Exercises

1. Prove that the Crank-Nicolson scheme (10.118) has accuracy $\mathcal{O}(h^2)$ provided that $r = \tau/h = \text{const}$.
2. Show that the Lax-Friedrichs scheme (10.85) is dissipative of order $2d = 2$, although not strictly in the sense of Definition 10.4. Prove that it rather satisfies inequality (10.114) for all $|\alpha| \le \pi - \varepsilon$, where $\varepsilon > 0$ can be arbitrary.
3. Use Theorem 10.5 to study stability of the scheme:

$$\frac{u_m^{p+1} - u_m^p}{\tau} - a(x_m)\frac{u_{m+1}^p - u_m^p}{h} = 0,$$
$$u_m^0 = \psi(x_m), \quad m = 0, \pm 1, \pm 2, \ldots, \quad p = 0, 1, 2, \ldots, [T/\tau] - 1,$$

   for the Cauchy problem:

$$\frac{\partial u}{\partial t} - a(x)\frac{\partial u}{\partial x} = 0, \quad -\infty < x < \infty, \quad 0 < t \le T,$$
$$u(x,0) = \psi(x), \quad -\infty < x < \infty,$$

   where $a(x)$ is a smooth function and $a_1 \ge a(x) \ge a_0 > 0$.
4. Show that the implicit downwind scheme (10.106) for the Cauchy problem (10.116) is dissipative of order $2d = 2$ when $r > 1$.
5. Show that the implicit central scheme (10.107) for the Cauchy problem (10.116) is dissipative of order $2d = 2$ in the same non-strict sense as outlined in Exercise 2.

10.5 Stability for Initial Boundary Value Problems

Instead of the Cauchy problem (10.108), let us now consider an initial boundary value problem for the heat equation formulated on the finite interval $0 \le x \le 1$:

$$\begin{gathered}\frac{u_m^{p+1} - u_m^p}{\tau} - a(x_m, t_p)\frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0,\\ u_m^0 = \psi(x_m), \quad \boldsymbol{l}_1 u_0^{p+1} = 0, \quad \boldsymbol{l}_2 u_M^{p+1} = 0,\\ m = 0, 1, 2, \ldots, M, \quad p \ge 0.\end{gathered} \tag{10.124}$$

In formula (10.124), we assume that the grid is uniform: $x_m = mh, m = 0, 1, 2, \ldots, M$, $M = 1/h, t_p = p\tau, p = 0, 1, 2, \ldots$, and denote by $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ the operators of the boundary conditions at the left and right endpoints of the interval, respectively.

10.5.1 The Babenko-Gelfand Criterion

To analyze stability of the difference problem (10.124), we will first develop a heuristic argument based on freezing the coefficients; this argument will further extend the previous considerations of Section 10.4.1. In Section 10.4.1, we have noticed that because of the continuity of the coefficient $a = a(x,t)$, its variation within

a fixed number of cells around any given point $(\tilde{x},\tilde{t})$ becomes smaller when the grid is refined. In the context of the initial boundary value problem (10.124), as opposed to the initial value problem (10.108), we supplement this consideration by another obvious observation. If the point $(\tilde{x},\tilde{t})$ lies inside the domain, then the distance from this point to either of the endpoints, $x = 0$ or $x = 1$, measured in the number of grid cells (of size h) will increase with no bound when $h \longrightarrow 0$. In other words, on fine grids the point $(\tilde{x},\tilde{t})$ can be considered to be located far away from the boundaries. Consequently, we can still claim that for small h the perturbations superimposed on the solution of problem (10.124) at the moment of time $t = \tilde{t}$ near any interior space location $x = \tilde{x}$ will evolve similarly to how the perturbations of the solution to the same "old" constant-coefficient equation (10.109) would have evolved. This, in turn, implies that stability of the scheme (10.109) for every $(\tilde{x},\tilde{t})$ inside the domain is still necessary for the overall stability of scheme (10.124).

The foregoing heuristic argument, however, becomes far less convincing if the point $(\tilde{x},\tilde{t})$ happens to lie precisely on one of the lateral boundaries: $x = 0$ or $x = 1$. For example, when we let $\tilde{x} = 0$, the distance from $(\tilde{x},\tilde{t})$ to any fixed location $x > 0$ (and in particular, to the right endpoint $x = 1$) measured in the number of grid cells will again increase with no bound as $h \longrightarrow 0$. Yet the number of grid cells to the left endpoint $x = 0$ will not change and will remain equal to zero. In other words, the point $(\tilde{x},\tilde{t})$ will never be far from the left boundary, no matter how fine the grid may be. Consequently, we can no longer expect that perturbations of the solution to problem (10.124) near $\tilde{x} = 0$ will behave similarly to perturbations of the solution to equation (10.109), as the latter is formulated on the grid infinite in both directions.

Instead, we shall rather expect that over the short periods of time the perturbations of the solution to problem (10.124) near the left endpoint $x = 0$ will develop analogously to perturbations of the solution to the following constant-coefficient problem:

$$\begin{aligned} &\frac{u_m^{p+1} - u_m^p}{\tau} - a(0,\tilde{t})\frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0, \\ &\mathit{l}_1 u_0^{p+1} = 0, \quad m = 0,1,2,\ldots, \quad p \geq 0. \end{aligned} \tag{10.125}$$

Problem (10.125) is formulated on the semi-infinite grid: $m = 0,1,2,\ldots$ (i.e., semi-infinite line $x \geq 0$). It is obtained from the original problem (10.124) by freezing the coefficient $a(x,t)$ at the left endpoint of the interval $0 \leq x \leq 1$ and by simultaneously "pushing" the right boundary off all the way to $+\infty$. Problem (10.125) shall be analyzed only for those grid functions $u^p = \{u_0^p, u_1^p, \ldots\}$ that satisfy:

$$u_m^p \longrightarrow 0, \quad \text{as} \quad m \longrightarrow +\infty. \tag{10.126}$$

Indeed, only in this case will the perturbation be concentrated near the left boundary $x = 0$, and only for the perturbations of this type will the problems (10.124) and (10.125) be similar to one another in the vicinity of $x = 0$.

Likewise, the behavior of perturbations to the solution of problem (10.124) near

the right endpoint $x = 1$ should resemble that for the problem:

$$\begin{aligned} &\frac{u_m^{p+1} - u_m^p}{\tau} - a(1,\tilde{t})\frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0, \\ &\boldsymbol{l}_2 u_M^{p+1} = 0, \quad m = \ldots, -2, -1, 0, 1, 2, \ldots, M, \quad p \geq 0, \end{aligned} \tag{10.127}$$

that has only one boundary at $m = M$. Problem (10.127) is derived from problem (10.124) by freezing the coefficient $a(x,t)$ at the right endpoint of the interval $0 \leq x \leq 1$ and by simultaneously pushing the left boundary off all the way to $-\infty$. It should be considered only for the grid functions $u^p = \{\ldots, u_{-1}^p, u_0^p, u_1^p, \ldots, u_M^p\}$ that satisfy:

$$u_m^p \longrightarrow 0, \quad \text{as} \quad m \longrightarrow -\infty. \tag{10.128}$$

The three problems: (10.109), (10.125), and (10.127), are easier to investigate than the original problem (10.124), because they are all h independent provided that $r = \tau/h^2 = \text{const}$, and they all have constant coefficients.

Thus, the issue of studying stability for the scheme (10.124), with the effect of the boundary conditions taken into account, can be addressed as follows. One needs to formulate three auxiliary problems: (10.109), (10.125), and (10.127). For each of these three h independent problems, one needs to find all those numbers λ (eigenvalues of the transition operator from u^p to u^{p+1}), for which solutions of the type

$$u_m^p = \lambda^p u_m^0 \tag{10.129}$$

exist. In doing so, for problem (10.109), the function $u^0 = \{u_m^0\}, m = 0, \pm 1, \pm 2, \ldots,$ has to be bounded on the grid. For problem (10.125), the grid function $u^0 = \{u_m^0\}$, $m \geq 0$, has to satisfy: $u_m^0 \longrightarrow 0$ as $m \longrightarrow +\infty$, and for problem (10.125), the grid function $u^0 = \{u_m^0\}$, $m \leq M$, has to satisfy: $u_m^0 \longrightarrow 0$ as $m \longrightarrow -\infty$. For scheme (10.124) to be stable, it is necessary that the overall spectrum of the difference initial boundary value problem, i.e., all eigenvalues of all three problems: (10.109), (10.125), and (10.127), belong to the unit disk: $|\lambda| \leq 1$, on the complex plane. This is the Babenko-Gelfand stability criterion. Note that problem (10.109) has to be considered for every fixed $\tilde{x} \in (0,1)$ and all $\tilde{t}$.

REMARK 10.1 Before we continue to study problem (10.124), let us present an important intermediate conclusion that can already be drawn based on the foregoing qualitative analysis. For stability of the pure Cauchy problem (10.108) that has no boundary conditions it is necessary that finite-difference equations (10.109) be stable in the von Neumann sense $\forall(\tilde{x},\tilde{t})$. This requirement remains necessary for stability of the initial boundary value problem (10.124) as well. Moreover, when boundary conditions are present, two more auxiliary problems: (10.125) and (10.127), have to be stable in a similar sense. Therefore, adding boundary conditions to a finite-difference Cauchy problem *will not, generally speaking, improve its stability.* Boundary conditions may either remain neutral or hamper the overall stability if, for example, problem

(10.109) appears stable but one of the problems (10.125) or (10.127) happens to be unstable. Later on, we will discuss this phenomenon in more detail. □

Let us now assume for simplicity that $a(x,t) \equiv 1$ in problem (10.124), and let us calculate the spectra of the three auxiliary problems (10.109), (10.125), and (10.127) for various boundary conditions $l_1 u_0^{p+1} = 0$ and $l_2 u_M^{p+1} = 0$.

Substituting the solution in the form $u_m^p = \lambda^p u_m$ into the finite-difference equation (10.109), we obtain:

$$(\lambda - 1)u_m - r(u_{m+1} - 2u_m + u_{m-1}) = 0, \quad r = \tau/h^2,$$

which immediately yields:

$$u_{m+1} - \frac{\lambda - 1 + 2r}{r} u_m + u_{m-1} = 0. \tag{10.130}$$

This is a second order homogeneous ordinary difference equation. To find the general solution of equation (10.130) we write down its algebraic characteristic equation:

$$q^2 - \frac{\lambda - 1 + 2r}{r} q + 1 = 0. \tag{10.131}$$

If q is a root of the quadratic equation (10.131), then the grid function

$$u_m^p = \lambda^p q^m$$

solves the homogeneous finite-difference equation:

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0. \tag{10.132}$$

If $|q| = 1$, i.e., if $q = e^{i\alpha}$, then the grid function

$$u_m^p = \lambda^p e^{i\alpha m},$$

which is obviously bounded for $m \longrightarrow +\infty$ and $m \longrightarrow -\infty$, yields the solution of equation (10.132), provided that

$$\lambda = 1 - 4r \sin^2 \frac{\alpha}{2}, \quad 0 \le \alpha < 2\pi,$$

see Example 6 of Section 10.3.3. These $\lambda = \lambda(\alpha)$ fill the interval $1 - 4r \le \lambda \le 1$ of the real axis, see Figure 10.8 on page 357. Therefore, interval $1 - 4r \le \lambda \le 1$ is the spectrum of problem (10.109) for $a(\tilde{x},\tilde{t}) = 1$, i.e., of problem (10.132). This problem has no eigenvalues that lie outside of the interval $1 - 4r \le \lambda \le 1$, because if the characteristic equation (10.131) does not have a root q with $|q| = 1$, then equation (10.130) may have no solution bounded for $m \longrightarrow \pm\infty$.

If λ does not belong to the interval $1 - 4r \le \lambda \le 1$, then the absolute values of both roots of the characteristic equation (10.131) differ from one. Their product, however,

is still equal to one, see equation (10.131). Consequently, the absolute value of the first root of equation (10.131) will be greater than one, while that of the second root will be less than one. Let us denote $|q_1(\lambda)| < 1$ and $|q_2(\lambda)| > 1$. The general solution of equation (10.130) has the form:

$$u_m = c_1 q_1^m + c_2 q_2^m,$$

where c_1 and c_2 are arbitrary constants. Accordingly, the general solution that satisfies additional constraint (10.126), i.e., that decays as $m \longrightarrow +\infty$, is written as

$$u_m = c_1 q_1^m, \quad |q_1| = |q_1(\lambda)| < 1,$$

and the general solution that satisfies additional constraint (10.128), i.e., that decays as $m \longrightarrow -\infty$, is given by

$$u_m = c_2 q_2^m, \quad |q_2| = |q_2(\lambda)| > 1.$$

To calculate the eigenvalues of problem (10.125), one needs to substitute $u_m^p = c_1 \lambda^p q_1^m$ into the left boundary condition $\boldsymbol{l}_1 u_0 = 0$ and find those q_1 and λ, for which it is satisfied. If, for example, $\boldsymbol{l}_1 u_0 \equiv u_0 = 0$, then $c_1 \lambda^p q_1^0 = 0$ implies $\lambda = 0$, because $c_1 = 0$ would mean a zero eigenfunction. Thus, $\lambda = 0$ is an eigenvalue provided that $r < 1/4$, because only in this case for $\lambda = 0$ we may have $|q_1| < 1$. Otherwise, problem (10.125) has no eigenvalues. Likewise, if $\boldsymbol{l}_1 u_0 \equiv u_1 - u_0 = 0$, then $c_1 \lambda^p (q_1 - q_1^0) = c_1 \lambda^p (q_1 - 1) = 0$ yields either $\lambda = 0$ for $r < 1/4$ or otherwise no eigenvalues because $c_1 \neq 0$ and $q_1 \neq 1$. If, however, $\boldsymbol{l}_1 u_0 \equiv 2u_1 - u_0 = 0$, then condition $c_1 \lambda^p (2q_1 - q_1^0) = c_1 \lambda^p (2q_1 - 1) = 0$ is satisfied for $c_1 \neq 0$ and $q_1 = 1/2 < 1$. Substituting $q_1 = 1/2$ into the characteristic equation (10.131) we find that

$$\lambda = 1 + r\left(q_1 - 2 + \frac{1}{q_1}\right) = 1 + \frac{r}{2}.$$

This is the only eigenvalue of problem (10.125). It does not belong to the unit disk on the complex plane, and therefore the necessary stability condition is violated.

The eigenvalues of the auxiliary problem (10.127) are calculated analogously. They are found from the equation $\boldsymbol{l}_2 u_M = 0$ when

$$u_m = c_2 q_2^m, \quad |q_2| = |q_2(\lambda)| > 1, \quad m = M, M-1, M-2, \ldots.$$

For stability, it is necessary that they all belong to the unit disk on the complex plane.

We can now provide more specific comments following Remark 10.1. When boundary condition $\boldsymbol{l}_1 u_0 \equiv 2u_1 - u_0 = 0$ is employed in problem (10.125) then the solution that satisfies condition (10.126) is found in the form $u_m^p = \lambda^p q_1^m$, where $q_1 = 1/2$ and $\lambda = 1 + r/2 > 1$. This solution is only defined for $m \geq 0$. If, however, we were to extend it to the region $m < 0$, we would have obtained an unbounded function: $u_m^p \longrightarrow \infty$ as $m \longrightarrow -\infty$. In other words, the function $u_m^p = \lambda^p q_1^m$ cannot be used in the framework of the standard von Neumann analysis of problem (10.109).

This consideration leads to a very simple explanation of the mechanism of instability. The introduction of a boundary condition merely expands the pool of candidate

functions, on which the instability may develop. In the pure von Neumann case, with no boundary conditions, we have only been monitoring the behavior of the harmonics $e^{i\alpha m}$ that are bounded on the entire grid $m = 0, \pm 1, \pm 2, \ldots$. With the boundary conditions present, we may need to include additional functions that are bounded on the semi-infinite grid, but unbounded if extended to the entire grid. These functions do not belong to the von Neumann category. If any of them brings along an unstable eigenvalue $|\lambda| > 1$, such as $\lambda = 1 + r/2$, then the overall scheme becomes unstable as well. We therefore re-iterate that if the scheme that approximates some Cauchy problem is supplemented by boundary conditions and thus transformed into an initial boundary value problem, then its stability will not be improved. In other words, if the Cauchy problem was stable, then the initial boundary value problem may either remain stable or become unstable. If, however, the Cauchy problem is unstable, then the initial boundary value problem will not become stable.

Our next example will be the familiar first order upwind scheme, but built on a finite grid: $x_m = mh$, $m = 0, 1, 2, \ldots, M$, $Mh = 1$, rather than on the infinite grid $m = 0, \pm 1, \pm 2, \ldots$:

$$\begin{gathered} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = 0, \\ m = 0, 1, 2, \ldots, M-1, \quad p = 0, 1, 2, \ldots, [T/\tau] - 1, \\ u_m^0 = \psi(x_m), \quad u_M^{p+1} = 0. \end{gathered} \tag{10.133}$$

Scheme (10.133) approximates the following first order hyperbolic initial boundary value problem:

$$\begin{gathered} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = 0, \quad 0 \le x \le 1, \quad 0 < t \le T, \\ u(x,0) = \psi(x), \quad u(1,t) = 0, \end{gathered}$$

on the interval $0 \le x \le 1$. To investigate stability of scheme (10.133), we will employ the Babenko-Gelfand criterion. In other words, we will need to analyze three auxiliary problems: A problem with no lateral boundaries:

$$\begin{gathered} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = 0, \\ m = 0, \pm 1, \pm 2, \ldots, \end{gathered} \tag{10.134}$$

a problem with only the left boundary:

$$\begin{gathered} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = 0, \\ m = 0, 1, 2, \ldots, \end{gathered} \tag{10.135}$$

and a problem with only the right boundary:

$$\begin{gathered} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = 0, \\ m = M-1, M-2, \ldots, 1, 0, -1, \ldots, \\ u_M^{p+1} = 0. \end{gathered} \tag{10.136}$$

Note that we do not set any boundary condition at the left boundary in problem (10.135) as we did not have any in the original problem (10.133) either.

We will need to find spectra of the three transition operators from u^p to u^{p+1} that correspond to the three auxiliary problems (10.134), (10.135), and (10.136), respectively, and determine under what conditions will all the eigenvalues belong to the unit disk $|\lambda| \le 1$ on the complex plane.

Substituting a solution of the type:

$$u_m^p = \lambda^p u_m$$

into the finite-difference equation:

$$u_m^{p+1} = (1-r)u_m^p + r u_{m+1}^p, \quad r = \tau/h,$$

that corresponds to all three problems (10.134), (10.135), and (10.136), we obtain the following first order ordinary difference equation for the eigenfunction $\{u_m\}$:

$$(\lambda - 1 + r)u_m - r u_{m+1} = 0. \tag{10.137}$$

Its characteristic equation:

$$\lambda - 1 + r - rq = 0 \tag{10.138}$$

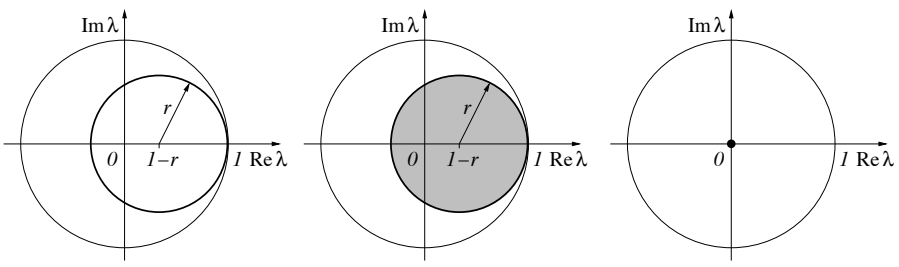
yields the relation between λ and q, so that the general solution of equation (10.137) can be written as

$$u_m = cq^m = c\left(\frac{\lambda - 1 + r}{r}\right)^m, \quad m = 0, \pm 1, \pm 2, \ldots, \quad c = \text{const}.$$

When $|q| = 1$, i.e., when $q = e^{i\alpha}$, $0 \le \alpha < 2\pi$, we have:

$$\lambda = 1 - r + re^{i\alpha}.$$

The point $\lambda = \lambda(\alpha)$ sweeps the circle of radius r centered at the point $(1-r, 0)$ on the complex plane. This circle gives the spectrum, i.e., the full set of eigenvalues, of the first auxiliary problem (10.134), see Figure 10.11(a). It is clearly the same spectrum as we have discussed in Section 10.3.2, see formula (10.77) and Figure 10.5(a).



(a) Problem (10.134) (b) Problem (10.135) (c) Problem (10.136)

FIGURE 10.11: Spectra of auxiliary problems for the upwind scheme (10.133).

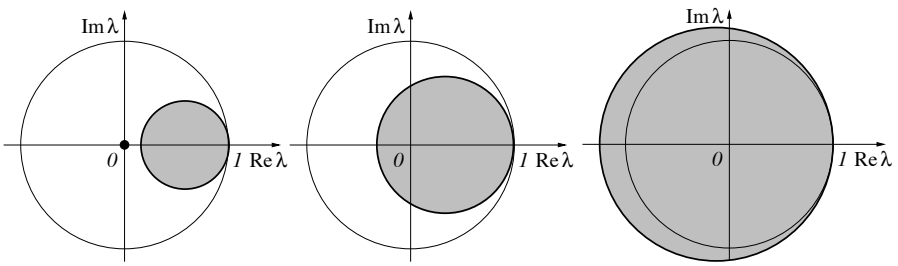
As far as the second auxiliary problem (10.135), we need to look for its non-trivial solutions that would decrease as $m \longrightarrow +\infty$, see formula (10.126). Such a solution, $u_m = c\lambda^p q^m$, obviously exists for any q: $|q| < 1$. The corresponding eigenvalues $\lambda = \lambda(q) = 1 - r + rq$ fill the interior of the disk bounded by the circle $\lambda = 1 - r + re^{i\alpha}$ on the complex plane, see Figure 10.11(b).

Solutions of the third auxiliary problem (10.136) that would satisfy (10.128), i.e., that would decay as $m \longrightarrow -\infty$, must obviously have the form: $u_m^p = c\lambda^p q^m$, where $|q| > 1$ and the relation between λ and q is, again, given by formula (10.138). The homogeneous boundary condition $u_M^{p+1} = 0$ of (10.136) implies that a non-trivial eigenfunction $u_m = cq^m$ may only exist when $\lambda = \lambda(q) = 0$, i.e., when $q = (r-1)/r$. The quantity q given by this expression may have its absolute value greater than one if either of the two inequalities holds:

$$\frac{r-1}{r} > 1 \quad \text{or} \quad \frac{r-1}{r} < -1.$$

The first inequality has no solutions. The solution to the second inequality is $r < 1/2$. Consequently, when $r < 1/2$, problem (10.136) has the eigenvalue $\lambda = 0$, see Figure 10.11(c).

In Figure 10.12, we are schematically showing the combined sets of all eigenvalues, i.e., combined spectra, for problems (10.134), (10.135), and (10.136) for the three different cases: $r < 1/2$, $1/2 < r < 1$, and $r > 1$.



(a) $r < 1/2$ (b) $1/2 < r < 1$ (c) $r > 1$

FIGURE 10.12: Combined spectra of auxiliary problems for scheme (10.133).

It is clear that the combined eigenvalues of all three auxiliary problems may only belong to the unit disk $|\lambda| \leq 1$ on the complex plane if $r \leq 1$. Therefore, condition $r \leq 1$ is necessary for stability of the difference initial boundary value problem (10.133).

Compared to the von Neumann stability condition of Section 10.3, the key distinction of the Babenko-Gelfand criterion is that it takes into account the boundary conditions for unsteady finite-difference equations on finite intervals. This criterion can also be generalized to systems of such equations. In this case, a scheme that may look perfectly natural and "benign" at a first glance, and that may, in particular,

satisfy the von Neumann stability criterion, could still be unstable because of a poor approximation of the boundary conditions. Consequently, it is important to be able to build schemes that are free of this shortcoming.

In [GR87], the spectral criterion of Babenko and Gelfand is discussed from a more general standpoint, using a special new concept of *the spectrum of a family of operators* introduced by Godunov and Ryaben'kii. In this framework, one can rigorously prove that the Babenko-Gelfand criterion is necessary for stability, and also that when it holds, stability cannot be disrupted too severely. In the next section, we reproduce key elements of this analysis, while referring the reader to [GR87, Chapters 13 & 14] and [RM67, § 6.6 & 6.7] for further detail.

10.5.2 Spectra of the Families of Operators. The Godunov-Ryaben'kii Criterion

In this section, we briefly describe a rigorous approach, due to Godunov and Ryaben'kii, for studying stability of evolution-type finite-difference schemes on finite intervals. In other words, we study stability of the discrete approximations to initial boundary value problems for hyperbolic and parabolic partial differential equations. This material is more advanced, and can be skipped during the first reading.

As we have seen previously, for evolution finite-difference schemes the discrete solution $u^{(h)} = \{u_m^p\}$, which is defined on a two-dimensional space-time grid:

$$(x_m, t_p) \equiv (mh, p\tau), \quad m = 0,1,\ldots,M, \quad p = 0,1,\ldots,[T/\tau],$$

gets naturally split or "stratified" into a collection of one-dimensional grid functions $\{u^p\}$ defined for individual time layers t_p, $p = 0,1,\ldots,[T/\tau]$. For example, the first order upwind scheme:

$$\begin{gathered} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = \varphi_m^p, \\ m = 0,1,2,\ldots,M-1, \quad p = 0,1,2,\ldots,[T/\tau]-1, \\ u_m^0 = \psi_m, \quad u_M^{p+1} = \chi^{p+1}, \end{gathered} \tag{10.139}$$

for the initial boundary value problem:

$$\begin{gathered} \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = \varphi(x,t), \quad 0 \le x \le 1, \quad 0 < t \le T, \\ u(x,0) = \psi(x), \quad u(1,t) = \chi(t), \end{gathered}$$

can be written as:

$$\begin{gathered} u_m^{p+1} = \left[(1-r)u_m^p + r u_{m+1}^p\right] + \tau\varphi_m^p, \quad m = 0,1,\ldots,M-1, \\ u_M^{p+1} = \chi^{p+1}, \quad u_m^0 = \psi_m, \quad m = 0,1,\ldots,M, \end{gathered} \tag{10.140}$$

where $r = \tau/h$. Form (10.140) suggests that the marching procedure for scheme (10.139) can be interpreted as consecutive computation of the grid functions:

$$u^0, u^1, \ldots, u^p, \ldots, u^{[T/\tau]},$$

defined on identical one-dimensional grids $m = 0, 1, \ldots, M$ that can all be identified with one and the same grid. Accordingly, the functions u^p, $p = 0, 1, \ldots, [T/\tau]$, can be considered elements of the linear space U_h' of functions $u = \{u_0, u_1, \ldots, u_M\}$ defined on the grid $m = 0, 1, \ldots, M$. We will equip this linear space with the norm, e.g.,

$$\|u\|_{U_h'} = \max_{0 \le m \le M} |u_m| \qquad \text{or} \qquad \|u\|_{U_h'} = \left[h \sum_{m}^{M} |u_m|^2 \right]^{\frac{1}{2}}.$$

We also recall that in the definitions of stability (Section 10.1.3) and convergence (Section 10.1.1) we employ the norm $\|u^{(h)}\|_{U_h}$ of the finite-difference solution $u^{(h)}$ on the entire two-dimensional grid. Hereafter, we will only be using norms that explicitly take into account the layered structure of the solution, namely, those that satisfy the equality:

$$\|u^{(h)}\|_{U_h} = \max_{0 \le p \le [T/\tau]} \|u^p\|_{U_h'}.$$

Having introduced the linear normed space U_h', we can represent any evolution scheme, in particular, scheme (10.139), in *the canonical form:*

$$\begin{aligned} &u^{p+1} = \boldsymbol{R}_h u^p + \tau \rho^p, \\ &u^0 \ \text{ is given.} \end{aligned} \tag{10.141}$$

In formula (10.141), $\boldsymbol{R}_h : U_h' \longmapsto U_h'$ is the transition operator between the consecutive time levels, and $\rho^p \in U_h'$. If we denote $v^{p+1} = \boldsymbol{R}_h u^p$, then formula (10.140) yields:

$$v_m^{p+1} = (1-r)u_m^p + r u_{m+1}^p, \quad m = 0, 1, \ldots M-1. \tag{10.142a}$$

As far as the last component $m = M$ of the vector v^{p+1}, a certain flexibility exists in the definition of the operator $\boldsymbol{R}_h$ for scheme (10.139). For example, we can set:

$$v_M^{p+1} = u_M^p, \tag{10.142b}$$

which would also imply:

$$\rho_m^p = \varphi_m^p, \quad m = 0, 1, \ldots M-1, \quad \text{and} \quad \rho_M^p = \frac{\chi^{p+1} - \chi^p}{\tau}, \tag{10.142c}$$

in order to satisfy the first equality of (10.141).

In general, the canonical form (10.141) for a given evolution scheme is not unique. For scheme (10.139), we could have chosen $v_M^{p+1} = 0$ instead of $v_M^{p+1} = u_M^p$ in formula (10.142b), which would have also implied $\rho_M^p = \frac{\chi^{p+1}}{\tau}$ in formula (10.142c). However, when building the operator $\boldsymbol{R}_h$, we need to make sure that the following

rather natural conditions hold that require certain correlation between the norms in the spaces U_h' and F_h:

$$\begin{gathered}\|\rho^p\|_{U_h'} \le K_1 \|f^{(h)}\|_{F_h}, \quad p = 0,1,\ldots,[T/\tau], \\ \|u^0\|_{U_h'} \le K_2 \|f^{(h)}\|_{F_h}.\end{gathered} \tag{10.143}$$

The constants K_1 and K_2 in inequalities (10.143) do not depend on h or on $f^{(h)}$. For scheme (10.139), if we define the norm in the space F_h as:

$$\|f^{(h)}\|_{F_h} = \max_{m,p} |\varphi_m^p| + \max_m |\psi_m| + \max_p \left| \frac{\chi^{p+1} - \chi^p}{\tau} \right|$$

and the norms of ρ^p and u_0 as $\|\rho^p\|_{U_h'} = \max\limits_m |\rho_m^p|$ and $\|u^0\|_{U_h'} = \max\limits_m |u_m^0|$, respectively, then conditions (10.143) obviously hold for the operator $\boldsymbol{R}_h$ and the source term ρ^p defined by formulae (10.142a)–(10.142c).

Let us now take an arbitrary $\hat{u}^0 \in U_h'$ and obtain $\hat{u}^1, \hat{u}^2, \ldots, \hat{u}^{[T/\tau]} \in U_h'$ using the recurrence formula $\hat{u}^{p+1} = \boldsymbol{R}_h \hat{u}^p$. Denote $\hat{u}^{(h)} = \{\hat{u}^p\}_{p=0}^{[T/\tau]}$ and evaluate $\hat{f}^{(h)} \stackrel{\text{def}}{=} \boldsymbol{L}_h \hat{u}^{(h)}$. Along with conditions (10.143), we will also require that

$$\|\hat{f}^{(h)}\|_{F_h} \le K_3 \|\hat{u}^0\|_{U_h'}, \tag{10.144}$$

where the constant K_3 does not depend on $\hat{u}^0 \in U_h'$ or on h.

In practice, inequalities (10.143) and (10.144) prove relatively non-restrictive.[7] These inequalities allow one to establish the following important theorem that provides a necessary and sufficient condition for stability in terms of the uniform boundedness of the powers of $\boldsymbol{R}_h$ with respect to the grid size h.

THEOREM 10.6

Assume that when reducing a given evolution scheme to the canonical form (10.141) the additional conditions (10.143) are satisfied. Then, for stability of the scheme in the linear sense (Definition 10.2) it is sufficient that

$$\|\boldsymbol{R}_h^p\| \le K, \quad p = 0,1,\ldots,[T/\tau], \tag{10.145}$$

where the constant K in formula (10.145) does not depend on h. If the third additional condition (10.144) is met as well, then estimates (10.145) are also necessary for stability.

Theorem 10.6 is proven in [GR87, § 41].

For scheme (10.139), estimates (10.145) can be established directly, provided that $r \le 1$. Indeed, according to formula (10.142a), we have for $m = 0,1,\ldots,M-1$:

$$|v_m^{p+1}| = |(1-r)u_m^p + r u_{m+1}^p| \le (1 - r + r) \max_m |u_m^p| = \|u^p\|_{U_h'},$$

[7] The first condition of (10.143) can, in fact, be further relaxed, see [GR87, § 42].

and according to formula (10.142b), we have for $m = M$:

$$|v_M^{p+1}| = |u_M^p| \leq \max_m |u_m^p| = \|u^p\|_{U_h'}.$$

Consequently,

$$\|\boldsymbol{R}_h u^p\|_{U_h'} = \|v^{p+1}\|_{U_h'} = \max_m |v_m^{p+1}| \leq \max_m |u_m^p| = \|u^p\|_{U_h'},$$

which means that $\|\boldsymbol{R}_h\| \leq 1$. Therefore, $\|\boldsymbol{R}_h^p\| \leq \|\boldsymbol{R}_h\|^p \leq 1$, and according to Theorem 10.6, scheme (10.139) is stable.

REMARK 10.2 We have already seen previously that the notion of stability for a finite-difference scheme can be reformulated as boundedness of powers for a family of matrices. Namely, in Section 10.3.6 we discussed stability of finite-difference Cauchy problems for systems of equations with constant coefficients (as opposed to scalar equations). We saw that finite-difference stability (Definition 10.2) was equivalent to stability of the corresponding family of amplification matrices. The latter, in turn, is defined as boundedness of their powers, and the Kreiss matrix theorem (Theorem 10.4) provides necessary and sufficient conditions for this property to hold.

Transition operators $\boldsymbol{R}_h$ can also be interpreted as matrices that operate on vectors from the space U_h'. In this perspective, inequality (10.145) implies uniform boundedness of all powers or stability of this family of operators (matrices). There is, however, a fundamental difference between the considerations of this section and those of Section 10.3.6. The amplification matrices that appear in the context of the Kreiss matrix theorem (Theorem 10.4) are parameterized by the frequency α and possibly the grid size h. Yet the dimension of all these matrices remains fixed and equal to the dimension of the original system, regardless of the grid size. In contradistinction to that, the dimension of the matrices $\boldsymbol{R}_h$ is inversely proportional to the grid size h, i.e., it grows with no bound as $h \longrightarrow 0$. Therefore, estimate (10.145) actually goes beyond the notion of stability for families of matrices of a fixed dimension (Section 10.3.6), as it implies stability (uniform bound on powers) for a family of matrices of increasing dimension. □

As condition (10.145) is equivalent to stability according to Theorem 10.6, then to investigate stability we need to see whether inequalities (10.145) hold. Let λ_h be an eigenvalue of the operator $\boldsymbol{R}_h$, and let $v^{(h)}$ be the corresponding eigenvector so that $\boldsymbol{R}_h v^{(h)} = \lambda_h v^{(h)}$. Then,

$$\|\boldsymbol{R}_h^p\| \|v^{(h)}\| \geq \|\boldsymbol{R}_h^p v^{(h)}\| = |\lambda_h|^p \|v^{(h)}\|$$

and consequently $\|\boldsymbol{R}_h^p\| \geq |\lambda_h|^p$. Since λ_h is an arbitrary eigenvalue, we have:

$$\|\boldsymbol{R}_h^p\| \geq [\max |\lambda_h|]^p, \quad p = 0, 1, \ldots [T/\tau],$$

where $[\max|\lambda_h|]$ is the largest eigenvalue of $\boldsymbol{R}_h$ by modulus. Hence, for the estimate (10.145) to hold, it is necessary that all eigenvalues λ of the transition operator $\boldsymbol{R}_h$ belong to the following disk on the complex plane:

$$|\lambda| \le 1 + c_1\tau, \tag{10.146}$$

where the constant c_1 does not depend on the grid size h (or τ). It means that inequality (10.146) must hold *with one and the same constant* c_1 for any given transition operator from the family $\{\boldsymbol{R}_h\}$ parameterized by h.

Inequality (10.146) is known as *the spectral necessary condition* for the uniform boundedness of the powers $\|\boldsymbol{R}_h^p\|$. It is called spectral because as long as the operators $\boldsymbol{R}_h$ can be identified with matrices of finite dimension, the eigenvalues of those matrices yield the spectra of the operators. This spectral condition is also closely related to the von Neumann spectral stability criterion for finite-difference Cauchy problems on infinite grids that we have studied in Section 10.3, see formula (10.81).

Indeed, instead of the finite-difference initial boundary value problem (10.139), consider a Cauchy problem on the grid that is infinite in space:

$$\begin{gathered}\frac{u_m^{p+1}-u_m^p}{\tau} - \frac{u_{m+1}^p - u_m^p}{h} = \varphi_m^p,\\ u_m^0 = \psi_m,\\ m = 0, \pm1, \pm2, \ldots, \quad p = 0, 1, 2, \ldots, [T/\tau] - 1.\end{gathered} \tag{10.147}$$

The von Neumann analysis of Section 10.3.2 has shown that for stability it is necessary that $r = \tau/h \le 1$. To apply the spectral criterion (10.146), we first reduce scheme (10.147) to the canonical form (10.141). The operator $\boldsymbol{R}_h : U_h' \longmapsto U_h'$, $\boldsymbol{R}_h u^p = v^{p+1}$, and the source term ρ^p are then given by [cf. formulae (10.142)]:

$$\begin{gathered}v_m^{p+1} = (1-r)u_m^p + ru_{m+1}^p, \quad \rho_m^p = \varphi_m^p,\\ m = 0, \pm1, \pm2, \ldots.\end{gathered}$$

The space U_h' contains infinite sequences $u = \{\ldots, u_{-m}, \ldots, u_{-1}, u_0, u_1, \ldots, u_m, \ldots\}$. We can supplement this space with the C norm: $\|u\| = \sup\limits_m |u_m|$. The grid functions $u = \{u_m\} = \{e^{i\alpha m}\}$ then belong to the space U_h' for all $\alpha \in [0, 2\pi)$ and provide eigenfunctions of the transition operator:

$$\boldsymbol{R}_h u = (1-r)e^{i\alpha m} + re^{i\alpha(m+1)} = [(1-r) + re^{i\alpha}]e^{i\alpha m} = \lambda(\alpha)u,$$

where the eigenvalues are given by:

$$\lambda(\alpha) = (1-r) + re^{i\alpha}. \tag{10.148}$$

According to the spectral condition of stability (10.146), all eigenvalues must satisfy the inequality: $|\lambda(\alpha)| \le 1 + c_1\tau$, which is the same as the von Neumann condition (10.78). As the eigenvalues (10.148) do not explicitly depend on the grid size, the spectral condition (10.146) reduces here to $|\lambda(\alpha)| \le 1$, cf. formula (10.79).

Let us also recall that as shown in Section 10.3.5, the von Neumann condition is not only necessary, but also sufficient for the l_2 stability of the two-layer (one-step) scalar finite-difference Cauchy problems, see formula (10.95). If, however, the space U_h' is equipped with the l_2 norm: $\|u\| = \left[h\sum_{m=-\infty}^{\infty}|u_m|^2\right]^{1/2}$ (as opposed to the C norm), then the functions $\{e^{i\alpha m}\}$ no longer belong to this space, and therefore may no longer be the eigenfunctions of $\boldsymbol{R}_h$. Nonetheless, we can show that the points $\lambda(\alpha)$ of (10.148) still belong to the spectrum of the operator $\boldsymbol{R}_h$, provided that the latter is defined as traditionally done in functional analysis, see Definition 10.7 on page 393.[8] Consequently, if we interpret λ in formula (10.146) as all points of the spectrum rather than just the eigenvalues of the operator $\boldsymbol{R}_h$, then the spectral condition (10.146) also becomes sufficient for the l_2 stability of the Cauchy problems (10.95) on an infinite grid $m = 0, \pm1, \pm2, \ldots$.

Returning now to the difference equations on finite intervals and grids (as opposed to Cauchy problems), we first notice that one can most easily verify estimates (10.145) when the matrices of all operators $\boldsymbol{R}_h$ happen to be normal: $\boldsymbol{R}_h\boldsymbol{R}_h^* = \boldsymbol{R}_h^*\boldsymbol{R}_h$. Indeed, in this case there is an orthonormal basis in the space U_h' composed of the eigenvectors of the matrix $\boldsymbol{R}_h$, see, e.g., [HJ85, Chapter 2]. Using expansion with respect to this basis, one can show that the spectral condition (10.146) is necessary and sufficient for the l_2 stability of an evolution scheme with normal operators $\boldsymbol{R}_h$ on a finite interval. More precisely, the following theorem holds.

THEOREM 10.7

Let the operators $\boldsymbol{R}_h$ in the canonical form (10.141) be normal, and let them all be uniformly bounded with respect to the grid: $\|R_h\| \le c_2$, where c_2 does not depend on h. Let also all norms be chosen in the sense of l_2. Then, for the estimates (10.145) to hold, it is necessary and sufficient that the inequalities be satisfied:

$$\max_n |\lambda_n| \le 1 + c_1\tau, \quad c_1 = \text{const}, \tag{10.149}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_N$ are eigenvalues of the matrix $\boldsymbol{R}_h$ and the constant c_1 in formula (10.149) does not depend on h.

One implication of Theorem 10.7, the necessity, coincides with the previous necessary spectral condition for stability that we have justified on page 389. The other implication, the sufficiency, is to be proven in Exercise 5 of this section. A full proof of Theorem 10.7 can be found, e.g., in [GR87, §43].

Unfortunately, in many practical situations the operators (matrices) $\boldsymbol{R}_h$ in the canonical form (10.141) are not normal. Then, the spectral condition (10.146) still remains necessary for stability. Moreover, we have just seen that in the special case of two-layer scalar constant-coefficient Cauchy problems it is also sufficient for stability and that sufficiency takes place regardless of whether or not $\boldsymbol{R}_h$ has a full sys-

[8] In general, the points $\lambda = \lambda(\alpha, h)$ given by Definition 10.3 on page 351 will be a part of the spectrum in the sense of its classical definition, see Definition 10.7 on page 393.

tem of orthonormal eigenfunctions. However, for general finite-difference problems on finite intervals the spectral condition (10.146) becomes pretty far detached from sufficiency and provides no adequate criterion for uniform boundedness of $\|\boldsymbol{R}_h^p\|$.

For instance, the matrix of the transition operator $\boldsymbol{R}_h$ defined by formulae (10.142a) and (10.142b) is given by:

$$\boldsymbol{R}_h = \begin{bmatrix} 1-r & r & 0 \cdots & 0 & 0 \\ 0 & 1-r & r \cdots & 0 & 0 \\ \vdots & \vdots & \vdots \ddots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & 1-r & r \\ 0 & 0 & 0 \cdots & 0 & 1 \end{bmatrix}. \tag{10.150}$$

Its spectrum consists of the eigenvalues $\lambda = 1$ and $\lambda = 1-r$ and as such, does not depend on h (or on τ). Consequently, for any $h > 0$ the spectrum of the operator $\boldsymbol{R}_h$ consists of only these two numbers: $\lambda = 1$ and $\lambda = 1-r$. This spectrum belongs to the unit disk $|\lambda| \le 1$ when $0 \le r \le 2$. However, for $1 < r \le 2$, scheme (10.139) violates the Courant, Friedrichs, and Lewy condition necessary for stability, and hence, there may be no stability $\|\boldsymbol{R}_h^p\| \le K$ for any reasonable choice of norms.

Thus, we have seen that the spectral condition (10.146) that employs the eigenvalues of the operators $\boldsymbol{R}_h$ and that is necessary for the uniform boundedness $\|\boldsymbol{R}_h^p\| \le K$ appears too rough in the case of non-normal matrices. For example, it fails to detect the instability of scheme (10.139) for $1 < r \le 2$.

To refine the spectral condition we will introduce a new concept. Assume, as before, that the operator $\boldsymbol{R}_h$ is defined on a normed linear space U_h'. We will denote by $\{\boldsymbol{R}_h\}$ the entire family of operators $\boldsymbol{R}_h$ for all legitimate values of the parameter h that characterizes the grid.[9]

DEFINITION 10.6 *A complex number λ is said to belong to* the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ *if for any $h_0 > 0$ and $\varepsilon > 0$ one can always find such a value of h, $h < h_0$, that the inequality*

$$\|\boldsymbol{R}_h u - \lambda u\|_{U_h'} < \varepsilon \|u\|_{U_h'}$$

will have a solution $u \in U_h'$. The set of all such λ will be called the spectrum of the family of operators $\{\boldsymbol{R}_h\}$.

The following theorem employs the concept of the spectrum of a family of operators from Definition 10.6 and provides *a key necessary condition for stability.*

THEOREM 10.8 (Godunov-Ryaben'kii)
If even one point λ_0 of the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ lies outside the unit disk on the complex plane, i.e., $|\lambda_0| > 1$, then there is no

[9] By the very nature of finite-difference schemes, h may assume arbitrarily small positive values.

common constant K such that the inequality

$$\|\boldsymbol{R}_h^p\| \le K$$

will hold for all $h > 0$ and all integer values of p from 0 till some $p = p_0(h)$, where $p_0(h) \longrightarrow \infty$ as $h \longrightarrow 0$.

PROOF Let us first assume that no such numbers $h_0 > 0$ and $c > 0$ exist that for all $h < h_0$ the following estimate holds:

$$\|\boldsymbol{R}_h\| \le c. \tag{10.151}$$

This assumption means that there is no uniform bound on the operators $\boldsymbol{R}_h$ themselves. As such, there may be no bound on the powers $\boldsymbol{R}_h^p$ either. Consequently, we only need to consider the case when there are such $h_0 > 0$ and $c > 0$ that for all $h < h_0$ inequality (10.151) is satisfied.

Let $|\lambda_0| = 1 + \delta$, where λ_0 is the point of the spectrum for which $|\lambda_0| > 1$. Take an arbitrary $K > 0$ and choose p and ε so that:

$$(1+\delta)^p > 2K,$$
$$1 - (1 + c + c^2 + \ldots + c^{p-1})\varepsilon > \frac{1}{2}.$$

According to Definition 10.6, one can find arbitrarily small h, for which there is a vector $u \in U_h'$ that solves the inequality:

$$\|\boldsymbol{R}_h u - \lambda_0 u\|_{U_h'} < \varepsilon \|u\|_{U_h'}.$$

Let u be the solution, and denote:

$$\boldsymbol{R}_h u = \lambda_0 u + z.$$

It is clear that $\|z\| < \varepsilon\|u\|$. Moreover, it is easy to see that

$$\boldsymbol{R}_h^p u = \lambda_0^p u + (\lambda_0^{p-1} z + \lambda_0^{p-2} \boldsymbol{R}_h z + \ldots + \boldsymbol{R}_h^{p-1} z).$$

As $|\lambda_0| > 1$, we have:

$$\|\lambda_0^{p-1} z + \lambda_0^{p-2} \boldsymbol{R}_h z + \ldots + \boldsymbol{R}_h^{p-1} z\| < |\lambda_0|^p (1 + \|\boldsymbol{R}_h\| + \|\boldsymbol{R}_h^2\| + \ldots + \|\boldsymbol{R}_h^{p-1}\|)\varepsilon\|u\|,$$

and consequently,

$$\begin{aligned}\|\boldsymbol{R}_h^p u\| &> |\lambda_0|^p [1 - \varepsilon(1 + c + c^2 + \ldots + c^{p-1})]\|u\| \\ &> (1+\delta)^p \frac{1}{2}\|u\| > 2K\frac{1}{2}\|u\| = K\|u\|.\end{aligned}$$

In doing so, the value of h can always be taken sufficiently small so that to ensure $p < p_0(h)$.

Since the value of K has been chosen arbitrarily, we have essentially proven that *for the estimate* $\|\boldsymbol{R}_h^p\| < K$ *to hold, it is necessary that all points of the spectrum of the family* $\{\boldsymbol{R}_h\}$ *belong to the unit disk* $|\lambda| \leq 1$ *on the complex plane*. ∎

Next recall that the following definition of the spectrum of an operator $\boldsymbol{R}_h$ (for a fixed h) is given in functional analysis.

DEFINITION 10.7 *A complex number* λ *is said to belong to the spectrum of the operator* $\boldsymbol{R}_h : U_h' \longmapsto U_h'$ *if for any* $\varepsilon > 0$ *the inequality*

$$\|\boldsymbol{R}_h u - \lambda u\|_{U_h'} < \varepsilon \|u\|_{U_h'}$$

has a solution $u \in U_h'$. *The set of all such* λ *is called the spectrum* $\boldsymbol{R}_h$.

At first glance, comparison of the Definitions 10.6 and 10.7 may lead one to thinking that the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ consists of all those and only those points on the complex plane that are obtained by passing to the limit $h \longrightarrow 0$ from the points of the spectrum of $\boldsymbol{R}_h$, when h approaches zero along all possible sub-sequences. However, this assumption is, generally speaking, not correct.

Consider, for example, the operator $\boldsymbol{R}_h : U_h' \longmapsto U_h'$ defined by formulae (10.142a) and (10.142b). It is described by the matrix (10.150) and operates in the $M+1$-dimensional linear space U_h', where $M = 1/h$. The spectrum of a matrix consists of its eigenvalues, and the eigenvalues of the matrix (10.150) are $\lambda = 1$ and $\lambda = 1 - r$. These eigenvalues do not depend on h (or on τ) and consequently, the spectrum of the operator $\boldsymbol{R}_h$ consists of only two points, $\lambda = 1$ and $\lambda = 1 - r$, for any $h > 0$. As, however, we are going to see (pages 397-402), the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ contains not only these two points, but also all points of the disk $|\lambda - 1 + r| \leq r$ of radius r centered at the point $(1-r, 0)$ on the complex plane, see Figure 10.12 on page 384. When $r \leq 1$, the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ belongs to the unit disk $|\lambda| \leq 1$, see Figures 10.12(a) and 10.12(b). However, when $r > 1$, this necessary spectral condition of stability does not hold, see Figure 10.12(c), and the inequality $\|\boldsymbol{R}_h^p\| \leq K$ can not be satisfied uniformly with respect to h.

Before we accurately compute the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ given by formula (10.150), let us qualitatively analyze the behavior of the powers $\|\boldsymbol{R}_h^p\|$ for $r > 1$ and also show that the necessary stability criterion given by Theorem 10.8 is, in fact, rather close to sufficient.

We first notice that for any $h > 0$ there is only one eigenvalue of the matrix $\boldsymbol{R}_h$ that has unit modulus: $\lambda = 1$, and that the similarity transformation $\boldsymbol{S}_h^{-1} \boldsymbol{R}_h \boldsymbol{S}_h$, where

$$\boldsymbol{S}_h = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{S}_h^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & -1 \\ 0 & 1 & 0 & \cdots & 0 & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

reduces this matrix to the block-diagonal form:

$$\boldsymbol{S}_h^{-1}\boldsymbol{R}_h\boldsymbol{S}_h = \begin{bmatrix} 1-r & r & 0 \cdots & 0 & 0 \\ 0 & 1-r & r \cdots & 0 & 0 \\ \vdots & \vdots & \vdots \ddots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & 1-r & 0 \\ 0 & 0 & 0 \cdots & 0 & 1 \end{bmatrix} \equiv \boldsymbol{B}_h.$$

When $1 < r < 2$, we have $|1-r| < 1$ for the magnitude of the diagonal entry $1-r$. Then, it is possible to prove that $\boldsymbol{B}_h^p \longrightarrow \mathrm{diag}\{0,0,\ldots,0,1\}$ as $p \longrightarrow \infty$ (see Theorem 6.2 on page 178). In other words, the limiting matrix of $\boldsymbol{B}_h^p$ for the powers p approaching infinity has only one non-zero entry equal to one at the lower right corner. Consequently,

$$\lim_{p\to\infty} \boldsymbol{R}_h^p = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

and as such, $\lim\limits_{p\to\infty} \|\boldsymbol{R}_h^p\| = 1$. We can therefore see that regardless of the value of h, the norms of the powers of the transition operator $\boldsymbol{R}_h$ approach one and the same finite limit. In other words, we can write $\lim\limits_{p\tau\to\infty} \|\boldsymbol{R}_h^p\| = 1$, and this "benign" asymptotic behavior of $\|\boldsymbol{R}_h^p\|$ for large $p\tau$ is indeed determined by the eigenvalues $\lambda = 1-r$ and $\lambda = 1$ that belong to the unit disk.

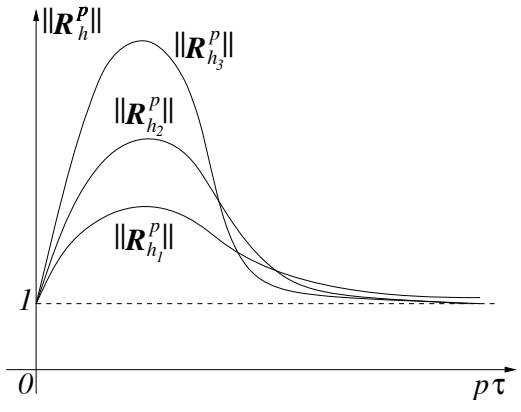


FIGURE 10.13: Schematic behavior of the powers $\|\boldsymbol{R}_h^p\|$ for $1 < r < 2$ and $h_3 < h_2 < h_1$.

The fact that the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ does not belong to the unit disk for $r > 1$ manifests itself in the behavior of $\|\boldsymbol{R}_h^p\|$ for $h \longrightarrow 0$ and for moderate (not so large) values of $p\tau$. The maximum value of $\|\boldsymbol{R}_h^p\|$ on the interval $0 < p\tau < T$, where T is an arbitrary positive constant, will rapidly grow as h decreases, see Figure 10.13. This is precisely what leads to the instability, whereas the behavior of $\|\boldsymbol{R}_h^p\|$ as $p\tau \longrightarrow \infty$, which is related to the spectrum of each individual operator $\boldsymbol{R}_h$, is not important from the standpoint of stability.

Let us also emphasize that even though from a technical point of view Theorem 10.8 only provides a necessary condition for stability, this condition, *is, in fact, not so distant from sufficient*. More precisely, the following theorem holds.

THEOREM 10.9

Let the operators $\boldsymbol{R}_h$ be defined on a linear normed space U_h' for each $h>0$, and assume that they are uniformly bounded with respect to h:

$$\|\boldsymbol{R}_h\| \le c. \tag{10.152}$$

Let also the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ belong to the unit disk on the complex plane: $|\lambda| \le 1$.

Then for any $\eta > 0$, the norms of the powers of operators $\boldsymbol{R}_h$ satisfy the estimate:

$$\|R_h^p\| \le A(\eta)(1+\eta)^p, \tag{10.153}$$

where $A = A(\eta)$ may depend on η, but does not depend on the grid size h.

Theorem 10.9 means that having the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ lie inside the unit disk is *not only necessary for stability, but it also guarantees us from a catastrophic instability*. Indeed, if the conditions of Theorem 10.9 hold, then the quantity $\max\limits_{1\le p\le [T/\tau]} \|\boldsymbol{R}_h^p\|$ either remains bounded as $h \longrightarrow 0$ or increases, but *slower than any exponential function,* i.e., slower than any $(1+\eta)^{[T/\tau]}$, where $\eta > 0$ may be arbitrarily small.

PROOF Let us first show that if the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ belongs to the disk $|\lambda| \le \rho$, then for any given λ that satisfies the inequality $|\lambda| \ge \rho + \eta$, $\eta > 0$, there are the numbers $A = A(\eta)$ and $h_0 > 0$ such that $\forall h < h_0$ and $\forall u \in U_h'$, $u \ne 0$, the following estimate holds:

$$\|\boldsymbol{R}_h u - \lambda u\|_{U_h'} > \frac{\rho+\eta}{A(\eta)} \|u\|_{U_h'}. \tag{10.154}$$

Assume the opposite. Then there exist: $\eta > 0$; a sequence of real numbers $h_k > 0$, $h_k \longrightarrow 0$ as $k \longrightarrow \infty$; a sequence of complex numbers λ_k, $|\lambda_k| > \rho + \eta$; and a sequence of vectors $u_{h_k} \in U_{h_k}'$ such that:

$$\|\boldsymbol{R}_{h_k} u_{h_k} - \lambda_k u_{h_k}\|_{U_{h_k}'} < \frac{\rho+\eta}{k} \|u_{h_k}\|_{U_{h_k}'}. \tag{10.155}$$

For sufficiently large values of k, for which $\frac{\rho+\eta}{k} < 1$, the numbers λ_k will not lie outside the disk $|\lambda| \le c+1$ by virtue of estimate (10.152), because outside this disk we have:

$$\|\boldsymbol{R}_{h_k} u_{h_k} - \lambda u_{h_k}\|_{U_{h_k}'} \ge \left(|\lambda| - \|\boldsymbol{R}_{h_k}\|\right) \|u_{h_k}\|_{U_{h_k}'} \ge \|u_{h_k}\|_{U_{h_k}'}.$$

Therefore, the sequence of complex numbers λ_k is bounded and as such, has a limit point $\tilde{\lambda}$, $|\tilde{\lambda}| \geq \rho + \eta$. Using the triangle inequality, we can write: $\|\boldsymbol{R}_{h_k} u_{h_k} - \lambda_k u_{h_k}\|_{U'_{h_k}} \geq \|\boldsymbol{R}_{h_k} u_{h_k} - \tilde{\lambda} u_{h_k}\|_{U'_{h_k}} - |\lambda_k - \tilde{\lambda}| \|u_{h_k}\|_{U'_{h_k}}$. Substituting into inequality (10.155), we obtain:

$$\|\boldsymbol{R}_{h_k} u_{h_k} - \tilde{\lambda} u_{h_k}\|_{U'_{h_k}} < \Big[\underbrace{\frac{\rho+\eta}{k} + |\lambda_k - \tilde{\lambda}|}_{\varepsilon}\Big] \|u\|_{U'_{h_k}}.$$

Therefore, according to Definition 10.6 the point $\tilde{\lambda}$ belongs to the spectrum of the family of operators $\{\boldsymbol{R}_h\}$. This contradicts the previous assumption that the spectrum belongs to the disk $|\lambda| \leq \rho$.

Now let $\boldsymbol{R}$ be a linear operator on a finite-dimensional normed space U, $\boldsymbol{R} : U \longmapsto U$. Assume that for any complex λ, $|\lambda| \geq \gamma > 0$, and any $u \in U$ the following inequality holds for some $a = \text{const} > 0$:

$$\|\boldsymbol{R}u - \lambda u\| \geq a\|u\|. \tag{10.156}$$

Then,

$$\|\boldsymbol{R}^p\| \leq \frac{\gamma^{p+1}}{a}, \quad p = 1, 2, \ldots. \tag{10.157}$$

Inequality (10.157) follows from the relation:

$$\boldsymbol{R}^p = -\frac{1}{2\pi i} \oint\limits_{|\lambda|=\gamma} \lambda^p (\boldsymbol{R} - \lambda \boldsymbol{I})^{-1} d\lambda \tag{10.158}$$

combined with estimate (10.156), because the latter implies that $\|(\boldsymbol{R} - \lambda \boldsymbol{I})^{-1}\| \leq \frac{1}{a}$. To prove estimate (10.153), we set $\boldsymbol{R} = \boldsymbol{R}_h$, $\rho = 1$ so that $|\lambda| \geq 1 + \eta = \gamma$, and use (10.154) instead of (10.156). Then estimate (10.157) coincides with (10.153).

It only remains to justify equality (10.158). For that purpose, we will use an argument similar to that used when proving Theorem 6.2. Define

$$u^{p+1} = \boldsymbol{R}u^p \quad \text{and} \quad w(\lambda) = \sum_{p=0}^{\infty} \frac{u^p}{\lambda^p},$$

where the series converges uniformly at least for all $\lambda \in \mathbb{C}$, $|\lambda| > c$, see formula (10.152). Multiply the equality $u^{p+1} = \boldsymbol{R}u^p$ by λ^{-p} and take the sum with respect to p from $p = 0$ to $p = \infty$. This yields:

$$\lambda w(\lambda) - \lambda u^0 = \boldsymbol{R} w(\lambda),$$

or alternatively,

$$(\boldsymbol{R} - \lambda \boldsymbol{I}) w(\lambda) = -\lambda u^0, \qquad w(\lambda) = -\lambda (\boldsymbol{R} - \lambda \boldsymbol{I})^{-1} u^0.$$

From the definition of $w(\lambda)$ it is easy to see that $-u^p$ is the residue of the vector-function $\lambda^{p-1}w(\lambda)$ at infinity:

$$u^p = \frac{1}{2\pi i}\oint\limits_{|\lambda|=\gamma} \lambda^{p-1}w(\lambda)d\lambda = -\frac{1}{2\pi i}\oint\limits_{|\lambda|=\gamma} \lambda^p(\boldsymbol{R}-\lambda\boldsymbol{I})^{-1}u^0 d\lambda.$$

As $u^p = \boldsymbol{R}^p u^0$, the last equality is equivalent to (10.158). ∎

Altogether, we have seen that the question of stability for evolution finite-difference schemes on finite intervals reduces to studying the spectra of the families of the corresponding transition operators $\{\boldsymbol{R}_h\}$. More precisely, we need to find out whether the spectrum for a given family of operators $\{\boldsymbol{R}_h\}$ belongs to the unit disk $|\lambda| \le 1$. *If it does, then the scheme is either stable or, in the worst case scenario, it may only develop a mild instability.*

Let us now show how we can actually calculate the spectrum of a family of operators. To demonstrate the approach, we will exploit the previously introduced example (10.142a), (10.142b). *It turns out that the algorithm for computing the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ coincides with the Babenko-Gelfand procedure described in Section 10.5.1.* Namely, we need to introduce three auxiliary operators: $\overleftrightarrow{\boldsymbol{R}}$, $\overrightarrow{\boldsymbol{R}}$, and $\overleftarrow{\boldsymbol{R}}$. The operator $\overleftrightarrow{\boldsymbol{R}}$, $v = \overleftrightarrow{\boldsymbol{R}}u$, is defined on the linear space of bounded grid functions $u = \{\ldots, u_{-1}, u_0, u_1, \ldots\}$ according to the formula:

$$v_m = (1-r)u_m + ru_{m+1}, \quad m = 0, \pm 1, \pm 2, \ldots, \tag{10.159}$$

which is obtained from (10.142a), (10.142b) by removing both boundaries. The operator $\overrightarrow{\boldsymbol{R}}$ is defined on the linear space of functions $u = \{u_0, u_1, \ldots, u_m, \ldots\}$ that vanish at infinity: $|u_m| \longrightarrow 0$ as $m \longrightarrow +\infty$. It is given by the formula:

$$v_m = (1-r)u_m + ru_{m+1}, \quad m = 0, 1, 2, \ldots, \tag{10.160}$$

which is obtained from (10.142a), (10.142b) by removing the right boundary. Finally, the operator $\overleftarrow{\boldsymbol{R}}$ is defined on the linear space of functions $\{\ldots, u_m, \ldots, u_0, \ldots, u_{M-1}, u_M\}$ that satisfy: $|u_m| \longrightarrow 0$ as $m \longrightarrow -\infty$. It is given by the formula:

$$\begin{aligned} &v_m = (1-r)u_m + ru_{m+1}, \quad m = \ldots, -1, 0, 1, \ldots, M-1, \\ &v_M = u_M, \end{aligned} \tag{10.161}$$

which is obtained from (10.142a), (10.142b) by removing the left boundary. Note that the spaces of functions for the operators $\overrightarrow{\boldsymbol{R}}$ and $\overleftarrow{\boldsymbol{R}}$ are defined on semi-infinite grids $m = 0, 1, 2, \ldots$ and $m = \ldots, -1, 0, 1, \ldots, M$, respectively.

None of the operators $\overleftrightarrow{\boldsymbol{R}}$, $\overrightarrow{\boldsymbol{R}}$, or $\overleftarrow{\boldsymbol{R}}$ depend on h. We will show that *the combination of all eigenvalues of these three auxiliary operators yields the spectrum of the family of operators $\{\boldsymbol{R}_h\}$*. In Section 10.5.1, we have, in fact, already computed the eigenvalues of the operators $\overleftrightarrow{\boldsymbol{R}}$ and $\overrightarrow{\boldsymbol{R}}$. For the operator $\overleftrightarrow{\boldsymbol{R}}$, the eigenvalues are all

those and only those complex numbers λ, for which the equation $\overleftrightarrow{\boldsymbol{R}}u = \lambda u$ has a bounded solution $u = \{u_m\}, m = 0, \pm 1, \pm 2, \ldots$. According to (10.159), this equation can be written as:

$$(1 - r - \lambda)u_m + ru_{m+1} = 0, \quad m = 0, \pm 1, \pm 2, \ldots,$$

and its general solution is $u_m = cq^m$, where q is a root of the characteristic equation: $(1 - r - \lambda) + rq = 0$. This solution is bounded as $|m| \longrightarrow \infty$ if and only if $|q| = 1$, i.e., $q = e^{i\alpha}$, $\alpha \in [0, 2\pi)$. The corresponding eigenvalues are given by:

$$\lambda = 1 - r + rq = 1 - r + re^{i\alpha}, \quad \alpha \in [0, 2\pi).$$

The curve $\lambda = \lambda(\alpha)$ is a circle of radius r on the complex plane centered at the point $(1 - r, 0)$, see Figure 10.11(a). We will denote this circle by $\overleftrightarrow{\Lambda}$.

The eigenvalues of the operator $\overrightarrow{\boldsymbol{R}}$ are all those and only those complex numbers λ, for which the equation $\overrightarrow{\boldsymbol{R}}u = \lambda u$ has a solution $u = \{u_0, u_1, \ldots, u_m, \ldots\}$ that satisfies $\lim\limits_{m \to +\infty} |u_m| = 0$. Recasting this equation with the help of formula (10.160), we have:

$$(1 - r - \lambda)u_m + ru_{m+1} = 0, \quad m = 0, 1, 2, \ldots.$$

The solution $u_m = cq^m$ may only be bounded as $m \longrightarrow +\infty$ if $|q| < 1$. The corresponding eigenvalues $\lambda = 1 - r + rq$ completely fill the interior of the disk of radius r centered at the point $(1 - r, 0)$, see Figure 10.11(b). We will denote this set by $\overrightarrow{\Lambda}$.

The eigenvalues of the operator $\overleftarrow{\boldsymbol{R}}$ are computed similarly. Using formula (10.161), we can write equation $\overleftarrow{\boldsymbol{R}}u = \lambda u$ as follows:

$$\begin{aligned} (1 - r - \lambda)u_m + ru_{m+1} &= 0, \quad m = \ldots, -1, 0, 1, \ldots, M - 1, \\ (1 - \lambda)u_M &= 0. \end{aligned}$$

The general solution of the first equation from this pair is $u_m = cq^m$, and the relation between λ and q is $\lambda = 1 - r + rq$. The solution $u_m = cq^m$ may only vanish as $m \longrightarrow -\infty$ if $|q| > 1$. The second equation provides an additional constraint $(1 - \lambda)q^M = 0$ so that $\lambda = 1$. However, for this particular λ we also have $q = 1$, which implies no decay as $m \longrightarrow -\infty$. We therefore conclude that the equation $\overleftarrow{\boldsymbol{R}}u = \lambda u$ has no solutions $u = \{u_m\}$ that satisfy $\lim\limits_{m \to -\infty} |u_m| = 0$, i.e., there are no eigenvalues: $\overleftarrow{\Lambda} = \emptyset$.

The combination of all eigenvalues $\Lambda = \overleftrightarrow{\Lambda} \cup \overrightarrow{\Lambda} \cup \overleftarrow{\Lambda}$ is the disk $|\lambda - (1 - r)| \le r$ on the complex plane; it is centered at $(1 - r, 0)$ and has radius r. We will now show that the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ coincides with the set Λ. This is equivalent to showing that every point $\lambda_0 \in \Lambda$ belongs to the spectrum of $\{\boldsymbol{R}_h\}$ and that this spectrum contains no other points.

According to Definition 10.6, to prove the first implication it is sufficient to demonstrate that for any $\varepsilon > 0$ the inequality

$$\|\boldsymbol{R}_h u - \lambda_0 u\|_{U'_h} < \varepsilon \|u\|_{U'_h} \tag{10.162}$$

has a solution $u \in U_h'$ for all sufficiently small $h > 0$. As $\lambda_0 \in \Lambda$, then $\lambda_0 \in \overleftrightarrow{\Lambda}$ or $\lambda_0 \in \overrightarrow{\Lambda}$, because $\overleftarrow{\Lambda} = \emptyset$. Note that when ε is small one may call the solution u of inequality (10.162) "almost an eigenvector" of the operator $\boldsymbol{R}_h$, since a solution to the equation $\boldsymbol{R}_h u - \lambda_0 u = 0$ is its genuine eigenvector.

Let us first assume that $\lambda_0 \in \overleftrightarrow{\Lambda}$. To construct a solution u of inequality (10.162), we recall that by definition of the set $\overleftrightarrow{\Lambda}$ there exists q_0: $|q_0| = 1$, such that $\lambda_0 = 1 - r + rq_0$ and the equation $(1 - r - \lambda_0)v_m + rv_{m+1} = 0$, $m = 0, \pm 1, \pm 2, \ldots$, has a bounded solution $v_m = q_0^m$, $m = 0, \pm 1, \pm 2, \ldots$. We will consider this solution only for $m = 0, 1, 2, \ldots, M$, while keeping the same notation v. It turns out that the vector:

$$v = [v_0, v_1, v_2, \ldots, v_M] = \left[1, q_0, q_0^2, \ldots, q_0^M\right]$$

almost satisfies the operator equation $\boldsymbol{R}_h v - \lambda_0 v = 0$ that we write as:

$$\begin{aligned} (1 - r - \lambda_0)v_m + rv_{m+1} &= 0, \quad m = 0, 1, 2, \ldots, M-1, \\ (1 - \lambda_0)v_M &= 0. \end{aligned}$$

The vector v would have completely satisfied the previous equation, which is an even stronger constraint than inequality (10.162), if it did not violate the last relation $(1 - \lambda_0)v_M = 0$.[10] This relation can be interpreted as a boundary condition for the difference equation:

$$(1 - r - \lambda_0)u_m + ru_{m+1} = 0, \quad m = 0, 1, 2, \ldots, M-1.$$

The boundary condition is specified at $m = M$, i.e., at the right endpoint of the interval $0 \le x \le 1$. To satisfy this boundary condition, let us "correct" the vector $v = \left[1, q_0, q_0^2, \ldots, q_0^M\right]$ by multiplying each of its components v_m by the respective factor $(M - m)h$. The resulting vector will be denoted $u = [u_0, u_1, \ldots, u_M]$, $u_m = (M - m)hq_0^m$. Obviously, the vector u has unit norm:

$$\|u\|_{U_h'} = \max_m |u_m| = \max_m |(M - m)hq_0^m| = Mh = 1.$$

We will now show that this vector u furnishes a desired solution to the inequality (10.162). Define the vector $w \stackrel{\text{def}}{=} \boldsymbol{R}_h u - \lambda_0 u$, $w = [w_0, w_1, \ldots, w_M]$. We need to estimate its norm. For the individual components of w, we have:

$$\begin{aligned} |w_m| &= |(1 - r - \lambda_0)(M - m)hq_0^m + r(M - m - 1)hq_0^{m+1}| \\ &= |[(1 - r - \lambda_0) + rq_0](M - m)hq_0^m - rhq_0^{m+1}| \\ &= |0 \cdot (M - m)hq_0^m - rhq_0^{m+1}| = rh, \quad m = 0, 1, \ldots, M-1, \\ |w_M| &= |u_M - \lambda_0 u_M| = |0 - \lambda_0 \cdot 0| = 0. \end{aligned}$$

Consequently, $\|w\|_{U_h'} = rh$, and for $h < \varepsilon/r$ we obtain: $\|w\|_{U_h'} = \|\boldsymbol{R}_h u - \lambda_0 u\|_{U_h'} < \varepsilon = \varepsilon \|u\|_{U_h'}$, i.e., inequality (10.162) is satisfied. Thus, we have shown that if $\lambda_0 \in \overleftrightarrow{\Lambda}$, then this point also belongs to the spectrum of the family of operators $\{\boldsymbol{R}_h\}$.

[10] Relation $(1 - \lambda_0)v_M = 0$ is violated unless $\lambda_0 = q_0 = 1$.

Next, let us assume that $\lambda_0 \in \overrightarrow{\Lambda}$ and show that in this case λ_0 also belongs to the spectrum of the family of operators $\{\boldsymbol{R}_h\}$. According to (10.160), equation $\overrightarrow{\boldsymbol{R}} v - \lambda_0 v = 0$ is written as:

$$(1 - r - \lambda_0)v_m + rv_{m+1} = 0, \quad m = 0, 1, 2, \ldots.$$

Since $\lambda_0 \in \overrightarrow{\Lambda}$, this equation has a solution $v_m = q_0^m$, $m = 0, 1, 2, \ldots$, where $|q_0| < 1$. We will consider this solution only for $m = 0, 1, 2, \ldots, M$:

$$u = [u_0, u_1, u_2, \ldots, u_M] = \left[1, q_0, q_0^2, \ldots, q_0^M\right], \qquad \|u\|_{U_h'} = 1.$$

As before, define $w \stackrel{\text{def}}{=} \boldsymbol{R}_h u - \lambda_0 u$. For the components of the vector w we have:

$$|w_m| = |(1 - r - \lambda_0)q_0^m + rq_0^{m+1}| = 0, \quad m = 0, 1, \ldots, M-1,$$
$$|w_M| = |1 - \lambda_0| \cdot |q_0^M|.$$

Consequently, $\|w\|_{U_h'} = |1 - \lambda_0| \cdot |q_0|^M = |1 - \lambda_0| \cdot |q_0|^{1/h}$. Since $|q_0| < 1$, then for any $\varepsilon > 0$ we can always choose a sufficiently small h so that $|1 - \lambda_0| \cdot |q_0|^{1/h} < \varepsilon$. Then, $\|w\|_{U_h'} = \|\boldsymbol{R}_h u - \lambda_0 u\|_{U_h'} < \varepsilon = \varepsilon \|u\|_{U_h'}$ and the inequality (10.162) is satisfied.

Note that if the set $\overleftarrow{\Lambda}$ were not empty, then proving that each of its elements belongs to the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ would have been similar. Altogether, we have thus shown that in our specific example given by equations (10.142) every $\lambda_0 \in \{\overleftrightarrow{\Lambda} \cup \overleftarrow{\Lambda} \cup \overrightarrow{\Lambda}\}$ is also an element of the spectrum of $\{\boldsymbol{R}_h\}$.

Now we need to prove that if $\lambda_0 \notin \{\overleftrightarrow{\Lambda} \cup \overleftarrow{\Lambda} \cup \overrightarrow{\Lambda}\}$ then it does not belong to the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ either. To that end, it will be sufficient to show that there is an h-independent constant A, such that for any $u = [u_0, u_1, \ldots, u_M]$ the following inequality holds:

$$\|\boldsymbol{R}_h u - \lambda_0 u\|_{U_h'} \geq A \|u\|_{U_h'}. \tag{10.163}$$

Then, for $\varepsilon < A$, inequality (10.162) will have no solutions, and therefore the point λ_0 will not belong to the spectrum. Denote $f = \boldsymbol{R}_h u - \lambda_0 u$, then inequality (10.163) reduces to:

$$\|f\|_{U_h'} \geq A \|u\|_{U_h'}. \tag{10.164}$$

Our objective is to justify estimate (10.164). Rewrite the equation $\boldsymbol{R}_h u - \lambda_0 u = f$ as:

$$(1 - r - \lambda_0)u_m + ru_{m+1} = f_m, \quad m = 0, 1, \ldots, M-1,$$
$$(1 - \lambda_0)u_M = f_M,$$

and interpret these relations as an equation with respect to the unknown $u = \{u_m\}$, whereas the right-hand side $f = \{f_m\}$ is assumed given. Let

$$u_m = v_m + w_m, \quad m = 0, 1, \ldots, M, \tag{10.165}$$

where v_m are components of the bounded solution $v = \{v_m\}$, $m = 0, \pm 1, \pm 2, \ldots$, to the following equation:

$$(1 - r - \lambda_0)v_m + rv_{m+1} = \hat{f}_m \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } m < 0, \\ f_m, & \text{if } m = 0, 1, \ldots, M-1, \\ 0, & \text{if } m \geq M. \end{cases} \tag{10.166}$$

Then because of the linearity, the grid function $w = \{w_m\}$ introduced by formula (10.165) solves the equation:

$$\begin{gathered} (1 - r - \lambda_0)w_m + rw_{m+1} = 0, \qquad m = 0, 1, \ldots, M-1, \\ (1 - \lambda_0)w_M = f_M - (1 - \lambda_0)v_M. \end{gathered} \tag{10.167}$$

Let us now recast estimate (10.164) as $|u_m| \leq A^{-1} \max\limits_m |f_m|$. According to (10.165), to prove this estimate it is sufficient to establish individual inequalities:

$$|v_m| \leq A_1 \max_m |f_m|, \tag{10.168a}$$

$$|w_m| \leq A_2 \max_m |f_m|, \tag{10.168b}$$

where A_1 and A_2 are constants. We begin with inequality (10.168a). Notice that equation (10.166) is a first order constant-coefficient ordinary difference equation:

$$av_m + bv_{m+1} = \hat{f}_m, \quad m = 0, \pm 1, \pm 2, \ldots,$$

where $a = 1 - r - \lambda_0$, $b = r$. Its bounded fundamental solution is given by

$$G_m = \begin{cases} \dfrac{1}{a}\left(-\dfrac{a}{b}\right)^m, & m \leq 0, \\ 0, & m \geq 1, \end{cases}$$

because $\lambda_0 \notin \{\overleftrightarrow{\Lambda} \cup \overleftarrow{\Lambda} \cup \overrightarrow{\Lambda}\}$, i.e., $|\lambda_0 - (1 - r)| > r$, and consequently $|a/b| > 1$. Representing the solution v_m in the form of a convolution: $v_m = \sum\limits_{k=-\infty}^{\infty} G_{m-k}\hat{f}_k$ and summing up the geometric sequence we arrive at the estimate:

$$|v_m| \leq \frac{\max\limits_m |\hat{f}_m|}{|a| - |b|} = \frac{\max\limits_m |f_m|}{|a| - |b|}.$$

Introducing the distance δ_0 between the point λ_0 and the set $\{\overleftrightarrow{\Lambda} \cup \overleftarrow{\Lambda} \cup \overrightarrow{\Lambda}\}$, we can obviously claim that $|a| - |b| > \delta_0/2$, which makes the previous estimate equivalent to (10.168a). Estimate (10.168b) can be obtained by representing the solution of equation (10.167) in the form:

$$w_m = \frac{f_M - (1 - \lambda_0)v_M}{1 - \lambda_0} q_0^{m-M}, \tag{10.169}$$

where q_0 is determined by the relation $(1-r-\lambda_0)+rq_0=0$. Our assumption is that $\lambda_0 \notin \{\overleftrightarrow{\Lambda} \cup \overleftarrow{\Lambda} \cup \overrightarrow{\Lambda}\}$, i.e., that λ_0 lies outside of the disk of radius r on the complex plane centered at $(1-r,0)$. In this case $|q_0|>1$. Moreover, we can say that $|1-\lambda_0|=\delta_1>0$, because if $\lambda_0=1$, then λ_0 would have belonged to the set $\{\overleftrightarrow{\Lambda} \cup \overleftarrow{\Lambda} \cup \overrightarrow{\Lambda}\}$. As such, using formula (10.169) and taking into account estimate (10.168a) that we have already proved, we obtain the desired estimate (10.168b):

$$
\begin{aligned}
|w_m| &= \left|\frac{f_M-(1-\lambda_0)v_M}{1-\lambda_0}\right| \cdot |q_0^{m-M}| \le \frac{|f_M|}{|1-\lambda_0|}+|v_M| \\
&\le \frac{\max\limits_m |f_m|}{\delta_1} + A_1 \max_m |f_m| = A_2 \max_m |f_m|.
\end{aligned}
$$

We have thus proven that the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ defined by formulae (10.142) coincides with the set $\{\overleftrightarrow{\Lambda} \cup \overleftarrow{\Lambda} \cup \overrightarrow{\Lambda}\}$ on the complex plane.

The foregoing algorithm for computing the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ is, in fact, quite general. We have illustrated it using a particular example of the operators defined by formulae (10.142). However, not only for this specific example but also for other scalar one-step finite-difference schemes with constant coefficients that do not explicitly depend on h, the spectrum of the family of operators $\{\boldsymbol{R}_h\}$ can be obtained by performing the same Babenko-Gelfand analysis of Section 10.5.1. *The key idea is to take into account other candidate modes that may be prone to developing the instability, besides the eigenmodes* $\{e^{i\alpha m}\}$ *of the pure Cauchy problem that are accounted for by the von Neumann analysis*.

For systems of finite-difference equations (as well as for scalar multi-step equations), the technical side of the procedure may become more involved. In this case, the computation of the spectrum of a family of operators can be reduced to studying uniform bounds for the solutions of certain ordinary difference equations with matrix coefficients. A necessary and sufficient condition has been obtained in [Rya64] for the existence of such uniform bounds. This condition is given in terms of the roots of the corresponding characteristic equation and also involves the analysis of some determinants originating from the matrix coefficients of the system. For further detail, we refer the reader to [GR87, § 4 & § 45] and [RM67, § 6.6 & § 6.7], as well as to the original journal publication by Ryaben'kii [Rya64].

10.5.3 The Energy Method

For some evolution finite-difference problems, one can obtain the l_2 estimates of the solution directly, i.e., without employing any special stability criteria, such as spectral. The corresponding technique is known as the method of energy estimates. It is useful for deriving *sufficient conditions of stability,* in particular, because it can often be applied to problems with variable coefficients on finite intervals. We illustrate the energy method with several examples.

In the beginning, let us analyze the continuous case. Consider an initial boundary

value problem for the first order constant-coefficient hyperbolic equation:

$$\begin{aligned} &\frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = 0, \quad 0 \le x \le 1, \quad 0 < t \le T, \\ &\qquad u(x,0) = \psi(x), \quad u(1,t) = 0. \end{aligned} \tag{10.170}$$

Note that both the differential equation and the boundary condition at $x = 1$ in problem (10.170) are homogeneous. Multiply the differential equation of (10.170) by $u = u(x,t)$ and integrate over the entire interval $0 \le x \le 1$:

$$\begin{aligned} &\frac{d}{dt}\int_0^1 \frac{u^2(x,t)}{2}dx - \int_0^1 \frac{\partial}{\partial x}\frac{u^2(x,t)}{2}dx \\ &\qquad = \frac{d}{dt}\frac{\|u(\cdot,t)\|_2^2}{2} - \frac{u^2(1,t)}{2} + \frac{u^2(0,t)}{2} = 0, \end{aligned}$$

where $\|u(\cdot,t)\|_2 \stackrel{\text{def}}{=} \left(\int_0^1 u^2(x,t)dx\right)^{1/2}$ is the L_2 norm of the solution in space for a given moment of time t. According to formula (10.170), the solution at $x = 1$ vanishes: $u(1,t) = 0$, and we conclude that $\frac{d}{dt}\|u(\cdot,t)\|_2^2 \le 0$, which means that $\|u(\cdot,t)\|_2$ is a non-increasing function of time. Consequently, we see that the L_2 norm of the solution will never exceed that of the initial data:

$$\|u(\cdot,t)\|_2 \le \|\psi\|_2, \quad t \ge 0. \tag{10.171}$$

Inequality (10.171) is the simplest energy estimate. It draws its name from the fact that the quantities that are quadratic with respect to the solution are often interpreted as energy in the context of physics. Note that estimate (10.171) holds for all $t \ge 0$ rather than only $0 \le t \le T$.

Next, we consider a somewhat more general formulation compared to (10.170), namely, an initial boundary value problem for the hyperbolic equation with a variable coefficient:

$$\begin{aligned} &\frac{\partial u}{\partial t} - a(x,t)\frac{\partial u}{\partial x} = 0, \quad 0 \le x \le 1, \quad 0 < t \le T, \\ &\qquad u(x,0) = \psi(x), \quad u(1,t) = 0. \end{aligned} \tag{10.172}$$

We are assuming that $\forall x \in [0,1]$ and $\forall t \ge 0:\ a(x,t) \ge a_0 > 0$, so that the characteristic speed is negative across the entire domain. Then, the differential equation renders transport from the right to the left. Consequently, setting the boundary condition $u(1,t) = 0$ at the right endpoint of the interval $0 \le x \le 1$ is legitimate.

Let us now multiply the differential equation of (10.172) by $u = u(x,t)$ and integrate over the entire interval $0 \le x \le 1$, while also applying integration by parts to

the spatial term:

$$\frac{d}{dt}\int_0^1 \frac{u^2(x,t)}{2}dx - \int_0^1 a(x,t)\frac{\partial}{\partial x}\frac{u^2(x,t)}{2}dx$$
$$= \frac{d}{dt}\frac{\|u(\cdot,t)\|_2^2}{2} - a(1,t)\frac{u^2(1,t)}{2} + a(0,t)\frac{u^2(0,t)}{2} + \int_0^1 a'_x(x,t)\frac{u^2(x,t)}{2}dx = 0.$$

Using the boundary condition $u(1,t)=0$, we find:

$$\frac{d}{dt}\frac{\|u(\cdot,t)\|_2^2}{2} = -a(0,t)\frac{u^2(0,t)}{2} - \int_0^1 a'_x(x,t)\frac{u^2(x,t)}{2}dx = 0.$$

The first term on the right-hand side of the previous equality is always non-positive. As far as the second term, let us denote $A = \sup_{(x,t)}[-a'_x(x,t)]$. Then we have:

$$\frac{d}{dt}\|u(\cdot,t)\|_2^2 \le A\|u(\cdot,t)\|_2^2,$$

which immediately yields:

$$\|u(\cdot,t)\|_2 \le e^{At/2}\|\psi\|_2, \quad t \ge 0.$$

If $A<0$, the previous inequality implies that the L_2 norm of the solution decays as $t \longrightarrow +\infty$. If $A=0$, then the norm of the solution stays bounded by the norm of the initial data. To obtain an overall uniform estimate of $\|u(\cdot,t)\|_2$ for $A \le 0$ and all $t \ge 0$, we need to select the maximum value of the constant: $\max_t e^{At/2} = 1$, and then the desired inequality will coincide with (10.171). For $A>0$, a uniform estimate can only be obtained for a given fixed interval $0 \le t \le T$, so that altogether we can write:

$$\|u(\cdot,t)\|_2 \le \begin{cases} \|\psi\|_2, & \text{if } A \le 0,\ t \ge 0, \\ e^{AT/2}\|\psi\|_2, & \text{if } A>0,\ 0 \le t \le T. \end{cases} \tag{10.173}$$

Similarly to inequality (10.171), the energy estimate (10.173) also provides a bound for the L_2 norm of the solution in terms of the L_2 norm of the initial data. However, when $A>0$ the constant in front of $\|\psi\|_2$ is no longer equal to one. Instead, $e^{AT/2}$ grows exponentially as the maximum time T elapses, and therefore estimate (10.173) for $A>0$ may only be considered on a finite interval $0 \le t \le T$ rather than for $t \ge 0$.

In problems (10.170) and (10.172) the boundary condition at $x=1$ was homogeneous. Let us now introduce yet another generalization and analyze the problem:

$$\begin{gathered} \frac{\partial u}{\partial t} - a(x,t)\frac{\partial u}{\partial x} = 0, \quad 0 \le x \le 1, \quad 0 < t \le T, \\ u(x,0) = \psi(x), \quad u(1,t) = \chi(t) \end{gathered} \tag{10.174}$$

that differs from (10.172) by its inhomogeneous boundary condition: $u(1,t) = \chi(t)$. Otherwise everything is the same; in particular, we still assume that $\forall x \in [0,1]$ and $\forall t \geq 0:\ a(x,t) \geq a_0 > 0$ and denote $A = \sup_{(x,t)}[-a'_x(x,t)]$. Multiplying the differential equation of (10.174) by $u(x,t)$ and integrating by parts, we obtain:

$$\frac{d}{dt}\frac{\|u(\cdot,t)\|_2^2}{2} = -a(0,t)\frac{u^2(0,t)}{2} + a(1,t)\frac{\chi^2(t)}{2} - \int_0^1 a'_x(x,t)\frac{u^2(x,t)}{2}dx = 0.$$

Consequently,

$$\frac{d}{dt}\|u(\cdot,t)\|_2^2 \leq A\|u(\cdot,t)\|_2^2 + a(1,t)\chi^2(t).$$

Multiplying the previous inequality by e^{-At}, we have:

$$\frac{d}{dt}\left[e^{-At}\|u(\cdot,t)\|_2^2\right] \leq e^{-At}a(1,t)\chi^2(t),$$

which, after integrating over the time interval $0 \leq \theta \leq t$, yields:

$$\|u(\cdot,t)\|_2^2 \leq e^{At}\|\psi\|_2^2 + e^{At}\int_0^t e^{-A\theta}a(1,\theta)\chi^2(\theta)d\theta.$$

As in the case of a homogeneous boundary condition, we would like to obtain a uniform energy estimate for a given interval of time. This can be done if we again distinguish between $A \leq 0$ and $A > 0$. When $A \leq 0$ we can consider all $t \geq 0$ and when $A > 0$ we can only have the estimate on some fixed interval $0 \leq t \leq T$:

$$\|u(\cdot,t)\|_2^2 \leq \begin{cases} \|\psi\|_2^2 + \int_0^\infty a(1,\theta)\chi^2(\theta)d\theta, & A \leq 0,\ t \geq 0, \\ e^{AT}\left[\|\psi\|_2^2 + \int_0^T a(1,\theta)\chi^2(\theta)d\theta\right], & A > 0,\ 0 \leq t \leq T. \end{cases} \tag{10.175}$$

When deriving inequalities (10.175), we obviously need to assume that the integrals on the right-hand side of (10.175) are bounded. These integrals can be interpreted as weighted L_2 norms of the boundary data $\chi(t)$. Clearly, energy estimate (10.175) includes the previous estimate (10.173) as a particular case for $\chi(t) \equiv 0$.

All three estimates (10.171), (10.173), and (10.175) indicate that the corresponding initial boundary value problem is *well-posed* in the sense of L_2. Qualitatively, well-posedness means that the solution to a given problem is only weakly sensitive to perturbations of the input data (such as initial and/or boundary data). In the case of linear evolution problems, well-posedness can, for example, be quantified by means of the energy estimates. These estimates provide a bound for the norm of the solution in terms of the norms of the input data. In the finite-difference context, similar estimates would have implied stability in the sense of l_2, provided that the corresponding constants on the right-hand side of each inequality can be chosen independent of the grid. We will now proceed to demonstrate how energy estimates can be obtained for finite-difference schemes.

Consider the first order upwind scheme for problem (10.170):

$$\begin{gathered}\frac{u_m^{p+1}-u_m^p}{\tau}-\frac{u_{m+1}^p-u_m^p}{h}=0,\\ m=0,1,\ldots,M-1,\quad p=0,1,\ldots,[T/\tau]-1,\\ u_m^0=\psi_m,\quad u_M^p=0.\end{gathered}\tag{10.176}$$

To obtain an energy estimate for scheme (10.176), let us first consider two arbitrary functions $\{u_m\}$ and $\{v_m\}$ on the grid $m=0,1,2,\ldots,M$. We will derive a formula that can be interpreted as a discrete analogue of the classical continuous integration by parts. In the literature, it is sometimes referred to as the summation by parts:

$$\begin{aligned}\sum_{m=0}^{M-1}u_m(v_{m+1}-v_m)&=\sum_{m=0}^{M-1}u_mv_{m+1}-\sum_{m=0}^{M-1}u_mv_m=\sum_{m=1}^{M}u_{m-1}v_m-\sum_{m=0}^{M-1}u_mv_m\\&=-\sum_{m=1}^{M}(u_m-u_{m-1})v_m+u_Mv_M-u_0v_0\\&=-\sum_{m=0}^{M-1}(u_{m+1}-u_m)v_{m+1}+u_Mv_M-u_0v_0\\&=-\sum_{m=0}^{M-1}(u_{m+1}-u_m)v_m-\sum_{m=0}^{M-1}(u_{m+1}-u_m)(v_{m+1}-v_m)\\&\quad+u_Mv_M-u_0v_0.\end{aligned}\tag{10.177}$$

Next, we rewrite the difference equation of (10.176) as

$$u_m^{p+1}=u_m^p+r(u_{m+1}^p-u_m^p),\quad r=\frac{\tau}{h}=\text{const},\quad m=0,1,\ldots,M-1,$$

square both sides, and take the sum from $m=0$ to $m=M-1$. This yields:

$$\sum_{m=0}^{M-1}\left(u_m^{p+1}\right)^2=\sum_{m=0}^{M-1}(u_m^p)^2+r^2\sum_{m=0}^{M-1}\left(u_{m+1}^p-u_m^p\right)^2+2r\sum_{m=0}^{M-1}u_m^p\left(u_{m+1}^p-u_m^p\right).$$

To transform the last term on the right-hand side of the previous equality, we apply formula (10.177):

$$\begin{aligned}\sum_{m=0}^{M-1}\left(u_m^{p+1}\right)^2&=\sum_{m=0}^{M-1}(u_m^p)^2+r^2\sum_{m=0}^{M-1}\left(u_{m+1}^p-u_m^p\right)^2+r\sum_{m=0}^{M-1}u_m^p\left(u_{m+1}^p-u_m^p\right)\\&\quad+r\left[-\sum_{m=0}^{M-1}\left(u_{m+1}^p-u_m^p\right)u_m^p-\sum_{m=0}^{M-1}\left(u_{m+1}^p-u_m^p\right)^2+\left(u_M^p\right)^2-\left(u_0^p\right)^2\right]\\&=\sum_{m=0}^{M-1}(u_m^p)^2+r(r-1)\sum_{m=0}^{M-1}\left(u_{m+1}^p-u_m^p\right)^2+r\left(u_M^p\right)^2-r\left(u_0^p\right)^2.\end{aligned}$$

Let us now assume that $r \le 1$. Then, using the conventional definition of the l_2 norm: $\|u\|_2 = \left[h\sum_0^M |u_m|^2\right]^{1/2}$ and employing the homogeneous boundary condition $u_M^p = 0$ of (10.176), we obtain the inequality:

$$\|u^{p+1}\|_2 \le \|u^p\|_2, \quad p = 0,1,2,\ldots,$$

which clearly implies the energy estimate:

$$\|u^p\|_2 \le \|\psi\|_2, \quad p = 0,1,2,\ldots. \tag{10.178}$$

The discrete estimate (10.178) is analogous to the continuous estimate (10.171).

To approximate the variable-coefficient problem (10.172), we use the scheme:

$$\begin{gathered} \frac{u_m^{p+1} - u_m^p}{\tau} - a_m^p \frac{u_{m+1}^p - u_m^p}{h} = 0, \\ m = 0,1,\ldots,M-1, \quad p = 0,1,\ldots,[T/\tau]-1, \\ u_m^0 = \psi_m, \quad u_M^p = 0, \end{gathered} \tag{10.179}$$

where $a_m^p \equiv a(x_m, t_p)$. Applying a similar approach, we obtain:

$$\begin{aligned} \sum_{m=0}^{M-1} \left(u_m^{p+1}\right)^2 = \sum_{m=0}^{M-1} \left(u_m^p\right)^2 + r^2 \sum_{m=0}^{M-1} \left(a_m^p\right)^2 \left(u_{m+1}^p - u_m^p\right)^2 + r \sum_{m=0}^{M-1} a_m^p u_m^p \left(u_{m+1}^p - u_m^p\right) \\ + r\left[- \sum_{m=0}^{M-1} \left(a_{m+1}^p u_{m+1}^p - a_m^p u_m^p\right) u_m^p \right. \\ \left. - \sum_{m=0}^{M-1} \left(a_{m+1}^p u_{m+1}^p - a_m^p u_m^p\right)\left(u_{m+1}^p - u_m^p\right) + a_M^p \left(u_M^p\right)^2 - a_0^p \left(u_0^p\right)^2 \right]. \end{aligned}$$

Next, we notice that $a_{m+1}u_{m+1} = a_m u_{m+1} + (a_{m+1} - a_m)u_{m+1}$. Substituting this expression into the previous formula, we again perform summation by parts, which is analogous to the continuous integration by parts, and which yields:

$$\begin{aligned} \sum_{m=0}^{M-1} \left(u_m^{p+1}\right)^2 = \sum_{m=0}^{M-1} \left(u_m^p\right)^2 + r^2 \sum_{m=0}^{M-1} \left(a_m^p\right)^2 \left(u_{m+1}^p - u_m^p\right)^2 + r \sum_{m=0}^{M-1} a_m^p u_m^p \left(u_{m+1}^p - u_m^p\right) \\ + r\left[- \sum_{m=0}^{M-1} \left(a_m^p \left(u_{m+1}^p - u_m^p\right) + \left(a_{m+1}^p - a_m^p\right) u_{m+1}^p\right) u_m^p \right. \\ - \sum_{m=0}^{M-1} \left(a_m^p \left(u_{m+1}^p - u_m^p\right) + \left(a_{m+1}^p - a_m^p\right) u_{m+1}^p\right)\left(u_{m+1}^p - u_m^p\right) \\ \left. + a_M^p \left(u_M^p\right)^2 - a_0^p \left(u_0^p\right)^2 \right] \\ = \sum_{m=0}^{M-1} \left(u_m^p\right)^2 + \sum_{m=0}^{M-1} \left(r^2 \left(a_m^p\right)^2 - r a_m^p\right) \left(u_{m+1}^p - u_m^p\right)^2 \end{aligned}$$

$$-r\sum_{m=0}^{M-1}\left(a_{m+1}^{p}-a_{m}^{p}\right)\left(u_{m+1}^{p}\right)^{2}+ra_{M}^{p}\left(u_{M}^{p}\right)^{2}-ra_{0}^{p}\left(u_{0}^{p}\right)^{2}.$$

Let us now assume that for all m and p we have: $ra_m^p \le 1$. Equivalently, we can require that $r \le [\sup_{(x,t)} a(x,t)]^{-1}$. Let us also introduce:

$$A=\sup_{(m,p)}\left\{-\frac{a_{m+1}^{p}-a_{m}^{p}}{h}\right\}.$$

Then, using the homogeneous boundary condition $u_M^p = 0$ and dropping the a priori non-positive term $\sum_{m=0}^{M-1} ra_m^p\left(ra_m^p-1\right)\left(u_{m+1}^p-u_m^p\right)^2$, we obtain:

$$\sum_{m=0}^{M-1}\left(u_{m}^{p+1}\right)^{2}\le\sum_{m=0}^{M-1}\left(u_{m}^{p}\right)^{2}+rA\sum_{m=0}^{M}h\left(u_{m}^{p}\right)^{2}-ra_{0}^{p}\left(u_{0}^{p}\right)^{2}-rAh\left(u_{0}^{p}\right)^{2}.$$

If $A > 0$, then for the last two terms on the right-hand side of the previous inequality we clearly have: $-r\left(u_0^p\right)^2\left(a_0^p+Ah\right)<0$. Even if $A \le 0$ we can still claim that $-r\left(u_0^p\right)^2\left(a_0^p+Ah\right)<0$ for sufficiently small h. Consequently, on fine grids the following inequality holds:

$$\|u^{p+1}\|_2^2 \le \|u^p\|_2^2+\tau A\|u^p\|_2^2=(1+A\tau)\|u^p\|_2^2,$$

which immediately implies:

$$\|u^p\|_2^2 \le (1+A\tau)^p\|\psi\|_2^2, \quad p=1,2,\ldots.$$

If $A \le 0$, the norm of the discrete solution will either decay or remain bounded as p increases. If $A > 0$, a uniform estimate of $\|u^p\|_2$ can only be obtained for $p = 0,1,\ldots,[T/\tau]$. Altogether, the solution $u^p = \{u_m^p\}$ to the finite-difference problem (10.179) satisfies the following energy estimate:

$$\|u^p\|_2 \le \begin{cases} \|\psi\|_2, & A\le 0,\ \ p=0,1,2,\ldots,\\ e^{AT/2}\|\psi\|_2, & A>0,\ \ p=0,1,2,\ldots,[T/\tau]. \end{cases} \tag{10.180}$$

The discrete estimate (10.180) is analogous to the continuous estimate (10.173).

Finally, for problem (10.174) we use the scheme:

$$\begin{gathered} \frac{u_m^{p+1}-u_m^p}{\tau}-a_m^p\frac{u_{m+1}^p-u_m^p}{h}=0,\\ m=0,1,\ldots,M-1, \quad p=0,1,\ldots,[T/\tau]-1,\\ u_m^0=\psi_m, \quad u_M^p=\chi^p. \end{gathered} \tag{10.181}$$

Under the same assumptions that we introduced when deriving estimate (10.180) for scheme (10.179), we can now write for scheme (10.181):

$$\sum_{m=0}^{M-1}\left(u_{m}^{p+1}\right)^{2}\le\sum_{m=0}^{M-1}\left(u_{m}^{p}\right)^{2}+rA\sum_{m=0}^{M}h\left(u_{m}^{p}\right)^{2}+ra_{M}^{p}\left(\chi^{p}\right)^{2}.$$

Denote $[\|u\|_2']^2 = h\sum_{m=0}^{M-1}|u_m|^2 = \|u\|_2^2 - hu_M^2$. Then the previous inequality implies:

$$[\|u^{p+1}\|_2']^2 \le (1+A\tau)[\|u^p\|_2']^2 + \tau(a_M^p + Ah)\,(\chi^p)^2, \quad p = 1,2,\ldots,$$

and consequently:

$$[\|u^p\|_2']^2 \le (1+A\tau)^p[\|\psi\|_2']^2 + \sum_{k=1}^{p}(1+A\tau)^{p-k}\tau(a_M^{k-1}+Ah)(\chi^{k-1})^2, \quad p = 1,2,\ldots.$$

We again need to distinguish between the cases $A \le 0$, $p = 0,1,2,\ldots$, and $A > 0$, $p = 0,1,2,\ldots,[T/\tau]$:

$$[\|u^p\|_2']^2 \le \begin{cases} [\|\psi\|_2']^2 + \sum_{k=1}^{\infty}\tau a_M^{k-1}(\chi^{k-1})^2, & A \le 0, \\ e^{AT}\left([\|\psi\|_2']^2 + \sum_{k=1}^{[T/\tau]}\tau(a_M^{k-1}+Ah)(\chi^{k-1})^2\right), & A > 0. \end{cases} \tag{10.182}$$

The discrete estimate (10.182) is analogous to the continuous estimate (10.175). To use the norms $\|\cdot\|_2$ instead of $\|\cdot\|_2'$ in (10.182), we only need to add a bounded quantity $h(\chi^0)^2 + h(\chi^p)^2$ on the right-hand side.

Energy estimates (10.178), (10.180), and (10.182) imply the l_2 stability of the schemes (10.176), (10.179), and (10.181), respectively, in the sense of the Definition 10.2 from page 312. Note that since the foregoing schemes are explicit, stability is not unconditional, and the Courant number has to satisfy $r \le 1$ for scheme (10.176) and $r \le [\sup_{(x,t)} a(x,t)]^{-1}$ for schemes (10.179) and (10.181).

In general, direct energy estimates appear helpful for studying stability of finite-difference schemes. Indeed, they may provide sufficient conditions for those difficult cases that involve variable coefficients, boundary conditions, and even multiple space dimensions. In addition to the scalar hyperbolic equations, energy estimates can be obtained for some hyperbolic systems, as well as for the parabolic equations. For detail, we refer the reader to [GKO95, Chapters 9 & 11], and to some fairly recent journal publications [Str94, Ols95a, Ols95b]. However, there is a key non-trivial step in proving energy estimates for finite-difference initial boundary value problems, namely, obtaining the discrete summation by parts rules appropriate for a given discretization [see the example given by formula (10.177)]. Sometimes, this step may not be obvious at all; otherwise, it may require using alternative norms based on specially chosen inner products.

10.5.4 A Necessary and Sufficient Condition of Stability. The Kreiss Criterion

In Section 10.5.2, we have shown that for stability of a finite-difference initial boundary value problem it is necessary that the spectrum of the family of transition operators $\boldsymbol{R}_h$ belongs to the unit disk on the complex plane. We have also shown, see Theorem 10.9, that this condition is, in fact, not very far from a sufficient one,

as it guarantees the scheme from developing a catastrophic exponential instability. However, it is not a fully sufficient condition, and there are examples of the schemes that satisfy the Godunov and Ryaben'kii criterion of Section 10.5.2, i.e., that have their spectrum of $\{\boldsymbol{R}_h\}$ inside the unit disk, yet they are unstable.

A comprehensive analysis of the necessary and sufficient conditions of stability for the schemes that approximate time-dependent problems on finite intervals is rather involved. In the literature, the corresponding series of results is commonly referred to as the Gustafsson, Kreiss, and Sundström (GKS) theory, and we refer the reader to the monograph [GKO95, Part II] for detail. A concise narrative of this theory can also be found in [Str04, Chapter 11]. All results of the GKS theory are formulated in terms of the l_2 norm. An important tool used for obtaining stability estimates is the Laplace transform in time.

Although a full account of (and even a self-contained introduction to) the GKS theory is beyond the scope of this text, its key ideas are easy to understand on the qualitative level and easy to illustrate with examples. The following material is essentially based on that of Section 10.5.2 and can be skipped during the first reading.

Let us consider an initial boundary value problem for the first order constant coefficient hyperbolic equation:

$$\begin{gathered}\frac{\partial u}{\partial t}-\frac{\partial u}{\partial x}=0,\quad 0\le x\le 1,\quad 0<t\le T,\\ u(x,0)=\psi(x),\quad u(1,t)=0.\end{gathered}\tag{10.183}$$

We introduce a uniform grid: $x_m=mh$, $m=0,1,\dots,M$, $h=1/M$; $t_p=p\tau$, $p=0,1,2,\dots$, and approximate problem (10.183) with the leap-frog scheme:

$$\begin{gathered}\frac{u_m^{p+1}-u_m^{p-1}}{2\tau}-\frac{u_{m+1}^p-u_{m-1}^p}{2h}=0,\\ m=1,2,\dots,M-1,\quad p=1,2,\dots,[T/\tau]-1,\\ u_m^0=\psi(x_m),\quad u_m^1=\psi(x_m+\tau),\quad m=0,1,\dots,M,\\ \boldsymbol{I}u_0^{p+1}=0,\quad u_M^{p+1}=0,\quad p=1,2,\dots,[T/\tau]-1.\end{gathered}\tag{10.184}$$

Notice that scheme (10.184) requires two initial conditions, and for simplicity we use the exact solution, which is readily available in this case, to specify u_m^1 for $m=0,1,\dots,M-1$. Also notice that the differential problem (10.183) does not require any boundary conditions at the "outflow" boundary $x=0$, but the discrete problem (10.184) requires an additional boundary condition that we symbolically denote $\boldsymbol{I}u_0^{p+1}=0$. We will investigate two different outflow conditions for scheme (10.184):

$$u_0^{p+1}=u_1^{p+1}\tag{10.185a}$$

and

$$u_0^{p+1}=u_0^p+r(u_1^p-u_0^p),\tag{10.185b}$$

where we have used our standard notation $r = \frac{\tau}{h} = \text{const}$.

Let us first note that scheme (10.184) is not a one-step scheme, which, in particular, renders the corresponding finite-difference Cauchy problem (mildly) unstable for $r = 1$, see Section 10.3.6. To reduce scheme (10.184) to the canonical form (10.141) so that to be able to investigate the spectrum of the family of operators $\{\boldsymbol{R}_h\}$, we would formally need to introduce additional variables (i.e., transform a scalar equation into a system) and then consider a one-step finite-difference equation, but with vector unknowns. However, it turns out that in this case the Babenko-Gelfand procedure of Section 10.5.1 applied to the resulting vector scheme is equivalent to the Babenko-Gelfand procedure applied directly to the scalar multi-step scheme (10.184). As such, we will skip the formal reduction of scheme (10.184) to the canonical form (10.141) and proceed immediately to computing the spectrum of the corresponding family of transition operators.

We need to analyze three model problems that follow from (10.184): A problem with no lateral boundaries:

$$\begin{gathered}\frac{u_m^{p+1}-u_m^{p-1}}{2\tau}-\frac{u_{m+1}^p-u_{m-1}^p}{2h}=0,\\ m=0,\pm1,\pm2,\ldots,\end{gathered} \tag{10.186}$$

a problem with only the left boundary:

$$\begin{gathered}\frac{u_m^{p+1}-u_m^{p-1}}{2\tau}-\frac{u_{m+1}^p-u_{m-1}^p}{2h}=0,\\ m=1,2,\ldots,\\ \boldsymbol{I}u_0^{p+1}=0,\end{gathered} \tag{10.187}$$

and a problem with only the right boundary:

$$\begin{gathered}\frac{u_m^{p+1}-u_m^{p-1}}{2\tau}-\frac{u_{m+1}^p-u_{m-1}^p}{2h}=0,\\ m=M-1,M-2,\ldots,1,0,-1,\ldots,\\ u_M^{p+1}=0.\end{gathered} \tag{10.188}$$

Substituting a solution of the type:

$$u_m^p=\lambda^p u_m$$

into the finite-difference equation:

$$u_m^{p+1}-u_m^{p-1}=r(u_{m+1}^p-u_{m-1}^p),\quad r=\tau/h,$$

that corresponds to all three problems (10.186), (10.187), and (10.188), we obtain the following second order ordinary difference equation for the eigenfunction $\{u_m\}$:

$$(\lambda-\lambda^{-1})u_m-r(u_{m+1}-u_{m-1})=0. \tag{10.189}$$

Its characteristic equation:

$$(\lambda - \lambda^{-1}) - r(q - q^{-1}) = 0 \tag{10.190a}$$

has two roots: $q_1 = q_1(\lambda)$ and $q_2 = q_2(\lambda)$, so that the general solution of equation (10.189) can be written as

$$u_m = c_1 q_1^m + c_2 q_2^m, \quad m = 0, \pm 1, \pm 2, \ldots, \quad c_1 = \text{const}, \quad c_2 = \text{const}.$$

It will also be convenient to recast the characteristic equation (10.190a) in an equivalent form:

$$q^2 - \frac{\lambda - \lambda^{-1}}{r} q - 1 = 0. \tag{10.190b}$$

From equation (10.190b) one can easily see that $q_1 q_2 = -1$ and consequently, unless both roots have unit magnitude, we always have $|q_1(\lambda)| < 1$ and $|q_2(\lambda)| > 1$.

The solution of problem (10.186) must be bounded: $|u_m| \leq \text{const}$ for $m = 0, \pm 1, \pm 2, \ldots$. We therefore require that for this problem $|q_1| = |q_2| = 1$, which means $q_1 = e^{i\alpha}$, $0 \leq \alpha < 2\pi$, and $q_2 = -e^{-i\alpha}$. The spectrum of this problem was calculated in Example 5 of Section 10.3.3:

$$\overleftrightarrow{\Lambda} = \left\{ \lambda(\alpha) = ir \sin\alpha \pm \sqrt{1 - r^2 \sin^2 \alpha} \,\middle|\, 0 \leq \alpha < 2\pi \right\}. \tag{10.191}$$

Provided that $r \leq 1$, the spectrum $\overleftrightarrow{\Lambda}$ given by formula (10.191) belongs to the unit circle on the complex plane.

For problem (10.188), we must have $u_m \longrightarrow 0$ as $m \longrightarrow -\infty$. Consequently, its general solution is given by:

$$u_m^p = c_2 \lambda^p q_2^m, \quad m = M, M-1, \ldots, 1, 0, -1, \ldots.$$

The homogeneous boundary condition $u_M^{p+1} = 0$ of (10.184) implies that a nontrivial eigenfunction $u_m = c_2 q_2^m$ may only exist if $\lambda = 0$. From the characteristic equation (10.190a) in yet another equivalent form $(\lambda^2 - 1)q - r\lambda(q^2 - 1) = 0$, we conclude that if $\lambda = 0$ then $q = 0$, which means that problem (10.188) has no eigenvalues:

$$\overleftarrow{\Lambda} = \emptyset. \tag{10.192}$$

To study problem (10.187), we first consider boundary condition (10.185a), known as the extrapolation boundary condition. The solution of problem (10.187) must satisfy $u_m \longrightarrow 0$ as $m \longrightarrow \infty$. Consequently, its general form is:

$$u_m^p = c_1 \lambda^p q_1^m, \quad m = 0, 1, 2, \ldots.$$

The extrapolation condition (10.185a) implies that a nontrivial eigenfunction $u_m = c_1 q_1^m$ may only exist if either $\lambda = 0$ or $c_1(1 - q_1) = 0$. However, we must have $|q_1| < 1$ for problem (10.187), and as such, we see that this problem has no eigenvalues either:

$$\overrightarrow{\Lambda} = \emptyset. \tag{10.193}$$

Combining formulae (10.191), (10.192), and (10.193), we obtain the spectrum of the family of operators:

$$\Lambda = \overleftrightarrow{\Lambda} \cup \overleftarrow{\Lambda} \cup \overrightarrow{\Lambda} = \overleftrightarrow{\Lambda}.$$

We therefore see that according to formula (10.191), the necessary condition for stability (Theorem 10.8) of scheme (10.184), (10.185a) is satisfied when $r \leq 1$.

However, scheme (10.184), (10.185a) still turns out unstable. The instability is not catastrophic, because according to Theorem 10.9, even if there is no uniform bound on the powers of the transition operators, their rate of growth should still be slower than any exponential function. Yet one can clearly see the instability in Figure 10.14, where we show the results of numerical integration of problem (10.183) with $\psi(x) = \cos 2\pi x$ and $u(1,t) = \cos 2\pi(1+t)$ so that $u(x,t) = \cos 2\pi(x+t)$, using scheme (10.184), (10.185a) with $r = 0.95$. (The actual proof of instability can be found, e.g., in [GKO95, Section 13.1] or in [Str04, Section 11.2].) Moreover, as $r < 1$, this instability cannot be attributed to the instability of the finite-difference Cauchy problem for the leap-frog scheme in the case $r = 1$, which is due to a multiple eigenvalue $|\lambda| = 1$, see Section 10.3.6.

In order to analyze what may have caused the instability of scheme (10.184), (10.185a), let us return to the proof of Theorem 10.9. If we were able to claim that the entire spectrum of the family of operators $\{\boldsymbol{R}_h\}$ lies strictly inside the unit disk, then a straightforward modification of that proof would immediately yield a uniform bound on the powers $\boldsymbol{R}_h^p$. This situation, however, is generally impossible. Indeed, in all our previous examples, the spectrum has always contained at least one point on the unit circle: $\lambda = 1$. It is therefore natural to assume that since the points λ inside the unit disk present no danger of instability according to Theorem 10.9, then the potential "culprits" should be sought on the unit circle.

As the finite-difference Cauchy problem (10.186) has no multiple eigenvalues $|\lambda| = 1$ for the case $r < 1$, let us revisit the problem with the left boundary (10.187). We have shown that this problem has no nontrivial eigenfunctions in the class $u_m \longrightarrow 0$ as $m \longrightarrow \infty$ and accordingly, it has no eigenvalues either, see formula (10.193). As such, it does not contribute to the overall spectrum of the family of operators. However, even though the boundary condition (10.185a) in the form $c_1(1-q_1) = 0$ is not satisfied by any function $u_m = c_1 q_1^m$, where $|q_1| < 1$, we see that it is "almost satisfied" if the root q_1 is close to one. Therefore, the function $u_m = c_1 q_1^m$ is "almost an eigenfunction" of problem (10.187), and the smaller the quantity $|1-q_1|$, the more of a genuine eigenfunction it becomes.

To investigate stability, we need to determine whether or not the foregoing "almost an eigenfunction" can bring along an unstable eigenvalue, or rather "almost an eigenvalue," $|\lambda| > 1$. By passing to the limit $q_1 \longrightarrow 1$, we find from equation (10.190a) that $\lambda = 1$ or $\lambda = -1$. We should therefore analyze the behavior of the quantities λ and q in a neighborhood of each of these two values of λ, when the relation between λ and q is given by equation (10.190a).

First recall that according to formula (10.191), if $|q| = 1$, then $|\lambda| = 1$ (provided that $r \leq 1$). Consequently, if $|\lambda| > 1$, then $|q| \neq 1$, i.e., there are two distinct roots: $|q_1| < 1$ and $|q_2| > 1$. In particular, when λ is near $(1,0)$, there are still two roots —

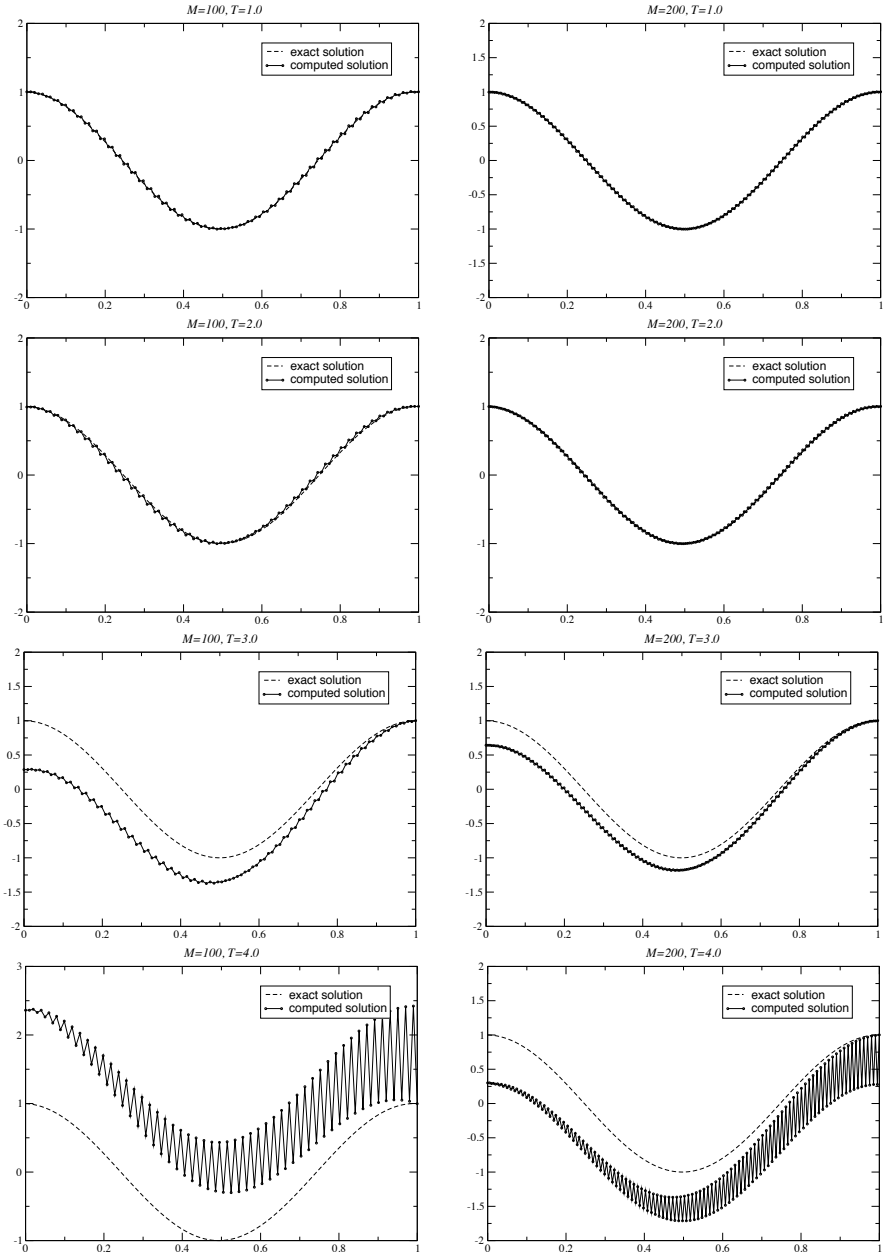


FIGURE 10.14: Solution of problem (10.183) with scheme (10.184), (10.185a).

one with the magnitude greater than one and the other with the magnitude less than one. When $|\lambda - 1| \longrightarrow 0$ we will clearly have $|q_1| \longrightarrow 1$ and $|q_2| \longrightarrow 1$. We, however,

don't know ahead of time which of the two possible scenarios actually takes place:

$$\lim_{|\lambda|>1,\,\lambda\to 1} q_1(\lambda) = 1, \qquad \lim_{|\lambda|>1,\,\lambda\to 1} q_2(\lambda) = -1 \tag{10.194a}$$

or

$$\lim_{|\lambda|>1,\,\lambda\to 1} q_1(\lambda) = -1, \qquad \lim_{|\lambda|>1,\,\lambda\to 1} q_2(\lambda) = 1. \tag{10.194b}$$

To find this out, let us notice that the roots $q_1(\lambda)$ and $q_2(\lambda)$ are continuous (in fact, analytic) functions of λ. Consequently, if we take λ in the form $\lambda = 1+\eta$, where $|\eta| \ll 1$, and if we want to investigate the root q that is close to one, then we can say that $q(\lambda) = 1+\zeta$, where $|\zeta| \ll 1$. From equation (10.190a) we then obtain:

$$2\eta + \mathscr{O}(\eta^2) = 2r\zeta + \mathscr{O}(\zeta^2). \tag{10.195}$$

Consider a special case of real $\eta > 0$, then ζ must obviously be real as well. From the previous equality we find that $\zeta > 0$ (because $r > 0$), i.e., $|q| > 1$. As such, we see that if $|\lambda| > 1$ and $\lambda \longrightarrow 1$, then

$$\{q = q(\lambda) \longrightarrow 1\} \Longrightarrow \{|q| > 1\}.$$

Indeed, for real η and ζ, we have $|q| = 1+\zeta > 1$; for other η and ζ the same result follows by continuity. Consequently, it is the root q_2 that approaches $(1,0)$ when $\lambda \longrightarrow 1$, and the true scenario is given by (10.194b) rather than by (10.194a).

We therefore see that when a potentially "dangerous" unstable eigenvalue $|\lambda| > 1$ approaches the unit circle at $(1,0)$: $\lambda \longrightarrow 1$, it is the grid function $u_m = c_2 q_2^m$, $|q_2| > 1$, that will almost satisfy the boundary condition (10.185a), because $c_2(1-q_2) \longrightarrow 0$. This grid function, however, does not satisfy the requirement $u_m \longrightarrow 0$ as $m \longrightarrow \infty$, i.e., it does not belong to the class of functions admitted by problem (10.187). On the other hand, the function $u_m = c_1 q_1^m$, $|q_1| > 1$, that satisfies $u_m \longrightarrow 0$ as $m \longrightarrow \infty$, will be very far from satisfying the boundary condition (10.185a) because $q_1 \longrightarrow -1$.

Next, recall that we actually need to investigate what happens when $q_1 \longrightarrow 1$, i.e., when $c_1 q_1^m$ is almost an eigenfunction. This situation appears opposite to the one we have analyzed. Consequently, when $q_1 \longrightarrow 1$ we will not have such a $\lambda(q_1) \longrightarrow 1$ where $|\lambda(q_1)| > 1$. Qualitatively, this indicates that there is no instability associated with "almost an eigenfunction" $u_m = c_1 q_1^m$, $|q_1| > 1$, of problem (10.187). In the framework of the GKS theory, this assertion can be proven rigorously.

Let us now consider the second case: $\lambda \longrightarrow -1$ while $|\lambda| > 1$. We need to determine which of the two scenarios holds:

$$\lim_{|\lambda|>1,\,\lambda\to -1} q_1(\lambda) = 1, \qquad \lim_{|\lambda|>1,\,\lambda\to -1} q_2(\lambda) = -1 \tag{10.196a}$$

or

$$\lim_{|\lambda|>1,\,\lambda\to -1} q_1(\lambda) = -1, \qquad \lim_{|\lambda|>1,\,\lambda\to -1} q_2(\lambda) = 1. \tag{10.196b}$$

Similarly to the previous analysis, let $\lambda = -1 + \eta$, where $|\eta| \ll 1$, then also $q(\lambda) = 1 + \zeta$, where $|\zeta| \ll 1$ (recall, we are still interested in $q \longrightarrow 1$). Consider a particular case of real $\eta < 0$, then equation (10.195) yields $\zeta < 0$, i.e., $|q| < 1$. Consequently, if $|\lambda| > 1$ and $\lambda \longrightarrow -1$, then

$$\{q = q(\lambda) \longrightarrow 1\} \Longrightarrow \{|q| < 1\}.$$

In other words, this time it is the root q_1 that approaches $(1,0)$ as $\lambda \longrightarrow -1$, and the scenario that gets realized is (10.196a) rather than (10.196b). In contradistinction to the previous case, this presents a potential for instability. Indeed, the pair (λ, q_1), where $|q_1| < 1$ and $|\lambda| > 1$, would have implied the instability in the sense of Section 10.5.2 if $c_1 q_1^m$ were a genuine eigenfunction of problem (10.187) and λ if were the corresponding genuine eigenvalue. As we know, this is not the case. However, according to the first formula of (10.196a), the actual setup appears to be a limit of the admissible yet unstable situation. In other words, the combination of "almost an eigenfunction" $u_m = c_1 q_1^m$, $|q_1| < 1$, that satisfies $u_m \longrightarrow 0$ as $m \longrightarrow \infty$ with "almost an eigenvalue" $\lambda = \lambda(q_1)$, $|\lambda| > 1$, is unstable. While remaining unstable, this combination becomes more of a genuine eigenpair of problem (10.187) as $\lambda \longrightarrow -1$. Again, a rigorous proof of the instability is given in the framework of the GKS theory using the technique based on the Laplace transform.

Thus, we have seen that two scenarios are possible when λ approaches the unit circle from the outside. In one case, there may be an admissible root q of the characteristic equation that almost satisfies the boundary condition, see formula (10.196a), and this situation is prone to instability. Otherwise, see formula (10.194b), there is no admissible root q that would ultimately satisfy the boundary condition, and as such, no instability will be associated with this λ.

In the unstable case exemplified by formula (10.196a), the corresponding limit value of λ is called *the generalized eigenvalue,* see [GKO95, Chapter 13]. In particular, $\lambda = -1$ is a generalized eigenvalue of problem (10.187). We re-emphasize that it is not a genuine eigenvalue of problem (10.187), because when $\lambda = -1$ then $q_1 = 1$ and the eigenfunction $u_m = cq_1^m$ does not belong to the admissible class: $u_m \longrightarrow 0$ as $m \longrightarrow \infty$. In fact, it is easy to see that $\|u\|_2 = \infty$. However, it is precisely this generalized eigenvalue that causes the instability even when the entire spectrum of the family of operators $\{\boldsymbol{R}_h\}$ belongs to the unit disk and $r < 1$.

Accordingly, *the Kreiss necessary and sufficient condition of stability* requires that the spectrum of the family of operators be confined to the unit disk as before, and additionally, that the scheme should have no generalized eigenvalues $|\lambda| = 1$. In the case of systems, the discrete Cauchy problem must also be stable in the sense of Theorem 10.4 (which, for the leap-frog scheme, means $r < 1$). Scheme (10.184), (10.185a) violates the Kreiss condition as it has a generalized eigenvalue $\lambda = -1$. Hence, it is unstable, see Figure 10.14.

Since, however, this instability is only due to a generalized eigenvalue with $|\lambda| = 1$, it is relatively mild, as expected. On the other hand, if we were to replace the marginally unstable boundary condition (10.185a) with a truly unstable one in the sense of Section 10.5.2, then the effect on the stability of the scheme would have

been much more drastic. Instead of (10.185a), consider, for example:

$$u_0^{p+1} = 1.05 \cdot u_1^{p+1}. \tag{10.197}$$

This boundary condition generates an eigenfunction $u_m = c_1 q_1^m$ of problem (10.187) with $q_1 = \frac{1}{1.05} < 1$. The corresponding eigenvalues are given by:

$$\lambda(q_1) = \frac{r}{2}\left(q_1 - \frac{1}{q_1}\right) \pm \sqrt{1 + \frac{r^2}{4}\left(q_1 - \frac{1}{q_1}\right)^2},$$

and for one of these eigenvalues we obviously have $|\lambda| > 1$. Therefore, the scheme is unstable according to Theorem 10.8, see also Figure 10.15.

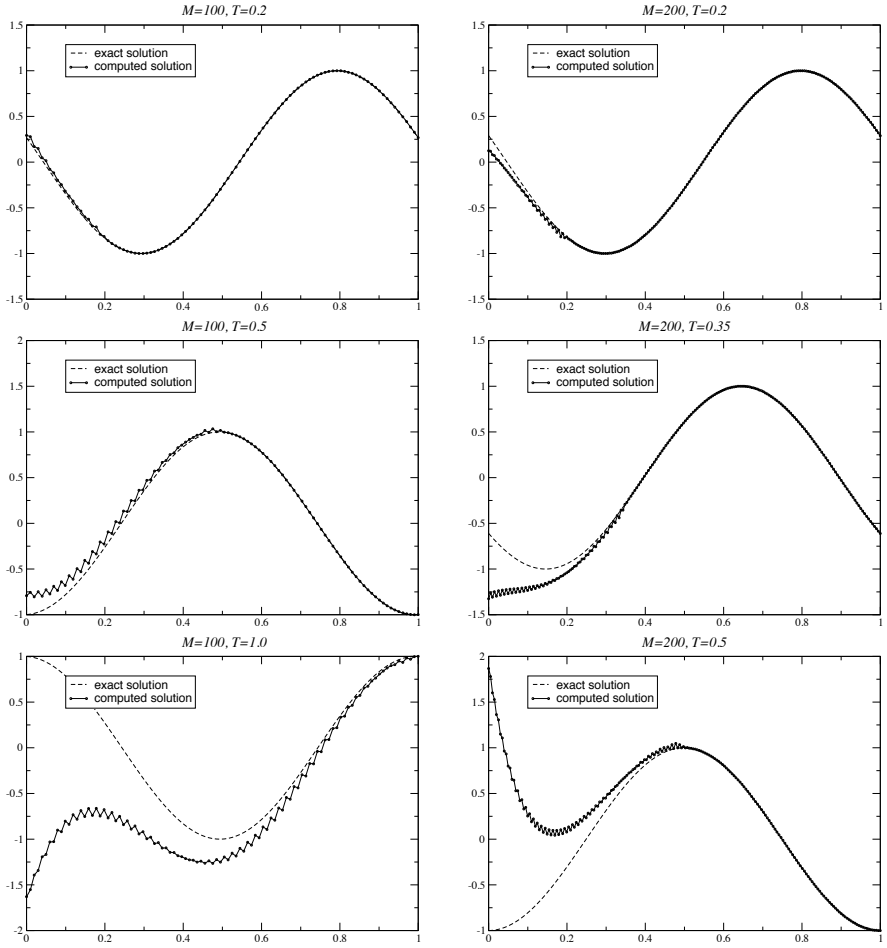


FIGURE 10.15: Solution of problem (10.183) with scheme (10.184), (10.197).

In Figure 10.15, we are showing the results of the numerical solution of problem (10.183) using the unstable scheme (10.184), (10.197). Comparing the plots in Figure 10.15 with those in Figure 10.14, we see that in the case of boundary condition (10.197) the instability develops much more rapidly in time. Moreover, comparing the left column in Figure 10.15 that corresponds to the grid with $M = 100$ cells with the right column in the same figure that corresponds to $M = 200$, we see that the instability develops more rapidly on a finer grid, which is characteristic of an exponential instability.

Let now now analyze the second outflow boundary condition (10.185b):

$$u_0^{p+1} = u_0^p + r(u_1^p - u_0^p).$$

Unlike the extrapolation-type boundary condition (10.185a), which to some extent is arbitrary, boundary condition (10.185b) merely coincides with the first order upwind approximation of the differential equation itself that we have encountered previously on multiple occasions. To study stability, we again need to investigate three model problems: (10.186), (10.187), and (10.188). Obviously, only problem (10.187) changes due to the new boundary condition, where the other two stay the same. Moreover, as the Cauchy problem (10.186) is not stable for $r = 1$, it is sufficient to analyze the boundary conditions only for $r < 1$.

To find λ and q for problem (10.187), we need to solve the characteristic equation (10.190a) along with a similar equation that comes from the boundary condition (10.185b):

$$\lambda = 1 - r + rq. \tag{10.198}$$

Substituting λ from equation (10.198) into equation (10.190a) and subsequently solving for q, we find that there is only one solution: $q = 1$. For the corresponding λ, we then have from equation (10.198): $\lambda = 1$. Consequently, for $r < 1$ problem (10.187) has no proper eigenfunctions/eigenvalues, which means that we again have $\overrightarrow{\Lambda} = \emptyset$. As far as the generalized eigenvalues, we only need to check one value of λ: $\lambda = 1$ (because $\lambda = -1$ does not satisfy equation (10.198) for $q = 1$). Let $\lambda = 1 + \eta$ and $q = 1 + \zeta$, where $|\eta| \ll 1$ and $|\zeta| \ll 1$. We then arrive at the same equation (10.195) that we obtained in the context of the previous analysis and conclude that $\lambda = 1$ does not violate the Kreiss condition, because $|\lambda| > 1$ implies $|q| > 1$. As such, the scheme (10.184), (10.185b) is stable when $r < 1$.

Exercises

1. For the scalar Lax-Wendroff scheme [cf. formula (10.83)]:

$$\frac{u_m^{p+1} - u_m^p}{\tau} - \frac{u_{m+1}^p - u_{m-1}^p}{2h} - \frac{\tau}{2}\frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = 0,$$
$$p = 0, 1, \ldots, [T/\tau] - 1, \quad m = 1, 2, \ldots, M-1, \quad Mh = 1,$$
$$u_m^0 = \psi(x_m), \quad m = 0, 1, 2, \ldots, M,$$
$$\frac{u_0^{p+1} - u_0^p}{\tau} - \frac{u_1^p - u_0^p}{h} = 0, \quad u_M^{p+1} = 0, \quad p = 0, 1, \ldots, [T/\tau] - 1,$$

that approximates the initial boundary value problem:

$$\frac{\partial u}{\partial t}-\frac{\partial u}{\partial x}=0,\quad 0\le x\le 1,\quad 0<t\le T,$$
$$u(x,0)=\psi(x),\quad u(1,t)=0,$$

on the uniform rectangular grid: $x_m=mh$, $m=0,1,\dots,M$, $Mh=1$, $t_p=p\tau$, $p=0,1,\dots,[T/\tau]$, find out when the Babenko-Gelfand stability criterion holds.

Answer. $r=\tau/h\le 1$.

2.* Prove Theorem 10.6.

a) Prove the sufficiency part.

b) Prove the necessity part.

3.* Approximate the acoustics Cauchy problem:

$$\frac{\partial \boldsymbol{u}}{\partial t}-\boldsymbol{A}\frac{\partial \boldsymbol{u}}{\partial x}=\boldsymbol{\varphi}(x,t),\quad -\infty\le x\le\infty,\quad 0<t\le T,$$
$$\boldsymbol{u}(x,0)=\boldsymbol{\psi}(x),\quad -\infty\le x\le\infty,$$
$$\boldsymbol{u}(x,t)=\begin{bmatrix}v(x,t)\\ w(x,t)\end{bmatrix},\quad \boldsymbol{\varphi}(x)=\begin{bmatrix}\varphi^{(1)}(x)\\ \varphi^{(2)}(x)\end{bmatrix},\quad \boldsymbol{\psi}(x)=\begin{bmatrix}\psi^{(1)}(x)\\ \psi^{(2)}(x)\end{bmatrix},\quad \boldsymbol{A}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},$$

with the Lax-Wendroff scheme:

$$\frac{\boldsymbol{u}_m^{p+1}-\boldsymbol{u}_m^p}{\tau}-\boldsymbol{A}\frac{u_{m+1}^p-u_{m-1}^p}{2h}-\frac{\tau}{2}\boldsymbol{A}^2\frac{\boldsymbol{u}_{m+1}^p-2\boldsymbol{u}_m^p+\boldsymbol{u}_{m-1}^p}{h^2}=\boldsymbol{\varphi}_m^p,$$
$$p=0,1,\dots,[T/\tau]-1,\quad m=0,\pm1,\pm2,\dots,$$
$$\boldsymbol{u}_m^0=\boldsymbol{\psi}(x_m),\quad m=0,\pm1,\pm2,\dots.$$

Define $\boldsymbol{u}^p=\{\boldsymbol{u}_m^p\}$ and $\boldsymbol{\varphi}^p=\{\boldsymbol{\varphi}_m^p\}$, and introduce the norms as follows:

$$\|\boldsymbol{u}^{(h)}\|_{U_h}=\max_p\|\boldsymbol{u}^p\|,\quad \|\boldsymbol{f}^{(h)}\|_{F_h}=\max\left[\|\boldsymbol{\psi}\|,\max_p\|\boldsymbol{\varphi}^p\|\right],$$

where

$$\|\boldsymbol{u}^p\|^2=\sum_m\left(|v_m^p|^2+|w_m^p|^2\right),\quad \|\boldsymbol{\psi}\|^2=\sum_m\left(|\psi^{(1)}(x_m)|^2+|\psi^{(2)}(x_m)|^2\right),$$
$$\|\boldsymbol{\varphi}^p\|^2=\sum_m\left(|\varphi^{(1)}(x_m,t_p)|^2+|\varphi^{(2)}(x_m,t_p)|^2\right).$$

a) Show that when reducing the Lax-Wendroff scheme to the canonical form (10.141), inequalities (10.143) and (10.144) hold.

b) Prove that when $r=\frac{\tau}{h}\le 1$ the scheme is l_2 stable, and when $r>1$ it is unstable.

Hint. To prove estimate (10.145) for the norms $\|\boldsymbol{R}_h^p\|$, first introduce the new unknown variables (called the Riemann invariants):

$$I_m^{(1)}=v_m+w_m\ \text{ and }\ I_m^{(2)}=v_m-w_m,$$

and transform the discrete system accordingly, and then employ the spectral criterion of Section 10.3.

4. Let the norm in the space U_h' be defined in the sense of l_2: $\|u\|_2 = \left[h \sum\limits_{m=-\infty}^{\infty} |u_m|^2\right]^{1/2}$.

   Prove that in this case all complex numbers $\lambda(\alpha) = 1 - r + re^{i\alpha}$, $0 \le \alpha < 2\pi$ [see formula (10.148)], belong to the spectrum of the transition operator $\boldsymbol{R}_h$ that corresponds to the difference Cauchy problem (10.147), where the spectrum is defined according to Definition 10.7.

   Hint. Construct the solution $u = \{u_m\}, m = 0, \pm 1, \pm 2, \ldots$, to the inequality that appears in Definition 10.7 in the form: $u_m = \begin{cases} q_1^m, & m \ge 0, \\ q_2^{-m}, & m < 0, \end{cases}$ where $q_1 = (1-\delta)e^{i\alpha}$, $q_2 = (1-\delta)e^{-i\alpha}$, and $\delta > 0$ is a small quantity.

5. Prove sufficiency in Theorem 10.7.

   Hint. Use expansion with respect to an orthonormal basis in U' composed of the eigenvectors of $\boldsymbol{R}_h$.

6. Compute the spectrum of the family of operators $\{\boldsymbol{R}_h\}$, $v = \boldsymbol{R}_h u$, given by the formulae:

$$\begin{aligned} v_m &= (1-r)u_m + ru_{m+1}, \quad m = 0, 1, \ldots, M-1, \\ v_M &= 0. \end{aligned}$$

   Assume that the norm is the maximum norm.

7. Prove that the spectrum of the family of operators $\{\boldsymbol{R}_h\}$, $v = \boldsymbol{R}_h u$, defined as:

$$\begin{aligned} v_m &= (1-r+\gamma h)u_m + ru_{m+1}, \quad m = 0, 1, \ldots, M-1, \\ v_M &= u_M, \end{aligned}$$

   does not depend on the value of γ and coincides with the spectrum computed in Section 10.5.2 for the case $\gamma = 0$. Assume that the norm is the maximum norm.

   Hint. Notice that this operator is obtained by adding $\gamma h \boldsymbol{I}'$ to the operator $\boldsymbol{R}_h$ defined by formulae (10.142a) & (10.142b), and then use Definition 10.6 directly. Here $\boldsymbol{I}'$ is a modification of the identity operator that leaves all components of the vector u intact except the last component u_M that is set to zero.

8. Compute the spectrum of the family of operators $\{\boldsymbol{R}_h\}$, $v = \boldsymbol{R}_h u$, given by the formulae:

$$\begin{aligned} v_m &= (1-r)u_m + r(u_{m-1} + u_{m+1})/2, \quad m = 1, 2, \ldots, M-1, \\ v_M &= 0, \quad av_0 + bv_1 = 0, \end{aligned}$$

   where $a \in \mathbb{R}$ and $b \in \mathbb{R}$ are known and fixed. Consider the cases $|a| > |b|$ and $|a| < |b|$.

9.* Prove that the spectrum of the family of operators $\{\boldsymbol{R}_h\}$, $v = \boldsymbol{R}_h u$, defined by formulae (10.142a) & (10.142b) and analyzed in Section 10.5.2:

$$\begin{aligned} v_m &= (1-r)u_m + ru_{m+1}, \quad m = 0, 1, \ldots, M-1, \\ v_M &= u_M, \end{aligned}$$

   will not change if the C norm: $\|u\| = \max_m |u_m|$ is replaced by the l_2 norm: $\|u\| = \left[h \sum_m u_m^2\right]^{1/2}$.

10. For the first order ordinary difference equation:

$$av_m + bv_{m+1} = f_m, \quad m = 0, \pm 1, \pm 2, \ldots,$$

the fundamental solution G_m is defined as *a bounded solution* of the equation:

$$aG_m + bG_{m+1} = \delta_m \equiv \begin{cases} 1, & m = 0, \\ 0, & m \neq 0. \end{cases}$$

a) Prove that if $|a/b| < 1$, then $G_m = \begin{cases} 0, & m \leq 0, \\ -\frac{1}{a}\left(-\frac{a}{b}\right)^m, & m \geq 1. \end{cases}$

b) Prove that if $|a/b| > 1$, then $G_m = \begin{cases} \frac{1}{a}\left(-\frac{a}{b}\right)^m, & m \leq 0, \\ 0, & m \geq 1. \end{cases}$

c) Prove that $v_m = \sum\limits_{k=-\infty}^{\infty} G_{m-k} f_k$.

11. Obtain energy estimates for the implicit first order upwind schemes that approximate problems (10.170), (10.172)*, and (10.174)*.

12.* Approximate problem (10.170) with the Crank-Nicolson scheme supplemented by one-sided differences at the left boundary $x = 0$:

$$\begin{gathered} \frac{u_m^{p+1} - u_m^p}{\tau} - \frac{1}{2}\left[\frac{u_{m+1}^{p+1} - u_{m-1}^{p+1}}{2h} + \frac{u_{m+1}^p - u_{m-1}^p}{2h}\right] = 0, \\ m = 1, 2, \ldots, M-1, \quad p = 0, 1, \ldots, [T/\tau] - 1, \\ \frac{u_0^{p+1} - u_0^p}{\tau} - \frac{1}{2}\left[\frac{u_1^{p+1} - u_0^{p+1}}{h} + \frac{u_1^p - u_0^p}{h}\right] = 0, \quad u_M^p = 0, \\ p = 0, 1, \ldots, [T/\tau] - 1, \\ u_m^0 = \psi_m, \quad m = 0, 1, 2, \ldots, M. \end{gathered} \tag{10.199}$$

a) Use an alternative definition of the l_2 norm: $\|u\|_2^2 = \frac{h}{2}(u_0^2 + u_M^2) + h\sum\limits_{m=1}^{M-1} u_m^2$ and develop an energy estimate for scheme (10.199).

Hint. Multiply the equation by $u_m^{p+1} + u_m^p$ and sum over the entire range of m.

b) Construct the schemes similar to (10.199) for the variable-coefficient problems (10.172) and (10.174) and obtain energy estimates.

13. Using the Kreiss condition, show that the leap-frog scheme (10.184) with the boundary condition:

$$u_0^{p+1} = u_1^p \tag{10.200a}$$

is stable, whereas with the boundary condition:

$$u_0^{p+1} = u_0^{p-1} + 2r(u_1^p - u_0^p) \tag{10.200b}$$

it is unstable.

14. Reproduce on the computer the results shown in Figures 10.14 and 10.15. In addition, conduct the computations using the leap-frog scheme with the boundary conditions (10.185b), (10.200a), and (10.200b), and demonstrate experimentally the stability and instability in the respective cases.

15.* Using the Kreiss condition, investigate stability of the Crank-Nicolson scheme applied to solving problem (10.183) and supplemented either with the boundary condition (10.185b) or with the boundary condition (10.200a).

10.6 Maximum Principle for the Heat Equation

Consider the following initial boundary value problem for a variable-coefficient heat equation, $a(x,t) > 0$:

$$\begin{gathered}\frac{\partial u}{\partial t} - a(x,t)\frac{\partial^2 u}{\partial x^2} = \varphi(x,t), \quad 0 \le x \le 1, \quad 0 \le t \le T, \\ u(x,0) = \psi(x), \quad 0 \le x \le 1, \\ u(0,t) = \vartheta(t), \quad u(1,t) = \chi(t), \quad 0 \le t \le T.\end{gathered} \tag{10.201}$$

To solve problem (10.201) numerically, we can use either an explicit or an implicit finite-difference scheme. We will analyze and compare both schemes. In doing so, we will see that quite often the implicit scheme has certain advantages compared to the explicit scheme, even though the algorithm of computing the solution with the help of an explicit scheme is simpler than that for the implicit scheme. The advantages of using an implicit scheme stem from its unconditional stability, i.e., stability that holds for any ratio between the spatial and temporal grid sizes.

10.6.1 An Explicit Scheme

We introduce a uniform grid on the interval $[0,1]$: $x_m = mh$, $m = 0,1,\ldots,M$, $Mh = 1$, and build the scheme on the four-node stencil shown in Figure 10.3(left) (see page 331):

$$\begin{gathered}\frac{u_m^{p+1} - u_m^p}{\tau_p} - a(x_m,t_p)\frac{u_{m+1}^p - 2u_m^p + u_{m-1}^p}{h^2} = \varphi(x_m,t_p), \\ m = 1,2,\ldots,M-1, \\ u_m^0 = \psi(x_m) \equiv \psi_m, \quad m = 0,1,\ldots,M, \\ u_0^{p+1} = \vartheta(t_{p+1}), \quad u_M^{p+1} = \chi(t_{p+1}), \quad p \ge 0, \\ t_0 = 0, \quad t_p = \tau_0 + \tau_1 + \ldots + \tau_{p-1}, \quad p = 1,2,\ldots.\end{gathered} \tag{10.202}$$

If the solution u_m^k, $m = 0,1,\ldots,M$, is already known for $k = 0,1,\ldots,p$, then, by virtue of (10.202), the values of u_m^{p+1} at the next time level $t = t_{p+1} = t_p + \tau_p$ can be