in the maximum and $L^2$ norms and estimate the order of convergence of the numerical to the exact solution. Plot the solutions as a function of $x$ at $t = 2$.

1.4. Repeat Part 3 using either the upwind or Lax-Friedrichs schemes. Compare results with those obtained using the Lax-Wendroff scheme.

### 3.4 The Lax Equivalence Theorem

In Section 3.3, we learned that every difference scheme should be consistent, convergent, and stable. The Lax Equivalence Theorem expresses a relationship between these three properties.

**Theorem 3.4.1. (Lax Equivalence Theorem).** Given a properly posed linear initial value problem and a consistent finite difference approximation of it, then stability is necessary and sufficient for convergence.

**Proof.** We will prove that a consistent and stable difference scheme is convergent. The proof that unstable schemes do not converge is more involved and we’ll defer it to Chapter 6. The entire proof appears in Richtmyer and Morton [3] and Strikwerda [5], Chapter 10.
Let $U^n$ be the solution of the difference equation (3.1.7) and let $\tau^n$ be the local discretization error. Then

$$u^{n+1} = L \Delta u^n + \Delta t \tau^n.$$ 

Let $e^n \equiv u^n - U^n$ be the global discretization error and subtract (3.1.7) to obtain

$$e^{n+1} = L \Delta e^n + \Delta t \tau^n.$$ 

For simplicity, assume that $L \Delta$ is independent of $n$ and iterate the above relation to obtain

$$e^n = (L \Delta)^n e^0 + \Delta t [L \Delta^{n-1} \tau^0 + L \Delta^{n-2} \tau^1 + \ldots + \tau^{n-1}].$$

Taking a norm and noting that $e^0 = 0$ in the absence of round-off errors,

$$\|e^n\| \leq \Delta t [\|L \Delta^{n-1}\| \|\tau^0\| + \|L \Delta^{n-2}\| \|\tau^1\| + \ldots + \|\tau^{n-1}\|]. \quad (3.4.1)$$

The difference scheme is consistent, so we can make $\tau^k$ arbitrarily small by making $\Delta x$ and $\Delta t$ sufficiently small. Let us choose a total time of interest $T$ and an $\epsilon > 0$ such that $\|\tau^k\| \leq \epsilon$ for all $k \Delta t \leq T$ whenever $\Delta x, \Delta t < \delta$. Furthermore, the scheme is stable, so there exists a constant $C$ such that $\|L \Delta^k\| \leq C$ for all $k \Delta t \leq T$. With (3.4.1), these considerations imply that

$$\|e^n\| \leq n \Delta t C \epsilon.$$ 

If $n \Delta t \leq T$

$$\|e^n\| = \|u^n - U^n\| \leq T C \epsilon,$$

and the scheme converges. \hfill \Box

**Remark 1.** The Lax Equivalence Theorem provides the link between stability and convergence that, perhaps, we expected based on the examples of Chapter 2.

**Remark 2.** In practice, convergence is difficult to establish, while consistency and stability are less so. Consistency merely requires the use of Taylor's series expansions and stability can be proven, at least, in simplified situations using von Neumann's approach. Thus, Lax's theorem provides very useful information.
Bibliography


