Chapter 2

Stability and Convergence

2.1 Convergence

2.1.1 The Concept of Convergence

The most basic property that a finite difference scheme must have in order to be useful is that its solution should approximate the solution of the corresponding partial differential equation and that the approximation improves as the grid spacings (h and k) are decreased. We call such a scheme a *convergent scheme.*

In order to formulate a definition of the concept of convergence, it should be noted that when $k \to 0$ the length of the time step decreases and the number of time steps needed to integrate up to a certain fixed time must increase. Also, when h is decreased, the number of grid points (M) in the x-direction must increase. Therefore we should focus our attention on a fixed point (x, t) and consider the approximate solution U_i^n when $k \to 0$ and $n \to \infty$ in such a way that $t_n = nk \to t$, and $h \to 0$ and $M \to \infty$ so that $x_i \to x$.

Definition 1 *(a) A finite difference scheme $L_i^n U_i^n = f_i^n$ is convergent to a solution u of the partial differential equation $\mathcal{L}u = f$ at a point (x, t) if*

$$U_i^n \to u(x, t)$$

when $(x_i, t_n) \to (x, t)$ if $h \to 0$ and $k \to 0$.
(b) The scheme is convergent of order (p, q) if

$$u(x, t) - U_i^n = O(h^p) + O(k^q)$$

when $(x_i, t_n) \to (x, t)$ if $h \to 0$ and $k \to 0$.

If a finite difference scheme is convergent the solution of the corresponding initial-boundary value problem can be approximated as closely as desired by choosing a sufficiently fine grid. A scheme is said to be *unconditionally convergent* if h and k can be chosen small independently of each other; however, for some schemes we have to restrict the manner in which h and k tend to zero, and such a scheme will be called *conditionally convergent.*

2.1.2 Convergence of the Explicit Method

It will now be proved that the explicit method for the heat equation $u_t = au_{xx}$ is conditionally convergent. The method was derived in Section 1.2.2:

$$U_i^{n+1} = rU_{i-1}^n + (1-2r)U_i^n + rU_{i+1}^n. \tag{2.1}$$

Since the scheme is accurate of order (2,1) (see Section 1.4) the solution $u(x,t)$ of the heat equation satisfies (writing $u(x_i, t_n) = u_i^n$)

$$\frac{1}{k}(u_i^{n+1} - u_i^n) = \frac{a}{h^2}\left[u_{i+1}^n - 2u_i^n + u_{i-1}^n\right] + O(h^2) + O(k)$$

or

$$u_i^{n+1} = ru_{i-1}^n + (1-2r)u_i^n + ru_{i+1}^n + O(kh^2) + O(k^2). \tag{2.2}$$

Denote the *discretization error* at the grid point (x_i, t_n) by

$$e_i^n = u_i^n - U_i^n$$

and subtract (2.1) from (2.2):

$$e_i^{n+1} = re_{i-1}^n + (1-2r)e_i^n + re_{i+1}^n + O(kh^2) + O(k^2).$$

If $r > 0$ and $1 - 2r \geq 0$ (or equivalently $0 < r \leq \frac{1}{2}$) it follows from the last equation and the triangle inequality that

$$\begin{aligned} |e_i^{n+1}| &\leq r\left|e_{i-1}^n\right| + (1-2r)\left|e_i^n\right| + r\left|e_{i+1}^n\right| + A(kh^2 + k^2) \\ &\leq rE^n + (1-2r)E^n + rE^n + A(kh^2 + k^2) \\ &= E^n + A(kh^2 + k^2) \end{aligned}$$

where A is a constant satisfying $O(k^2) + O(kh^2) \leq A(kh^2 + k^2)$ and E^n is the maximum discretization error on time level t_n:

$$E^n = \max_i \left|e_i^n\right|.$$

Since this holds for every point on time level t_{n+1}, we have that

$$E^{n+1} \leq E^n + A(kh^2 + k^2).$$

Applying this inequality repeatedly yields

$$\begin{aligned} E^n &\leq E^{n-1} + A(kh^2 + k^2) \\ &\leq E^{n-2} + 2A(kh^2 + k^2) \\ &\ \ \vdots \\ &\leq E^0 + nA(kh^2 + k^2) \\ &\leq TA(h^2 + k) \end{aligned}$$

since $E^0 = 0$ and $nk = t_n \leq T$ for all n, where t_n is restricted to the time interval $[0, T]$. We have thus shown that

$$u(x_i, t_n) - U_i^n = O(h^2) + O(k)$$

and hence that $U_i^n \to u(x, t)$ if h and $k \to 0$ in such a way that $(x_i, t_n) \to (x, t)$, provided that $0 < r \leq \frac{1}{2}$.

It should be noted that the assumption $0 < r \leq \frac{1}{2}$ was important in the convergence proof. We have not proved anything about the convergence (or otherwise) if r does not satisfy this condition. Numerical computations show that the explicit method is not convergent when $r > \frac{1}{2}$; this will be confirmed theoretically in Section 2.3.

2.2 Stability and the Theorem of Lax

2.2.1 Definition of Stability

A finite difference scheme is stable if small errors in the initial conditions cause small errors in the solution. An unstable scheme allows errors to grow explosively and swamp the true solution. Since rounding errors always enter a calculation, an unstable scheme cannot be used to calculate approximate solutions.

Definition 2 *A finite difference scheme $L_i^n U_i^n = f_i^n$ is stable if there exists a constant K such that the following condition holds:*
If U_i^n is calculated starting with initial values U_i^0, and $\tilde{U}_i^n$ is calculated starting with initial conditions $\tilde{U}_i^0$, then

$$\left|U_i^n - \tilde{U}_i^n\right| \leq K \left|U_i^0 - \tilde{U}_i^0\right|$$

for all i, and for all n such that $nk \leq T$.

Stability implies that the differences between the calculated values are small if the differences between the initial values are small.

2.2.2 The Equivalence Theorem of Lax

Consistency of a finite difference scheme connects the finite difference equation with the partial differential equation, while stability is a property of the solution of the finite difference equation. Convergence relates the solution of the finite difference equation to that of the PDE.

The following important theorem, proved by PD Lax and RD Richtmyer in 1956, provides a connection between these concepts.

Theorem 1 *A consistent finite difference scheme for a well-posed initial-boundary value problem is convergent if and only if it is stable.*

It is usually difficult to discover whether a finite difference scheme is convergent or not. In the next Section a method will be discussed which will allow us to determine when a scheme is stable. This, together with the consistency of the scheme, will then enable us to conclude whether the scheme is convergent or divergent.

The theorem of Lax refers to a *well-posed* initial-boundary value problem. This concept is defined as follows:

> An initial-boundary value problem is well-posed if (a) it has a unique solution, and (b) the solution depends continuously on the initial conditions; in other words, if the initial conditions are perturbed slightly, so is the solution.

Most initial-boundary value problems arising in practice are well-posed; if not, they would not be of much use in modelling physical processes.

We shall not attempt to prove the theorem of Lax here; the proof requires a precise definition of a finite difference scheme as well as some advanced concepts from the theory of Functional Analysis.

2.3 Von Neuman Analysis of Stability

The famous mathematician, theoretical physicist and pioneer computer scientist John von Neuman developed a stability test for finite difference schemes during the last years of World War II. In this section his technique will be discussed and illustrated with a number of examples. First we have to consider finite Fourier expansions.

2.3.1 Finite Fourier Transforms

If $\mathbf{u} = (u_0, u_1, \ldots, u_M)$ and $\mathbf{v} = (v_0, v_1, \ldots, v_M)$ are vectors with $M+1$ components (which may be real or complex numbers) the *inner product* of $\mathbf{u}$ and $\mathbf{v}$ is defined as

$$(\mathbf{u}, \mathbf{v}) = \sum_{i=0}^{M} u_i \bar{v}_i$$

and the *norm* of $\mathbf{u}$ is

$$\|\mathbf{u}\| = (\mathbf{u}, \mathbf{u})^{1/2} = \sqrt{\sum_{i=0}^{M} |u_i|^2}.$$

Let ϕ be the complex number

$$\phi = e^{I\xi} \quad \text{where } I = \sqrt{-1} \quad \text{and } \xi = \frac{2\pi}{M+1}$$

and define the vectors

$$\begin{aligned} \boldsymbol{\phi}_\ell &= (1, \phi^\ell, \phi^{2\ell}, \ldots, \phi^{M\ell}) \\ &= (1, e^{I\ell\xi}, e^{I2\ell\xi}, \ldots, e^{IM\ell\xi}), \qquad \ell = 0, 1, \ldots, M. \end{aligned}$$

Theorem 2

$$(\boldsymbol{\phi}_\ell, \boldsymbol{\phi}_m) = (M+1)\delta_{\ell m} = \begin{cases} M+1, & \ell = m \\ 0, & \ell \neq m. \end{cases}$$

Proof:

$$\begin{aligned} (\boldsymbol{\phi}_\ell, \boldsymbol{\phi}_m) &= \sum_{i=0}^{M} e^{Ii\ell\xi} e^{-Iim\xi} = \sum_{i=0}^{M} e^{Ii(\ell-m)\xi} \\ &= \sum_{i=0}^{M} q^i \qquad \text{where } q = e^{I(\ell-m)\xi} \\ &= \begin{cases} \dfrac{1-q^{M+1}}{1-q}, & q \neq 1 \\ M+1, & q = 1. \end{cases} \end{aligned}$$

Now $q = e^0 = 1$ if $\ell = m$ and $q^{M+1} = e^{2\pi I(\ell-m)} = 0$ if $\ell \neq m$, and the theorem follows.

Theorem 3 *(Finite Fourier Transform)*
Let $\mathbf{w} = (w_0, w_1, \ldots, w_M)$*. If*

$$\hat{w}_j = \sum_{i=0}^{M} e^{-Iij\xi} w_i, \qquad j = 0, 1, \ldots, M$$

then

$$w_i = \frac{1}{M+1} \sum_{j=0}^{M} e^{Iij\xi} \hat{w}_j, \qquad i = 0, 1, \ldots, M.$$

Proof:

$$\begin{aligned}\sum_{j=0}^{M} e^{Iij\xi}\hat{w}_j &= \sum_{j=0}^{M} \phi^{ij}\hat{w}_j = \sum_{j=0}^{M} \phi^{ij} \sum_{\ell=0}^{M} \bar{\phi}^{\ell j} w_\ell \\ &= \sum_{\ell=0}^{M} w_\ell \sum_{j=0}^{M} \phi^{ij}\bar{\phi}^{\ell j} = \sum_{\ell=0}^{M} w_\ell(\phi_i, \phi_\ell) \\ &= (M+1)w_i.\end{aligned}$$

■

Theorem 4 *(Finite Parseval relation)*

$$\|\mathbf{w}\|^2 = \frac{1}{M+1}\|\hat{\mathbf{w}}\|^2 .$$

Proof:

$$\begin{aligned}\|\hat{\mathbf{w}}\|^2 &= (\hat{\mathbf{w}}, \hat{\mathbf{w}}) = \sum_{i=0}^{M} \hat{w}_i\overline{\hat{w}_i} \\ &= \sum_{i=0}^{M}\sum_{\ell=0}^{M} e^{-Ii\ell\xi} w_\ell \sum_{m=0}^{M} e^{Iim\xi}\bar{w}_m = \sum_{\ell=0}^{M} w_\ell \sum_{m=0}^{M} \bar{w}_m(\phi_m, \phi_\ell) \\ &= (M+1)\sum_{\ell=0}^{M} w_\ell\bar{w}_\ell = (M+1)\|\mathbf{w}\|^2 .\end{aligned}$$

■

2.3.2 An Example: Stability of FTFS

We shall now illustrate the application of a finite Fourier transform to the analysis of the stability of the FTFS difference scheme for the transport equation. The scheme is

$$U_i^{n+1} = (1+p)U_i^n - pU_{i+1}^n$$

starting with initial values U_i^0. Suppose that we start with perturbed initial values $\tilde{U}_i^0$ and compute

$$\tilde{U}_i^{n+1} = (1+p)\tilde{U}_i^n - p\tilde{U}_{i+1}^n.$$

Let $w_i^n = U_i^n - \tilde{U}_i^n$. Subtracting the last two equations, we find

$$w_i^{n+1} = (1+p)w_i^n - pw_{i+1}^n.$$

Now find the finite Fourier transform of w_i^{n+1}:

$$\hat{w}_j^{n+1} = \sum_{i=0}^{M} e^{-Iij\xi} w_i^{n+1}$$

$$
\begin{aligned}
&= (1+p)\sum_{i=0}^{M} e^{-Iij\xi} w_i^n - p\sum_{i=0}^{M} e^{-Iij\xi} w_{i+1}^n \qquad \text{if } w_{M+1}^n = 0 \\
&= (1+p)\hat{w}_i^n - p\sum_{i=1}^{M} e^{-I(i-1)j\xi} w_i^n \\
&= (1+p)\hat{w}_i^n - pe^{Ij\xi}\hat{w}_i^n \qquad \text{if } w_0^n = 0
\end{aligned}
$$

or

$$\hat{w}_j^{n+1} = g\hat{w}_j^n \qquad \text{where } g = 1 + p - pe^{Ij\xi}$$

and hence

$$\hat{w}_j^n = g^n \hat{w}_j^0.$$

g is called the *amplification factor* of the finite difference scheme. If $|g| \leq 1$ it is clear that $|\hat{w}_j^n|$ will remain bounded when $n \to \infty$, and from $\|\hat{\mathbf{w}}\|^2 = (M+1)\|\mathbf{w}\|^2$ then follows that $|w_i^n|$ will also remain bounded and that the scheme will then be stable. Now

$$g = 1 + p - p(\cos\theta + I\sin\theta) \qquad \text{where } \theta = j\xi$$

and after some simplification we find that

$$|g|^2 = 1 + 2p(1+p)(1-\cos\theta).$$

The condition $|g| \leq 1$ will hold for all θ only if $p(1+p) \leq 0$, or

$$-1 \leq p \leq 0.$$

We have shown that if p satisfies this condition, the FTFS scheme is stable. In the next subsection it will be shown that this is also a necessary condition for stability, and hence for convergence, of the scheme.

2.3.3 Von Neuman's Stability Criterion

Suppose that we are computing with the difference scheme

$$L_i^n U_i^n = f_i^n \tag{2.3}$$

and that the values U_i^n and $\tilde{U}_i^n$ are computed starting with initial values U_i^0 and $\tilde{U}_i^0$ respectively. Let $w_i^n = U_i^n - \tilde{U}_i^n$ and suppose that the finite Fourier transform of the vector $\mathbf{w}^n$ is such that

$$\hat{w}_j^{n+1} = g\hat{w}_j^n$$

or

$$\hat{w}_j^n = g^n \hat{w}_j^0.$$

Theorem 5 *The finite difference scheme (2.3) is stable if and only if there is a constant K such that*

$$\|\hat{\mathbf{w}}^n\| \le K \left\|\hat{\mathbf{w}}^0\right\|.$$

Proof: By definition, the scheme is stable if and only if there is a constant K such that

$$|w_i^n| \le K \left|w_i^0\right|.$$

Therefore

$$\begin{aligned}\|\hat{\mathbf{w}}^n\|^2 &= (M+1)\|\mathbf{w}^n\|^2 \\ &\le (M+1)K^2\left\|\mathbf{w}^0\right\|^2 \\ &= K^2\left\|\hat{\mathbf{w}}^0\right\|^2.\end{aligned}$$

In exactly the same way it can be shown that, if $\|\hat{\mathbf{w}}^n\|^2 \le K^2\left\|\hat{\mathbf{w}}^0\right\|^2$, then $\|\mathbf{w}^n\|^2 \le K^2\left\|\mathbf{w}^0\right\|^2$. ■

Theorem 6 *(Von Neuman's Stability Criterion)*
A linear finite difference scheme is stable if and only if its amplification factor g satisfies

$$|g| \le 1 + O(k)$$

where k is the length of the time step.

Proof: (a) Suppose that $|g| \le 1 + O(k)$, or

$$|g| \le 1 + Ak \qquad \text{for some constant } A.$$

Then

$$|g^n| = |g|^n \le (1+Ak)^n \le (e^{Ak})^n = e^{Ank} \le e^{AT} \qquad \text{if } nk \le T$$

since $e^x = 1 + x + x^2 + \ldots \ge 1 + x$ if $x \ge 0$. From $\hat{w}_j^n = g^n \hat{w}_j^0$ then follows that

$$\|\hat{\mathbf{w}}^n\| \le K\left\|\hat{\mathbf{w}}^0\right\| \qquad \text{where } K = e^{AT}$$

and the scheme is stable.

(b) Suppose that the scheme is stable, then it follows from $\|\hat{\mathbf{w}}^n\| \le K\left\|\hat{\mathbf{w}}^0\right\|$ and $\|\hat{\mathbf{w}}^n\| = |g^n|\left\|\hat{\mathbf{w}}^0\right\|$ that

$$|g^n| \le K \qquad \text{if } n \to \infty \text{ and } nk \le T.$$

If $K \le 1$ then $|g^n| \le 1$ and $|g| \le 1$. If $K > 1$ then

$$\begin{aligned}|g| &\le K^{1/N} \qquad \text{where } kN = T \\ &= K^{k/T} \\ &\le 1 + K\frac{k}{T} \qquad \text{(see Fig. 2.1)}\end{aligned}$$

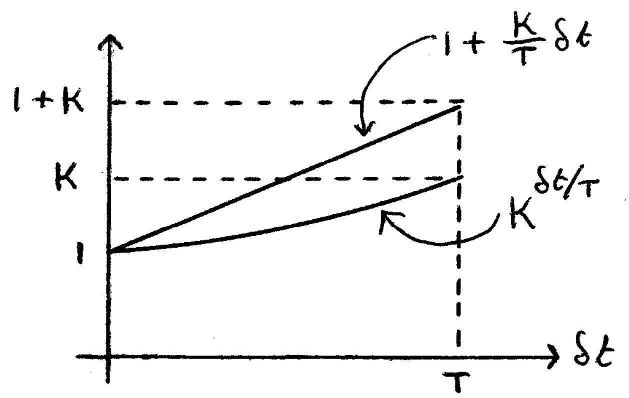

Fig. 2.1

and therefore

$$|g| \leq 1 + O(k).$$

This completes the proof.

The Von Neuman method for analyzing the stability of a finite difference scheme is strictly speaking only valid for pure initial value problems on the real line with periodic initial conditions. The method can also be employed for initial-boundary values with Dirichlet boundary values (i.e. the value of the solution $u(x,t)$ is prescribed on the boundaries) as was shown in the example on the FTFS method. For other types of boundary conditions the Von Neuman criterion can be shown to be necessary for stability, but not always sufficient.

Since the error w_i^n satisfies the same difference equation as does the approximate solution U_i^n, the finite Fourier transform of U_i^{n+1} can be computed to obtain the amplification factor g. It is not even necessary to calculate the finite Fourier transform; a simpler and equivalent procedure is to replace U_i^n in the scheme by $g^n e^{Ii\theta}$ and solve the resulting equation for g.

Example 1.

The explicit difference scheme for the heat equation is

$$U_i^{n+1} = rU_{i-1}^n + (1-2r)U_i^n + rU_{i+1}^n.$$

Substituting $U_i^n = g^n e^{Ii\theta}$ we have

$$g^{n+1}e^{Ii\theta} = rg^n e^{I(i-1)\theta} + (1-2r)g^n e^{Ii\theta} + rg^n e^{I(i+1)\theta}$$

and after cancelling $g^n e^{Ii\theta}$ we find

$$g \;=\; re^{-I\theta} + (1-2r) + re^{I\theta}$$

$$\begin{aligned} &= 1 - 2r + 2r\cos\theta \\ &= 1 - 4r\sin^2\frac{\theta}{2} \qquad (\cos 2\alpha = 1 - 2\sin^2\alpha) \end{aligned}$$

and the condition $|g| \leq 1$ will hold if and only if

$$0 \leq r \leq \frac{1}{2\sin^2\frac{\theta}{2}}.$$

This will be true for all θ if and only if $0 \leq r \leq 1/2$. We have thus confirmed that the explicit method is convergent if $0 < r \leq 1/2$ and divergent if $r > 1/2$. ■

It is usually sufficient to use the simple form $|g| \leq 1$ of the Von Neuman criterion as was done in this example. When a PDE contains lower order terms the more general form $|g| \leq 1 + O(k)$ must be used. Also note that non-homogeneous terms (terms not containing U_i^n) in a difference equation can be dropped when calculating the amplification factor, since the equation for the error w_i^n is actually used for finding g.

Example 2.

For the heat equation with lower order term bu and source term f

$$u_t = au_{xx} + bu + f$$

the explicit difference scheme is

$$U_i^{n+1} = rU_{i-1}^n + (1-2r)U_i^n + rU_{i+1}^n + kbU_i^n + kf_i^n$$

and the amplification factor is found to be (ignoring the non-homogeneous term kf_i^n):

$$g = 1 - 4r\sin^2\frac{\theta}{2} + bk.$$

The stability condition $|g| \leq 1 + O(k)$ will again hold if $0 < r \leq 1/2$. ■

We have derived the Von Neuman criterion for two-level difference schemes. It can be shown that the condition $|g| \leq 1 + O(k)$ is also necessary for the stability of multi-level schemes, but in some cases it may not be sufficient.

Example 3.

The leapfrog scheme for the transport equation is

$$U_i^{n+1} = U_i^{n-1} - p(U_{i+1}^n - U_{i-1}^n).$$

Substitute $U_i^n = g^n e^{Ii\theta}$ and cancel terms:

$$g^{n+1}e^{Ii\theta} = g^{n-1}e^{Ii\theta} - p(g^n e^{I(i+1)\theta} - g^n e^{I(i-1)\theta})$$

or

$$g^2 + 2Ip\sin\theta\, g - 1 = 0.$$

The roots of this equation are

$$g = -Ip\sin\theta \pm \sqrt{1 - p^2\sin^2\theta}.$$

If $p^2 \leq 1$, then $1 - p^2\sin^2\theta \geq 0$ and

$$|g|^2 = 1 - p^2\sin^2\theta + (p\sin\theta)^2 = 1.$$

If $p^2 > 1$ then for $\theta = \pi/2$ one of the roots satisfies

$$|g| = |p| + \sqrt{p^2 - 1} > 1$$

and the scheme is unstable. A more sophisticated analysis which we shall not undertake here, shows that the scheme is actually unstable when $|p| = 1$; the stability condition is therefore

$$-1 < p < 1.$$