

## SYMMETRIZABLE FINITE DIFFERENCE OPERATORS

BRUCE A. WADE

**ABSTRACT.** We introduce the notion of a symmetrizable finite difference operator and prove that such operators are stable. We then present some sufficient conditions for symmetrizability. One of these extends H.-O. Kreiss' theorem on dissipative difference schemes for hyperbolic equations to a more general case with full  $(x, t)$ -variable coefficients.

### 1. INTRODUCTION

The problem of finding useful sufficient conditions for the stability of linear, variable-coefficient finite difference operators (for hyperbolic problems) has not yet been satisfactorily resolved since existing results make significant limiting assumptions on the symbol of the operator. In this work we extend and unify the various sufficient conditions for stability, e.g., those of Kreiss [4] (also Parlett [9]), Lax and Nirenberg [6], Michelson [7, Theorem 1.2], Shintani and Tomoeda [11], and Strikwerda and Wade [12]. In the process, we simplify the proof of stability for variable-coefficient operators. We consider multistep systems of finite difference equations with  $(x, t)$ -variable coefficients and only minimal assumptions on the symbol.

Primarily, the results of this paper center around the works of Kreiss [4] and Michelson [7, §6]. In [4], stability is proved under some very restrictive assumptions, namely, that there is no  $t$ -dependence in the operator and that both the differential and difference operators have Hermitian coefficients. We eliminate these restrictions, and so address the conjecture in [4, p. 337], in which it is stated that properties of the eigenvalues could possibly replace the special assumptions made there. Michelson's theorem for the pure Cauchy problem [7, Theorem 1.2], concerning finite difference equations for strictly hyperbolic partial differential equations, is a special case of our theory; however, we simplify the proof of stability by using only the weak Gårding inequality, in which the symbol is positive definite, instead of the sharp Gårding inequality. Since the weak Gårding inequality is much easier to prove, we thus obtain a more general result with less machinery.

---

Received January 27, 1989.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 65M10.

*Key words and phrases.* Symmetrizer, stability.

This work was partly supported by the U. S. Army Research Office through the Mathematical Sciences Institute of Cornell University.

Strikwerda and Wade [12] have recently introduced a condition in the Kreiss Matrix Theorem, called a symmetrizer condition, and have shown that the symmetrizer condition ( $[N]$  in [12]) implies stability for variable-coefficient problems in a certain norm involving the Laplace transform in the  $t$ -variable. The symmetrizer condition of [12] is a direct extension of the Lax-Nirenberg non-negative real part condition arising in [6, Corollary 1.2], where the symmetrizer matrix happens to be the identity.

In [12] it is proven that conditions  $[H]$  and  $[N]$  in the Kreiss Matrix Theorem are equivalent, and that the matrix  $N$  can always be taken equal to the matrix  $H$ . However, the converse is not true; in §4 we give an example of a family of matrices which satisfies condition  $[N]$  with the identity as  $N$ , even though the matrix  $H$  cannot be taken to be the identity. To conclude  $[H]$  from  $[N]$ , one would have to go completely around the circle of conditions in the Kreiss Matrix Theorem. For variable-coefficient problems this creates a difficulty because the construction of the matrix  $H$  in the Kreiss Matrix Theorem (which we would like to use as a model), cf. [10], does not produce a smooth  $H$  as a function of the elements of the family of matrices, and smoothness is essential for our pseudodifference operator machinery to go through. Therefore, condition  $[H]$  seems to be somehow stronger than  $[N]$ . For this reason we adopt here a variation of Kreiss' condition  $[H]$  in [4] for our definition of a symmetrizable finite difference operator, rather than the condition  $[N]$ , which was called a symmetrizer condition in [12]. Through the weak Gårding inequality (and condition  $[H]$  as a model) we are able to now prove the same results as those which came out of the sharp Gårding inequality and condition  $[N]$  in [12]. The difference arises only in the variable-coefficient case.

Some work is still needed to answer the natural question of whether the stability estimate resulting from condition  $[N]$  in [12] is equivalent to that from this paper (Theorem 3.1). So far, we can only assert that there is equivalence in the constant-coefficient case, and that the result from [12] may be weaker than that in this work.

The novelty of our method for proving stability consists in the notion of a symmetrizable finite difference operator (one which parallels the already established theory for pseudodifferential operators, cf. [2 or 14]), in our method of proving stability, and also in our method of constructing the symmetrizer. The symmetrizer property given in §3 is basically the same as Kreiss' condition  $[H]$  in [4], but differs in specific details relating to the pseudodifference operator symbol class. Our method of proving stability does not rely on the operator  $H$  as simply a means of changing the norm to obtain a family of contractions, which does not help in the  $t$ -dependent case because the same norm,  $\langle H\cdot, \cdot \rangle$ , cannot work for all time levels; rather, we utilize the operator  $H$  in the spirit of a Lyapunov function to allow an energy method to go through for the full  $(x, t)$ -variable coefficient case. We separate out the question of proving stability and the actual construction of the symmetrizer; this approach allows a

unification of the various existing conditions for stability as special cases of our theory.

We have organized this work as follows. Section 2 contains a brief description of the pseudodifference operator theory. Section 3 contains the first mention of the type of finite difference operators to be considered, a definition of symmetrizability, and a proof that symmetrizable operators are stable. Section 4 is devoted to the question of constructing a symmetrizer for various classes of finite difference operators, which is the most difficult part. We present two theorems on the existence of a symmetrizer, one of which is related to the Kreiss condition of dissipation and accuracy in [4]. Each of these has hypotheses which are useful in practice.

### 2. PSEUDODIFFERENCE OPERATORS

We now briefly discuss the theory of pseudodifference operators, but we omit proofs since we consider only a careful description of the symbol class and the relevant results to be necessary. The reader should consult [1 or 7, §4] for rigorous details.

We take  $M$  to be the collection of complex-valued,  $m \times m$  matrices with norm induced by  $\langle x, y \rangle := y * x$  for  $x, y \in \mathbb{C}^m$ . If  $\alpha \in \mathbb{N}^d$  is a multi-index, we let  $|\alpha| := \sum \alpha_j$ . We assume given a grid parameter  $h \in (0, h_0)$ , for some fixed  $h_0 > 0$ , and we have a quasi-uniform grid  $\mathbb{R}_h^d$  defined to be  $\{x \in \mathbb{R}^d : x_j \in h_j \mathbb{Z}\}$ , where the  $h_j$  satisfy  $c^{-1}h \leq h_j \leq ch$  for  $1 \leq j \leq d$  and all  $h \in (0, h_0)$ , with  $c \geq 1$  fixed. If  $\omega \in \mathbb{R}^d$ ,  $\omega h$  will denote the element of  $\mathbb{R}^d$  with components  $\omega_j h_j$ . We define

$$\Gamma_h := \{\omega \in \mathbb{R}^d : h_j |\omega_j| \leq \pi\}$$

and

$$\Lambda_h(\omega) := \left( 1 + \sum_{j=1}^d h_j^{-2} |1 - e^{-i\omega_j h_j}|^2 \right)^{1/2}.$$

Our discrete function spaces are built around

$$L_h^2 := \left\{ \varphi : \mathbb{R}_h^d \rightarrow \mathbb{C}^m : h^d \sum_{x \in \mathbb{R}_h^d} |\varphi(x)|^2 < \infty \right\}.$$

The discrete Fourier transform is

$$\hat{\varphi}(\omega) := (2\pi)^{-d/2} h^d \sum_{x \in \mathbb{R}_h^d} e^{-ix \cdot \omega} \varphi(x), \quad \omega \in \Gamma_h,$$

and the inversion formula is

$$\varphi(x) := (2\pi)^{-d/2} \int_{\Gamma_h} e^{ix \cdot \omega} \hat{\varphi}(\omega) d\omega,$$

cf. [7, 15, or 16]. We shall utilize the following discrete Sobolev spaces:

$$H_{h,\mu} := \{\varphi: \mathbb{R}^d \rightarrow \mathbb{C}^m : \|\Lambda_h^\mu \hat{\varphi}\|_{L^2(\Gamma_h)} < \infty\}.$$

For pseudodifference operators we follow [7, §4], except that we do not have a particular variable singled out via a Laplace transform. (One simply takes the real part of the Laplace transform dual variable to be zero.) For each  $\mu \in \mathbb{R}$  and  $h \in (0, h_0)$  we define the symbol class of pseudodifference operators with order  $\mu \in \mathbb{R}$  to be  $S_h^\mu$ , taken to be the collection of  $p_h \in C^\infty(\mathbb{R}^d \times \Gamma_h, M)$  which satisfy

$$\|(1 + |x|)^\gamma \partial_x^\alpha (p_h(x, \omega) - p(\infty, \omega))\| \leq c_{\alpha,\gamma} \quad \forall x \in \mathbb{R}^d, \omega \in \Gamma_h,$$

where  $p_h(\infty, \cdot) \in C^\infty(\Gamma_h, M)$ , and also

$$(2.1) \quad \begin{aligned} &\|(1 + |x|)^\gamma \partial_x^\alpha \partial_\omega^\beta p_h(x, \omega)\| \\ &\leq c_{\alpha,\beta,\gamma} \Lambda_h^{\mu-|\beta|}(\omega) \quad \forall x \in \mathbb{R}^d \cup \{\infty\}, \omega \in \Gamma_h, \end{aligned}$$

where  $\gamma \in \mathbb{N}$ ,  $\alpha, \beta \in \mathbb{N}^d$ , and all constants are independent of  $h$ . (Note that we require  $p_h(\infty, \omega)$  to also satisfy (2.1).)

We shall also need a special subclass of pseudodifference operator symbols with a slightly different property relative to the parameter  $h$ . Essentially, this class arises from symbols which are bounded functions of  $\xi = \omega h$ , whereby differentiating with respect to  $\omega$  yields successively higher powers of  $h$  instead of lower powers of  $\Lambda_h(\cdot)$ . For instance, a cutoff function in the  $\xi$  variable satisfies the conditions of this special class of symbols.

For each  $\mu \in \mathbb{R}$  and  $h \in (0, h_0)$  we define the symbol class  $\tilde{S}_h^\mu$  to be the collection of  $p_h \in S_h^\mu$  satisfying

$$(2.2) \quad \begin{aligned} &\|(1 + |x|)^\gamma \partial_x^\alpha \partial_\omega^\beta p_h(x, \omega)\| \\ &\leq c_{\alpha,\beta,\gamma} h^{|\beta|} \Lambda_h^\mu(\omega) \quad \forall x \in \mathbb{R}^d \cup \{\infty\}, \omega \in \Gamma_h. \end{aligned}$$

For  $p_h \in S_h^\mu$  we define the corresponding pseudodifference operator  $P_h: H_{h,\mu} \rightarrow H_{h,0}$  by

$$P_h \varphi(x) := (2\pi)^{-d/2} \int_{\Gamma_h} p_h(x, \omega) e^{ix \cdot \omega} \hat{\varphi}(\omega) d\omega,$$

and we let  $\sigma(P_h) := p_h$  be the symbol of  $P_h$ . We take  $OPS_h^\mu$  to be the collection of such operators. (That  $P_h: H_{h,\mu} \rightarrow H_{h,0}$  is actually proved in [7, Theorem 4.1].)

We now state the relevant results on pseudodifference operators, without giving proofs. We assume that the reader is familiar with [7, §4].

**Proposition 2.1.** *For each  $P_h \in OPS_h^\mu$  and  $\nu \in \mathbb{R}$  there is  $c > 0$ , independent of  $h$ , such that*

$$\|P_h \varphi\|_{h,\nu} \leq c \|\varphi\|_{h,\mu+\nu} \quad \forall \varphi \in H_{h,\mu+\nu}.$$

We denote by  $P_h^*$  the  $H_{h,0}$ -adjoint and by  $P_h^{(*)}$  the operator whose symbol is  $\sigma(P_h)^*$  (the matrix adjoint).

**Proposition 2.2.** *Let  $P_h \in OPS_h^\mu$  and  $Q_h \in OPS_h^\nu$ , and let  $R_h$  be the pseudo-difference operator whose symbol is  $\sigma(P_h)\sigma(Q_h)$ . Then*

- (i)  $P_h \circ Q_h, R_h \in OPS_h^{\mu+\nu}$ ,
- (ii)  $P_h \circ Q_h - R_h \in OPS_h^{\mu+\nu-1}$ ,
- (iii)  $P_h^* - P_h^{(*)} \in OPS_h^{\mu-1}$ .

**Proposition 2.3.** *Let  $P_h \in OPS_h^{\tilde{\mu}}$  and  $Q_h \in OPS_h^\nu$ , and let  $R_h$  be the operator whose symbol is  $\sigma(P_h)\sigma(Q_h)$ . Then there are  $A_h \in OPS_h^{\mu+\nu}$  and  $B_h \in OPS_h^\mu$  such that  $P_h \circ Q_h - R_h = hA_h$  and  $P_h^* - P_h^{(*)} = hB_h$ .*

Proposition 2.3 is the same as [7, Theorem 4.2a, 4.3a or 6, Lemma 1.1]. Comments on how to easily modify the standard proof of Proposition 2.2 to the case with symbols in  $\tilde{S}_h^\mu$  are contained in [7, p. 32]. Intuitively, the only difference between Propositions 2.2 and 2.3 is that  $OPS_h^{\tilde{\mu}}$  provides a different meaning to the phrase ‘lower-order terms’.

Next, we give a special case of the weak Gårding inequality involving operators in the restricted class  $OPS_h^{\tilde{0}}$ . This version is all that we will need in the next section.

**Proposition 2.4.** *Suppose  $P_h \in OPS_h^{\tilde{0}}$  and there is  $c_0 > 0$  such that  $\text{Re } \sigma(P_h)(\cdot) \geq c_0 I$ ; then there is  $c_1 \geq 0$ , independent of  $h$ , such that*

$$\text{Re} \langle P_h \varphi, \varphi \rangle_{h,\nu} \geq (\frac{1}{2}c_0 - c_1 h) \|\varphi\|_{h,\nu}^2 \quad \forall \varphi \in H_{h,\nu}.$$

We note that the conclusion of Proposition 2.4 is essentially the same as that of [6, Theorem 1.1], except that our assumption  $P_h \in OPS_h^{\tilde{0}}$  greatly simplifies the proof because the standard derivation of the weak Gårding inequality, cf. [7], goes through.

### 3. SYMMETRIZABILITY AND STABILITY

In this section we introduce our class of finite difference equations, define the concept of a symmetrizable finite difference operator, and prove that such operators are stable.

We consider finite difference equations of the following general type:

$$(3.1) \quad \begin{aligned} q(t, h, x, T_x, T_t)v_h(t, x) &= f_h(t, x) \quad \forall (t, x) \in \mathbb{N}_k \times \mathbb{R}_h^d, \\ v_h(\sigma k, x) &= g_{h,\sigma}(x), \quad 0 \leq \sigma \leq \sigma_0 - 1, \end{aligned}$$

where

$$q(t, h, x, T_x, T_t) := \sum_{\sigma=0}^{\sigma_0} q_\sigma(t, h, x, T_x) T_t^\sigma.$$

We have used  $T_x$  and  $T_t$  to indicate the forward translation operators,  $k = \lambda h$  for some fixed  $\lambda > 0$  and  $\mathbb{N}_k := k\mathbb{N}$ . The following assumptions are made:  $k^{-1}q_0(t, h, x, T_x)^{-1} \in OPS_h^0$ ,  $\{q_\sigma(t, \cdot)\}_{\sigma=0}^{\sigma_0} \subset C^1(\mathbb{R}_+ \cdot S_h^\mu)$  for some  $\mu > 0$ ,

and  $\{q_0^{-1}(t, \cdot)q_\sigma(t, \cdot)\}_{\sigma=0}^{\sigma_0} \subset C^1(\mathbb{R}_+, S_h^0)$ , where  $q_0^{-1}(t, \cdot)$  denotes the symbol of the inverse of  $q_0(t, h, x, T_x)$ .

These assumptions are certainly not very restrictive—they include as special cases any of the finite difference equations of [4; 6; 7, Theorem 1.2; 9; or 11]. We require only that  $f_n$  and  $\{g_{h,\sigma}\}_{\sigma=0}^{\sigma_0-1}$  be grid functions, and we have no differential equation in sight. The conditions on the inverse of the operator  $q_0(t, h, x, T_x)$  are somewhat troublesome because they are not easy to check, unless one has a constant-coefficient operator. However, we need the invertibility of  $q_0(t, h, x, T_x)$  even to know that there exists a solution to (3.1), but we do not desire here to deal with the problem of finding conditions on the symbol of  $q_0(t, h, x, e^{i\omega h})$  which guarantee the invertibility of the operator. Also, these assumptions implicitly force a relationship between  $k$  and the  $q_\sigma(\cdot)$ , which is the natural one for hyperbolic problems.

We now fix our definition of stability, which is the usual one.

**Definition 3.1.** The finite difference operator  $q(t, h, x, T_x, T_t)$  in (3.1) is said to be stable if for every  $T > 0$  there exist  $h_0, c > 0$  such that any solution satisfies

$$(3.2) \quad \|v_h(t)\|_{h,0}^2 \leq c \left( \sum_{\sigma=0}^{\sigma_0-1} \|g_{h,\sigma}\|_{h,0}^2 + k \sum_{\tau \in [0,t] \cap \mathbb{N}_k} \|f_h(\tau)\|_{h,0}^2 \right),$$

for  $t \in [0, T] \cap \mathbb{N}_k$  and  $h \in (0, h_0)$ .

For our stability proof to go through, we need to reduce the multistep operator in (3.1) to a single-step canonical form—called the reduced operator—by employing the standard method, cf. [10, Chapter 7]. We go over to capital letters. Let  $V_h$  be the vector of size  $m\sigma_0$ , defined in terms of  $\sigma_0$  blocks of length  $m$ , as follows:

$$[V_h]_j := T_t^{\sigma_0-j} v_h, \quad 1 \leq j \leq \sigma_0,$$

where the square brackets indicate an  $m$ -block. Similarly, let

$$[G_h]_j := g_{h,\sigma_0-j}, \quad 1 \leq j \leq \sigma_0,$$

and

$$[F_h]_j := \begin{cases} k^{-1}q_0(t, h, x, T_x)^{-1}f_h, & j = 1; \\ 0, & \text{else.} \end{cases}$$

Finally, we define a block matrix of operators consisting of a  $\sigma_0 \times \sigma_0$  matrix of  $m \times m$  blocks as follows:

$$[Q]_{ij} := \begin{cases} q_0(t, h, x, T_x)^{-1}q_j(t, h, x, T_x), & i = 1; \\ I, & i > 1, j = i - 1; \\ 0, & \text{else,} \end{cases}$$

for  $1 \leq i, j \leq \sigma_0$ . Our assumptions from (3.1) imply that  $Q \in OPS_h^0$  and that  $\|F_h\|_{h,0} \leq c\|f_h\|_{h,0}$ .

The resulting reduced equation, equivalent to (3.1), is

$$(3.3) \quad \begin{aligned} (T_t - Q(t, h, x, T_x))V_h(t, x) &= kF_h(t, x), \\ V_h(0, x) &= G_h(x) \quad \forall (t, x) \in \mathbb{N}_k \times \mathbb{R}_h^d. \end{aligned}$$

Stability for the reduced equation becomes: for every  $T > 0$  there are  $h_0, c > 0$  such that any solution satisfies

$$(3.4) \quad \|V_h(t)\|_{h,0}^2 \leq c \left( \|G_h\|_{h,0}^2 + k \sum_{\tau \in [0, t] \cap \mathbb{N}_k} \|F_h(\tau)\|_{h,0}^2 \right),$$

for  $t \in [0, T] \cap \mathbb{N}_k$  and  $h \in (0, h_0)$ .

We now introduce the notion of a symmetrizable finite difference operator, one which parallels that for pseudodifferential operators, cf. [14].

**Definition 3.2.** The reduced finite difference operator is said to be *symmetrizable* if for every  $T > 0$  there are  $h_0, \eta_0, c > 0$  and a family of Hermitian matrices  $H(t, \cdot) \in C^1([0, T], \tilde{S}_h^0)$  satisfying  $c^{-1}I \leq H(\cdot) \leq cI, \partial_t H(t, \cdot) \leq cI$ , and

$$(3.5) \quad Q^*(t, h, x, e^{i\omega h})H(t, h, x, \omega)Q(t, h, x, e^{i\omega h}) \leq e^{\eta_0 k} H(t, h, x, \omega),$$

for  $h \in (0, h_0)$  and  $(t, x, \omega) \in [0, T] \times \mathbb{R}_h^d \times \Gamma_h$ .

The difference between this definition and Kreiss' condition [4, Theorem 1, #4] is the requirement of smoothness and that  $H(t, \cdot)$  must be in the special class of pseudodifference operator symbols  $\tilde{S}_h^0$ . We postpone the question of actually constructing the symmetrizer until the next section.

We shall need a convenient form of the Gronwall lemma. Suppose  $\varphi, \psi: \mathbb{N}_k \rightarrow \mathbb{R}_+$  satisfy  $\delta_- \varphi(t) - c\varphi(t) \leq \psi(t)$ , for  $t \in \mathbb{N}_k$ , where  $\delta_- := k^{-1}(I - T_t^{-1})$ . Then there is  $h_0 > 0$  such that

$$(3.6) \quad \varphi(t) \leq e^{ct} \left( \varphi(0) + 2k \sum_{\tau \in (0, t] \cap \mathbb{N}_k} e^{-c\tau} \psi(\tau) \right),$$

for  $h \in (0, h_0)$  and  $t \in \mathbb{N}_k$ .

To prove (3.6), we introduce the summation factor  $I(\tau) := (1 - ck)^{\tau k^{-1}}$ , where  $h_0$  is chosen small enough to insure that  $ck < 1$ . Since  $\delta_-(I(\tau)\varphi(\tau)) = I(\tau - k)\delta_-\varphi(\tau) - cI(\tau - k)\varphi(\tau)$ , we see that  $\delta_-(I(\tau)\varphi(\tau)) \leq I(\tau - k)\psi(\tau)$ . Summation and the telescoping property yield

$$I(t)\varphi(t) - I(0)\varphi(0) \leq k \sum_{\tau \in (0, t] \cap \mathbb{N}_k} I(\tau - k)\psi(\tau),$$

which gives (3.6) after a simple computation.

**Theorem 3.1.** *If the reduced finite difference operator is symmetrizable, then it is stable.*

*Proof.* Given the symbol  $H(\cdot)$ , let  $H_h$  denote the corresponding pseudodifference operator. We will compute  $\delta_- \langle H_h(t)V_h(t), V_h(t) \rangle_{h,0}$ , dropping the subscripts for convenience.

But first we change variables by letting  $\tilde{V}_h(t) := e^{-\eta t} V_h(t)$ , where  $\eta \geq 0$  is to be fixed later. The reduced equation, (3.3), then goes over to

$$\begin{aligned} (T_t - e^{-\eta k} Q(t, h, x, T_x, T_t)) \tilde{V}_h(t, x) &= k \tilde{F}_h(t, x), \\ \tilde{V}_h(0, x) &= G_h(x). \end{aligned}$$

However, we now drop the tildes (for simplicity), keeping in mind that the new reduced operator is actually  $e^{-\eta k} Q(t, h, x, T_x)$ .

Our first step consists in adding and subtracting

$$k^{-1} \langle H(t-k)V(t), V(t) \rangle$$

to obtain

$$\begin{aligned} \delta_- \langle H(t)V(t), V(t) \rangle &= k^{-1} \langle H(t-k)V(t), V(t) \rangle \\ &\quad - k^{-1} \langle H(t-k)V(t-k), V(t-k) \rangle \\ &\quad + \langle (\delta_- H(t))V(t), V(t) \rangle. \end{aligned}$$

Now we use the reduced equation, above, to find that this equals

$$\begin{aligned} &k^{-1} e^{-2\eta k} \langle H(t-k)Q(t-k)V(t-k), Q(t-k)V(t-k) \rangle \\ &\quad - k^{-1} \langle H(t-k)V(t-k), V(t-k) \rangle \\ &\quad + \langle H(t-k)F(t-k), V(t) \rangle + e^{-\eta k} \langle H(t-k)Q(t-k)V(t-k), F(t-k) \rangle \\ &\quad + \langle (\delta_- H(t))V(t), V(t) \rangle. \end{aligned}$$

Using the Cauchy-Schwarz inequality and the fact that  $H(\cdot)$ ,  $\delta_- H(\cdot)$ ,  $Q(\cdot) \in OPS_h^0$ , we arrive at the upper bound

$$\begin{aligned} &k^{-1} \langle (e^{-2\eta k} Q^*(t-k)H(t-k)Q(t-k) - H(t-k))V(t-k), V(t-k) \rangle \\ &\quad + c(\|V(t-k)\|^2 + \|V(t)\|^2 + \|F(t-k)\|^2), \end{aligned}$$

where  $c > 0$  is independent of the parameters.

The symmetrizer property, (3.5),  $H(\cdot) \in OPS_h^0$ , and Proposition 2.3 yield

$$\begin{aligned} \delta_- \langle H(t)V(t), V(t) \rangle &\leq c(k^{-1}(e^{-2(\eta-\eta_0)k} - 1) + 1)\|V(t-k)\|^2 \\ &\quad + c\|V(t)\|^2 + c\|F(t-k)\|^2, \end{aligned}$$

where  $\eta_0$  comes from Definition 3.2. Choosing  $\eta$  large enough to make  $e^{-2(\eta-\eta_0)k} - 1 < 0$  and using the property  $C^{-1}I \leq H(\cdot) \leq CI$ , we find that

$$\begin{aligned} \delta_- \langle H(t)V(t), V(t) \rangle &\leq c \langle H(t-k)V(t-k), V(t-k) \rangle \\ &\quad + c \langle H(t)V(t), V(t) \rangle + c\|F(t-k)\|^2. \end{aligned}$$

The discrete Gronwall inequality (3.6) applied to

$$\varphi(t) := \langle H(t)V(t), V(t) \rangle, \quad \psi(t) := \|F(t-k)\|^2,$$



then gives

$$\langle H(t)V(t), V(t) \rangle \leq e^{ct} \left( \langle H(0)V(0), V(0) \rangle + 2k \sum_{\tau \in [0, t] \cap \mathbb{N}_k} e^{-c\tau} \|F(\tau)\|^2 \right),$$

where now  $V(t)$  represents the original solution, and the  $\eta$ -dependence is implicit in the constant  $c$ .

By choosing  $h_0$  sufficiently small and applying the weak Gårding inequality, we find that

$$\begin{aligned} \frac{1}{4}c_0 \|V(t)\|^2 &\leq (\frac{1}{2}c_0 - c_1 h) \|V(t)\|^2 \leq \langle H(t)V(t), V(t) \rangle \\ &\leq ce^{ct} \left( \|V(0)\|^2 + k \sum_{\tau \in [0, t] \cap \mathbb{N}_k} \|F(\tau)\|^2 \right), \end{aligned}$$

for some fixed  $c > 0$  (depending on  $T$ ), where  $c_0$  and  $c_1$  come from Proposition 2.4. This completes the proof.  $\square$

#### 4. CONSTRUCTION OF A SYMMETRIZER

There remains the problem of actually constructing the symmetrizer, which is nontrivial. In this section we present two theorems on the constructibility of a symmetrizer matrix, each based on ‘easily verifiable’ sufficient conditions. Then we finish up with a few comments about remaining difficulties in this theory.

Our method utilizes eigenprojection and total projection operators derived from integrating the resolvent over certain specially chosen contours in the complex plane. This method, which is based on [8], relies on the resolvent condition of the Kreiss Matrix Theorem and a certain uniformity property of the eigenvalues. No differential equation appears at first.

Our second result extends the Kreiss theory of [4] concerning accurate and dissipative difference schemes. This well-known theory requires the orders of accuracy and dissipation to match, and applies only to an unnaturally restricted class of finite difference schemes, mainly because of the method of proof. We prove in this section that one can eliminate the restriction in [4] to the case of Hermitian coefficients without  $t$ -dependence. For simplicity we consider in Theorem 4.2 only explicit, single-step schemes, but the method obviously generalizes in the same way as in Widlund’s work [17] to the multistep case, provided extra assumptions are added to control the spurious eigenvalues arising from the multistep nature. Michelson’s theorem [7, Theorem 1.2] concerning dissipative difference schemes for strictly hyperbolic partial differential equations is also a special case of our first theorem of this section.

In the constant-coefficient case there would be no trouble constructing a symmetrizer, by condition  $[H]$  in the Kreiss Matrix Theorem, because one would not need the pseudodifference operator theory, hence no smoothness properties, and Theorem 3.1 would go through directly. In the variable-coefficient case this method breaks down because all known proofs of the Kreiss Matrix Theorem

utilize the eigenvalues explicitly, and these could be at most continuous functions of the parameters; therefore, the pseudodifference operator results would not apply. Essentially, the problem at hand is to construct the symmetrizer to satisfy (3.5) while still being in  $\tilde{S}_h^0$ .

We are given the operator symbol  $Q(t, h, x, e^{i\omega h})$ , which we denote by  $A(t, h, x, \xi)$ ,  $\xi := \omega h$ , for notational convenience. We need to construct a symmetrizer  $H(t, h, x, \xi)$  to be a symbol of order zero; in particular, it must be  $C^1$  in  $t$ ,  $C^\infty$  in  $(x, \xi)$ , and its derivatives must have the right behavior to be in  $\tilde{S}_h^0$ . For convenience of notation we let  $X := \mathbb{R}_+ \times (0, h_0) \times \mathbb{R}^d \times [-\pi, \pi]$ , the resolvent of  $A$  is  $R_z(A) := (zI - A)^{-1}$ , and we take  $B_\varepsilon(\chi_0)$  to mean the  $\varepsilon$ -ball about  $\chi_0$  intersected with  $X$ .

From the Kreiss Matrix Theorem, cf. [10], we recall the resolvent condition for this situation: there is  $c > 0$  such that for  $|z| > 1$  and  $\chi \in X$

$$(4.1) \quad \|R_z(A(\chi))\| \leq c(|z| - 1)^{-1}.$$

Although it is not necessarily easy to check in practice, we assume until the completion of Theorem 4.1 that the resolvent condition holds. This amounts to a pointwise condition on the symbol. It is easy to see that the resolvent condition implies that all eigenvalues of  $A(\cdot)$  have modulus less than or equal to one, and those on the unit circle are simple poles of  $R_z(A)$ . Difficulties in constructing the symmetrizer arise because the resolvent condition does not necessarily restrict the eigenvalues near the unit circle to be smooth.

Following [3, Chapter 1], we consider the eigenprojections as follows. Let  $\lambda(\cdot) \in \sigma(A(\cdot))$ , the spectrum of  $A(\cdot)$ , and let  $\Gamma$  be any unit-index, rectifiable contour in  $\mathbb{C}$  containing at least  $\lambda$  out of the spectrum and not intersecting it. The corresponding eigenprojection operator is

$$P_\Gamma := (2\pi i)^{-1} \int_\Gamma R_z(A) dz.$$

Consulting [3], we see that  $P_\Gamma^2 = P_\Gamma$  and  $P_{\Gamma_\mu} P_{\Gamma_\nu} = 0$  if  $\Gamma_\mu$  and  $\Gamma_\nu$  contain no common element of the spectrum. We say that  $\lambda$  is simple if  $\dim(P_{\Gamma_\lambda} \mathbb{C}^m) = 1$  whenever  $\Gamma_\lambda$  contains only  $\lambda$  out of the spectrum, and we say that  $\lambda$  is semisimple if  $(A - \lambda I)P_{\Gamma_\lambda} = 0$ .

We now present an estimate on the resolvent, along certain contours, which will be useful in the proofs of Theorems 4.1 and 4.2. Miller's method from [8] is basically the means for the following estimate. Suppose  $\lambda \in \sigma(A)$  with  $|\lambda| < 1$  and let  $\Gamma$  be any unit-index contour in  $\mathbb{C}$  containing at least  $\lambda$  out of the spectrum, not intersecting the spectrum, and contained inside the circle centered at  $\lambda$  with radius  $1 - |\lambda|$ . Set  $\tilde{\lambda} := (2 - |\lambda|)\lambda$ , i.e., the reflection over the unit circle  $U$ , and let  $\tilde{\Gamma}$  be obtained in the same manner.

Consider the matrix polynomial of degree at most  $m - 1$  ( $A$  is  $m \times m$ )

$$q(z) := R_z(A) \prod_{\mu \in \sigma(A)} (z - \mu).$$

This yields

$$q(z) = \sum_{j=0}^{m-1} (z - \tilde{\lambda})^j (2\pi i)^{-1} \int_{\tilde{\Gamma}} q(\zeta) (\zeta - \tilde{\lambda})^{-(j+1)} d\zeta,$$

which is straightforward to verify, and therefore

$$(4.2) \quad R_z(A) = (2\pi i)^{-1} \int_{\tilde{\Gamma}} \varphi(\zeta, z) R_\zeta(A) d\zeta,$$

where

$$\varphi(\zeta, z) := (\zeta - \tilde{\lambda})^{-1} \prod_{\mu \in \sigma(A)} \left( \frac{\zeta - \mu}{z - \mu} \right) \sum_{j=1}^{m-1} \left( \frac{z - \tilde{\lambda}}{\zeta - \tilde{\lambda}} \right)^j.$$

From (4.2), for  $z \in \Gamma$ ,

$$\|R_z(A)\| \leq c|\Gamma| \sup_{\zeta \in \tilde{\Gamma}} \{|\varphi(\zeta, z)| (|\zeta| - 1)^{-1}\},$$

where  $|\Gamma|$  is the length of  $\Gamma$  and  $c$  depends only on the resolvent constant. Therefore, for  $z \in \Gamma$ ,

$$\|R_z(A)\| \leq c|\Gamma| \text{dist}(\Gamma, U)^{-1} \sup_{\zeta \in \tilde{\Gamma}} \{|\varphi(\zeta, z)|\}.$$

The following inequalities can easily be verified for  $z \in \Gamma$ ,  $\zeta \in \tilde{\Gamma}$ , and  $\mu \in \sigma(A)$ :

$$\begin{aligned} |\zeta - \tilde{\lambda}|^{-1} &\leq \text{dist}(\Gamma, \lambda)^{-1}, \\ |z - \tilde{\lambda}| |\zeta - \tilde{\lambda}|^{-1} &\leq 3 \text{dist}(\Gamma, \lambda)^{-1} (1 - |\lambda|), \\ |\zeta - \mu| |z - \mu|^{-1} &\leq \text{dist}(\Gamma, \mu)^{-1} (4(1 - |\lambda|) + \text{dist}(\Gamma, \mu)). \end{aligned}$$

We therefore obtain

$$(4.3) \quad \sup_{z \in \Gamma} \|R_z(A)\| \leq c|\Gamma| \text{dist}(\Gamma, U)^{-1} (1 - |\lambda|)^{m-1} \times \text{dist}(\Gamma, \lambda)^{-m} (1 + (1 - |\lambda|) \text{dist}(\Gamma, \sigma(A))^{-1})^m,$$

where  $c$  depends on  $m$  and the resolvent constant.

In (4.3) we have a delicate balance between the various expressions which depend on  $\Gamma$  and  $\sigma(A)$ ; we will utilize this estimate in our next theorems.

Our symmetrizer matrix must be in  $\tilde{S}_h^0$ , and we will need some additional assumptions on the family  $\{A(\cdot)\}_X$  to be able to construct such an object. For our first theorem we simply assume a certain uniformity property on the eigenvalues, which we now describe.

Fix  $\chi_0 \in X$ ; let  $\Lambda$  denote the set of unit-modulus eigenvalues of  $A(\chi_0)$  and let  $G := \sigma(A(\chi_0)) \setminus \Lambda$ . Let  $\rho := (1 + \max\{|\mu| : \mu \in G\})/2$  and  $\delta := \min\{|\mu - \nu|, 1 - \rho\}/2$ , taken over  $\mu, \nu \in \Lambda$ . By continuity, there is  $\varepsilon(\chi_0) > 0$  so that on  $B_\varepsilon(\chi_0)$  all eigenvalues of  $A(\cdot)$  remain strictly inside, and never on, one of the  $\{\Gamma_\lambda\}_{\lambda \in \Lambda}$  or  $\Gamma_0$ , where  $\Gamma_0$  is the positively oriented circle with

radius  $\rho$  centered at the origin and  $\Gamma_\lambda$  is the positively oriented circle centered at  $\lambda \in \Lambda$  with radius  $\delta$ . We call the set of eigenvalues  $\{\mu(\cdot)\}_{B_\varepsilon(\chi_0)}$  which satisfy  $|\mu - \lambda| < \delta$  on  $B_\varepsilon(\chi_0)$  the  $\lambda$ -group.

Now we get more specialized. Suppose that we could find a finite number of points  $\{\chi_j\}_{j \in J} \subset X$  such that the above  $\varepsilon_j$  have the property that  $\bigcup_j B_{\varepsilon_j}(\chi_j) = X$  and for each  $\lambda \in \Lambda_j$  (unit-modulus eigenvalues of  $A(\chi_j)$ ) the  $\lambda$ -group is semisimple; then we say that the family  $\{A(\cdot)\}_X$  is uniformly semisimple at the unit circle. This situation would occur for instance (by a continuity argument) if  $A(t, h, x, \xi)$  were independent of  $h$  and constant in  $(t, x)$  outside some large ball, a common situation.

To fix notation, then, we have a finite set  $J$  and  $\{\chi_j\}_{j \in J} \subset X$  such that  $\bigcup_j B_{\varepsilon_j}(\chi_j) = X$ , and a collection of unit-modulus eigenvalues  $\Lambda_j$  for each  $j \in J$ . Further, for each  $\lambda \in \Lambda_j$  we have a contour  $\Gamma_{\lambda,j}$  of radius  $\delta_j$ , and also we have  $\rho_j \in (0, 1)$  and a contour  $\Gamma_j$  of radius  $\rho_j$ . These expressions are utilized in the proof of our next theorem.

**Theorem 4.1.** *Suppose the resolvent condition holds and the family  $\{A(\cdot)\}_X$  is uniformly semisimple at the unit circle. Then there exists a symmetrizer.*

*Proof.* We shall construct the symmetrizer locally, and then use a partition of unity subordinate to the finite cover  $\{B_{\varepsilon_j}(\chi_j)\}_{j \in J}$ . For each  $j \in J$  and  $\lambda \in \Lambda_j$ , let

$$A_j(\cdot) := (2\pi i)^{-1} \int_{\Gamma_j} z R_z(A(\cdot)) dz,$$

and

$$P_{\lambda,j}(\cdot) := (2\pi i)^{-1} \int_{\Gamma_{\lambda,j}} R_z(A(\cdot)) dz,$$

defined on  $B_{\varepsilon_j}(\chi_j)$ . The symmetrizer can be constructed as

$$H(\cdot) := \sum_{j \in J} \varphi_j(\cdot) H_j(\cdot),$$

where the  $\varphi_j$  are cutoff functions and the  $H_j(\cdot)$  are defined next. The functions  $\varphi_j$  depend on  $\xi = \omega h$  rather than  $\omega$ , and this is crucial to  $H(\cdot)$  being in  $\tilde{S}_h^0$ . We define

$$(4.4) \quad H_j(\cdot) := I + \sum_{n=1}^{\infty} (A_j^*(\cdot))^n (A_j(\cdot))^n + \sum_{\lambda \in \Lambda_j} P_{\lambda,j}^*(\cdot) P_{\lambda,j}(\cdot),$$

where

$$A_j^n = (2\pi i)^{-1} \int_{\Gamma_j} z^n R_z(A) dz.$$

Clearly,  $H(\cdot) \geq I$  and  $H(\cdot)$  is Hermitian. It is easy to verify that  $\|A_j^n(\cdot)\| \leq \rho_j^{n+1} \sup_{z \in \Gamma_j} \|R_z(\cdot)\| \leq c \rho_j^{n+1}$ , with  $c$  independent of all parameters (depending only on the resolvent constant). Therefore, the infinite series in (4.4) is

uniformly and absolutely convergent on  $B_{\varepsilon_j}(\chi_j)$ , and likewise, so is each appropriate derivative.

Next, we consider any particular  $j \in J$  and  $\lambda \in \Lambda_j$ , and we want to estimate the smooth function  $P_{\lambda,j}(\cdot)$ , on  $B_{\varepsilon_j}(\chi_j)$ , and also its derivatives. To do this, we shall need to classify the  $\lambda$ -group eigenvalues into clusters, in the same manner as in [8]. For each  $\mu \in \lambda$ -group with  $|\mu| = 1$  we define the cluster  $C_\mu := \{\mu\}$ . Of the remaining elements of the  $\lambda$ -group, choose any one with largest modulus, say  $\nu$  ( $|\nu| < 1$ ). We define the cluster  $C_\nu$  successively by the following procedure. First, put into  $C_\nu$  the eigenvalues  $\nu$  and any  $\mu \in \lambda$ -group with  $|\mu| < 1$  and  $|\mu - \nu| < (1 - |\nu|)/2m$ . Next, put into  $C_\nu$  any remaining  $\lambda$ -group eigenvalues  $\eta$  with  $|\eta| < 1$  and  $|\eta - \mu| < (1 - |\nu|)/2m$  for any  $\mu$  already in  $C_\nu$ , and continue. Eventually, we would exhaust the  $\lambda$ -group with a collection of clusters  $\{C_i\}_{i \in I}$  each built around some  $\mu_i \in \lambda$ -group with  $|\mu_i| < 1$  and  $\{C_k\}_{k \in K}$ , each built around the singleton  $\{\mu_k\}$  with  $|\mu_k| = 1$ . (This introduces the index sets  $I$  and  $K$ , which are disjoint.) This detailed classification of the  $\lambda$ -group allows us to apply the estimate (4.3).

As a reminder, at this point we have a fixed  $j \in J$  and  $\lambda \in \Lambda_j$ , and the above clusters depend on these  $j$  implicitly. We define the contours  $\{\gamma_k\}_{k \in K}$  to be any positively oriented circles centered at  $\mu_k$  and containing only  $\mu_k$  out of the spectrum. Next, it is clear that we can find unit-index contours,  $\{\gamma_i\}_{i \in I}$ , each surrounding only  $C_i$  out of the spectrum, respectively, and such that  $\text{dist}(\gamma_i, \sigma(A)) \geq (1 - |\mu_i|)/4m$ ,  $\text{dist}(\gamma_i, \mu_i) = (1 - |\mu_i|)/4m$ ,  $|\gamma_i| \leq 1 - |\mu_i|$ , and  $\text{dist}(\gamma_i, U) \geq (1 - |\mu_i|)/2$ , where  $U$  is the unit circle.

The operator  $P_{\lambda,j}(\cdot)$  is smooth on  $B_{\varepsilon_j}(\chi_j)$  since it is defined over a fixed contour  $\Gamma_{\lambda,j}$ . However, on  $B_{\varepsilon_j}(\chi_j)$ , this contour integral can be written in the form  $P_{\lambda,j}(\cdot) = \sum_{i \in I \cup K} Q_i(\cdot)$ , where  $Q_i(\cdot) := (2\pi i)^{-1} \int_{\gamma_i} R_z(A(\cdot)) dz$ . We will estimate the  $Q_i(\cdot)$  separately.

First, if  $i \in K$ , then the fact that  $R_z(A)$  has a simple pole at  $\mu_i$  yields  $\|Q_i\| = \|\lim_{\eta \rightarrow 0^+} \eta R_{(1+\eta)\mu_i}(A)\| \leq c$ , where  $c$  is the resolvent constant. For  $i \in I$  we utilize (4.3) and a straightforward computation to conclude that  $\|Q_i\| \leq |\gamma_i| \sup_{z \in \gamma_i} \|R_z(A)\| \leq c$ , where  $c$  depends on  $m, \varepsilon_j$ , and the resolvent constant. Thus  $P_{\lambda,j}(\cdot)$  is uniformly bounded on  $B_{\varepsilon_j}(\chi_j)$ . Clearly, appropriate derivatives of  $P_{\lambda,j}(\cdot)$  can also be estimated by the same method. One would only have to differentiate under the integral over the fixed contour  $\Gamma_{\lambda,j}$  and then split the integral over the separate contours  $\{\gamma_i\}_{i \in I \cup K}$ .

We have therefore shown that there is  $c > 0$  such that  $I \leq H(\cdot) \leq cI$ ,  $\partial_t H(t, \cdot) \leq cI$ , and also that  $H(t, \cdot) \in \tilde{S}_h^0$ .

It only remains to prove (3.5). We shall prove that  $H(\cdot)$  satisfies (3.5) by proving that each  $H_j(\cdot)$  from (4.4) does. On  $B_{\varepsilon_j}(\chi_j)$  we can represent  $A(\cdot)$  by

$$A(\cdot) = A_j(\cdot) + \sum_{\lambda \in \Lambda_j} \lambda P_{\lambda,j}(\cdot),$$

because of the semisimplicity at the unit circle. This last property is crucial to our proof. Thus, on  $B_{\varepsilon_j}(\chi_j)$ ,

$$\begin{aligned} A^* H_j A &= \left( A_j^* + \sum_{\lambda \in \Lambda_j} \bar{\lambda} P_{\lambda,j}^* \right) H_j \left( A_j + \sum_{\lambda \in \Lambda_j} \lambda P_{\lambda,j} \right) \\ &= \sum_{n=1}^{\infty} (A_j^*)^n (A_j)^n + \sum_{\lambda \in \Lambda_j} |\lambda|^2 P_{\lambda,j}^* P_{\lambda,j} \\ &\leq I + \sum_{n=1}^{\infty} (A_j^*)^n (A_j)^n + \sum_{\lambda \in \Lambda_j} P_{\lambda,j}^* P_{\lambda,j} = H_j. \end{aligned}$$

The proof is now complete.  $\square$

The preceding construction of the symmetrizer could be applied in cases where there is no differential equation in sight. However, in most situations we are interested in finite difference approximations for well-posed partial differential equations, in particular hyperbolic equations. The work of Kreiss [4] shows that stability follows from a match between the orders of accuracy and dissipation, under restrictions on the types of equations.

Michelson, in [7, Theorem 1.2], then showed that no match is needed between the accuracy and dissipation if the differential equation is strictly hyperbolic, i.e., possesses uniformly distinct (hence simple) eigenvalues. The result of Michelson is already a special case of Theorem 4.1. This is easy to see because the strict hyperbolicity forces the eigenvalues to be distinct near the unit circle.

If one would trace through the proof of Theorem 1.2 in [7], one would find techniques which are quite similar to ours, though definitely not the same. However, we have replaced the sharp Gårding inequality with the weak version and have also framed the whole result in such a way as to include much more general types of hyperbolic equations.

We now prove that Kreiss' condition can be utilized to construct the pseudodifference operator symmetrizer, even in our more general setting, thereby extending the original result to the full  $(x, t)$ -variable coefficient case and allowing (general) hyperbolic partial differential equations instead of only symmetric hyperbolic ones. Our method relies on properties of the eigenvalues of the amplification matrix and so could be extended to the multistep case simply by adding assumptions on the spurious eigenvalues in the same manner as in [17]. We prefer now to simplify the form of the finite difference operator in order to allow our proof to be more comprehensible to the reader. The construction of the symmetrizer given next differs from that in [4], although it is conjectured in [4, p. 337] that one could possibly attack the problem this way.

We now consider the following partial differential equation:

$$(4.5) \quad \begin{aligned} (\partial_t - p(t, x, \partial_x))u(t, x) &= f(t, x) \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d, \\ u(t, x) &= g(x), \end{aligned}$$

where  $\partial_x := (\partial_{x_1}, \dots, \partial_{x_d})^t$  and  $p(t, x, \partial_x) := \sum_{|\alpha| \leq p, \alpha \in \mathbb{N}^d} p_\alpha(t, x) \partial_x^\alpha$ . We assume that the symbol  $p(t, \cdot) \in C^1(\mathbb{R}_+, S^1)$ , where  $S^1$  represents the symbol class of pseudodifferential operators, say, from [14, Chapter 2]. However, we shall need to have the behavior as  $|x| \rightarrow \infty$  of symbols in  $S^1$  to be consistent with that of our symbols  $S_h^1$  (no specific asymptotic behavior is assumed in [14]). We let the reader fill in this small detail, and we shall not dwell on this point.

We assume that  $p(t, x, \partial_x)$  is hyperbolic in the following sense: there is  $T(t, \cdot) \in C^1(\mathbb{R}_+, S^0)$  such that  $T^{-1}(t, \cdot) \in C^1(\mathbb{R}_+, S^0)$  and

$$\operatorname{Re}(T(t, x, \omega)p(t, x, i\omega)T^{-1}(t, x, \omega)) = 0 \quad \forall (t, x, \omega) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d.$$

The finite difference operator  $q(t, h, x, T_x, T_t)$  of (3.1) is said to be dissipative with order  $\rho > 0$  if there are  $\alpha, c_0 > 0$  such that each root (in  $z$ ), say  $\lambda(\cdot)$  where  $\lambda$  depends on the parameters, of  $q(t, h, x, e^{i\xi}, z) = 0$  satisfies  $|\lambda(\cdot)| \leq e^{\alpha k}(1 - c_0|\xi|^\rho)$ , for all  $(t, h, x, \xi) \in \mathbb{R}_+ \times (0, h_0) \times \mathbb{R}^d \times [-\pi, \pi]^d$ .

The finite difference operator  $q(t, h, x, T_x, T_t)$  of (3.1) is said to be accurate with order  $\rho \geq 0$  if

$$(4.6) \quad \|sk - kp(t, x, i\omega) - kq(t, h, x, e^{i\omega h}, e^{sk})\| \leq c(t)(|\omega h|^{\rho+1} + |sk|^{\rho+1}),$$

for  $s \in \mathbb{C}$  with  $\operatorname{Re} s \geq 0$  and  $(t, h, x, \omega) \in \mathbb{R}_+ \times (0, h_0) \times \mathbb{R}^d \times \Gamma_h$ . We note that the variable  $s$  represents the Laplace transform dual variable with respect to  $t$ , as in [12], although we are not here transforming in  $t$ .

Our definitions of accuracy and dissipation are insensitive to whether  $q$  is multi- or single-step, and generally, the order of dissipation is even.

To ease the details, we now assume, as in [4], that the finite difference operator  $q(t, h, x, T_x, T_t)$  is single-step and equals  $k^{-1}T_t + q_1(t, h, x, T_x)$ , in which case  $A(t, h, x, \xi) = -kq_1(t, h, x, e^{i\xi})$ . We are interested in the question of whether the matching condition between accuracy and dissipation could yield a symmetrizer. It seems that in general the answer is no; there are many possibilities for pathological behavior of the eigenvalues of  $A(\cdot)$  near  $\xi = 0$ . To get around these problems, we now assume that none of the eigenvalues of  $A(\cdot)|_{\xi=0}$  are exceptional points, that is (cf. [3, 2.1.1]), a point of  $X$  where the number of eigenvalues changes (i.e., collisions at the unit circle). This assumption allows us to define smooth eigenprojection and eigennilpotent operators. One could most likely modify the next theorem to allow for exceptional points with more work, but we choose not to pursue that line because the results would be quite specialized.

**Theorem 4.2.** *Suppose the single-step finite difference operator is accurate with order  $\rho - 1$ , dissipative with order  $\rho$ ,  $\rho > 0$ , and there are no exceptional points at the unit circle. Then there exists a symmetrizer.*

*Proof.* The family  $\{A(\cdot)\}_X$  is no longer uniformly semisimple at the unit circle, so we cannot utilize the decomposition of  $A(\cdot)$  given at the end of the proof of

Theorem 4.1. We will modify the construction of  $H(\cdot)$  for  $\xi$  near 0 in a manner inspired by [8]. Accuracy and dissipation allow us to break up the set  $X$  into two pieces, one being  $X_\varepsilon := \{(t, h, x, \xi) \in \mathbb{R}_+ \times (0, h_0) \times \mathbb{R}^d \times B_\varepsilon(0)\}$  and the other being  $X \setminus X_\varepsilon$ , for  $\varepsilon > 0$  arbitrarily small. There is no problem constructing the local symmetrizer on  $X \setminus X_\varepsilon$ , by the dissipation assumption and the previous method for handling eigenvalues bounded away from the unit circle. In fact, the match between accuracy and dissipation allows enough control over the location of the eigenvalues so that the finite partition  $X_\varepsilon$  and  $X \setminus X_\varepsilon$  can be applied. We only need to construct the local symmetrizer on  $X_\varepsilon$ .

From (4.6), using the flexibility of the choice of  $s$  there, we see that the eigenvalues of  $A(\cdot)$  on  $X_\varepsilon$  ( $\varepsilon$  sufficiently small) must be contained in the region  $G(\xi) := \{z \in \mathbb{C} : |z - 1| < c_1|\xi|^\rho \text{ and } |z| \leq 1 - c_0|\xi|^\rho\}$ , for some  $c_0, c_1 > 0$ . For each  $\lambda(\cdot) \in \sigma(A(\cdot))$ , let  $\Gamma_\lambda$  be the positively oriented circle centered at  $\lambda$  and containing only  $\lambda$  out of the spectrum. We define the local symmetrizer on  $X_\varepsilon$  to be

$$H(\cdot) := \sum_{\lambda \in \sigma(A(\cdot))} H_\lambda, \quad H_\lambda := P_\lambda^* P_\lambda + \sum_{n=1}^\infty (D_\lambda^*)^n (D_\lambda)^n,$$

where  $P_\lambda := (2\pi i)^{-1} \int_{\Gamma_\lambda} R_z(A) dz$  and

$$D_\lambda := (2\pi i)^{-1} (1 - |\lambda|)^{-1} \int_{\Gamma_\lambda} (z - \lambda) R_z(A) dz.$$

We must first check that  $H$  is well defined. The assumption of no exceptional points means that these operators are smooth functions of  $\chi \in X_\varepsilon$ , cf. [3, 2.1]. We need to classify  $\sigma(A)$  on  $X_\varepsilon$  by the cluster method and then utilize (4.3). First we note that the accuracy and dissipation assumptions guarantee that the resolvent condition holds (pointwise); this was shown in [9, Theorem 1 or 4, Theorem 4], and amounts to using the hyperbolicity of symbol  $p(t, x, i\omega)$  together with  $s = 0$  in (4.6). The resolvent condition is all that is needed for the estimate (4.3) to go through.

For each  $\chi \in X_\varepsilon$  we can change our chain  $\bigcup_{\lambda \in \sigma(A(\lambda))} \Gamma_\lambda$  to a certain chain of contours by using the cluster method of the last theorem. It is easy to see that we can divide the spectrum  $\sigma(A(\chi))$  into clusters  $\{C_\mu\}_{\mu \in M(\chi)}$ , for some set  $M(\chi) \subset \sigma(A(\chi))$ , such that there are unit-index contours  $\{\gamma_\mu\}_{\mu \in M(\chi)}$  which satisfy  $\sup_{z \in \gamma_\mu} |z - \lambda|(1 - |\lambda|)^{-1} < 1/2$ , for  $\lambda \in C_\mu$ . Therefore, by changing the chain of contours, we have

$$\begin{aligned} \sum_{\lambda \in \sigma(A(\chi))} D_\lambda^n &= \sum_{\mu \in M(\chi)} \sum_{\lambda \in C_\mu} D_\lambda^n \\ &= \sum_{\mu \in M(\chi)} (2\pi i)^{-1} \int_{\gamma_\mu} \sum_{\lambda \in C_\mu} (z - \lambda)^n (1 - |\lambda|)^{-n} R_z(A) dz. \end{aligned}$$



We obtain from (4.3)

$$\left\| \sum_{\mu \in M(\chi)} \sum_{\lambda \in C_\mu} D_\lambda^n \right\| \leq \sum_{\mu \in M(\chi)} c_m \left( \sup_{\substack{z \in \tilde{\gamma}_\mu \\ \lambda \in C_\mu}} \{|z - \lambda|(1 - |\lambda|)^{-1}\} \right)^n \leq c_m (1/2)^n.$$

Thus, the series defining  $H$  are uniformly convergent.

The only property of Definition 3.2 which is not completely straightforward to check is (3.5). The family can be decomposed as

$$A = \sum_{\lambda \in \sigma(A)} (\lambda P_\lambda + (1 - |\lambda|)D_\lambda).$$

We have:

$$\begin{aligned} A^*HA &= \sum_{\lambda \in \sigma(A)} (\bar{\lambda}P_\lambda^* + (1 - |\lambda|)D_\lambda^*)H_\lambda(\lambda P_\lambda + (1 - |\lambda|)D_\lambda) \\ &\leq \sum_{\lambda \in \sigma(A)} (|\lambda|^2 H_\lambda + 2 \operatorname{Re}(\bar{\lambda}H_\lambda(1 - |\lambda|)D_\lambda) + (1 - |\lambda|)^2 H_\lambda). \end{aligned}$$

Since  $0 \leq (\lambda|\lambda|^{-1}P_\lambda - D_\lambda)^*H_\lambda(\lambda|\lambda|^{-1}P_\lambda - D_\lambda)$  and  $D_\lambda^*H_\lambda D_\lambda \leq H_\lambda$ , we immediately obtain  $2 \operatorname{Re}(\bar{\lambda}H_\lambda(1 - |\lambda|)D_\lambda) \leq 2|\lambda|(1 - |\lambda|)H_\lambda$ . This implies

$$A^*HA \leq \sum_{\lambda \in \sigma(A)} (|\lambda| + 1 - |\lambda|)^2 H_\lambda = H,$$

and completes the proof.  $\square$

We conclude with some comments. The problem of constructing a symmetrizer has not been completely solved here because there is a need for examples in which a smooth symmetrizer cannot be constructed in order to sharpen our understanding of stability theory for general types of equations. Our machinery provides only sufficient conditions, which are practical, but it seems that there is the possibility for rather pathological behavior of the eigenvalues of  $A(\cdot)$  even in the presence of matching dissipation and accuracy.

Regarding condition  $[N]$  from [12], it seems that it may not be the case that  $H$  can always be constructed to be smooth, even if the matrix  $N$  could be. This dilemma stems both from the fact that the resolvent condition in the Kreiss Matrix Theorem does not necessarily imply the existence of a smooth symmetrizer  $H$  (condition  $[S]$  causes trouble, cf. [10]), and also from the following (quite elementary) example. Let

$$A := \begin{pmatrix} 1/2 & 1 \\ 0 & 1/2 \end{pmatrix};$$

then  $A$  satisfies condition  $[N]$  of [12] with the matrix  $N$  taken to be the identity. However, the 2-norm of  $A$  is strictly greater than one, and so  $H$

must be different from the identity matrix. One can prove these claims by noting that the numerical radius of  $A$  equals the spectral radius of its real part, a fact which follows from the nonnegativity of the entries of  $A$ ; details are in [16]. Since the spectral radius of  $\operatorname{Re} A$  is one, it is easy to check that the matrix  $N$  from [12] can be taken as the identity (it is convenient here to first use Tadmor's condition, cf. [13]). However,  $\|A\| > 1$ , so  $H$  cannot be the identity.

So there is a curious question before us: For variable coefficients, are the conditions  $[H]$  of Definition 3.2 and  $[N]$  of [12] (suitably modified for the variable-coefficient context) equivalent?

#### ACKNOWLEDGMENT

Professor John C. Strikwerda, the author's thesis advisor, has exerted a significant influence on this work, especially in the proof of Theorem 4.1. His contribution is greatly appreciated.

#### BIBLIOGRAPHY

1. K. P. Bube and J. C. Strikwerda, *Interior regularity estimates for elliptic systems of difference equations*, SIAM J. Numer. Anal. **20** (1983), 653–670.
2. J. Chazarain and A. Piriou, *Introduction to the theory of linear partial differential equations*, North-Holland, New York, 1982.
3. T. Kato, *Perturbation theory for linear operators*, second ed., Springer-Verlag, New York, 1976.
4. H.-O. Kreiss, *On difference approximations of the dissipative type for hyperbolic differential equations*, Comm. Pure Appl. Math. **17** (1964), 335–353.
5. —, *Über sachgemäße Cauchyprobleme*, Math. Scand. **13** (1963), 109–128.
6. P. D. Lax and L. Nirenberg, *On stability for difference schemes: a sharp form of Gårding's inequality*, Comm. Pure Appl. Math. **19** (1966), 473–492.
7. D. Michelson, *Stability theory of difference approximations for multidimensional initial-boundary value problems*, Math. Comp. **40** (1983), 1–45.
8. J. J. H. Miller, *On power bounded operators and operators satisfying a resolvent condition*, Numer. Math. **10** (1967), 389–396.
9. B. Parlett, *Accuracy and dissipation in difference schemes*, Comm. Pure Appl. Math. **19** (1966), 111–123.
10. R. D. Richtmyer and K. W. Morton, *Difference methods for initial value problems*, second ed., Wiley-Interscience, New York, 1967.
11. H. Shintani and K. Tomoeda, *Stability of difference schemes for nonsymmetric linear hyperbolic systems with variable coefficients*, Hiroshima J. Math. **7** (1977), 309–378.
12. J. C. Strikwerda and B. A. Wade, *An extension of the Kreiss Matrix Theorem*, SIAM J. Numer. Anal. **25** (1988), 1272–1278.
13. E. Tadmor, *The equivalence of  $L_2$ -stability, the resolvent condition, and strict  $H$ -stability*, Linear Algebra Appl. **41** (1981), 151–159.
14. M. E. Taylor, *Pseudodifferential operators*, Princeton Univ. Press, Princeton, N. J., 1984.
15. R. Vichnevetsky and J. B. Bowles, *Fourier analysis of numerical approximations of hyperbolic equations*, Studies in Applied Math., vol. 5, SIAM, Philadelphia, Pa., 1982.

16. B. A. Wade, *Stability and sharp convergence estimates for symmetrizable difference operators*, Ph. D. dissertation, University of Wisconsin-Madison, 1987.
17. O. B. Widlund, *On the stability of parabolic difference schemes*, *Math. Comp.* **19** (1965), 1–13.

MATHEMATICAL SCIENCES INSTITUTE, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853-2602

*Current address:* Department of Mathematical Sciences, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, Wisconsin 53201