

Contractive Difference Schemes for Symmetric Hyperbolic Systems

By Philip Brenner and Vidar Thomée

Abstract. Consider the initial-value problem for a constant-coefficient symmetric hyperbolic system with initial-values vanishing in a half-space. Consider also a finite difference operator consistent with the system. Conditions are given in terms of the orders of dissipation and accuracy which ensure that the solution of the discrete problem tends to zero exponentially with the mesh-width in half-spaces where the solution of the continuous problem vanishes.

1. Introduction. Consider the Cauchy problem

$$(1.1) \quad \frac{\partial u}{\partial t} = \sum_{i=1}^d A_i \frac{\partial u}{\partial x_i}, \quad t > 0,$$

$$(1.2) \quad u(x, 0) = v(x),$$

where A_i are constant hermitian $N \times N$ matrices and u and v are N -vector valued functions. As is well known, this problem is correctly posed in L_2 . If $\mathcal{F} = \mathcal{F}v$ denotes the Fourier transform on L_2 normalized to coincide for $v \in L_1$ with

$$\hat{v}(\xi) = \int_{\mathbb{R}^d} v(x) \exp(-i\langle x, \xi \rangle) dx, \quad \langle x, \xi \rangle = \sum_{i=1}^d x_i \xi_i,$$

then the solution operator corresponding to (1.1), (1.2) can be written as

$$(E(t)v)(x) = u(x, t) = \mathcal{F}^{-1}(\exp(itP(\xi))\mathcal{F}v),$$

where $P(\xi) = \sum_{i=1}^d A_i \xi_i$. The correctness in L_2 is then a consequence of the fact that $\exp(itP(\xi))$ is a unitary matrix for real ξ and $t > 0$.

An important feature of the solution operator is that the value $u(x, t)$ at the point (x, t) of the solution only depends on the initial-values v on a compact set and that the values of these initial-values on a compact set on $t = 0$ only influence $u = E(t)v$ on a corresponding compact set for fixed positive t . More precisely, if for an arbitrary hermitian $N \times N$ matrix A we let $\lambda_{\min}(A)$ denote its minimal eigenvalue, we have the following well-known result (cf., e.g. Lax [11]).

THEOREM 1.1. *If $v \in L_2$ vanishes for $\langle x, \eta \rangle > K$, then $u(x, t) = 0$ a.e. for $\langle x, \eta \rangle + t\lambda_{\min}(P(\eta)) > K$.*

Proof. The essential part of the argument is that of the proof of the Paley-Wiener theorem. To emphasize the connections with the subsequent analysis, we give some

Received November 2, 1970.

AMS 1970 subject classifications. Primary 65M10; Secondary 65M10.

Key words and phrases. Symmetric hyperbolic system, initial-value problem, difference scheme, contractive, dissipative, stability, accuracy, Fourier transform.

of the details. We first prove the following lemma. Here and below, $|\cdot|$ denotes euclidean length and matrix norm and (\cdot, \cdot) the corresponding inner product.

LEMMA 1.1. *For any real t, ξ , and η , we have*

$$|\exp(itP(\xi + i\eta))| \leq \exp(-t\lambda_{\min}(P(\eta))).$$

Proof. Let v be an arbitrary N -vector. Then, $u(t) = \exp(itP(\xi + i\eta))v$ is the solution of the following initial-value problem, namely,

$$\frac{du}{dt} = iP(\xi)u - P(\eta)u, \quad t > 0, \quad u(0) = v.$$

We hence obtain

$$\begin{aligned} \frac{d}{dt} |u|^2 &= 2 \operatorname{Re} \left(\frac{du}{dt}, u \right) \\ &= 2 \operatorname{Re} [i(P(\xi)u, u) - (P(\eta)u, u)] = -2(P(\eta)u, u), \end{aligned}$$

since $P(\xi)$ and $P(\eta)$ are hermitian. It follows that

$$\frac{d}{dt} |u|^2 \leq -2\lambda_{\min}(P(\eta))|u|^2,$$

or, taking into account the initial condition,

$$|u(t)|^2 \leq \exp(-2t\lambda_{\min}(P(\eta)))|v|^2,$$

which clearly proves the result.

We can now complete the proof of the theorem. By a trivial transformation of variables, it is sufficient to consider the case $K = 0$. For any η , we have

$$e^{(x, \eta)} E(t)v = \mathfrak{F}^{-1}(\exp(itP(\xi + i\eta))\mathfrak{F}(e^{(x, \eta)}v)).$$

It therefore follows by Parseval's relation and Lemma 1.1, for $s > 0$,

$$\|\exp(s\langle x, \eta \rangle + t\lambda_{\min}(P(\eta)))E(t)v\|_2 \leq \|e^{s\langle x, \eta \rangle}v\|_2 \leq \|v\|_2.$$

(We denote the norm on $L_p(R^d)$ by $\|\cdot\|_p$.) If $E(t)v$ did not vanish a.e. when $\langle x, \eta \rangle + t\lambda_{\min}(P(\eta)) > 0$, then we would obtain a contradiction by letting $s \rightarrow \infty$.

We shall consider explicit difference operators

$$E_h v(x) = \sum_{\beta} a_{\beta} v(x + \beta h),$$

where h is a (small) positive number, $\beta = (\beta_1, \dots, \beta_d)$ are multi-integers, a_{β} are constant $N \times N$ matrices, and the summation is over a finite set of β . Introducing the symbol of E_h , $\hat{E}(h\xi) = \sum_{\beta} a_{\beta} \exp(i\langle \beta, \xi \rangle)$, we can write

$$E_h v = \mathfrak{F}^{-1}(\hat{E}(h\xi)\mathfrak{F}v).$$

The operator E_h is consistent with (1.1) if

$$\hat{E}(\xi) = \exp(iP(\xi) + o(|\xi|)) \quad \text{as } \xi \rightarrow 0,$$

and accurate of order ν if

$$\hat{E}(\xi) = \exp(iP(\xi) + O(|\xi|^{\nu+1})) \quad \text{as } \xi \rightarrow 0.$$

We have chosen here to take the time step equal to the mesh-width in space which is clearly no restriction of generality since, in the general case, this can be accomplished by a change of scale in t .

For $t = nh$, we may introduce

$$E_h(t)v = E_h^n v = \mathcal{F}^{-1}(\hat{E}(h\xi)^n \mathcal{F}v),$$

and it is obvious, by the definition of $E_h(t)$, that the following analogue of Theorem 1.1 holds, namely: If $v = 0$ for $\langle x, \eta \rangle > 0$, then $E_h(t)v(x) = 0$ for

$$(1.3) \quad \langle x, \eta \rangle + t \min_{\alpha\beta \neq 0} \langle \beta, \eta \rangle > 0.$$

If E_h is consistent with (1.1) and stable in L_2 which latter property is equivalent to the uniform boundedness of the set of matrices $\hat{E}(\xi)^n$ for real ξ and positive n , then it is well known that for any positive t and $v \in L_2$,

$$\lim_{h \rightarrow 0; nh = t} \|E_h(t)v - E(t)v\|_2 = 0.$$

Consequently, if $v = 0$ for $\langle x, \eta \rangle > 0$, then by Theorem 1.1,

$$(1.4) \quad \lim_{h \rightarrow 0; nh = t} \|\chi_{\eta, t} E_h(t)v\|_2 = 0,$$

where $\chi_{\eta, t}$ is the characteristic function of the half-space

$$(1.5) \quad \langle x, \eta \rangle + t\lambda_{\min}(P(\eta)) > 0.$$

This is, in a certain sense, a stronger statement than the above. In fact, by consistency and stability in L_2 (actually one only has to assume weak stability, i.e. $\rho(\hat{E}(\xi)) \leq 1$ where $\rho(A)$ denotes the spectral radius of A), one may conclude the Courant-Friedrichs-Lewy condition

$$\min_{\alpha\beta \neq 0} \langle \beta, \eta \rangle \leq \lambda_{\min}(P(\eta)),$$

(cf. [11] or [15]) so that (1.3) implies (1.5).

We shall be concerned with a stronger property than the above, namely, when the decay in (1.4) is exponential. Thus let, for t, δ positive and η a real N -vector, $\chi_{\eta, t, \delta}$ be the characteristic function of the half-space $\{x; \langle x, \eta \rangle + t\lambda_{\min}(P(\eta)) \geq \delta\}$. We shall say that E_h is contractive with respect to η if for given positive t, δ there are positive constants γ and n_0 such that for v bounded and vanishing for $\langle x, \eta \rangle > 0$,

$$\|\chi_{\eta, t, \delta} E_h(t)v\|_\infty \leq \exp(-\gamma n) \|v\|_\infty, \quad nh = t, \quad n \geq n_0.$$

We shall be able to prove below the following characterization of contractive difference operators.

THEOREM 1.2. *Assume that $\rho(\hat{E}(\xi)) \leq 1$. Then E_h is contractive with respect to η if and only if*

$$(1.6) \quad \sup_{\xi} \rho(\hat{E}(\xi + is\eta)) \leq \exp(-s\lambda_{\min}(P(\eta)) + o(s)) \quad \text{as } s \rightarrow 0, s > 0.$$

In order to find criteria for this condition to be satisfied, and also to obtain more precise estimates, we shall introduce the following concept. For $\mu > 1, \tau > 1$, we say that E_h is strongly dissipative-contractive of orders (μ, τ) if there are positive

constants c, C with the property that for any $\xi \in Q = \{\xi; |\xi_j| \leq \pi, j = 1, \dots, d\}$ there is a hermitian matrix $H(\xi)$ such that for $\xi \in Q$ and $\eta \in R^d$,

$$(1.7) \quad C^{-1}I \leq H(\xi) \leq CI, \\ |\hat{E}(\xi + i\eta)|_{H(\xi)} \leq \exp(-c|\xi|^\mu - \lambda_{\min}(P(\eta)) + C|\eta|^\tau).$$

Here, for any $N \times N$ matrices A and H with H positive definite, we have denoted by $|A|_H$ the matrix norm of A induced by the vector norm $(Hu, u)^{1/2}$. Obviously, if E_h is strongly dissipative-contractive, it follows by Theorem 1.2 that E_h is contractive with respect to any η . We shall prove the following more precise estimate.

THEOREM 1.3. *Let $1 \leq p \leq \infty$ and assume that E_h is strongly dissipative-contractive of orders (μ, τ) . Then, there are positive constants c and C such that for any η with $|\eta| = 1$ and any $v \in L_p$ with $v = 0$ for $\langle x, \eta \rangle > 0$, we have*

$$(1.8) \quad \|\chi_{\eta, t, \delta} E_h(t)v\|_p \leq Cn^{d(1-1/\mu)|1/2-1/p|} \exp(-cn\delta_t^{\tau/(\tau-1)})\|v\|_p,$$

where $\delta_t = \delta t^{-1}$.

The proofs of Theorems 1.2 and 1.3 will be given in Section 3 below. In Section 4, we then give criteria for difference operators to be dissipative-contractive based on the behavior of $\hat{E}(\xi)$ for real ξ and relating the order of contractivity to the orders of dissipation and accuracy. In particular, it will be proved that if E_h has hermitian coefficients, is dissipative of order μ , and accurate of order $\mu - 1$, then E_h is strongly dissipative-contractive of orders (μ, μ) . In the case of strictly hyperbolic systems, we shall obtain orders $(\mu, \frac{1}{2}\mu)$ under the reduced requirement of order of accuracy $\mu - 2$. The technical work in Sections 3 and 4 is preceded by Section 2 where a number of matrix lemmas are collected.

Contractive difference operators were first discussed by Kreiss and Lundqvist [10] who used a similar notion in studying the domain of influence of "wrong" boundary values in the case of scalar one-dimensional problems. These results were generalized to the case of hyperbolic systems in more dimensions by Osher [12]. Again, for scalar equations and one space variable, Apelkrans [1] used analogous techniques to study the propagation of jump discontinuities in the initial-values. That same problem had also been attacked previously by Hedstrom [7] who obtained very precise estimates by using refined methods for appraising the coefficients in E_h^n . In a subsequent paper [3], we show how these results can be obtained directly by the methods of the present paper and how these techniques can be used to treat more general discontinuities. In particular, it is then possible to improve the estimate (1.8) in some cases by deleting the power of n in front of the exponential. Two reasons why these results are difficult to generalize to systems are that the operators E_h and $E(t)$ in general do not commute and that $E(t)$ is bounded only in L_2 and not in $L_p, p \neq 2$.

2. Some Matrix Lemmas. In this section, we collect some lemmas on $N \times N$ matrices most of which are well known and can be found elsewhere.

LEMMA 2.1. *Given N , there is a constant C such that for any $N \times N$ matrix A with spectral radius ρ ,*

$$\rho^n \leq |A^n| \leq C\rho^{n-N+1}(n|A|)^{N-1}.$$

In particular,

$$\rho = \lim_{n \rightarrow \infty} |A^n|^{1/n}.$$

Proof. See e.g. [14, Lemma 3.2].

We shall need the following version of Carlson’s inequality. This has been used previously in connection with difference schemes in [2], [3] (cf. also [5], [6]).

LEMMA 2.2. *Given $k > d/2$ and N . Then, there is a constant C such that for all $N \times N$ matrix trigonometric polynomials $A(\xi) = \sum_{\beta} a_{\beta} e^{i(\beta \cdot \xi)}$ the following inequality is valid, namely,*

$$\sum_{\beta} |a_{\beta}| \leq C \|A\|_{W_2^k(Q)}^{d/2k} \|A\|_{L_2(Q)}^{1-d/2k},$$

where $\|A\|_{W_2^k(Q)} = (\sum_{|\alpha| \leq k} \|D^{\alpha} A\|_{L_2(Q)}^2)^{1/2}$.

Proof. Since all norms for $N \times N$ matrices are equivalent, one easily finds that it is sufficient to prove the inequality for $N = 1$. By Parseval’s relation,

$$A_k = \left(\sum_{j=0}^k \sum_{\beta} |\beta|^{2j} |a_{\beta}|^2 \right)^{1/2}$$

is equivalent to the norm in $W_2^k(Q)$. Setting $A'_j = (\sum_{\beta \neq 0} |\beta|^{2j} |a_{\beta}|^2)^{1/2}$, we obtain by Schwarz’ inequality

$$\begin{aligned} \sum_{\beta \neq 0} |a_{\beta}| &\leq \sum_{0 < |\beta| \leq \omega} |a_{\beta}| + \sum_{|\beta| > \omega} |a_{\beta}| \\ &\leq C \left\{ \left(\omega^d \sum_{\beta \neq 0} |a_{\beta}|^2 \right)^{1/2} + C \sum_{|\beta| \geq \omega} |\beta|^{-2k} \sum_{\beta} |\beta|^{2k} |a_{\beta}|^2 \right\}^{1/2} \\ &\leq C \{ \omega^{d/2} A'_0 + \omega^{d/2-k} A'_k \}. \end{aligned}$$

Consequently, choosing $\omega = (A'_k/A'_0)^{1/k}$, we obtain

$$\sum_{\beta \neq 0} |a_{\beta}| \leq C (A'_k)^{d/2k} (A'_0)^{1-d/2k} \leq C A_k^{d/2k} A_0^{1-d/2k}.$$

Since

$$|a_0| = (2\pi)^{-d} \left| \int_Q A(\xi) d\xi \right| \leq C A_0 \leq C A_k^{d/2k} A_0^{1-d/2k},$$

the proof is complete.

The following is the version in which we shall use Lemma 2.2. Here and below, we denote the operator norm of a linear operator in $L_p(R^d)$ by $\|\cdot\|_p$.

LEMMA 2.3. *If the difference operator E_h has the symbol \hat{E} , we have, for $nh = t$,*

$$\|E_h(t)\|_{\infty} \leq C n^{d/2} \|\hat{E}^{n-d}\|_{L_2(Q)}.$$

Proof. Setting $E_h(t)v(x) = E_h^n v(x) = \sum_{\beta} a_{n\beta} v(x + \beta h)$, we have

$$\|E_h(t)\|_{\infty} = \sum_{\beta} |a_{n\beta}|.$$

Therefore, if we apply Lemma 2.2 with $k = d$ to the symbol $\hat{E}(\xi)^n$ of $E_h(t)$ and notice that $\|\hat{E}^n\|_{W_2^d(Q)} \leq C n^d \|\hat{E}^{n-d}\|_{L_2(Q)}$, the result follows.

For a $N \times N$ matrix $A = (a_{jk})$, we shall denote by A_0 the diagonal matrix $(\delta_{jk} a_{jk})$. Further, we shall denote by $\text{sp } A$ the set of eigenvalues of A , counting multiple eigenvalues the appropriate number of times. A statement like $\text{sp } A = \text{sp } B + o(\epsilon)$ as $\epsilon \rightarrow 0$ will mean that one can number the eigenvalues $\lambda_j(A)$ and $\lambda_j(B)$ in $\text{sp } A$ and $\text{sp } B$ in such a way that $\lambda_j(A) = \lambda_j(B) + o(\epsilon)$ as $\epsilon \rightarrow 0$ for $j = 1, \dots, N$. We shall

formulate three lemmas where this notation will be used, the first of which is the well-known Gerschgorin theorem.

LEMMA 2.4. Let $A = (a_{jk})$ be a $N \times N$ matrix and let Γ_j be the disk in the complex plane with center a_{jj} and radius $\rho_j = \sum_{k \neq j} |a_{jk}|$. Then, $\text{sp } A \subset \bigcup_{j=1}^N \Gamma_j$. Moreover, if the union Γ of p of the disks is disjoint from the remaining $N - p$ disks, then Γ contains exactly p of the eigenvalues of A , counting multiplicities.

Proof. See e.g. [4].

LEMMA 2.5. Let D be a diagonal matrix with distinct diagonal elements and let $\frac{1}{2} < \alpha < 1$. If $F_\epsilon = D + O(\epsilon^\alpha)$ when $\epsilon \rightarrow 0$, then

$$\text{sp } F_\epsilon = \text{sp } (F_\epsilon)_0 + o(\epsilon) \quad \text{as } \epsilon \rightarrow 0.$$

Proof. Evidently, $\text{sp } F_\epsilon = \text{sp } D + o(1)$ as $\epsilon \rightarrow 0$. We shall have to prove that the eigenvalue $\lambda_j(\epsilon)$ of F_ϵ for which $\lambda_j(\epsilon) = d_j + o(1)$ when $\epsilon \rightarrow 0$ actually has the form

$$\lambda_j(\epsilon) = (f_\epsilon)_{jj} + o(\epsilon) \quad \text{as } \epsilon \rightarrow 0.$$

By permuting rows and columns, we find that we may restrict ourselves to consider the case $j = 1$. Let P_ϵ be the diagonal matrix with diagonal $(1, \sqrt{\epsilon}, \dots, \sqrt{\epsilon})$. Then,

$$\text{sp } F_\epsilon = \text{sp } (P_\epsilon F_\epsilon P_\epsilon^{-1}) \quad \text{and} \quad P_\epsilon F_\epsilon P_\epsilon^{-1} = \begin{pmatrix} (f_\epsilon)_{11} & o(1) \\ o(\epsilon) & D_2 + o(1) \end{pmatrix} \quad \text{as } \epsilon \rightarrow 0,$$

where D_2 is the diagonal matrix with diagonal (d_2, \dots, d_N) . The result now follows by Lemma 2.4, since for small ϵ the disk corresponding to the upper left-hand element is distinct from the remaining disks.

LEMMA 2.6. Let D be the diagonal matrix $\text{diag } (d_1, \dots, d_N)$ and let $Q = (q_{jk})$ have the property that $q_{jk} = 0$ when $d_j = d_k$. If $F_\epsilon = D + \epsilon Q + o(\epsilon)$ as $\epsilon \rightarrow 0$, then $\text{sp } F_\epsilon = \text{sp } (F_\epsilon)_0 + o(\epsilon)$ as $\epsilon \rightarrow 0$.

Proof. Consider an eigenvalue d_j of D and assume that it has multiplicity p . As before, we may assume that

$$F_\epsilon = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} + \epsilon \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} + o(\epsilon) \quad \text{as } \epsilon \rightarrow 0,$$

where D_1 has the diagonal (d_1, \dots, d_1) with p elements, D_2 has the diagonal (d_{p+1}, \dots, d_N) and Q_{11} is diagonal. Analogously to above, let P_ϵ be the diagonal matrix with the p first diagonal elements 1 and the remaining diagonal elements $\sqrt{\epsilon}$. We then have

$$P_\epsilon F_\epsilon P_\epsilon^{-1} = \begin{pmatrix} D_1 + \epsilon Q_{11} + o(\epsilon) & o(1) \\ o(\epsilon) & D_2 + o(1) \end{pmatrix} \quad \text{as } \epsilon \rightarrow 0,$$

from which the result again follows by Lemma 2.4.

We conclude this section with a version of the well-known Kreiss stability theorem. For a $N \times N$ matrix A , we denote by $\Lambda(A)$ the maximum of the real parts of the eigenvalues of A and set $\text{Re } A = \frac{1}{2}(A + A^*)$.

LEMMA 2.7. Assume that $F(\xi)$ has the property that

$$\sup_{t>0} |\exp(tF(\xi))| \leq C, \quad \xi \in \mathbb{D},$$

for some set $\Omega \subset R^d$. Then, $\Lambda(F(\xi)) \leq 0$ for $\xi \in \Omega$ and there is a positive constant C_1 and for any $\xi \in \Omega$ a hermitian matrix $H(\xi)$ such that

$$C_1^{-1}I \leq H(\xi) \leq C_1I \quad \text{and} \quad \text{Re}(H(\xi)F(\xi)) \leq \frac{1}{2}\Lambda(F(\xi))H(\xi).$$

Proof. See [8], [14].

3. Proofs of the Main Results. In this section, we shall prove Theorems 1.2 and 1.3. We begin with a lemma.

LEMMA 3.1. E_h is contractive with respect to η if and only if for any positive δ and t there are positive γ and n_0 such that for $nh = t$, $n \geq n_0$,

$$(3.1) \quad \sum_{-h\langle\beta, \eta\rangle + t\lambda_{\min}(P(\eta)) \geq \delta} |a_{n\beta}| \leq \exp(-\gamma n),$$

where as in Lemma 2.3, the $a_{n\beta}$ denote the coefficients in $E_h(t)$.

Proof. Assume that the condition is satisfied and that v is bounded and vanishes for $\langle x, \eta \rangle > 0$. We obtain

$$(3.2) \quad |E_h(t)v(x)| \leq \|v\|_\infty \sum_{\langle x+\beta h, \eta \rangle \leq 0} |a_{n\beta}|.$$

Hence, if x is such that $\langle x, \eta \rangle + t\lambda_{\min}(P(\eta)) \geq \delta$ this implies for $n \geq n_0$,

$$|E_h(t)v(x)| \leq \sum_{-h\langle\beta, \eta\rangle + t\lambda_{\min}(P(\eta)) \geq \delta} |a_{n\beta}| \leq \exp(-\gamma n),$$

so that E_h is contractive.

On the other hand, assume that E_h is contractive with respect to η and let δ, t be positive and x satisfy $\langle x, \eta \rangle + t\lambda_{\min}(P(\eta)) = \delta$. Clearly, it is possible to choose v with $\|v\|_\infty = 1$ vanishing for $\langle x, \eta \rangle > 0$ such that we have equality in (3.2). But this then proves (3.1).

Proof of Theorem 1.2. We shall first prove that if E_h is contractive with respect to η with $|\eta| = 1$ then (1.6) holds. We have, by the triangle inequality for $s > 0$,

$$(3.3) \quad \exp(n\lambda_{\min}(P(s\eta))) |\hat{E}(\xi + ts\eta)^n| \leq \sum_{\beta} |a_{n\beta}| \exp(-s\langle\beta, \eta\rangle + ns\lambda_{\min}(P(\eta))).$$

Let $B = \max_{\alpha\beta \neq 0} |\beta|$ and $B_1 = B + \sum_i |A_i|$. Then, $a_{n\beta} = 0$ for $|\beta| > Bn$ and hence for $-\langle\beta, \eta\rangle + n\lambda_{\min}(P(\eta)) > B_1n$. Let now $\delta > 0$ be arbitrary. We obtain by Lemma 3.1,

$$(3.4) \quad \sum_{-h\langle\beta, \eta\rangle + t\lambda_{\min}(P(\eta)) \geq \delta} |a_{n\beta}| \exp(-s\langle\beta, \eta\rangle + ns\lambda_{\min}(P(\eta))) \leq \exp(n(B_1s - \gamma)) \leq 1,$$

for $n \geq n_1, s \leq s_0 = \gamma/B_1$. On the other hand, by Lemmas 2.1 and 2.3, we obtain for $nh = 1$,

$$(3.5) \quad \sum_{-h\langle\beta, \eta\rangle + t\lambda_{\min}(P(\eta)) \leq \delta} |a_{n\beta}| \exp(-s\langle\beta, \eta\rangle + ns\lambda_{\min}(P(\eta))) \leq \exp(s\delta h^{-1}) \sum_{\beta} |a_{n\beta}| \leq Cn^{N-1+d/2} \exp(s\delta n).$$

Together, (3.3), (3.4), and (3.5) imply for $s \leq s_0$,

$$|\hat{E}(\xi + is\eta)^n| \leq Cn^{N-1+d/2} \exp(-ns\lambda_{\min}(P(\eta)) + s\delta n).$$

The result now follows by Lemma 2.1 by taking n th roots and letting n tend to infinity.

We shall now prove that, conversely, if $|\eta| = 1$ and (1.6) is satisfied then E_h is contractive with respect to η . Let δ, t be positive and let s be positive and so small that with $\delta_t = \delta/t$,

$$(3.6) \quad \rho(\hat{E}(\xi + is\eta)) \leq \exp(-s\lambda_{\min}(P(\eta)) + \frac{1}{2}s\delta_t), \quad \xi \in Q.$$

We then obtain with $s_h = s/h$,

$$\begin{aligned} R_\delta &= \|\chi_{\eta, t, \delta} E_h(t)v\|_\infty \\ &\leq \|\exp(s_h \langle x, \eta \rangle + t\lambda_{\min}(P(\eta)) - \delta) E_h(t)v\|_\infty \\ &= \exp(sn\lambda_{\min}(P(\eta)) - ns\delta_t) \|\exp(s_h \langle x, \eta \rangle) E_h(t)v\|_\infty. \end{aligned}$$

Introducing the operators

$$(3.7) \quad E_{h, \eta} v = e^{\langle x, \eta \rangle} E_h(e^{-\langle x, \eta \rangle} v),$$

$$(3.8) \quad E_{h, \eta}(t) = E_{h, \eta}^n, \quad t = nh,$$

and noticing that $E_{h, \eta}$ has the symbol $\hat{E}(\xi + ih\eta)$, we get

$$\begin{aligned} \|\exp(s_h \langle x, \eta \rangle) E_h(t)v\|_\infty &\leq \|E_{h, s_h \eta}(t)\|_\infty \|\exp(s_h \langle x, \eta \rangle)v\|_\infty \\ &\leq \|E_{h, s_h \eta}(t)\|_\infty \|v\|_\infty, \end{aligned}$$

since $\langle x, \eta \rangle$ is nonpositive when v is nonzero. By Lemmas 2.1 and 2.3 and by (3.6), we now obtain

$$\begin{aligned} \|E_{h, s_h \eta}(t)\|_\infty &\leq Cn^{N-1+d/2} \rho(\hat{E}(\xi + is\eta)^{n-d}) \\ &\leq Cn^{N-1+d/2} \exp(-ns\lambda_{\min}(P(\eta)) + \frac{1}{2}ns\delta_t), \end{aligned}$$

so that

$$R_\delta \leq Cn^{N-1+d/2} \exp(-\frac{1}{2}ns\delta_t) \|v\|_\infty,$$

which proves our statement.

We shall now prove Theorem 1.3. The main step is in the following lemma.

LEMMA 3.2. *If E_h is strongly dissipative-contractive of orders (μ, τ) and $1 \leq p \leq \infty$, then there is a constant C such that the operator $E_{h, \eta}$ defined in (3.7), (3.8) satisfies*

$$(3.9) \quad \|E_{h, \eta}(t)\|_p \leq Cn^{d(1-1/\mu) |1/2-1/p|} \exp(-t\lambda_{\min}(P(\eta)) + Cth^{\tau-1}|\eta|^\tau).$$

Proof. By assumption, we have for $\xi \in Q$,

$$(3.10) \quad |\hat{E}(\xi + ih\eta)^n| \leq C \exp(-cn|\xi|^\mu - nh\lambda_{\min}(P(\eta)) + Cnh^\tau|\eta|^\tau),$$

and hence,

$$\|\hat{E}(\cdot + ih\eta)^n\|_{L_2(Q)} \leq Cn^{-d/2\mu} \exp(-t\lambda_{\min}(P(\eta)) + Cth^{\tau-1}|\eta|^\tau).$$

The result for $p = \infty$, therefore, follows by Lemma 2.3. For $p = 2$, (3.9) is an immediate consequence of (3.10) since

$$\|E_{h, \eta}(t)\|_2 = \sup_{\xi \in Q} |\hat{E}(\xi + ih\eta)^n|.$$

The statement for general p with $1 \leq p \leq \infty$ is now a consequence of the Riesz-Thorin interpolation theorem.

Proof of Theorem 1.3. We have for any $s > 0$,

$$\begin{aligned} R_\delta &= \|\chi_{\eta, t, \delta} E_h(t)v\|_p \\ &\leq \|\exp(s\langle x, \eta \rangle + t\lambda_{\min}(P(\eta)) - \delta)E_h(t)v\|_p, \end{aligned}$$

and hence

$$R_\delta \leq \exp(s(t\lambda_{\min}(P(\eta)) - \delta)) \|E_{h, s\eta}(t)\|_p \|\exp(s\langle x, \eta \rangle)v\|_p.$$

Using the fact that $v = 0$ for $\langle x, \eta \rangle > 0$ and Lemma 3.2, we obtain

$$R_\delta \leq Cn^{d(1-1/\mu) 11/2-1/p} \exp(-s\delta + Cts^\tau h^{\tau-1}) \|v\|_p.$$

The result now follows by choosing

$$(3.11) \quad hs = ((2C)^{-1} \delta_t)^{1/(\tau-1)}.$$

4. Criteria for Strong Contractivity. In this section, we shall give sufficient conditions for operators E_h to be strongly dissipative-contractive. We recall that E_h is said to be dissipative of order μ if there is a positive c such that

$$\rho(\hat{E}(\xi)) \leq \exp(-c|\xi|^\mu), \quad \xi \in Q.$$

We shall assume that $\hat{E}(\xi)$ is nonsingular so that we may write $\hat{E}(\xi) = \exp F(\xi)$. By continuity, this holds also in a neighborhood of the real line and $F(\xi)$ is analytic near the origin. We shall then introduce a somewhat stronger condition than dissipativity: We say that E_h is strongly dissipative of order μ if there are positive constants c, C and for $\xi \in Q$ a hermitian matrix $H(\xi)$ such that

$$(4.1) \quad C^{-1}I \leq H(\xi) \leq CI,$$

$$(4.2) \quad |H(\xi) - I| \leq C|\xi|^{\mu-1},$$

$$(4.3) \quad \text{Re}(H(\xi)F(\xi)) \leq -c|\xi|^\mu.$$

We shall give below a number of conditions for operators to be strongly dissipative. However, before we do so we shall see that this condition implies that of strong dissipativity-contractivity.

THEOREM 4.1. *Assume that E_h is strongly dissipative of order μ and accurate of order ν . Then, E_h is strongly dissipative-contractive of orders (μ, τ) where $\tau = \mu(\mu - \nu)^{-1}$.*

Proof. Since E_h is explicit and $\tau > 1$, it is clearly sufficient to prove (1.7) for small $\eta, |\eta| \leq \epsilon$, say. For such η and $\xi \in Q$, we want to estimate for $t = 1, \exp(tF(\xi + i\eta))v$ which is the solution of

$$du/dt = F(\xi + i\eta)u, \quad t > 0, \quad u(0) = v.$$

With $R(\xi) = F(\xi) - iP(\xi)$, we have

$$F(\xi + i\eta) = F(\xi) - P(\eta) + Q(\xi, \eta),$$

where $Q(\xi, \eta) = R(\xi + i\eta) - R(\xi)$. Since E_h is accurate of order ν , we have $R(\xi) = O(|\xi|^{\nu+1})$ as $\xi \rightarrow 0$ and hence

$$Q(\xi, \eta) \leq C(|\eta| |\xi|^\nu + |\eta|^{\nu+1}), \quad |\eta| \leq \epsilon.$$

We find for $U = (H(\xi)u, u) = |u|_{H(\xi)}^2$ that

$$\begin{aligned} \frac{dU}{dt} &= 2 \operatorname{Re} \left(\frac{du}{dt}, u \right)_{H(\xi)} \\ &= 2 \operatorname{Re} (F(\xi)u, u)_{H(\xi)} - 2 \operatorname{Re} (P(\eta)u, u)_{H(\xi)} + 2 \operatorname{Re} (Q(\xi, \eta)u, u)_{H(\xi)}. \end{aligned}$$

Now, by our definitions,

$$\begin{aligned} 2 \operatorname{Re} (F(\xi)u, u)_{H(\xi)} &\leq -2c|\xi|^\mu U, \\ -2 \operatorname{Re} (P(\eta)u, u)_{H(\xi)} &\leq -2(P(\eta)u, u) + C|\eta| |\xi|^{\mu-1} U \\ &\leq [-2\lambda_{\min}(P(\eta)) + C|\eta| |\xi|^{\mu-1}] U. \end{aligned}$$

Consequently,

$$\frac{dU}{dt} \leq [-2c|\xi|^\mu - 2\lambda_{\min}(P(\eta)) + C(|\eta| |\xi|^\tau + |\eta|^{\tau+1})] U.$$

Since $\nu < \mu$, we obtain

$$|\eta| |\xi|^\tau \leq \epsilon |\xi|^\mu + C_\epsilon |\eta|^\tau, \quad \tau = \mu/(\mu - \nu),$$

and since $\nu + 1 \geq \tau$, we therefore have

$$\frac{dU}{dt} \leq [-c|\xi|^\mu - 2\lambda_{\min}(P(\eta)) + C|\eta|^\tau] U,$$

or

$$U = |u|_{H(\xi)}^2 \leq \exp(-c|\xi|^\mu - 2\lambda_{\min}(P(\eta)) + C|\eta|^\tau) |v|_{H(\xi)}^2,$$

which clearly proves the theorem.

We shall first give a criterion for strong dissipativity in which the matrix $H(\xi)$ can be chosen as the identity matrix. In this case, (4.1), (4.2), (4.3) reduce to

$$(4.4) \quad \operatorname{Re} F(\xi) \leq -c|\xi|^\mu I.$$

THEOREM 4.2. *Assume that E_n is consistent, dissipative of order μ and such that $\hat{E}(\xi)$ is a normal matrix. Then, E_n is strongly dissipative of order μ .*

Proof. By the spectral mapping theorem, we have

$$\Lambda(F(\xi)) \leq -c|\xi|^\mu, \quad \xi \in \mathcal{Q},$$

and hence (4.4) follows by the normality of $F(\xi)$.

We shall now give two results in which the construction of the matrix $H(\xi)$ in (4.1), (4.2), (4.3) is the essential part of the proofs. These results and their proofs are modelled after Parlett's proofs [13] of similar estimates of Kreiss [9]. In fact, our results contain those previous results; it is quite easy to see that (4.1), (4.2), (4.3) imply, with the same $H(\xi)$,

$$|\hat{E}(\xi)|_{H(\xi)} \leq \exp(-c|\xi|^\mu), \quad \xi \in \mathcal{Q}.$$

In fact, this follows, as in the proof of Theorem 4.1, by noticing that $\hat{E}(\xi)v = (\exp F(\xi))v$ is the solution at $t = 1$ with initial-values v at $t = 0$ of the differential equation $du/dt = F(\xi)u$.

THEOREM 4.3. *Assume that E_h has hermitian coefficients, is dissipative of order μ , and accurate of order $\mu - 1$. Then, E_h is strongly dissipative of order μ .*

Proof. For any $\delta > 0$, we have for $|\xi| > \delta, \xi \in Q$, that $\rho(\hat{E}(\xi)) \leq \gamma < 1$ and hence

$$\sup \{ |\hat{E}(\xi)^n| ; |\xi| > \delta, \xi \in Q \} < \infty .$$

With $\hat{E}(\xi) = \exp F(\xi)$, this implies that

$$\sup \{ |\exp (tF(\xi))| ; |\xi| > \delta, \xi \in Q, t > 0 \} < \infty ,$$

and hence, by Lemma 2.7, there exist for these ξ hermitian matrices $H(\xi)$ such that (4.1) holds and

$$\operatorname{Re} (H(\xi)F(\xi)) \leq -cI .$$

Consequently, it remains only to consider a neighborhood of $\xi = 0$.

Setting $\xi = \sigma\xi_0, \sigma = |\xi|, |\xi_0| = 1$, we have

$$(4.5) \quad F(\xi) = iP(\xi) + R(\xi),$$

where $R(\xi) = \sum_{j=\mu}^{\infty} \sigma^j R_j(\xi_0)$. Notice that $R_\mu(\xi_0)$ is hermitian. We shall consider hermitian matrices $H(\xi)$ of the form

$$(4.6) \quad H(\xi) = I + \sigma^{\mu-1} H_0(\xi_0).$$

For such $H(\xi)$ we have for $\eta = \sigma\eta_0, |\eta_0| = 1$,

$$\begin{aligned} \operatorname{Re} (H(\xi)F(\eta)) &= \operatorname{Re} \{ (I + \sigma^{\mu-1} H_0(\xi_0))(i\sigma P(\eta_0) + \sigma^\mu R_\mu(\eta_0)) \} + o(\sigma^\mu) \\ &= \sigma^\mu B(\xi_0, \eta_0, \sigma), \end{aligned}$$

where

$$(4.7) \quad B(\xi_0, \eta_0, \sigma) = B(\xi_0, \eta_0) + o(1) \quad \text{as } \sigma \rightarrow 0,$$

$$(4.8) \quad B(\xi_0, \eta_0) = R_\mu(\eta_0) + \frac{1}{2}i[H_0(\xi_0)P(\eta_0) - P(\eta_0)H_0(\xi_0)].$$

We shall prove that, for fixed $\xi_0, H_0 = H_0(\xi_0)$ can be chosen so that $B(\xi_0, \xi_0) \leq -2\delta I$ for some positive δ . If this is the case, then, by continuity, there are positive $\epsilon(\xi_0)$ and $\sigma(\xi_0)$ such that $B(\xi_0, \eta_0, \sigma) \leq -\delta I$ for $\sigma \leq \sigma(\xi_0)$ and

$$(4.9) \quad |\eta_0 - \xi_0| < \epsilon(\xi_0),$$

so that the same $H(\xi)$ can be used for a whole cone with vertex at the origin. By compactness, we can find a finite set of such ξ_0 so that the corresponding subsets (4.9) cover the unit sphere. Consequently, $H(\xi)$ can be defined in such a way that for all $\sigma < \sigma_0, |\xi_0| = 1$,

$$|H(\xi) - I| \leq C\sigma^{\mu-1}, \quad \operatorname{Re} (H(\xi)F(\xi)) \leq -\delta\sigma^\mu I.$$

It remains to prove that, for fixed $\xi_0, B(\xi_0, \xi_0)$ can be made negative definite by a proper choice of $H_0 = H_0(\xi_0)$. Since the property of being negative definite is invariant under unitary transformations, it is no restriction of generality to assume that $P(\xi_0)$ is diagonal, $P(\xi_0) = D = \operatorname{diag} (d_1, \dots, d_N)$, and that $R_\mu = R_\mu(\xi_0) = (r_{\mu, jk})$ has the property that $r_{jk} = 0$ when $d_j \neq d_k$. Hence, we want to prove that, given D and R_μ with these properties, we can find H_0 hermitian such that $B = R_\mu + \frac{1}{2}i(H_0 D - D H_0)$

is negative definite. But clearly, $H_0 = (h_{ik})$ can be chosen so that B is diagonal; we only have to take

$$\begin{aligned} h_{ik} &= \frac{2r_{\mu, ik}}{i(d_i - d_k)}, & d_i &\neq d_k, \\ &= 0, & d_i &= d_k. \end{aligned}$$

Since $H_0D - DH_0$ has zero diagonal elements, it remains to prove that the diagonal elements of R_μ are negative.

By the assumption of dissipation, we know that for small σ ,

$$(4.10) \quad \Lambda(F(\xi)) = \sigma\Lambda(iD + \sigma^{\mu-1}R_\mu + o(\sigma^{\mu-1})) \leq -c\sigma^\mu.$$

Setting $\epsilon = \sigma^{\mu-1}$, the assumptions in Lemma 2.6 are satisfied, and consequently, we may conclude that the eigenvalues of $F(\xi)$ are of the form

$$i\sigma d_i + \sigma^\mu r_{\mu, ii} + o(\sigma^\mu), \quad \text{as } \sigma \rightarrow 0.$$

But (4.10) then implies that the $r_{\mu, ii}$ are negative. This concludes the proof.

We recall that (1.1) is called strictly hyperbolic if the eigenvalues of $P(\xi)$ are distinct for all real $\xi \neq 0$.

THEOREM 4.4. *Assume that (1.1) is strictly hyperbolic, that E_h has hermitian coefficients, is dissipative of order $\mu \geq 4$ and accurate of order $\mu - 2$. Then, E_h is strongly dissipative of order μ .*

Proof. This time, we have (4.5) where $R(\xi) = \sum_{j=\mu-1}^\infty \sigma^j R_j(\xi_0)$. Here $R_{\mu-1}$ and R_μ are skew-hermitian and hermitian, respectively. Setting again (4.6), we have now

$$\begin{aligned} \text{Re}(H(\xi)F(\eta)) &= \text{Re} \{ (I + \sigma^{\mu-1}H_0(\xi_0))(i\sigma P(\eta_0) + \sigma^{\mu-1}R_{\mu-1}(\eta_0) + \sigma^\mu R_\mu(\eta_0)) \} + o(\sigma^\mu) \\ &= \sigma^\mu B(\xi_0, \eta_0, \sigma), \end{aligned}$$

where, as before, (4.7) and (4.8) hold. The rest of the reasoning goes as before up to the point where we want to prove that the diagonal elements of R_μ are negative. This time, by dissipation, we have for small σ ,

$$\Lambda(F(\xi)) = \sigma\Lambda(iD + \sigma^{\mu-2}R_{\mu-1} + \sigma^{\mu-1}R_\mu + o(\sigma^{\mu-1})) \leq -c\sigma^\mu.$$

Setting $\epsilon = \sigma^{\mu-1}$, we notice that $\sigma^{\mu-2} = \epsilon^{(\mu-2)/(\mu-1)}$ and since $(\mu - 2)/(\mu - 1) > \frac{1}{2}$, the assumptions of Lemma 2.5 are satisfied so that the eigenvalues of $F(\xi)$ are of the form

$$i\sigma d_i + \sigma^{\mu-1}r_{\mu-1, ii} + \sigma^\mu r_{\mu, ii} + o(\sigma^\mu) \quad \text{as } \sigma \rightarrow 0.$$

Since $r_{\mu-1, ii}$ are purely imaginary, we obtain as before that $r_{\mu, ii}$ are negative. This completes the proof.

Remark. Our main result, Theorem 1.3, admits a generalization to implicit schemes. In that case, the inequality (1.7) which is the conclusion in Theorem 4.1 can be proved only for small values of η , $|\eta| \leq \epsilon$, say, and the choice of hs in (3.11) can only be made when $hs \leq \epsilon$. The exponent in the estimate (1.8), therefore, has to be replaced by $-cn \min \{ \delta_i^{r_i/(r-1)}, \delta_i \}$ (cf. [13]).

Department of Mathematics
Chalmers University of Technology
Göteborg, Sweden

Department of Mathematics
University of Göteborg
Göteborg, Sweden

1. M. Y. T. APELKRANS, "On difference schemes for hyperbolic equations with discontinuous initial values," *Math. Comp.*, v. 22, 1968, pp. 525–539. MR 38 #1848.
2. PH. BRENNER & V. THOMÉE, "Stability and convergence rates in L_p for certain difference schemes," *Math. Scand.*, v. 27, 1970, pp. 5–23.
3. PH. BRENNER & V. THOMÉE, "Estimates near discontinuities for some difference schemes," *Math. Scand.* (To appear.)
4. S. GERSCHGORIN, "Über die Abgrenzung der Eigenwerte einer Matrix," *Izv. Akad. Nauk SSSR*, v. 7, 1931, pp. 749–754.
5. G. W. HEDSTROM, "Norms of powers of absolutely convergent Fourier series," *Michigan Math. J.*, v. 13, 1966, pp. 393–416. MR 34 #3193.
6. G. W. HEDSTROM, "Norms of powers of absolutely convergent Fourier series in several variables," *Michigan Math. J.*, v. 14, 1967, pp. 493–495. MR 36 #5599.
7. G. W. HEDSTROM, "The rate of convergence of some difference schemes," *SIAM J. Numer. Anal.*, v. 5, 1968, pp. 363–406. MR 37 #6051.
8. H. O. KREISS, "Über Matrizen die beschränkte Halbgruppen erzeugen," *Math. Scand.*, v. 7, 1959, pp. 71–80.
9. H. O. KREISS, "On difference approximations of dissipative type for hyperbolic equations," *Comm. Pure Appl. Math.*, v. 17, 1964, pp. 335–353. MR 29 #4210.
10. H. O. KREISS & E. LUNDQVIST, "On difference approximations with wrong boundary values," *Math. Comp.*, v. 22, 1968, pp. 1–12. MR 37 #3777.
11. P. D. LAX, "Differential equations, difference equations and matrix theory," *Comm. Pure Appl. Math.*, v. 11, 1958, pp. 175–194. MR 20 #4572.
12. S. OSHER, "On systems of difference equations with wrong boundary conditions," *Math. Comp.*, v. 23, 1969, pp. 567–572.
13. B. PARLETT, "Accuracy and dissipation in difference schemes," *Comm. Pure Appl. Math.*, v. 19, 1966, pp. 111–123. MR 33 #5141.
14. V. THOMÉE, "Parabolic difference operators," *Math. Scand.*, v. 19, 1966, pp. 77–107. MR 35 #590.
15. V. THOMÉE, "Stability theory for partial difference operators," *SIAM Rev.*, v. 11, 1969, pp. 152–195. MR 40 #3739.