it follows that the above relation cannot hold. Thus, there is no eigenvalue for $\mathrm{Re}\,\tilde{s} \geq 0$, $\tilde{s} \neq 0$.

For the eigenvalue problem in its original form, $\tilde{s} = 0$ is a generalized eigenvalue corresponding to $\kappa_1 = 1$. Therefore, we have to modify it to the form shown in Eq. (12.6.20). The boundary conditions become

$$\psi_0 + \psi_{-1} = 0, \qquad D_+^2 \psi_{-1} = 0. \tag{12.6.24}$$

Both φ_j and ψ_j are linear combinations of κ_1^j and κ_2^j if $\kappa_1 \neq \kappa_2$. Thus, in the neighborhood of $\tilde{s} = 0$, the general solution of Eq. (12.6.20) can be written as

$$\varphi_j \approx -\sigma_1 (1 - \tilde{s}^{1/2})^j + \sigma_2 \frac{\tilde{s}^{1/2}}{\kappa_2(0) - 1} \kappa_2^j(0).$$

Therefore,

$$\tilde{s}^{1/2} \psi_j = \varphi_{j+1} - \varphi_j \approx \sigma_1 \tilde{s}^{1/2} (1 - \tilde{s}^{1/2})^j + \sigma_2 \tilde{s}^{1/2} \kappa_2^j(0);$$

that is,

$$\psi_j \approx \sigma_1 (1 - \tilde{s}^{1/2})^j + \sigma_2 \kappa_2^j(0).$$

(The choice of coefficients in the representation of φ_j was made to get a simple normalized form for ψ_j.) Substituting this expression into the boundary conditions (12.6.24) shows that $\sigma_1 = \sigma_2 = 0$. Thus, $\tilde{s} = 0$ is not a generalized eigenvalue and the approximation is stable in the generalized sense.

The above results can be extended to very general parabolic systems. We can neglect all lower order terms and we can prove the following theorem.

Theorem 12.6.3. *The approximation for general parabolic systems is stable in the generalized sense if it is dissipative for the Cauchy problem and if the modified eigenvalue problem has no eigenvalues or generalized eigenvalues for* $\mathrm{Re}\,\tilde{s} \geq 0$.

EXERCISES

12.6.1. Prove that the estimates (12.6.14) are sharp.

12.6.2. Consider the matrix $M_0 + \tilde{s}^{1/2} M_1$ in Eq. (12.6.16a). Prove that its eigenvalues are the roots κ given by the characteristic equation given in Lemma 12.6.1.

12.6.3. Prove that the estimates (12.6.18) are sharp.

12.6.4. Introduce nonzero data in the boundary conditions of Example 2 above and prove that the approximation is strongly stable in the generalized sense.

12.7. THE CONVERGENCE RATE

In Section 11.3, the basic principles for obtaining error estimates were discussed in connection with the energy method for stability analysis. In this section, we present a more detailed discussion, which also includes the more general stability analysis based on the Laplace transform.

The stability estimate obtained for a certain approximation is the key to the proper error estimates. This was demonstrated in Part I for the pure initial value problem and the same principles also hold when boundary conditions are involved. However, to obtain optimal estimates, one has to be a little more careful.

The basic idea is to insert the true solution $u(x,t)$ into the approximation. This generally introduces truncation errors as inhomogeneous terms in the difference approximation, the initial conditions, and the boundary conditions. The stability estimate then gives the desired error estimate, because small data only can give small solutions.

The error $e_j(t) = v_j(t) - u(x_j, t)$ satisfies the equations

$$\frac{de_j}{dt} = Qe_j + h^{q_1} F_j, \qquad j = 1, 2, \ldots, \tag{12.7.1a}$$

$$e_j(0) = h^{q_2} f_j, \tag{12.7.1b}$$

$$L_0 e_0 = h^{q_3} g, \tag{12.7.1c}$$

where F, f, and g are smooth functions. (For convenience, we use the same notation for these functions as in the original approximation for $v(t)$. Furthermore, we consider only the quarter-space problem even when the energy method is used.) The integers q_1, q_2, and q_3 are not necessarily equal. Equations (12.7.1) are based on the fact that there is a smooth solution $u(x,t)$ to the continuous problem, and we recall that this requires certain compatibility conditions on the initial and boundary data and on the forcing function if there is one.

Let us first consider the case that Q is a semibounded operator, so that the energy method can be applied. This requires homogeneous boundary conditions, and our recipe above was to subtract a certain smooth function $\varphi_j(t)$ that satisfies the boundary condition. Assuming that this is possible and that $\varphi_j(t) = h^{q_3} \psi_j(t)$, where $\psi_j(t)$ is smooth, we have

$$\frac{d\varphi_j}{dt} = \mathcal{O}(h^{q_3}), \qquad (12.7.2a)$$
$$Q\varphi_j = \mathcal{O}(h^{q_3}), \qquad (12.7.2b)$$

which yields, for $\tilde{e}_j(t) = e_j(t) - \varphi_j(t)$,

$$\begin{aligned} \frac{d\tilde{e}_j}{dt} &= Q\tilde{e}_j + (h^{q_1} + h^{q_3})\tilde{F}_j, \qquad j = 1,2,\ldots, \\ \tilde{e}_j(0) &= (h^{q_2} + h^{q_3})\tilde{f}_j, \\ L_0\tilde{e}_0 &= 0. \end{aligned} \qquad (12.7.3)$$

The energy estimate yields a bound on any finite time interval $[0, T]$,

$$\|\tilde{e}(t)\|_h \leq \text{constant}\,(h^{q_1} + h^{q_2} + h^{q_3}),$$

and by construction

$$\|e(t)\|_h \leq \text{constant}\,h^q, \qquad q = \min(q_1, q_2, q_3). \qquad (12.7.4)$$

[This estimate corresponds to Theorem 11.3.1.] We often have $q_2 = \infty$, and q_1 is given by the order of the approximation at inner points. It is, therefore, natural to choose boundary conditions such that $q_3 = q_1$.

The crucial issue is the construction of $\varphi_j(t)$. Let us first assume that all the boundary conditions in Eq. (12.7.1c) are approximations of the boundary conditions for the differential equation. For example, for a scalar parabolic equation we could approximate $au_x(0,t) + bu(0,t) = 0$ by

$$a\,\frac{v_1 - v_0}{h} + b\,\frac{v_1 + v_0}{2} = 0. \qquad (12.7.5)$$

If the grid is located such that $x_0 = -h/2$, then we have $q_3 = 2$ in Eq. (12.7.1c) with g being a combination of u derivatives. Hence, φ can be constructed such that it satisfies Eq. (12.7.2). For the equation $u_t + u_x = 0$, $u(0,t) = 0$, the same conclusion holds with $a = 0$, and $b = 1$ in Eq. (12.7.5). This is a general principle: As long as Eq. (12.7.1c) does not contain any extra boundary conditions, the error estimate follows immediately when there is an energy estimate. (One can also show that $q_3 \geq q_1$ is a necessary condition for an h^{q_1} estimate for this type of boundary conditions.)

Let us next consider the case that there are extra boundary conditions. We use the familiar example $u_t = u_x$ with extrapolation at the boundary

$$v_0 - 2v_1 + v_2 = 0.$$

The error satisfies

$$e_0 - 2e_1 + e_2 = -h^2 u_{xx}(0,t) + \mathcal{O}(h^3) =: h^2 g(t),$$

and we need a function $\varphi_j(t) = h^2\psi_j(t)$, where

$$\psi_0 - 2\psi_1 + \psi_2 = g(t). \qquad (12.7.6)$$

But if ψ is smooth, we have

$$\psi_0 - 2\psi_1 + \psi_2 \approx h^2\psi_{xx}(0,t),$$

and since g is, in general, not small, Eq. (12.7.6) is impossible. [If the solution happens to satisfy $u_{xx}(h,t) = 0$ the construction would be possible.]

Now consider the usual centered second-order approximation for inner points. As noted earlier, the linear extrapolation condition at $j = 0$ is equivalent to

$$\frac{dv_0}{dt} = D_+v_0, \qquad (12.7.7)$$

where the grid function index has been shifted one step, $j \to j-1$. By defining

$$Qv_j = \begin{cases} D_0v_j, & j = 1,2,\ldots, \\ D_+v_0, & j = 0, \end{cases}$$

we can write the approximation as

$$\begin{aligned} \frac{dv_j}{dt} &= Qv_j, \qquad j = 0,1,\ldots, \\ v_j(0) &= f_j \end{aligned} \qquad (12.7.8)$$

without any boundary condition. Because Q is only first-order accurate at $x = 0$, the error equation is

$$\begin{aligned} \frac{de_j}{dt} &= Qe_j + F_j, \qquad j = 0,1,\ldots, \\ e_j(0) &= 0, \end{aligned} \qquad (12.7.9)$$

where

$$F_j = \begin{cases} \mathcal{O}(h), & j = 0, \\ \mathcal{O}(h^2), & j = 1,2,\ldots, \end{cases}$$
$$\|F\|_{1,\infty} = \mathcal{O}(h^2).$$

By stability and Duhamel's principle, it follows that

$$\|e(t)\|_{0,\infty}^2 \leq \text{constant}\, \|F(t)\|_{0,\infty}^2 = \mathcal{O}(h^3), \qquad 0 \leq t \leq T.$$

However, this estimate is not optimal. Since the approximation satisfies the Kreiss condition, we can do better. Returning to the original formulation with an explicit boundary condition, the error equation is

$$\begin{aligned} \frac{de_j}{dt} &= D_0 e_j + h^2 F_j, \qquad j = 1,2,\ldots, \\ e_j(0) &= h^2 f_j, \\ e_0 - 2e_1 + e_2 &= h^2 g. \end{aligned} \tag{12.7.10}$$

We split the error into two parts $e = e^{(1)} + e^{(2)}$, where

$$\begin{aligned} \frac{de_j^{(1)}}{dt} &= D_0 e_j^{(1)} + h^2 F_j, \qquad j = 1,2,\ldots, \\ e_j^{(1)}(0) &= h^2 f_j, \\ e_0^{(1)} - 2e_1^{(1)} + e_2^{(1)} &= 0, \end{aligned} \tag{12.7.11}$$

$$\begin{aligned} \frac{de_j^{(2)}}{dt} &= D_0 e_j^{(2)}, \qquad j = 1,2,\ldots, \\ e_j^{(2)}(0) &= 0, \\ e_0^{(2)} - 2e_1^{(2)} + e_2^{(2)} &= h^2 g. \end{aligned} \tag{12.7.12}$$

This approximation was shown to be stable in Section 11.1, and we immediately get

$$\|e^{(1)}(t)\|_h \leq \text{constant}\, h^2,$$

where the norm is based on the scalar product

$$(v, w)_h = \frac{h}{2} v_1 w_1 + \sum_{j=2}^{\infty} v_j w_j h$$

for real grid functions v and w.

To estimate $e^{(2)}$, we Laplace transform Eq. (12.7.12), use the fact that the Kreiss condition is satisfied, and transform back again. Because D_0 is semibounded for the Cauchy problem, we use the same procedure as in Section 12.2 to obtain the estimate

$$\|e^{(2)}(t)\|_{1,\infty}^2 \leq \text{constant} \int_0^t |h^2 g(\tau)|^2 \, d\tau, \tag{12.7.13}$$

(see Theorem 12.2.2). This implies the final estimate

$$\|e(t)\|_h \leq \text{constant}\, h^2, \qquad 0 \leq t \leq T. \tag{12.7.14}$$

The technique we have used for deriving the optimal estimate (12.7.14) is similar to the one used to derive strong stability in Section 12.2. In fact, recalling that the approximation in our example is strongly stable, the result follows immediately from the error equation (12.7.10).

Considering the method in the form (12.7.8), the approximation of $\partial/\partial x$ is only first order accurate at $x = 0$. Still we have shown that there is an overall h^2 accuracy. This is possible because the lower order approximation is applied only at one point. This is the background for the expression "one order less accuracy at the boundary is allowed." This statement is only valid if it refers to "extra" boundary conditions. The "physical" boundary conditions must always be approximated to the same order as the differential operator at inner points.

Let us next consider the problem

$$u_t = A u_x + F, \qquad 0 \leq x < \infty, \quad t \geq 0, \tag{12.7.15a}$$

$$u(x,0) = f(x), \tag{12.7.15b}$$

$$u^{II}(0,t) = R^I u^I(0,t) + g(t), \tag{12.7.15c}$$

and the fourth-order approximation

$$\frac{dv_j}{dt} = A\left(\frac{4}{3} D_0(h) - \frac{1}{3} D_0(2h)\right) v_j + F_j, \qquad j = 1,2,\ldots, \tag{12.7.16a}$$

$$v_j(0) = f_j, \tag{12.7.16b}$$

where

$$A = \begin{bmatrix} \Lambda^I & 0 \\ 0 & \Lambda^{II} \end{bmatrix}, \qquad \Lambda^I > 0, \quad \Lambda^{II} < 0.$$

This is a generalization of the example in Section 12.3. By differentiating the boundary conditions (12.7.15c) twice with respect to t and using the differential equation (12.7.15a) we get

$$u_{xx}^{II}(0,t) = S^I u_{xx}^I(0,t) + \tilde{g}(t),$$

where

$$\begin{aligned} S^I &= (\Lambda^{II})^{-2} R^I (\Lambda^I)^2, \\ \tilde{g}(t) &= (\Lambda^{II})^{-2}(R^I(\Lambda^I F_x^I(0,t) + F_t^I(0,t)) - \Lambda^{II} F_x^{II}(0,t) - F_t^{II}(0,t) + g_{tt}(t)). \end{aligned} \tag{12.7.17}$$

As boundary conditions for the approximation, we use

$$v_0^{II}(t) = R^I v_0^I(t) + g(t), \tag{12.7.16c}$$
$$D_+D_-v_0^{II}(t) = S^I D_+D_-v_0^I(t) + \tilde{g}(t), \tag{12.7.16d}$$
$$D_+^4 v_0^I(t) = 0, \tag{12.7.16e}$$
$$D_+^4 v_{-1}^I(t) = 0. \tag{12.7.16f}$$

Let $u(x,t)$ be a smooth solution of Eq. (12.7.15), and consider the truncation error for the extra boundary conditions (12.7.16d,e,f). We have for $e = u - v$

$$\begin{aligned} D_+D_-e_0^{II} &= S^I D_+D_-e_0^I + \mathcal{O}(h^2), \\ D_+^4 e_0^I &= \mathcal{O}(1), \\ D_+^4 e_{-1}^I &= \mathcal{O}(1). \end{aligned}$$

However, the normalized form corresponding to (12.1.12) is

$$\begin{aligned} h^2 D_+D_-e_0^{II}(t) &= S^I h^2 D_+D_-e_0^I(t) + \mathcal{O}(h^4), \\ (hD_+)^4 e_0^I(t) &= \mathcal{O}(h^4), \\ (hD_+)^4 e_{-1}^I(t) &= \mathcal{O}(h^4), \end{aligned} \tag{12.7.18}$$

which yields the right order of truncation error. The other error equations are

$$\begin{aligned} \frac{de_j(t)}{dt} &= A\left(\tfrac{4}{3} D_0(h) - \tfrac{1}{3} D_0(2h)\right) e_j(t) + \mathcal{O}(h^4), \qquad j = 1,2,\ldots, \\ e_j(0) &= 0, \\ e_0^{II}(t) &= R^I e_0^I(t). \end{aligned} \tag{12.7.19}$$

The approximations (12.7.18) and (12.7.19) were analyzed in Section 12.3 and shown to be strongly stable. Thus, we get the error estimate

$$\|e(t)\|_h \le \text{constant}\, h^4. \tag{12.7.20}$$

Note that Eq. (12.7.16d) is only a second-order approximation of Eq. (12.7.17), and we still have an h^4 error in the solution. This is possible, since it is an extra numerical boundary condition. The condition (12.7.17) is also "extra" in the sense that is is not required for defining a unique solution of the differential equation.

In summary, when the extra boundary conditions are written in the normalized form as in Eq. (12.1.12), the error term must be of the same order as the truncation error at inner points.

Now consider the case where the approximation is strongly stable in the generalized sense. If there is no error in the initial data, then the error estimate follows immediately from Eq. (12.4.1b). For Eq. (12.7.1), we get

$$\begin{aligned} \int_0^\infty e^{-2\eta t} \|e(t)\|_h^2\, dt &\le K(\eta) \int_0^\infty e^{-2\eta t} (h^{2q_1} \|F(t)\|_h^2 + h^{2q_3} |g(t)|^2)\, dt, \\ &= \mathcal{O}(h^{2q_1} + h^{2q_3}). \end{aligned} \tag{12.7.21}$$

Assume that there is an initial error

$$e_j(0) = h^{q_2} f_j,$$

where f_j is smooth; that is,

$$|D_+^\nu f_j| \le \text{constant}, \qquad \nu = 0,1,\ldots.$$

$$\tilde{e}_j(t) = e_j(t) - h^{q_1}\varphi_j(t), \qquad j = 1,2,\ldots. \tag{12.7.22}$$

We obtain

$$\begin{aligned} \frac{d\tilde{e}_j}{dt} &= Q\tilde{e}_j + h^{q_1}F_j + h^{q_2}\tilde{F}_j, \qquad j = 1,2,\ldots, \\ \tilde{e}_j(0) &= 0, \\ L_0\tilde{e}_0(t) &= h^{q_3}g(t) + h^{q_2}\tilde{g}(t), \end{aligned} \tag{12.7.23}$$

which yields the estimate

$$\begin{aligned} \int_0^\infty e^{-2\eta t}\|e(t)\|_h^2\,dt &\le 2\int_0^\infty e^{-2\eta t}(\|\tilde{e}(t)\|_h^2 + h^{2q_2}\|\varphi(t)\|_h^2)\,dt, \\ &= \mathcal{O}(h^{2q_1} + h^{2q_2} + h^{2q_3}). \end{aligned} \tag{12.7.24}$$

Note that we don't have the same difficulty when making the initial condition homogeneous as we have in some cases when making the boundary conditions homogeneous. The only requirement is that f_j be smooth.

Next assume that the approximation is stable and that the Kreiss condition is not satisfied. Then the error estimate does not follow directly from the stability estimate because it does not permit nonzero boundary data. The procedure of splitting the error into $e = e^{(1)} + e^{(2)}$, as demonstrated above, cannot be used either, because we need the Kreiss condition to estimate $e^{(2)}$. One alternative is to subtract a suitable function that satisfies the inhomogeneous boundary condition. Another alternative is to eliminate the boundary values $v_{-r+1}, v_{-r+2}, \ldots, v_0$ and modify the difference operator Q near the boundary. However, as we have already seen above, we may lose accuracy in this process.

Finally, we consider the case where the approximation is stable in the generalized sense and the Kreiss condition is not satisfied. Now we must construct a function that satisfies both the inhomogeneous initial and boundary conditions. This may be tricky, because we cannot in general expect compatibility at the corner $x = 0$, $t = 0$, even if the solution $u(x,t)$ is smooth. The reason for this is that the truncation error in the boundary conditions is different from the truncation error in the initial condition, the latter one typically being zero. And even if such a function exists, we may lose accuracy in the subtraction process, just as in the previous case.

The most general theory based on the Laplace transform method for obtaining optimal error estimates in this case is given in Gustafsson (1981). The essential condition is that there be no eigenvalue or generalized eigenvalue at $s = 0$. (The theory is given for the fully discrete case where $s = 0$ corresponds to $z = 1$.)

EXERCISES

12.7.1. Prove that the error $\|v_j(t) - u(x_j,t)\|_h$ of the approximation (11.4.3) and (11.4.5) is $\mathcal{O}(h^4)$.

12.7.2. Use the result of Exercise 12.6.4 to prove that the approximation (12.6.4) and (12.6.23) gives a fourth-order accurate solution.

BIBLIOGRAPHIC NOTES

The first general stability theory for semidiscrete approximations based on the Laplace transform technique was given by Strikwerda (1980). The stability concept there corresponds to strong stability in the generalized sense for hyperbolic problems. Strikwerda also proves stability for a fourth-order approximation of $u_t = \pm u_x$, but with different boundary conditions than ours.

The method of lines has been used extensively in applications, for example in fluid dynamics. However, very little analysis has been done. In Gustafsson and Kreiss (1983) and Johansson (1993), semidiscrete approximations of model problems corresponding to incompressible flow are analyzed.

Gustafsson and Oliger (1982) derived stable boundary conditions for a number of implicit time discretizations of a centered second-order in space approximation of the Euler equations.

Regarding optimal error estimates for approximations that are stable in the generalized sense, see Notes on Chapter 13.