Stability Theory for Difference Approximations of Mixed Initial Boundary Value Problems. I*

By Heinz-Otto Kreiss

0. Introduction. Consider a first-order hyperbolic system of partial differential equations

$$(0.1) \qquad \partial u/\partial t = A\,\partial u/\partial x$$

with constant coefficients in the quarter space, $x \geqq 0$, $t \geqq 0$. Here $u(x,t)' = (u^{(1)}(x,\,t),\ \cdots,\ u^{(n)}(x,\,t))$** is a vector function of the real variables $(x,\,t)$, and A is a constant matrix of order n. Without restriction we can assume that A has diagonal form. Furthermore we make the

Assumption 1. A is nonsingular. The unknowns $u^{(\nu)}(x,\,t)$ are ordered so that A has the form

$$(0.2)\ \ A = \begin{pmatrix} A^{\mathrm{I}} & 0 \\ 0 & A^{\mathrm{II}} \end{pmatrix}, \quad A^{\mathrm{I}} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & 0 \cdots 0 \\ \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\ 0 & \cdots & 0 & a_l \end{pmatrix}, \quad A^{\mathrm{II}} = \begin{pmatrix} a_{l+1} & 0 & \cdots & 0 \\ 0 & a_{l+2} & 0 \cdots 0 \\ \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\ 0 & 0 \cdots 0 & a_n \end{pmatrix}$$

with

$$(0.3) \qquad a_1 \leqq a_2 \leqq \cdots \leqq a_l < 0 < a_{l+1} \leqq \cdots \leqq a_n \,.$$

We will study difference methods for solving initial boundary value problems for (0.1). Therefore we suppose that initial values $u(x,\,0) = f(x)$ and boundary conditions

$$(0.4) \qquad\qquad u^{\mathrm{I}}(0,\,t) = S u^{\mathrm{II}}(0,\,t)$$

are given. u^{I} and u^{II} are defined according to the partition of A, i.e. $u^{\mathrm{I}} = (u^{(1)},\ \cdots,\ u^{(l)})'$, $u^{\mathrm{II}} = (u^{(l+1)},\ \cdots,\ u^{(n)})'$, and S is a given constant rectangular matrix.

It is well known that the above problem is correctly posed in L_2 (see for example Thomée [4]). The present treatment of the case when A is a constant matrix can be extended, as in [1], to the case when A depends on $(x,\,t)$ in a sufficiently smooth fashion.

In the earlier paper [1], we considered the case when the coefficient matrices of the difference schemes were diagonal. The same class of problems has also been treated in an interesting paper by Osher [2]. The assumption of diagonality would

** If y is a vector then y' denotes its transpose, y^* its adjoint (conjugate complex transpose), and $|y| = (\sum |y_i|^2)^{1/2}$ its Euclidean norm. Similar notations hold for matrices, for example $|A| = \sup |Ay|/|y|,\ y \neq 0$.

be adequate for practical purposes, if we were interested only in the one-dimensional case. In a forthcoming paper we will consider approximations for hermitian systems of partial differential equations

$$\partial u/\partial t = A\,\partial u/\partial x_1 + \sum_{j=2}^{m} B_j \partial u/\partial x_j = P(\partial/\partial x)u$$

with constant coefficients in the quarter space $x_1 \geqq 0$, $-\infty < x_j < \infty$, $j = 2, 3, \cdots, m$; $t \geqq 0$. Fourier transformation of the difference approximation with respect to the "tangential-variables" $x_2, \cdots, x_m$ will lead us to difference equations of the kind discussed in this paper, where the coefficient matrices depend on $m - 1$ parameters $\xi_2, \cdots, \xi_m$ and are nondiagonal matrices. We will establish the stability of such difference schemes by making all our estimates uniform with respect to $\xi_2, \cdots, \xi_m$.

1. Statement of Results. We want to solve the initial boundary value problem by using a difference approximation. We therefore introduce a time-step $k > 0$ and a mesh-width $h > 0$ and divide the x-axis into intervals of length h. As usual we assume that $k/h = \lambda = $ const. Using the notation $x_\nu = \nu h$, $v_\nu(t) = v(x_\nu, t)$, we approximate the differential equation for $x > 0$ by a consistent difference scheme

$$(1.1) \qquad \begin{aligned} v_\nu(t + k) &= Qv_\nu(t)\,, & \nu = 1, 2, \cdots, \\ v_\nu(0) &= f_\nu\,, \end{aligned}$$

where

$$(1.2) \qquad Q = \sum_{j=-r}^{p} A_j E^j\,, \qquad Ev_\nu = v_{\nu+1}\,,$$

and the A_j are constant matrices of order n.

For convenience only we make

Assumption 2. $p \geqq 1$ and A_p, A_r are nonsingular.

Let

$$(1.3) \qquad \hat{Q}(\xi) = \sum_{j=-r}^{p} A_j e^{ij\xi}\,, \qquad \xi \text{ real}\,,$$

denote the symbol (or amplification matrix) of Q. We now require that $\hat{Q}(\xi)$ satisfies the following two assumptions:

Assumption 3. There exist a constant $\delta > 0$ and a natural number $2s > 0$ such that for all ξ with $0 \leqq |\xi| \leqq \pi$ the eigenvalues $\mu(\xi)$ of $\hat{Q}(\xi)$ satisfy the estimate

$$(1.4) \qquad |\mu(\xi)| \leqq 1 - \delta|\xi|^{2s}\,,$$

i.e. the approximation is dissipative.

Assumption 4. $|\hat{Q}(\xi)| \leqq 1$.

Assumption 4 guarantees that if the scheme (1.1) were used to solve the pure initial-value problem (i.e. for $-\infty < \nu < +\infty$) then the scheme would be stable. Stability for the pure initial value problem is obviously a necessary condition for the stability of the mixed problem. Assumption 3 is introduced to make sure that high frequencies (i.e. ξ bounded away from zero) have no influence.

From the nature of Q, defined by (1.2), we see that the solution of (1.1) can be carried out only if we specify boundary conditions to eliminate the values of v_ν at $\nu = 0, -1, \cdots, -r + 1$. These shall be of the type

$$(1.5) \qquad v_\mu(t) = \sum_{j=1}^{s} C_{j\mu} v_j(t) , \qquad \mu = 0, -1, \cdots, -r + 1 ,$$

where the $C_{j\mu}$ are constant square matrices of order n.

Our aim is to derive algebraic stability conditions. Let us denote by H the space of all grid-functions w_ν, defined for $\nu > -r$, which fulfill the boundary conditions (1.5) and for which

$$\sum_{\nu=1}^{\infty} |w_\nu|^2 h < \infty .$$

H is a Hilbert space if norm and scalar product are defined by

$$(u, v)_h = \sum_{\nu=1}^{\infty} u_\nu^* v_\nu h , \qquad \|u\|_h^2 = (u, u)_h .$$

(Observe that $u \in H$, $\|u\|_h = 0$ implies that $0 = u_{-r+1} = u_{-r+2} = \cdots = u_0 = u_1 \cdots$.) Stability is defined in the usual way:

Definition 1. The difference approximation is stable if there is a constant K, independent of k, such that

$$\|v(t)\|_h \leq K \|v(0)\|_h$$

for all $t = mk$ and all initial values $v(x, 0)$ in H.

We may write the difference approximation in operator form

$$(1.6) \qquad v(t + k) = \mathfrak{G}v(t) , \qquad v(t), v(t + k) \in H ,$$

where $\mathfrak{G}$ is a bounded operator in H defined by (1.1) and (1.2). It is now easy to derive a necessary stability condition:

LEMMA 1. *A necessary condition for stability is that $\mathfrak{G}$ has no eigenvalues z_0 with* $|z_0| > 1$.

Proof. Assume that $\mathfrak{G}$ has an eigenvalue z_0 with $|z_0| > 1$, i.e. there is a $g \in H$ with

$$(1.7) \qquad z_0 g = \mathfrak{G}g , \qquad g \in H .$$

Then $z_0^{t/k} g$ is a solution of (1.6) with initial values belonging to H and which increases exponentially with the number of time-steps.

Our sufficient conditions for stability will be stated in terms of the eigenvalues and (generalized) eigenvectors of $\mathfrak{G}$. We therefore have to understand how one determines the eigenvalues and eigenfunctions of $\mathfrak{G}$. We start with

LEMMA 2. *Consider the (generalized) matrix eigenvalue problem*

$$\sum_{j=-r}^{p} A_j \kappa^j \phi = z\phi ,$$

and let Assumption 3 be fulfilled. Then for every z with $|z| \geq 1$, $z \neq 1$ there are precisely rn eigenvalues κ_j with $|\kappa_j| < 1$ and np eigenvalues with $|\kappa_j| > 1$, i.e. there are no eigenvalues with $|\kappa_j| = 1$.

Proof. The eigenvalues are the solutions of the characteristic equation

$$(1.8) \qquad \text{Det} \left| \sum_{j=-r}^{p} A_j \kappa^j - zI \right| = 0 .$$

If $\kappa = e^{i\xi} \neq 1$ then $\sum_{j=-r}^{p} A_j \kappa^j = \hat{Q}(\xi)$ and by Assumption 3 the eigenvalues $\mu(\xi)$ of $\hat{Q}(\xi)$ cannot be equal to z, with $|z| \geq 1$. By consistency $\sum_{j=-r}^{p} A_j = I$, and therefore $\kappa = 1$ is ruled out for $z \neq 1$. Since the eigenvalues are continuous functions of z, we can determine the number of eigenvalues κ_j with $|\kappa_j| < 1$ by considering large values of $|z|$. For $|z| \to \infty$ these eigenvalues converge to zero and therefore the leading term of $\sum A_j \kappa^j$ is $A_{-r} \kappa^{-r}$. By assumption, A_{-r} is nonsingular and the lemma follows without difficulties.

Let z_0 with $|z_0| \geq 1$, $z_0 \neq 1$ be a fixed value. We want to determine whether (1.7) has a corresponding eigensolution $g \in H$. Equation (1.7) is equivalent to

$$(1.9) \qquad (Q - z_0)g_\nu = 0 , \qquad \nu = 1, 2, \cdots ,$$

$$(1.10) \qquad g_\mu = \sum_{j=1}^{s} C_{j\mu} g_j , \qquad \mu = 0, -1, \cdots, -r+1 .$$

Equation (1.9) is an ordinary difference equation with constant coefficients and its most general solution in H can be written in the form:

$$(1.11) \qquad g_\nu = g_\nu(z_0) = \sum_{|\kappa_j|<1} P_j(\nu)\kappa_j^{\nu} = \sum_{|\kappa_j|<1} P_j(\nu, z_0)(\kappa_j(z_0))^{\nu} .$$

Here κ_j are the solution of (1.8) with $|\kappa_j| < 1$ and $P_j(\nu)$ are polynomials in ν with vector coefficients. The degree of $P_j(\nu)$ is one less than the multiplicity of the corresponding κ_j. There are precisely nr such linearly independent solutions, thus (1.11) depends on nr parameters σ_j, $j = 1, 2, \cdots, nr$. (This is obvious if all κ_j are different. Then $g_\nu = \sum \sigma_j P_j \kappa_j^\nu$ where P_j are constant vectors. In the same way as in the scalar case the general formula (1.11) can then be obtained by a continuity argument.)

Inserting (1.11) into the boundary conditions (1.10) we get a system of nr linear homogeneous equations in the nr parameters σ_j which we may write as

$$(1.12) \qquad E(z_0)\sigma = 0 , \qquad \sigma = (\sigma_1, \cdots, \sigma_{nr})' ,$$

where $E(z_0)$ is a matrix of order nr. This gives us

LEMMA 3. z_0 *with* $|z_0| \leq 1$, $z_0 \neq 1$ *is an eigenvalue of* $\mathfrak{S}$ *if and only if* Det $E(z_0) = 0$.

Now consider (1.11) for $z_0 \to 1$. The general solution given by (1.11) converges to a solution

$$(1.13) \qquad g_\nu(1) = \sum P_j(\nu, 1)(\kappa_j(1))^\nu$$

which again depends on nr parameters σ_j. However, in general, $g_\nu(1)$ does not belong to H because (as we shall show) there are precisely l indices j with $\kappa_j(1) = 1$. Inserting (1.13) into the boundary conditions (1.10) the parameters σ_j are determined by a linear system of equations:

$$(1.12') \qquad E(1)\sigma = 0 .$$

This leads us to

Definition 2. $z_0 = 1$ is a generalized eigenvalue of $\mathfrak{G}$ if (1.12′) has a nontrivial solution, i.e. if Det $E(1) = 0$. The corresponding solution (1.13) is called a generalized eigenfunction.

We can now formulate our

MAIN THEOREM (THEOREM 0). *The approximation to the initial boundary value problem is stable if*

(1) *the Assumptions 1–4 are fulfilled.*

(2) $z_0 = 1$ *is not a generalized eigenvalue of* $\mathfrak{G}$.

(3) $\mathfrak{G}$ *has no eigenvalue z_0 with $|z_0| \geqq 1$, $z_0 \neq 1$.*

A proof of this theorem is given in Section 3.

Remark. In [1] we have treated the case that the A_j are diagonal matrices. In that case the above conditions are equivalent with those assumed there. For examples, we refer to [1].

Conditions (1) and (2) of the Main Theorem are relatively easy to check by analytic methods. To decide whether (3) is also satisfied requires further study. We shall show that if $z_0 = 1$ is not a generalized eigenvalue of $\mathfrak{G}$, then there is a constant $\rho > 0$ such that $\mathfrak{G}$ has no eigenvalues z_0 with $|z_0 - 1| \geqq \rho$ and $|z_0| \geqq 1$. However it is often very difficult to find out by analytic methods whether a value of z, bounded away from 1, is an eigenvalue or not. To decide this matter, we consider the following reduced problem:

$$(1.14) \qquad u_\nu(t + k) = Q u_\nu(t) , \qquad \nu = 1, 2, \cdots, N - 1 ,$$

with boundary conditions

$$(1.15) \qquad u_\mu(t) = \sum C_{j\mu} u_j(t) , \qquad \mu = 0, -1, \cdots, -r + 1 ,$$

$$(1.16) \qquad u_N(t) = u_{N+1}(t) = \cdots = u_{N+p-1}(t) = 0 ,$$

with N some sufficiently large natural number.

We write this scheme as an operator equation:

$$u(t + k) = \mathfrak{G}_N u(t)$$

where $\mathfrak{G}_N$ is a matrix of order $n(N - 1)$. We shall prove the following theorem:

THEOREM 2. *For every $\rho > 0$ there exist constants $\epsilon > 0$; $K_j > 0$, $j = 1, 2$; and τ with $0 < \tau < 1$ all independent of h such that:*

(1) *If $\mathfrak{G}$ has no eigenvalue with $|z| \geqq 1$, $|z - 1| \geqq \rho$ then $\mathfrak{G}_N$ has no eigenvalue for $|z| \geqq 1 - \epsilon$, $|z - 1| \geqq \rho$ provided $N \geqq K_1 |\log (1 - \tau)|$.*

(2) *If $\mathfrak{G}$ has an eigenvalue z_0 of multiplicity p with $|z_0| \geqq 1 - \epsilon$, $|z_0 - 1| \geqq \rho$ then $\mathfrak{G}_N$ has an eigenvalue λ_N with $|\lambda_N - z_0| \leqq K_2 \tau^{N/p}$.*

A proof of Theorem 2 is given in Section 4.

With the help of Theorem 2, one can detect eigenvalues of $\mathfrak{G}$ by computing the eigenvalues of $\mathfrak{G}_N$ for an increasing sequence of values of N. If any eigenvalue λ_N of $\mathfrak{G}_N$ converges to a value z_0 with $|z_0| \geqq 1$, $z_0 \neq 1$ then z_0 is an eigenvalue of $\mathfrak{G}$.

2. The Resolvent of $\mathfrak{G}$. It is well known (see for example Kato [3]) that we can write the solution of (1.6) in the form:

$$(2.1) \qquad v(t) = - \frac{1}{2\pi i} \oint_\Gamma z^n (\mathfrak{G} - zI)^{-1} dz\, v(0) , \qquad t = nk ,$$

where Γ is any contour (for example $|z| = |\mathfrak{G}| + 1$) which includes the spectrum of $\mathfrak{G}$ in its interior. We are therefore going to study $(\mathfrak{G} - zI)^{-1}$, the resolvent of $\mathfrak{G}$. Let us explicitly compute

$$f = (\mathfrak{G} - zI)^{-1}v$$

which is equivalent to finding the solution of

(2.2) $(Q - zI)f_\nu = v_\nu , \qquad \nu = 1, 2, \cdots ,$

(2.3) $f \in H ,$

i.e. f fulfills the boundary conditions (1.5).

By using Assumption 2, we write (2.2) as

$$f_{\nu+p} = -A_r^{-1}\left(\sum_{j=-r}^{p-1} A_j f_{\nu+j} - zf_\nu - v_\nu\right).$$

Upon introducing the vector

$$y_\nu = (f_{\nu+p-1}, f_{\nu+p-2}, \cdots , f_{\nu-r})'$$

we get a one-step formula

(2.4) $y_{\nu+1} = My_\nu + g_\nu , \qquad \nu = 1, 2, 3, \cdots ,$

where

$$M = -\begin{pmatrix} A_p^{-1} A_{p-1} & \cdots & A_p^{-1}(A_0 - zI) & \cdots & A_p^{-1} A_{-r} \\ -I & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & -I & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & -I & 0 \end{pmatrix}, \quad g_\nu = A_p^{-1}\begin{pmatrix} v_\nu \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}.$$

Furthermore the boundary conditions (2.3) can be written as nr linear relations between the components of $y_1, \cdots , y_{s+1}$ and we write them formally as

(2.5) $L_j y = 0 , \qquad j = 1, 2, \cdots , rn .$

The eigenvalues κ determined by the characteristic equation (1.8) are also the eigenvalues of M and the eigenvectors of M have the form $(\kappa^{p+r-1}\phi, \ \kappa^{p+r-2}\phi, \cdots , \phi)'$ where ϕ is a solution of

(2.6) $\sum A_j \kappa^j \phi = z\phi .$

The eigenvalues $\kappa_j = \kappa_j(z)$ of M are functions of z. For $|z| \geq 1$, $z \neq 1$ they form by Lemma 2 two separate sets: S_1, S_2, where S_1 contains the κ_j with $|\kappa_j| < 1$ and S_2 the κ_j with $|\kappa_j| > 1$. Therefore we may use the following well-known lemma (see Kato [3, Chapter 2]).

LEMMA 4. *There exist a nonsingular matrix $T = T(z)$ which is analytic in z for $|z| \geq 1$, $z \neq 1$ such that*

(2.7) $T(z)MT^{-1}(z) = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix}.$

Here M_{11} is of order nr and has eigenvalues $|\kappa_j| < 1$ and M_{22} is of order np and has eigenvalues $|\kappa_j| > 1$.

Introduce now in (2.4)

(2.8) $$w_\nu = T(z)y_\nu$$

as a new variable. Then we get

(2.9) $$w_{\nu+1} = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix} w_\nu + Tg_\nu$$

and its general solution w_ν with $||w||_h < \infty$ is given by

(2.10)
$$w_j{}^{\mathrm{I}} = \sum_{\nu=1}^{j-1} M_{11}^{j-\nu-1}(Tg_\nu)^{\mathrm{I}} + M_{11}^{j-1}w_1{}^{\mathrm{I}}, \qquad j = 1, 2, 3, \cdots.$$
$$w_j{}^{\mathrm{II}} = -\sum_{\nu=j}^{\infty} M_{22}^{j-\nu-1}(Tg_\nu)^{\mathrm{II}},$$

w^{I}, w^{II} are defined according to the partition of TMT^{-1}.

By introducing (2.8) into the boundary conditions (1.5) and eliminating $w_j{}^{\mathrm{I}}$, $j = 2, 3, \cdots, s$, with help of (2.10) we get

(2.11) $$C_1(z)w_1{}^{\mathrm{I}} = \sum_{j=1}^{s+1} F_j(z)w_j{}^{\mathrm{II}} + G_j(z)g_j.$$

Here $C_1(z)$, $F_j(z)$, $G_j(z)$ are analytic for $|z| \geqq 1$, $z \neq 1$. By using the representation (2.8)–(2.11) of $(\mathfrak{G} - zI)^{-1}$ we will establish

LEMMA 5. *Let* $\mathfrak{N}$ *denote the set of complex numbers* z *with* $|z| \geqq 1$, $z \neq 1$. *If* $z_0 \in \mathfrak{N}$ *and* z_0 *is not an eigenvalue of* $\mathfrak{G}$, *then* $(\mathfrak{G} - z_0I)^{-1}$ *exists in all of* H *and is bounded.*

Proof. If $C_1(z_0)^{-1}$ exists, then w is uniquely determined by (2.10) and (2.11) and by Lemma 4 the resolvent $(\mathfrak{G} - z_0I)^{-1}$ exists in all of H and is bounded. If $C_1(z_0)^{-1}$ does not exist then the set of equations $C_1(z_0)w_1{}^{\mathrm{I}} = 0$ has a nontrivial solution $\tilde{w}_1{}^{\mathrm{I}}$ and

$$w_j{}^{\mathrm{II}} = 0, \qquad w_j{}^{\mathrm{I}} = M_{11}^j \tilde{w}_1{}^{\mathrm{I}},$$

is a solution of (2.10) with $Tg \equiv 0$. Therefore the homogeneous equation (2.9) has a nontrivial solution which by Lemma 4 defines a solution of the homogeneous equations (2.2), (2.3), and z_0 is an eigenvalue of $\mathfrak{G}$. From Lemma 5 we get immediately

LEMMA 6. *Assume that* $\mathfrak{G}$ *has no eigenvalue* $z \in \mathfrak{N}$; *then for any constant* $\rho > 0$ *there exists an* $\epsilon > 0$ *such that* $|(\mathfrak{G} - zI)^{-1}|$ *is uniformly bounded for* $|z| \geqq 1 - \epsilon$, $|z - 1| \geqq \rho$.

Proof. If $(\mathfrak{G} - z_0I)^{-1}$ exists and is bounded then $(\mathfrak{G} - zI)^{-1}$ exists and is bounded in a whole neighborhood of z_0.

3. Proof of the Main Theorem. We start with a variant of Lemma 4.

LEMMA 7. *There exist a* $\rho > 0$ *and a nonsingular matrix* $T(z)$ *which is analytic for* $|z - 1| \leqq \rho$ *such that*

(3.1) $$T(z)MT^{-1}(z) = \begin{pmatrix} M_{11}^{\hat{}} & 0 \\ 0 & M_{22}^{\hat{}} \end{pmatrix}$$

where the $M_{ii}^{\hat{}}$ *have, for* $|z| > 1$, *the properties of the* M_{ii} *of Lemma 4.*

Proof. As we know, the eigenvalues κ of M are the solutions of the eigenvalue problem

$$(3.2) \qquad \sum_{j=-r}^{p} A_j \kappa^j \phi = z\phi .$$

Consistency implies that this equation can be written in the form

$$(3.3) \qquad (\lambda A(\kappa - 1) + (\kappa - 1)^2 D((\kappa - 1)))\phi = (z - 1)\phi , \qquad \lambda = k/h ,$$

where D is a rational function in $(\kappa - 1)$ with matrix coefficients. Therefore there are precisely n eigenvalues κ_j with $\lim \kappa_j = 1$ for $z \to 1$. Introduce a new variable $\mu = (\kappa - 1)/(z - 1)$. Then (3.3) becomes

$$(\lambda A \mu + \mu^2 (z - 1) D(\mu(z - 1)))\phi = \phi$$

which implies that the eigenvalues of the last equation are

$$(3.4) \qquad \mu_j = (\lambda a_j)^{-1} + O((z - 1)^{1/n})$$

and therefore

$$(3.5) \qquad \kappa_j = 1 + (z - 1)\mu_j = 1 + (z - 1)(\lambda a_j)^{-1} + O((z - 1)^{1+1/n}) .$$

Observing that $(\lambda a_j)^{-1} < 0$ for $j = 1, 2, \cdots, l$ we get for $z = 1 + \delta, \delta > 0$ and sufficiently small, that

$$(3.6a) \qquad |\kappa_j| < 1 \quad \text{for } j = 1, 2, \cdots, l$$

and correspondingly

$$(3.6b) \qquad |\kappa_j| > 1 \quad \text{for } j = l + 1, \cdots, n .$$

By Lemma 2 these inequalities hold for all z with $|z| > 1$.

Now, consider M in a neighborhood of $z = 1$. There is certainly an analytic transformation $T_1(z)$ which transforms M into

$$T_1(z) M T_1(z)^{-1} = \begin{pmatrix} M_{11} & 0 & 0 \\ 0 & M_{22} & 0 \\ 0 & 0 & M_{33} \end{pmatrix},$$

where the eigenvalues of M_{11} (M_{33}) are strictly smaller (larger) than 1 in absolute value; M_{22} is of order n and all its eigenvalues converge to 1 for $z \to 1$, i.e. they are given by (3.5). Furthermore $M_{22}(1) = I$ because for $z = 1$ and $\kappa = 1$ the Eq. (3.2) has n linearly independent eigenvectors. (Observe that consistency implies that $\sum A_j = I$.) Therefore we can write the eigenvalue problem for M_{22} in the form

$$(3.7) \quad \tilde{M}_{22}\psi = \mu\psi, \text{ with } \mu = (\kappa - 1)/(z - 1) , \qquad \tilde{M}_{22} = (M_{22} - I)/(z - 1)$$

and the eigenvalues of $\tilde{M}_{22}$ are given by (3.4). These eigenvalues therefore split into two separate groups: $\{\mu_j \text{ with } j = 1, 2, \cdots, l\}$ and $\{\mu_j \text{ with } j = l + 1, \cdots, n\}$. Hence there exists in a neighborhood of $z = 1$ an analytic transformation T_2 which transforms $\tilde{M}_{22}$ into block diagonal form and which separates these two sets.

Since M_{11} is of order $a = nr - l$ and M_{33} is of order $b = np - n + l$, we define

$$T_3 = \begin{pmatrix} I_a & 0 & 0 \\ 0 & T_2 & 0 \\ 0 & 0 & I_b \end{pmatrix},$$

where I_a is a unit matrix of order a. We now set $T = T_3 T_1$ and observe that the proof is complete.

Assume that $\mathfrak{G}$ has no eigenvalue for $|z| \geqq 1$, $z \neq 1$ then by Lemma 6 we get the solution of the difference equation in the form

$$v(t) = \mathfrak{G}^n v(0) = -\frac{1}{2\pi i} \int_{\Gamma_1 + \Gamma_2} z^n (\mathfrak{G} - zI)^{-1} v(0) dz$$

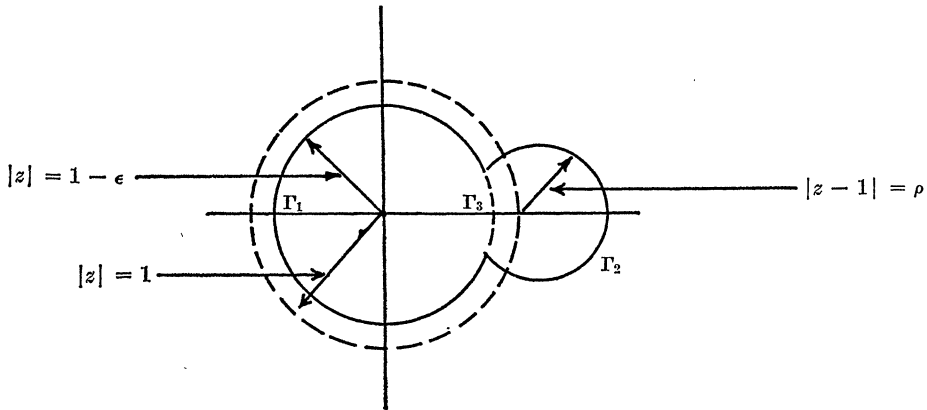if we choose the contour as shown in the following figure:

$|z| = 1 - \epsilon$ Γ_1 Γ_3 $|z - 1| = \rho$

$|z| = 1$ Γ_2

FIGURE 1

Again by Lemma 6 there exists a constant K (depending on ρ only), such that

$$\left\| -\frac{1}{2\pi i} \int_{\Gamma_1} z^n (\mathfrak{G} - zI)^{-1} v(0) dz \right\|_h \leqq K(1 - \epsilon)^n \|v(0)\|_h .$$

Therefore we have only to estimate the integral over Γ_2. Consider now initial values which are different from zero only in a finite number of points near the boundary, i.e. there is a natural number q such that

(3.8) $$|v_\nu(0)| = 0 \quad \text{for } \nu > q .$$

By Lemma 7 we can use the representation of $(\mathfrak{G} - zI)^{-1}$ derived in the preceding section in a neighborhood of $z = 1$. We find

$$w_j{}^{\mathrm{I}} = \sum_{\nu=1}^{j-1} M_{11}^{j-\nu-1}(Tg_\nu)^{\mathrm{I}} + M_{11}^{j-1} w_1{}^{\mathrm{I}}, \quad j = 1, 2, \cdots,$$

(3.9) $$w_j{}^{\mathrm{II}} = -\sum_{\nu=q}^{j} M_{22}^{j-\nu-1}(Tg_\nu)^{\mathrm{II}} \quad \text{for } j \leqq q ,$$

$$w_j{}^{\mathrm{II}} = 0 \quad \text{for } j > q ,$$

where $w_1{}^{\mathrm{I}}$ is determined by the boundary conditions (2.11) when $C_1(z)$ has an inverse. Assume now that $z = 1$ is not a generalized eigenvalue of $\mathfrak{G}$. Then $C_1(z)^{-1}$ exists in a neighborhood of $z = 1$. Therefore for every fixed j: $((\mathfrak{G} - zI)^{-1} v(0))_j$ is analytic in a neighborhood of $z = 1$. This is so because the representation (3.8)

consists, for every fixed j, of only a finite number of terms and all functions and transformations are analytic. Therefore we get (see Fig. 1)

$$\left| \int_{\Gamma_2} z^n ((\mathfrak{G} - zI)^{-1} v(0))_j \, dz \right| = \left| \int_{\Gamma_3} z^n ((\mathfrak{G} - zI)^{-1} v(0))_j \, dz \right|$$

$$\leq \text{const } (1 - \epsilon)^n \sum_{r=1}^{q-1} |v_\nu(0)| .$$

For initial values satisfying (3.8) we get the estimate

$$(3.10) \qquad\qquad |v_j(t)| \leq K_j (1 - \epsilon)^n \sum_{r=1}^{q-1} |v_\nu(0)| , \qquad t = nk ,$$

i.e. the solution decreases exponentially near the boundary.

It should be pointed out that the estimate (3.10) does not depend on Assumption 4. We have therefore proved

THEOREM 1. *If the Assumptions $1-3$ are fulfilled, then the estimate* (3.10) *holds for all initial values of type* (3.8).

In Theorem 3 of [1],* we have shown that a difference scheme, which

(a) has the property described in Theorem 1 (i.e. estimate (3.10) holds for special initial value problems) and which

(b) also satisfies Assumption 4, must be stable for the initial boundary value problem given in (1.1), (1.2) and (1.5).

This completes the proof of the Main Theorem.

In a later paper we shall generalize the above theorem of [1] by weakening Assumption 4. This enables us to get more general stability theorems for the initial boundary value problem.

4. Proof of Theorem 2.

We start with

LEMMA 8. *Consider the difference approximation* (1.5) *with boundary conditions*

$$v_\mu(t) = 0 , \qquad \mu = 0, -1, \cdots, -r + 1 ,$$

and assume that the Assumptions 1–4 are fulfilled. Then for every $\rho > 0$ there exists an $\epsilon > 0$ such that the corresponding operator $\mathfrak{G}_1$ has no eigenvalue z with $|z| \geq 1 - \epsilon$, $|z - 1| > \rho$.

Proof. Assume there is an eigenvalue z with $|z| \geq 1$, $z \neq 1$ and denote by g the corresponding eigenfunction. Consider now the Cauchy problem for (1.1) with initial values

$$v_\nu(0) = g_\nu \quad \text{for } \nu > 0 ,$$

$$= 0 \quad \text{for } \nu \leq 0 .$$

Let ν_0 denote the smallest index for which $g_\nu \neq 0$, i.e.

$$g_\nu = 0 \quad \text{for } \nu < \nu_0 , \qquad g_{\nu_0} \neq 0 .$$

By remembering that A_p^{-1} exists, we find, $\nu_0 \leq p$. Compute now one step of the Cauchy problem, then there is certainly a $v_\nu(k) \neq 0$ for $\nu \leq 0$. Therefore, we get from Assumption 4 and the observation that $v_\nu(k) = zg_\nu$ for $\nu = 1, 2, \cdots$ the inequality

---

* See also Osher's [2] abstract version of the theorem in [1].

$$|z|^2\|g\|_h^2 < \sum_{\nu=-\infty}^{\infty} |v_\nu(h)|^2 h \leqq \sum_{\nu=-\infty}^{\infty} |v(0)|^2 h = \|g\|_h^2 \,,$$

which is impossible. Therefore there is no eigenvalue z with $|z| \geqq 1$, $z \neq 1$. By the same argument as in Lemma 6 we get the above lemma.

Now we can prove the first part of Theorem 2: Assume that $\mathfrak{G}_N$ has an eigenvalue z with $|z| \geqq 1 - \epsilon$, $|z - 1| \geqq \rho$ and let n be the corresponding eigenfunction:

$$\mathfrak{G}_N u = zu \,.$$

This equation is equivalent with (see (2.4))

$$y_{\nu+1} = My_\nu \,, \qquad \nu = 1, 2, \cdots, N - 1 \,,$$

where y has to fulfill boundary conditions corresponding to (1.15) and (1.16). Upon introducing again the variable $w = Ty$, we get

(4.1) $$w_{\nu+1} = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix} w_\nu \qquad \nu = 1, 2, \cdots, N - 1 \,,$$

with boundary conditions of the form

(4.2)
$$C_1(z)w_1^{\mathrm{I}} = \sum_{j=1}^{s+1} F_j w_j^{\mathrm{II}} \,,$$
$$C_2(z)w_{N-1}^{\mathrm{II}} = Gw_{N-1}^{\mathrm{I}} \,.$$

The general solution of (4.1) is

(4.3) $$w_\nu^{\mathrm{I}} = M_{11}^{\nu-1}w_1^{\mathrm{I}} \,, \qquad w_\nu^{\mathrm{II}} = M_{22}^{\nu-N+1}w_{N-1}^{\mathrm{II}} \,.$$

If $\epsilon > 0$ is sufficiently small, then the absolute value of the eigenvalues of M_{11} are strictly smaller than one and those of M_{22} are strictly larger than one. Furthermore, by assumption $C_1^{-1}(z)$ exists and from Lemma 8 it follows that $C_2^{-1}(z)$ exists as well. If we insert (4.3) into the boundary conditions (4.2) then we can derive homogeneous systems of linear equations for w_1^{I}, w_{N-1}^{II} which are nonsingular for sufficiently large N. Therefore the first part of the theorem follows without difficulties.

Let us now prove the second part: Assume that $\mathfrak{G}$ has an eigenvalue z_0 with $|z_0| \geqq 1 - \epsilon$ and let v be a corresponding eigenfunction: i.e. $\mathfrak{G}v = z_0 v$.

This equation is equivalent with $y_{\nu+1} = My_\nu$ which after transformation can be written as

$$w_\nu^{\mathrm{I}} = M_{11}^{\nu-1}w_1^{\mathrm{I}} \,, \qquad w_\nu^{\mathrm{II}} = 0 \,, \qquad C_1(z_0)w_1^{\mathrm{I}} = 0 \,, \qquad \nu = 1, 2, \cdots \,.$$

The eigenvalues κ_j of M_{11} fulfill the inequality $|\kappa_j| \leqq \tau < 1$. Therefore the eigenfunction v_ν decreases exponentially with ν and we can find a constant d_1 such that

(4.4) $$(\mathfrak{G}_N - z_0 I)v^{(N)} = f^{(N)} \text{ with } |f^{(N)}| \leqq d_1 \tau^N |v^{(N)}| \,.$$

Here $v^{(N)}$ is the function which we get from v by setting $v_\nu = 0$ for $v = N, N + 1, \cdots$. Assume now that $\mathfrak{G}_N - zI$ is nonsingular for $|z - z_0| \leqq \sigma$ and that $\mathfrak{G}$ has no other eigenvalue than z_0 in this neighborhood. Then $|C_1^{-1}(z)| \leqq d_2 \sigma^{-p}$ for $|z - z_0| = \sigma$. Here p is the multiplicity of the root z_0 of $\operatorname{Det} C_1(z) = 0$ which

is equal to the multiplicity of the eigenvalue. Furthermore, by Lemma 9, $C_2^{-1}(z)$ exists and the eigenvalues of M_{22} are strictly larger than 1 in absolute value. Therefore

$$|(\mathfrak{G}_N - zI)^{-1}| \leqq d_3\sigma^{-p} \quad \text{for } |z - z_0| = \sigma \, .$$

This gives

$$|(\mathfrak{G}_N - z_0I)^{-1}| = \left| \frac{1}{2\pi i} \oint_{|z-z_0|=\sigma} \frac{(\mathfrak{G}_N - zI)^{-1}}{z - z_0} \, dz \right| \leqq d_3\sigma^{-p} \, ,$$

and by (4.4)

$$|v^N| \leqq d_1 d_3 \sigma^{-p} \cdot \tau^N \cdot |v^N| \, ,$$

i.e. $\sigma \leqq (d_1 d_3)^{1/p} \cdot \tau^{N/p} = \kappa_2 \tau^{N/p}$. Therefore $\mathfrak{G}_N$ has an eigenvalue λ_N with $|\lambda_N - z_0| \leqq K_2 \tau^{N/p}$. This proves the theorem.

University of Uppsala
Computer Science Department
Uppsala, Sweden

1. H.-O. KREISS, "Difference approximations for the initial-boundary value problem for hyperbolic differential equations," *Numerical Solutions of Nonlinear Differential Equations*, Proc. Adv. Sympos., Madison, Wis., 1966, Wiley, New York, 1966. MR **35** #5156.
2. S. OSHER, *Systems of Difference Equations with General Homogeneous Boundary Conditions*, BNL 11246, Brookhaven National Laboratory, Appl. Math. Dept.
3. T. KATO, *Perturbation Theory for Linear Operators*, Die Grundlehren der math. Wiss., Band 132, Springer-Verlag, New York, 1966. MR **34** #3324.
4. V. THOMÉE, "Estimates of the Friedrichs-Lewy type for mixed problems in the theory of linear hyperbolic differential equations in two independent variables," *Math. Scand.*, v. 5, 1957, pp. 93–113. MR **19**, 1059.