

## HERMITE METHODS FOR HYPERBOLIC INITIAL-BOUNDARY VALUE PROBLEMS

JOHN GOODRICH, THOMAS HAGSTROM, AND JENS LORENZ

**ABSTRACT.** We study arbitrary-order Hermite difference methods for the numerical solution of initial-boundary value problems for symmetric hyperbolic systems. These differ from standard difference methods in that derivative data (or equivalently local polynomial expansions) are carried at each grid point. Time-stepping is achieved using staggered grids and Taylor series. We prove that methods using derivatives of order  $m$  in each coordinate direction are stable under  $m$ -independent CFL constraints and converge at order  $2m + 1$ . The stability proof relies on the fact that the Hermite interpolation process generally decreases a seminorm of the solution. We present numerical experiments demonstrating the resolution of the methods for large  $m$  as well as illustrating the basic theoretical results.

### 1. INTRODUCTION

Challenging problems in large scale scientific computing typically involve multiple spatial and temporal scales. For hyperbolic problems this entails the propagation of short waves over many wavelengths. In such cases, one expects high-order methods to be particularly efficient. In recent years, a number of high-order finite-element methods have been developed and applied to wave propagation problems [3, 4, 13]. In contrast, the application of high-order nodal based finite-difference methods has lagged. A possible explanation for this is difficulties associated with the application of boundary conditions. (See, e.g., [18] for a somewhat complicated approach to dealing with this issue.)

In this work we develop the basic theory for a different class of node-based methods of arbitrary order. The ingredients of our method, which could be applied independently, are:

- i:** The association of derivative data, or equivalently local polynomials, with each grid point. Using Hermite interpolation, this allows us to derive methods of arbitrary order using  $2\mu$ -point stencils in  $\mu$  space dimensions.

---

Received by the editor January 7, 2004 and, in revised form, February 4, 2005.

2000 *Mathematics Subject Classification.* Primary 35G20.

*Key words and phrases.* High-order methods, hyperbolic problems, stability.

The second author was supported in part by NSF Grants DMS-9971772, DMS-0306285, NASA Contract NAG3-2692, and the Institute for Computational Mechanics in Propulsion (ICOMP), NASA Glenn Research Center, Cleveland, OH. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of NSF or NASA.

The third author was supported in part by DOE Grant DE-FG03-98ER25235 and the Institute for Computational Mechanics in Propulsion (ICOMP), NASA Glenn Research Center.

- ii: The use of high-order Taylor series in time to evolve the solution on a staggered space-time mesh. Employing the Cauchy-Kowalewskaya recursion (e.g., [6, 12]), time derivatives of a local solution are calculated using the Hermite interpolant and the partial differential equation itself. For constant coefficient linear systems in any number of space dimensions, the Taylor series truncates for piecewise polynomial data except at the break-points. Then the evolution is exact so long as a CFL-type condition is satisfied. Here the use of a staggered mesh is crucial, as it allows us to evolve the data at a point where the piecewise polynomial is smooth.

Certainly these ingredients have been used elsewhere in different contexts: methods for hyperbolic systems using piecewise cubic Hermite interpolation have been proposed by Yabe and coworkers in [19, 20, 21] and studied further in [17]; high-order Taylor series methods for ordinary differential equations are discussed in [15]; and the use of staggered grids to facilitate time-integration of hyperbolic equations is a key part of the so-called central difference schemes for hyperbolic conservation laws [14]. However, we believe their combination with arbitrary-order Hermite interpolation to solve hyperbolic systems and the convergence analysis we provide is new.

At first glance Hermite methods might appear less efficient than high-order methods requiring only function data; after all we are introducing many additional degrees of freedom. However, we will demonstrate that the Hermite methods allow significant subgrid resolution, so that the number of degrees of freedom required per wavelength is comparable to other techniques. An advantage of our approach is that the time-step stability restrictions are determined by the cell size independent of the order. Thus much larger time steps can be taken.

The outline of the remainder of the paper is as follows. In section 2 we describe in detail the basic method for space-periodic problems. In section 3 we analyze our method for constant coefficient systems in one space dimension. The convergence proof is based on the observation that Hermite interpolation generally decreases a seminorm of the solution. In sections 4 and 5 we generalize the analysis to systems with variable coefficients and nonperiodic boundary conditions, which is conceptually straightforward but technically detailed. Again we emphasize that the boundary treatment is completely natural, requiring no changes in the basic discretization. In section 6 we consider systems in multiple space dimensions using tensor-product grids. Here we must generalize the seminorm in an interesting way to obtain the convergence theorem. Finally, in section 7 we present a number of simple numerical experiments which demonstrate the resolving power of the methods as well as the basic theoretical results. An appendix collecting various facts about Hermite interpolation and solutions of hyperbolic systems which are used in the analysis is also included. Proofs of these facts or proof outlines are given for completeness.

We note that the methods are applicable to nonlinear hyperbolic-parabolic systems. Extension of the theory to cover that case as well as applications of the method to the study of compressible flows will be reported elsewhere.

## 2. HERMITE-TAYLOR METHODS FOR PERIODIC PROBLEMS

To begin we describe the method for the simple model problem

$$(1) \quad u_t = au_x, \quad u(x, 0) = f(x),$$

where  $f(x)$  is a smooth  $2\pi$ -periodic function and  $a$  is a real constant. For reasons detailed below, we use staggered space-time grids and our grid data includes not only approximate function values but scaled space derivatives of orders 1 through  $m$ . Precisely let:

$$(2) \quad 0 < x_{\frac{1}{2}} < x_1 < x_{\frac{3}{2}} < x_2 < \dots < x_{N-\frac{1}{2}} < x_N < x_{\frac{1}{2}} + 2\pi = x_{N+\frac{1}{2}}.$$

Set

$$\begin{aligned} \tilde{G}_h &= \left\{ x_{j+\frac{1}{2}} : j \in \mathbb{Z} \right\}, & x_{j+\frac{1}{2}+N} &= x_{j+\frac{1}{2}} + 2\pi, \\ G_h &= \left\{ x_j : j \in \mathbb{Z} \right\}, & x_{j+N} &= x_j + 2\pi, \end{aligned}$$

$$(3) \quad h_{j+\frac{1}{2}} = x_{j+1} - x_j, \quad h_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}.$$

Starting at time  $t_n$ , a full step of the method proceeds as follows (see Figure 1):

- i:** On each cell  $[x_j, x_{j+1}]$  (i.e., each interval bounded by adjacent points on the current grid), construct the Hermite interpolant of the function and derivative data at the nodes  $x_j, x_{j+1}$ . This will be a polynomial of degree  $2m + 1$ . We express it as an expansion in local coordinates centered at the staggered grid point,  $x_{j+\frac{1}{2}}$ :

$$(4) \quad p_{j+\frac{1}{2}}(x, t_n) = \sum_{l=0}^{2m+1} c_{l0} \left( x - x_{j+\frac{1}{2}} \right)^l.$$

- ii:** At each point on the staggered grid,  $x_{j+\frac{1}{2}}$ , compute a temporal Taylor series,

$$(5) \quad p_{j+\frac{1}{2}}^n(x, t) = \sum_{l=0}^{2m+1} \sum_{s=0}^{2m+1-l} c_{ls} \left( x - x_{j+\frac{1}{2}} \right)^l (t - t_n)^s.$$

To compute the coefficients  $c_{ls}$  we note that time derivatives of a smooth solution to (1) may be directly expressed in terms of space derivatives.

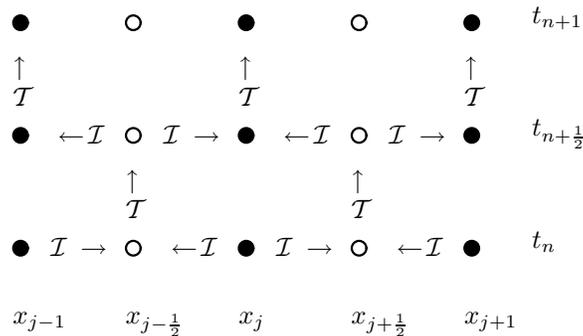


FIGURE 1. Schematic of the numerical process for a full time step. Solid circles represent the base mesh and open circles represent the dual mesh.  $\mathcal{I}$  is the Hermite interpolation operator, and  $\mathcal{T}$  is the Taylor evolution operator.

Writing this relationship recursively:

$$(6) \quad \frac{\partial^s u}{\partial t^s} = a \frac{\partial^s u}{\partial t^{s-1} \partial x}.$$

Applying (6) to  $p_{j+\frac{1}{2}}$  we derive a simple recursion for the coefficients  $c_{ls}$ ,  $s = 1, \dots, 2m + 1$  (recall we have computed the  $c_{l0}$ ):

$$(7) \quad c_{ls} = a \frac{(l+1)}{s} c_{l+1, s-1}, \quad l = 0, \dots, 2m + 1 - s.$$

We note that we have included all nonzero terms in (5) as computed by (7).

iii: Evaluate the series at  $t_{n+\frac{1}{2}} = t_n + \frac{k}{2}$  to derive:

$$(8) \quad \frac{1}{l!} \frac{\partial^l u}{\partial x^l}(x_{j+\frac{1}{2}}, t_{n+\frac{1}{2}}) \approx \sum_{s=0}^{2m+1-l} c_{ls} \left(\frac{k}{2}\right)^s, \quad l = 0, \dots, m.$$

Since we have included all terms in the Taylor series, the right-hand side of (8) represents the exact values of the solution of the governing equation with the piecewise polynomial data so long as the domain of dependence of the solution at  $(x_{j+\frac{1}{2}}, t_{n+\frac{1}{2}})$  on data at  $t_n$  is included in the cell, that is, if the CFL-type condition

$$(9) \quad a \frac{k}{2} < \min(x_{j+1} - x_{j+\frac{1}{2}}, x_{j+\frac{1}{2}} - x_j)$$

is satisfied.

iv: Repeat the process on the dual grid starting at  $t_{n+\frac{1}{2}}$  producing data on the original grid at time  $t_{n+1}$ .

Functionally, the method requires two procedures: an interpolation procedure,  $\mathcal{I}$ , which, on each cell, transforms the derivative data at the endpoints to the local expansion coefficients,  $c_{l0}$ , of (4) and an evolution procedure,  $\mathcal{T}$ , which applies the recursion (7) to compute the space-time expansion coefficients  $c_{ls}$  and updates the data using (8). Generalizations of the method to more complex problems require generalizations of the evolution procedure,  $\mathcal{T}$ , while the interpolation procedure remains unchanged. We note that by evolving the piecewise polynomial data on the dual grid we are able to consider a point where the approximate solution is smooth, and for constant coefficient problems, satisfaction of the CFL constraint results in the evolution being exact on the grid.

**2.1. Extensions.** We now consider how to generalize the method to systems, variable coefficients, and multiple dimensions. The treatment of boundary conditions will be discussed in section 5.

**2.1.1. Constant coefficient systems.** We begin by looking at the constant coefficient system:

$$(10) \quad u_t = Au_x, \quad A = A^T \in R^{d \times d},$$

$$(11) \quad u(x, 0) = f(x), \quad f(x + 2\pi) = f(x).$$

In this case the data consists of approximate function and scaled derivative values for each of the  $d$  components of the solution on the grid. The Hermite interpolation step leading to (4) is carried out componentwise and is thus identical to the scalar

case. The result is an  $R^d$ -valued function each of whose components is a polynomial of degree  $2m + 1$ . From the perspective of the data the coefficients  $c_{l0}$  are now  $d$ -vectors.

Similarly, the space-time polynomial in (5) is vector-valued, described by coefficients  $c_{ls} \in R^d$ . They can be recursively computed using:

$$(12) \quad \frac{\partial^s u}{\partial t^s} = A \frac{\partial^s u}{\partial t^{s-1} \partial x},$$

which implies the vector recursion:

$$(13) \quad c_{ls} = \frac{(l+1)}{s} A c_{l+1, s-1}.$$

As in the scalar case, it is easy to show that (5) contains all nonzero terms in a temporal Taylor series for a solution of (10) with polynomial data. The update is exact on the dual grid if the appropriate CFL condition, determined by the spectral radius of  $A$ ,  $\rho(A)$ , is satisfied. This is stated in the following lemma.

**Lemma 2.1.** *Suppose  $p(x, t_n)$  is an  $R^d$ -valued degree  $2m + 1$  piecewise polynomial function of  $x$  with breaks at  $x_0 < x_1 < \dots < x_N$ . If  $p(x, t)$  satisfies (10) for  $t \geq t_n$  and the time step,  $k$ , satisfies the CFL condition*

$$\rho(A) \frac{k}{2} < \min \left( x_{j+1} - x_{j+\frac{1}{2}}, x_{j+\frac{1}{2}} - x_j \right),$$

then  $p(x_{j+\frac{1}{2}}, t_{n+\frac{1}{2}})$  is given by (5), where the coefficients satisfy (13).

*Proof.* The result follows directly from (13) and standard domain of dependence considerations [16, Ch. 3,6]. □

2.1.2. *Variable coefficients.* We now consider the spatially-periodic, forced, symmetric hyperbolic system:

$$(14) \quad u_t = A(x, t)u_x + B(x, t)u + F(x, t), \quad A = A^T, \quad u(x, 0) = f(x),$$

$$(15) \quad f(x + 2\pi) = f(x), \quad F(x + 2\pi, t) = F(x, t),$$

$$(16) \quad A(x + 2\pi, t) = A(x, t), \quad B(x + 2\pi, t) = B(x, t).$$

Here  $u, f, F \in R^d$ ,  $A, B \in R^{d \times d}$ .

As in the constant coefficient case, we compute  $R^d$ -valued Hermite interpolants on each cell; i.e., we compute the coefficient vectors  $c_{l0}$ . However, a terminating local space-time expansion of the form (5) no longer exists. Thus we give up on exact local evolution of the data, specifying instead the degree of the Taylor series in time we wish to use. To be precise, denote by  $p(x, t)$ ,  $t_n \leq t \leq t_{n+\frac{1}{2}}$ , the solution of (14) with initial data  $p(x, t_n)$ , which is a piecewise polynomial function of degree  $2m + 1$  with breaks  $x_j$  as before. Let  $x_j < x_{j+1/2} < x_{j+1}$ . Then for a time step satisfying the CFL condition

$$(17) \quad \max_{x \in [x_j, x_{j+1}], t \in [t_n, t_{n+\frac{1}{2}}]} \rho(A) \frac{k}{2} < \min \left( x_{j+1} - x_{j+\frac{1}{2}}, x_{j+\frac{1}{2}} - x_j \right),$$

we choose  $q \geq 2m + 1$  and replace (8) by the approximate expression

$$(18) \quad \frac{1}{l!} \frac{\partial^l u}{\partial x^l} \left( x_{j+\frac{1}{2}}, t_{n+\frac{1}{2}} \right) \approx \sum_{s=0}^{q-l} c_{ls} \left( \frac{k}{2} \right)^s \equiv \frac{1}{l!} \mathcal{T}^{(q,l)} p_{j+\frac{1}{2}}^n.$$

Recursions for the space-time coefficients  $c_{ls}$  follow from differentiating (14) in space and time and applying Leibniz' rule (see, e.g., [8, Ch. 9]):

$$(19) \quad c_{ls} = \frac{1}{s} \sum_{i=0}^l \sum_{j=0}^{s-1} (i+1) A^{[l-i, s-1-j]} c_{i+1, j} + \frac{1}{s} \sum_{i=0}^l \sum_{j=0}^{s-1} B^{[l-i, s-1-j]} c_{ij} + \frac{1}{s} F^{[l, s-1]}.$$

Here, for any function  $G$ , vector or matrix-valued, we are using the notation of [8]:

$$(20) \quad G^{[i, j]}(x, t) = \frac{1}{i!j!} \frac{\partial^{i+j} G}{\partial x^i \partial t^j}(x, t).$$

In (19) all functions are evaluated at  $(x_{j+\frac{1}{2}}, t_n)$ . We note that the right-hand side of (19) only involves  $t$ -derivatives of order up to  $s - 1$ , and thus the recursion may be initiated by space derivatives of  $p$  alone. If we choose  $q > 2m + 1$ , then space derivatives of  $p$  of order greater than  $2m + 1$  appear. Although these are zero initially, their time derivatives generally are not. We finally note that the degree of the Taylor series in (18) decreases as the order of the spatial derivative increases. In some implementations for nonlinear systems it may be advantageous to use the same degree for all derivatives. Such issues will be discussed in a later work.

2.1.3. *Multiple dimensions.* We now consider a symmetric hyperbolic system in  $R^\mu$ :

$$(21) \quad u_t = \sum_{j=1}^{\mu} A_j(x, t) u_{x_j} + B(x, t) u + F(x, t), \quad A_j = A_j^T, \quad u(x, 0) = f(x),$$

under the assumption that all functions are  $2\pi$ -periodic  $d$ -vectors or  $d \times d$  matrices in all space variables. We introduce tensor-product grids of type (2) with  $N_i$  points in each coordinate direction:

$$(22) \quad \tilde{G}_h = \bigotimes_{i=1}^s \tilde{G}_{h, i}, \quad G_h = \bigotimes_{i=1}^s G_{h, i},$$

$$\tilde{G}_{h, i} = \left\{ x_{i, j+\frac{1}{2}} : j \in \mathbb{Z} \right\}, \quad x_{i, j+\frac{1}{2}+N_i} = x_{i, j+\frac{1}{2}} + 2\pi,$$

$$G_{h, i} = \left\{ x_{i, j} : j \in \mathbb{Z} \right\}, \quad x_{i, j+N_i} = x_{i, j} + 2\pi.$$

We set

$$(23) \quad h_i = \max_j \max((x_{i, j+1} - x_{i, j}), (x_{i, j+\frac{1}{2}} - x_{i, j-\frac{1}{2}}))$$

and

$$(24) \quad h = \max_i h_i.$$

Thus a cell on the base grid takes the form

$$(25) \quad \bigotimes_{i=1}^{\mu} [x_{i,j_i}, x_{i,j_i+1}],$$

and the coordinates of its interior dual grid point are

$$(26) \quad (x_{1,j_1+\frac{1}{2}}, x_{2,j_2+\frac{1}{2}}, \dots, x_{\mu,j_\mu+\frac{1}{2}}).$$

On these grids one can define tensor-product Hermite interpolants of degree  $\mu(2m + 1)$ , that is, polynomials of degree  $\mu(2m + 1)$  that are sums of products of polynomials of degree  $2m + 1$  in each variable  $x_i$ . The nodal data used to define the interpolants is

$$(27) \quad \left( \prod_{i=1}^{\mu} \frac{1}{\nu_i!} D_i^{\nu_i} \right) u, \quad 0 \leq \nu_i \leq m.$$

(Here we introduce the notation  $D_i = \partial/\partial x_i$ .) Obviously, the computation of each interpolant can be reduced to the composition of one-dimensional Hermite interpolation operators along coordinate directions. For example, in two space dimensions one might first interpolate in the  $x_1$ -direction along the top and bottom of each cell for each order,  $\nu_2$ , of the  $x_2$ -derivative. That is, one would compute  $2m + 2$  polynomials centered at  $x_{1,j_1+\frac{1}{2}}$ , half located at the top and half located at the bottom. Thus to each power,  $(x_1 - x_{1,j_1+\frac{1}{2}})^l$ , one associates  $2m + 2$  coefficients computed with  $D_2^{\nu_2} u$  at  $x_{2,j_2}$  and  $x_{2,j_2+1}$ . Applying the one-dimensional interpolation operator to these produces the analogue of (4):

$$(28) \quad p_{j_1+\frac{1}{2},j_2+\frac{1}{2}}(x_1, x_2, t_n) = \sum_{l_1=0}^{2m+1} \sum_{l_2=0}^{2m+1} c_{l_1 l_2 0} (x_1 - x_{1,j_1+\frac{1}{2}})^{l_1} (x_2 - x_{2,j_2+\frac{1}{2}})^{l_2},$$

or in  $R^\mu$ :

$$(29) \quad p_{j_1+\frac{1}{2},\dots,j_\mu+\frac{1}{2}}(x_1, \dots, x_\mu, t_n) = \sum_{l_1=0}^{2m+1} \cdots \sum_{l_\mu=0}^{2m+1} c_{l_1 \dots l_\mu 0} \prod_{i=1}^{\mu} (x_i - x_{i,j_i+\frac{1}{2}})^{l_i}.$$

To evolve these polynomials in time, we look at the special case  $A_j = \text{constant}$ ,  $B = 0$ ,  $F = 0$ . Then, if we insist on exact evolution, the Taylor series in time must in general be of degree  $\mu(2m + 1)$ , which is of significantly higher degree than required simply to attain the order of accuracy  $2m + 1$ . We note, however, that carrying the series to higher degree does not require any additional spatial data. However, for variable coefficient problems, it does lead to additional derivatives of coefficients to be evaluated. In the final section we will present some numerical experiments related to the degree of the temporal Taylor series. Thus in general we will as above choose a temporal degree  $q \geq 2m + 1$  and approximate the updated data by the analogue of (18):

$$(30) \quad \left( \prod_{i=1}^{\mu} \frac{1}{\nu_i!} D_i^{\nu_i} \right) u(x_{1,j_1+\frac{1}{2}}, \dots, x_{\mu,j_\mu+\frac{1}{2}}, t_{n+\frac{1}{2}}) \approx \sum_{s=0}^q c_{\nu_1 \dots \nu_\mu s} \left( \frac{k}{2} \right)^s.$$

(For  $q$  sufficiently large some of the terms in the sum above will be identically zero, but we include them for ease of notation.)

In the constant coefficient case the coefficients again satisfy a simple recursion:

$$(31) \quad c_{l_1 \dots l_\mu s} = \sum_{j=1}^{\mu} \frac{(l_j + 1)}{s} A_j c_{l_1 \dots l_{j-1} l_j + 1 l_{j+1} \dots l_\mu s - 1},$$

with a more complex formula generalizing (19) holding in general. In either instance one generates the coefficients for increasing  $s$  starting with (29), the output of the Hermite interpolation.

To complete our discussion, we must also consider the domain of dependence of the updated solution at a mesh point over a half-step. Given a unit vector  $e \in R^\mu$  the maximum wave speed in the direction  $e$  is given by

$$(32) \quad c(e) = \rho \left( \sum_{j=1}^{\mu} e_j A_j \right),$$

leading to the CFL constraint

$$(33) \quad e_i c(e) \frac{k}{2} < \min_j \left( x_{i, j_i + 1} - x_{i, j_i + \frac{1}{2}}, x_{i, j_i + \frac{1}{2}} - x_{i, j_i} \right), \quad i = 1, \dots, \mu.$$

Note that for an isotropic problem on a uniform mesh this reduces to the one-dimensional CFL condition (17). Also, in the case  $A_j$  constant and  $B = 0$ ,  $F = 0$  we have the generalization of Lemma 2.1:

**Lemma 2.2.** *Suppose  $p(x_1, \dots, x_\mu, t_n)$  is an  $R^d$ -valued degree  $\mu(2m + 1)$  tensor-product piecewise polynomial function of  $x$  with breaks at the edges of the tensor-product grid described above. If  $p$  satisfies (21) with  $A_j$  constant and  $B = 0$ ,  $F = 0$  for  $t \geq t_n$  and the time step,  $k$ , satisfies the CFL condition (33), then  $p(x_{1, j_1 + \frac{1}{2}}, \dots, x_{\mu, j_\mu + \frac{1}{2}}, t_{n + \frac{1}{2}})$  and its derivatives are given by (30) with  $q = \mu(2m + 1)$  and coefficients satisfying (31).*

**2.2. Algorithm.** We are now in a position to give a complete description of the algorithm for space-periodic problems. Note that the details of the evolution steps depend on the particular case as described above. Of course many variations on the algorithm, such as variable time-stepping, are possible.

**Initialize:** Determine parameters and initial conditions.

- 1:** Choose a primal and dual grid, as well as a time step,  $k$ , satisfying the appropriate CFL restrictions.
- 2:** Choose the maximal derivative order in each direction,  $m$ , to be carried on the grid as well as the order,  $q$ , of the temporal Taylor series.
- 3:** Initialize the data on the primal grid. That is, compute and store the required function and derivative values from the initial condition,  $f$ , at the nodes of the primal grid.

**Time Step:** For  $n = 0, \dots, N_{\text{steps}} - 1$

- 1:** For all cells on the primal grid
  - i:** Use Hermite interpolation to transform the data at the cell vertices to the Taylor coefficients,  $c_{l_1 \dots l_\mu 0}$ , of the interpolating polynomial at the node of the dual grid contained within the cell.
  - ii:** Use the recursion implied by the partial differential equation to compute the space-time Taylor coefficients,  $c_{l_1 \dots l_\mu s}$ , of the solution of the PDE with the piecewise polynomial initial data, centered at the node of the dual grid contained within the cell.

- iii: Evaluate the series at  $t = t_n + \frac{k}{2}$  to obtain the required function and derivative data at the included dual grid node.
- 2: For all cells on the dual grid
  - i: Use Hermite interpolation to transform the data at the cell vertices to the Taylor coefficients,  $c_{l_1 \dots l_\mu 0}$ , of the interpolating polynomial at the node of the primal grid contained within the cell.
  - ii: Use the recursion implied by the PDE to compute the space-time Taylor coefficients,  $c_{l_1 \dots l_\mu s}$ , of the solution of the PDE with the piecewise polynomial initial data, centered at the node of the primal grid contained within the cell.
  - iii: Evaluate the series at  $t = t_{n+1} = t_n + k$  to obtain the required function and derivative data at the included primal grid node.

### 3. CONVERGENCE ANALYSIS

#### FOR CONSTANT COEFFICIENT SYSTEMS IN ONE SPACE DIMENSION

In this section we prove the basic convergence result in the simplest case: constant coefficient systems in one space dimension (10)–(11). The proof is facilitated by considering the discrete evolution of the interpolants rather than the evolution of the grid data that determines the interpolants. Thus we express the algorithm in terms of the interpolation operators  $\mathcal{I}$ ,  $\tilde{\mathcal{I}}$ , defined on the primal and dual grids respectively, and the exact solution operator,  $\mathcal{S}$ , over a half step  $k/2$ . Recalling Lemma 2.1 and assuming that the time step satisfies the CFL condition hypothesized there, the evolution of the grid data by Taylor series followed by interpolation is equivalent to:

- i: Let  $p^0 = \mathcal{I}f$ ;
- ii: For  $n = 0, 1, \dots$ , let

$$(34) \quad p^{n+\frac{1}{2}} = \tilde{\mathcal{I}}\mathcal{S}p^n,$$

$$(35) \quad p^{n+1} = \mathcal{I}\mathcal{S}p^{n+\frac{1}{2}}.$$

Obviously, the stability and accuracy of the proposed method are determined by the properties of Hermite interpolation. Denote the  $L_2$ -inner product and norm of  $2\pi$ -periodic functions by

$$(f, g) = \int_{-\pi}^{\pi} f(x)^T g(x) dx, \quad \|g\|^2 = (g, g),$$

and the Sobolev norms on  $H_{\text{per}}^q$ , the space of  $2\pi$ -periodic functions with  $q$  weak derivatives in  $L_2$ , by

$$\|f\|_{H^q}^2 = \sum_{\nu=0}^q \|D^\nu f\|^2.$$

In some instances we will use these norms, as well as the uniform norms  $\|\cdot\|_{C^q}$ , on subintervals  $(a, b)$ . In such cases the subinterval will be explicitly included in the index:  $\|f\|_{H^q(a,b)}$ ,  $\|f\|_{C^q(a,b)}$ . Also, let

$$(36) \quad h = \max(\max_j h_j, \max_j h_{j+\frac{1}{2}}).$$

We begin with standard results on the interpolation error which follow directly from the Peano kernel representation. Such results are proven in more generality in [2] in one and two space dimensions, and extensions to higher dimensions via

induction arguments are straightforward. Note that interpolation on each of the meshes has the same properties, so we state results for  $\mathcal{I}$  only.

**Lemma 3.1.** *The Hermite interpolation operators,  $\mathcal{I}$  and  $\tilde{\mathcal{I}}$ , satisfy:*

$$(37) \quad \|g - \mathcal{I}g\| \leq Ch^{2m+2} \|D^{2m+2}g\| \quad \text{for } g \in H_{\text{per}}^{2m+2},$$

$$(38) \quad \|D^{m+1}(g - \mathcal{I}g)\| \leq Ch^{m+1} \|D^{2m+2}g\| \quad \text{for } g \in H_{\text{per}}^{2m+2},$$

$$(39) \quad \|g - \mathcal{I}g\| \leq Ch^{m+1} \|D^{m+1}g\| \quad \text{for } g \in H_{\text{per}}^{m+1}.$$

If one sets  $g = q - \mathcal{I}q$  in the last equation and observes that  $\mathcal{I}g = 0$  for this  $g$ , then one obtains

**Corollary 3.1.**

$$(40) \quad \|q - \mathcal{I}q\| \leq Ch^{m+1} \|D^{m+1}(q - \mathcal{I}q)\| \quad \text{for } q \in H_{\text{per}}^{m+1}.$$

Stability of the algorithm depends on the stability of the interpolation process. A direct analysis in  $L_2$  fails because the operator norms,  $\|\mathcal{I}\|$  and  $\|\tilde{\mathcal{I}}\|$ , are *not* bounded by  $1 + O(h)$ . Instead we make use of the following crucial *Orthogonality lemma*, which is closely related to the well-known integral relation from the theory of splines (e.g., [11, Ch. 6.2]).

**Lemma 3.2.** *For all  $f, g \in H_{\text{per}}^{m+1}$ ,*

$$\left( D^{m+1}\mathcal{I}f, D^{m+1}(g - \mathcal{I}g) \right) = 0.$$

*Proof.* As this lemma is central to the stability of the algorithm, we give its simple proof. Noting that  $\mathcal{I}f$  restricted to each interval  $(x_j, x_{j+1})$  is a polynomial of degree  $2m + 1$  and integrating by parts we have

$$(41) \quad \begin{aligned} & (D^{m+1}\mathcal{I}f, D^{m+1}(g - \mathcal{I}g)) \\ &= \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} D^{m+1}\mathcal{I}f^T D^{m+1}(g - \mathcal{I}g) dx \\ &= \sum_{j=0}^{N-1} \left[ \sum_{k=0}^m (-1)^{(m-k)} D^{2m+1-k}\mathcal{I}f^T D^k(g - \mathcal{I}g)|_{x_j}^{x_{j+1}} \right. \\ & \quad \left. + (-1)^{(m+1)} \int_{x_j}^{x_{j+1}} D^{2m+2}\mathcal{I}f^T(g - \mathcal{I}g) dx \right]. \end{aligned}$$

However, since by interpolation,

$$D^k(g - \mathcal{I}g) = 0, \quad x = x_j, x_{j+1}, \quad 0 \leq k \leq m,$$

and since  $\mathcal{I}f$  is a polynomial of degree no greater than  $2m + 1$ ,

$$D^{2m+2}\mathcal{I}f = 0,$$

all terms on the right-hand side of equation (41) are zero. Thus the lemma is proved.  $\square$

If we define the semi-inner product

$$(p, q)_{m+1} = (D^{m+1}p, D^{m+1}q), \quad p, q \in H_{\text{per}}^{m+1},$$

then the Orthogonality lemma says that any interpolant,  $\mathcal{I}f$ , is semi-orthogonal to any interpolation error,  $g - \mathcal{I}g$ . Clearly, for every  $f \in H_{\text{per}}^{m+1}$ ,

$$\begin{aligned} f &= \mathcal{I}f + (f - \mathcal{I}f), \\ D^{m+1}f &= D^{m+1}\mathcal{I}f + D^{m+1}(f - \mathcal{I}f). \end{aligned}$$

Applying Lemma 3.2 with  $g = f$  we obtain

$$(42) \quad \|D^{m+1}f\|^2 = \|D^{m+1}\mathcal{I}f\|^2 + \|D^{m+1}(f - \mathcal{I}f)\|^2, \quad f \in H_{\text{per}}^{m+1}.$$

In particular,

$$(43) \quad \|D^{m+1}\mathcal{I}f\| \leq \|D^{m+1}f\|, \quad f \in H_{\text{per}}^{m+1}.$$

In this sense the interpolation process is stable w.r.t. the seminorm,  $\|\cdot\|_{m+1}$ , corresponding to the above semi-inner product. This is the key fact that allows us to prove convergence.

Lastly, we make a technical assumption concerning the grids: namely, that for some fixed positive constants  $c_1, c_2$ ,

$$(44) \quad c_1 \max(h_j, h_{j+1}) \leq h_{j+\frac{1}{2}} \leq c_2 \min(h_j, h_{j+1}).$$

**3.1. Convergence theorem.** Let  $u$  be the solution to (10)–(11) and set

$$u^n = u(\cdot, nk) \quad \text{and} \quad u^{n+\frac{1}{2}} = u(\cdot, (n + \frac{1}{2})k).$$

Denote the corresponding approximations by  $p^n(x)$  and  $p^{n+\frac{1}{2}}(x)$ . We measure the error in the  $L_2$ -norm. Thus we consider

$$\|u^n - p^n\| \quad \text{and} \quad \|u^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\|.$$

Precisely we prove that the error in the propagated piecewise polynomials is one order less than would be obtained by the direct piecewise Hermite interpolation of the solution. The strategy in the proof is to first estimate the error using the semi-inner product defined above, where stability automatically follows from (43). We then bootstrap the result to estimate the error in  $L_2$ . Here Corollary 3.1 plays a key role, as its right-hand side is estimated in the first step.

**Theorem 3.1.** *Let  $\lambda = k/h > 0$  and let  $T > 0$  be fixed. Suppose  $f \in (H_{\text{per}}^{2m+2})^d$  and that the hypotheses of Lemma 2.1 hold. Then there is a constant  $C$ , independent of  $h$ , so that*

$$\|u^n - p^n\| + \|u^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\| \leq Ch^{2m+1} \|D^{2m+2}f\| \quad \text{for } 0 \leq nk \leq T.$$

*Proof.* Substituting the exact solution into the equations (34), (35) defines the local truncation errors,  $\eta^n$  and  $\eta^{n+\frac{1}{2}}$ ,

$$(45) \quad u^{n+\frac{1}{2}} = \tilde{\mathcal{I}}\mathcal{S}u^n + \eta^n,$$

$$(46) \quad u^{n+1} = \mathcal{I}\mathcal{S}u^{n+\frac{1}{2}} + \eta^{n+\frac{1}{2}}.$$

Since  $\mathcal{S}u^n = u^{n+\frac{1}{2}}$  and  $\mathcal{S}u^{n+\frac{1}{2}} = u^{n+1}$ , the local truncation errors are in fact interpolation errors,

$$(47) \quad \eta^n = u^{n+\frac{1}{2}} - \tilde{\mathcal{I}}u^{n+\frac{1}{2}},$$

$$(48) \quad \eta^{n+\frac{1}{2}} = u^{n+1} - \mathcal{I}u^{n+1},$$

and we obtain from Lemma 3.1,

$$(49) \quad \|\eta^n\| + \|\eta^{n+\frac{1}{2}}\| \leq C_m h^{2m+2} \|D^{2m+2}f\|.$$

Let

$$e^n = u^n - p^n, \quad e^{n+\frac{1}{2}} = u^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}$$

denote the errors. By subtracting the equations (34) and (35) from the equations (45) and (46), we obtain the error equations

$$(50) \quad e^{n+\frac{1}{2}} = \tilde{\mathcal{I}}\mathcal{S}e^n + \eta^n,$$

$$(51) \quad e^{n+1} = \mathcal{I}\mathcal{S}e^{n+\frac{1}{2}} + \eta^{n+\frac{1}{2}}.$$

Also,

$$e^0 = u^0 - p^0 = f - \mathcal{I}f$$

is an interpolation error; thus,

$$(52) \quad \|e^0\| \leq C_m h^{2m+2} \|D^{2m+2} f\|.$$

As mentioned above, we cannot directly use these error equations due to the fact that the  $L_2$ -operator norms of  $\mathcal{I}$  and  $\tilde{\mathcal{I}}$  are not bounded by  $1 + O(h)$ . Therefore, the above considerations are not sufficient to prove the theorem.

If instead one measures the local truncation error in the semi-norm,  $\|\cdot\|_{m+1}$ , one obtains an estimate  $\|\eta^n\|_{m+1} = \|D^{m+1}\eta^n\| \leq Ch^{m+1}$ . Combining this fact with (43) we can prove convergence, but we do not get sharp bounds on  $\|e^n\|$ . For this reason, the remainder of the proof uses a combination of the  $L_2$ -norm,  $\|\cdot\|$ , in which the local truncation error is  $\mathcal{O}(h^{2m+2})$ , and the semi-norm,  $\|\cdot\|_{m+1}$ , in which the interpolation is stable.

**Estimates in the semi-norm  $\|\cdot\|_{m+1}$ .** Apply  $D^{m+1}$  to the error equation (50) and obtain

$$D^{m+1}e^{n+\frac{1}{2}} = D^{m+1}\tilde{\mathcal{I}}\mathcal{S}e^n + D^{m+1}\eta^n.$$

Since  $\eta^n = u^{n+\frac{1}{2}} - \tilde{\mathcal{I}}u^{n+\frac{1}{2}}$  we can apply Lemma 3.2 and obtain

$$(53) \quad \|D^{m+1}e^{n+\frac{1}{2}}\|^2 = \|D^{m+1}\tilde{\mathcal{I}}\mathcal{S}e^n\|^2 + \|D^{m+1}\eta^n\|^2.$$

For simplicity, if we estimate the local truncation error, we will only write the correct powers of  $h$ , but we will not write how the estimate depends on  $f$ . Thus we have

$$(54) \quad \|D^{m+1}e^{n+\frac{1}{2}}\|^2 = \|D^{m+1}\tilde{\mathcal{I}}\mathcal{S}e^n\|^2 + \mathcal{O}(h^{2m+2}).$$

Next we consider the first term on the right-hand side of the above equation. Write

$$D^{m+1}\mathcal{S}e^n = D^{m+1}\tilde{\mathcal{I}}\mathcal{S}e^n + D^{m+1}(\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n).$$

Using Lemma 3.2 again one obtains

$$(55) \quad \|D^{m+1}\mathcal{S}e^n\|^2 = \|D^{m+1}\tilde{\mathcal{I}}\mathcal{S}e^n\|^2 + \|D^{m+1}(\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n)\|^2.$$

Let us set

$$\varepsilon_n := \|D^{m+1}e^n\|, \quad \delta_n := \|D^{m+1}(\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n)\|.$$

Then, using (54), (55), and the fact that  $\mathcal{S}$  preserves the semi-norm we find

$$(56) \quad \varepsilon_{n+\frac{1}{2}}^2 = \varepsilon_n^2 - \delta_n^2 + \mathcal{O}(h^{2m+2}).$$

From the above equation one obtains the recursive error estimate

$$\varepsilon_{n+\frac{1}{2}}^2 \leq \varepsilon_n^2 + \mathcal{O}(h^{2m+2}).$$

It turns out that the resulting error estimate,  $\|D^{m+1}e^n\| \leq Ch^{m+\frac{1}{2}}$  for  $nk \leq T$ , is not good enough to prove the theorem. Therefore we proceed differently and first obtain an estimate of the quantities  $\delta_n = \|D^{m+1}(\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n)\|$ .

From (56) we obtain

$$\delta_n^2 = \varepsilon_n^2 - \varepsilon_{n+\frac{1}{2}}^2 + \mathcal{O}(h^{2m+2}).$$

Therefore, noting that inequalities analogous to those we derived also apply at the even half-steps,

$$\begin{aligned} \sum_{n=0}^J (\delta_n^2 + \delta_{n+\frac{1}{2}}^2) &\leq \varepsilon_0^2 + CJh^{2m+2} \\ (57) \qquad \qquad \qquad &\leq Ch^{2m+1} \quad \text{for } Jk \leq T. \end{aligned}$$

**Error estimates in the  $L_2$ -norm.** Having established the bound (57) we now treat the error equations (50), (51) in the  $L_2$ -norm:

$$\begin{aligned} \|e^{n+\frac{1}{2}}\| &\leq \|\tilde{\mathcal{I}}\mathcal{S}e^n\| + Ch^{2m+2} \\ &\leq \|\mathcal{S}e^n\| + \|\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n\| + Ch^{2m+2} \\ &= \|e^n\| + \|\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n\| + Ch^{2m+2}. \end{aligned}$$

Here we have, using Corollary 3.1,

$$\begin{aligned} \|\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n\| &\leq Ch^{m+1}\|D^{m+1}(\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n)\| \\ &= Ch^{m+1}\delta_n. \end{aligned}$$

Therefore,

$$\|e^{n+\frac{1}{2}}\| \leq \|e^n\| + Ch^{m+1}\delta_n + Ch^{2m+2},$$

and, consequently, using the analogous inequality for the even half-steps:

$$(58) \quad \|e^{J+1}\| + \|e^{J+\frac{1}{2}}\| \leq Ch^{m+1} \sum_{n=0}^J (\delta_n + \delta_{n+\frac{1}{2}}) + Ch^{2m+1} \quad \text{for } 0 \leq (J+1)k \leq T.$$

Finally,

$$\begin{aligned} \sum_{n=0}^N (\delta_n + \delta_{n+\frac{1}{2}}) &\leq C \left( \sum_{n=0}^N 1 \right)^{1/2} \left( \sum_{n=0}^N (\delta_n^2 + \delta_{n+\frac{1}{2}}^2) \right)^{1/2} \\ &\leq Ch^{-1/2}h^{m+\frac{1}{2}} \\ &= Ch^m. \end{aligned}$$

Using this bound in (58), the theorem is proved. □

#### 4. CONVERGENCE FOR VARIABLE COEFFICIENT SYSTEMS

We now extend our convergence analysis to approximations to (14). The main difference in this case is that the local evolution by a truncated Taylor series (18) is not exact. Thus we must specify the degree,  $q$ , of the Taylor expansion. Our notation will be

$$(59) \qquad \qquad \qquad p^{n+\frac{1}{2}} = \tilde{\mathcal{I}}\mathcal{T}^{(q)}p^n,$$

$$(60) \qquad \qquad \qquad p^{n+1} = \mathcal{I}\mathcal{T}^{(q)}p^{n+\frac{1}{2}}.$$

This is a slight abuse of notation as the Taylor series in time is not defined at the break points; in fact, we only use it on the dual grid. In our analysis the operator

$\mathcal{T}^{(q)}$  is followed by an interpolation, so its values on the dual grid are all that is needed to make sense of the formulas.

In this case we will also need pointwise estimates of the piecewise Hermite interpolants and bounds on the solution of (14) in terms of local properties of the data and coefficients. These essentially standard estimates are given in Lemma 9.1 and Lemma 9.2 in the Appendix.

**4.1. Convergence theorem.** We now state and prove a generalization of Theorem 3.1. The new feature in the result is the necessity to account for the truncation error in the temporal Taylor series.

**Theorem 4.1.** *Suppose  $q \geq 2m + 1$ ,  $A$ ,  $B$  and  $F$  satisfy the assumptions of Lemma 9.2 with  $r = q + 1$ , and  $f \in C_{\text{per}}^{2m+2}$ . Let  $u$  be the solution of (14), and  $p$  the numerical approximation to  $u$  computed via the degree  $(2m + 1, q)$  Hermite-Taylor method with  $q \geq 2m + 1$ . Then for any  $T > 0$  there exists  $C$  depending on  $T$ ,  $F$ ,  $f$ ,  $A$  and  $B$  but independent of the mesh such that:*

$$\|u^n - p^n\| + \|u^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\| \leq Ch^\kappa \quad \text{for } 0 \leq nk \leq T,$$

where

$$\kappa = \min(2m + 1, q - \frac{1}{2}).$$

*Remark.* Clearly if we choose the minimal value of  $q = 2m + 1$  our error estimate is one-half order lower than in the constant coefficient case. We recover the earlier estimate by taking  $q \geq 2m + 2$ . However in practice we have seen an advantage to choosing  $q > 2m + 1$  only for marginally resolved cases with  $m$  large. Some comparisons will be made in the final section.

*Proof.* We copy the proof of Theorem 3.1 but with the additional time truncation error term. Precisely we write

$$\begin{aligned} e^{n+1/2} &= \tilde{\mathcal{I}}(\mathcal{S}u^n - \mathcal{S}p^n) + \tilde{\mathcal{I}}(\mathcal{S}p^n - \mathcal{T}^{(q)}p^n) + \mathcal{S}u^n - \tilde{\mathcal{I}}\mathcal{S}u^n \\ (61) \quad &= \tilde{\mathcal{I}}\mathcal{S}_h e^n + \gamma^n + \eta^n, \end{aligned}$$

where  $\mathcal{S}_h$  denotes the solution operator for the homogeneous problem ( $F = 0$ ) and  $\gamma^n, \eta^n$  are the truncation error terms:

$$\gamma^n = \tilde{\mathcal{I}}(\mathcal{S}p^n - \mathcal{T}^{(q)}p^n),$$

and  $\eta^n$  is defined as in (45). We begin with the equality for the error semi-norm:

$$(62) \quad \|D^{m+1}e^{n+1/2}\|^2 = \|D^{m+1}\tilde{\mathcal{I}}\mathcal{S}_h e^n + D^{m+1}\gamma^n\|^2 + \|D^{m+1}\eta^n\|^2.$$

From Lemmas 3.1 and 9.2 we then have for any  $c > 0$ :

$$(63) \quad \begin{aligned} \|D^{m+1}e^{n+1/2}\|^2 &\leq (1 + ch)\|D^{m+1}\tilde{\mathcal{I}}\mathcal{S}_h e^n\|^2 \\ &\quad + (1 + (ch)^{-1})\|D^{m+1}\gamma^n\|^2 + Ch^{2m+2}. \end{aligned}$$

To proceed further we must estimate the Taylor truncation error term, which we accomplish in the next lemma. The idea behind the estimate is to express  $p^n$  in terms of  $u^n$  and  $e^n$ .

**Lemma 4.1.** *Under the assumptions of the theorem we have for some constant  $C$  independent of  $h$ :*

$$\begin{aligned} \|D^{m+1}\gamma^n\|^2 &\leq Ch^{2q-4m} (h^{2m} + \|D^{m+1}e^n\|^2 + \|e^n\|^2), \\ \|\gamma^n\|^2 &\leq Ch^{2q-2m+2} (h^{2m} + \|D^{m+1}e^n\|^2 + \|e^n\|^2). \end{aligned}$$

*Proof.* By Taylor's Theorem we have for any space derivative of order  $l$ ,  $0 \leq l \leq m$ , evaluated at  $(x_{i+\frac{1}{2}}, t_{n+\frac{1}{2}})$ :

$$(64) \quad \left| \left( D_x^l \mathcal{S}p^n - \mathcal{T}^{(q,l)} p^n \right) \right|^2 \leq Ch^{2q+2-2l} \max_{t_n \leq t \leq t_{n+1/2}} \left| \frac{\partial^{q+1} \mathcal{S}p^n}{\partial t^{q+1-l} \partial x^l} (x_{i+1/2}, t) \right|^2.$$

From Lemma 9.2 and using the fact that  $p^n$  is piecewise polynomial of degree  $2m+1$  we bound the right-hand side of (64) by

$$Ch^{2q+2-2l} \left( 1 + \sum_{j=0}^{2m+1} \max_{x \in (x_i, x_{i+1})} \left( \frac{\partial^j p^n}{\partial x^j} (x, t_n) \right)^2 \right).$$

We write

$$p^n = \mathcal{I}p^n = \mathcal{I}u^n - \mathcal{I}e^n.$$

Then by Lemma 9.1 we bound the right-hand side of (64) by

$$Ch_{i+\frac{1}{2}}^{2q+2-2l} \left( 1 + h_{i+\frac{1}{2}}^{-(2m+1)} \|e^n\|_{H^{m+1}(x_i, x_{i+1})}^2 \right).$$

Again applying Lemma 9.1 to bound  $\|D^{m+1}\gamma^n\|^2$  and recalling (44) we obtain

$$\begin{aligned} \|D^{m+1}\gamma^n\|^2 &\leq C \sum_{i=1}^N \sum_{l=0}^m h_i^{2l-2m-1} \sum_{\nu=0}^1 \left| \left( D^l \mathcal{S}p^n - \mathcal{T}^{(q,l)} p^n \right) (x_{i+\nu-1/2}, t_{n+1/2}) \right|^2 \\ &\leq C \sum_{i=0}^{N-1} h_{i+\frac{1}{2}}^{2q-2m+1} \left( 1 + h_{i+\frac{1}{2}}^{-(2m+1)} \|e^n\|_{H^{m+1}(x_i, x_{i+1})}^2 \right) \\ &\leq Ch^{2q-4m} (h^{2m} + \|e^n\|_{H^{m+1}}^2). \end{aligned}$$

Applying the well-known inequality

$$\|e^n\|_{H^{m+1}}^2 \leq C (\|D^{m+1}e^n\|^2 + \|e^n\|^2),$$

we reach the first result. Similarly

$$\begin{aligned} \|\gamma^n\|^2 &\leq C \sum_{i=1}^N \sum_{l=0}^m h_i^{2l+1} \sum_{\nu=0}^1 \left| \left( D^l \mathcal{S}p^n - \mathcal{T}^{(q,l)} p^n \right) (x_{i+\nu-1/2}, t_{n+1/2}) \right|^2 \\ &\leq C \sum_{i=0}^{N-1} h_{i+\frac{1}{2}}^{2q+3} \left( 1 + h_{i+\frac{1}{2}}^{-(2m+1)} \|e^n\|_{H^{m+1}(x_i, x_{i+1})}^2 \right) \\ &\leq Ch^{2q-2m+2} (h^{2m} + \|D^{m+1}e^n\|^2 + \|e^n\|^2). \end{aligned}$$

This completes the proof of Lemma 4.1.  $\square$

Returning to the proof of Theorem 4.1 we now have

$$\begin{aligned} \|D^{m+1}e^{n+1/2}\|^2 &\leq (1+ch) \|D^{m+1}\tilde{\mathcal{I}}\mathcal{S}_h e^n\|^2 \\ &\quad + Ch^{2q-4m-1} (\|D^{m+1}e^n\|^2 + \|e^n\|^2) + Ch^{\bar{\kappa}}, \end{aligned}$$

where  $\bar{\kappa} = \min(2q-2m-1, 2m+2)$ . Defining as before

$$\epsilon_n = \|D^{m+1}e^n\|, \quad \delta_n = \|D^{m+1}(\mathcal{S}_h e^n - \tilde{\mathcal{I}}\mathcal{S}_h e^n)\|,$$

we use Lemma 3.2 to deduce

$$\epsilon_{n+\frac{1}{2}}^2 \leq (1+ch) \|D^{m+1}\mathcal{S}_h e^n\|^2 - \delta_n^2 + Ch^{2q-4m-1} (\epsilon_n^2 + \|e^n\|^2) + Ch^{\bar{\kappa}}.$$

By Lemma 9.2 we conclude that

$$\|D^{m+1}\mathcal{S}_h(e^n)\|^2 \leq (1 + Ch)\epsilon_n^2 + Ch\|e^n\|^2.$$

Since  $q \geq 2m + 1$  we finally have

$$(65) \quad \epsilon_{n+\frac{1}{2}}^2 \leq (1 + Ch)\epsilon_n^2 - \delta_n^2 + Ch\|e^n\|^2 + Ch^{\bar{\kappa}}.$$

From (65) we conclude

$$(66) \quad \epsilon_n^2 + \epsilon_{n+\frac{1}{2}}^2 \leq Ch^{\bar{\kappa}-1} + Ch \sum_{j=0}^n \left( \|e^j\|^2 + \|e^{j+\frac{1}{2}}\|^2 \right),$$

$$(67) \quad \sum_{j=0}^n \left( \delta_j^2 + \delta_{j+\frac{1}{2}}^2 \right) \leq Ch^{\bar{\kappa}-1} + Ch \sum_{j=0}^n \left( \|e^j\|^2 + \|e^{j+\frac{1}{2}}\|^2 \right).$$

Finally we estimate the  $L_2$ -norm directly. In particular, we have

$$\|e^{n+\frac{1}{2}}\| \leq \|\tilde{\mathcal{I}}\mathcal{S}_h e^n\| + \|\gamma^n\| + \|\eta^n\|.$$

Estimating the truncation errors as above we have

$$\|\eta^n\| \leq Ch^{2m+2},$$

and by Lemma 4.1 and (66)

$$\|\gamma^n\| \leq Ch^{q+1} + Ch^{q-m+\frac{3}{2}} \left( \sum_{j=0}^n \left( \|e^j\|^2 + \|e^{j+\frac{1}{2}}\|^2 \right) \right)^{1/2} + Ch^{q-m+1}\|e^n\|.$$

Finally we write as in the proof of Theorem 3.1:

$$\begin{aligned} \|\tilde{\mathcal{I}}\mathcal{S}_h e^n\| &\leq \|\mathcal{S}_h e^n\| + Ch^{m+1}\delta_n \\ &\leq (1 + Ch)\|e^n\| + Ch^{m+1}\delta_n. \end{aligned}$$

Set

$$E^n = \max_{j \leq n} (\|e^j\|, \|e^{j-\frac{1}{2}}\|), \quad E^{n+\frac{1}{2}} = \max_{j \leq n} (\|e^j\|, \|e^{j+\frac{1}{2}}\|).$$

Then combining the last few inequalities and recalling that  $q \geq 2m + 1$  we find

$$E^{n+\frac{1}{2}} \leq (1 + Ch)E^n + Ch^{m+1}\delta_n + Ch^{2m+2}.$$

Using (67) this implies for  $nk \leq T$ :

$$\begin{aligned} E^{n+1} + E^{n+\frac{1}{2}} &\leq Ch^{2m+1} + Ch^{m+1} \sum_{j=0}^n \left( \delta_j + \delta_{j+\frac{1}{2}} \right) \\ &\leq Ch^{2m+1} + Ch^{m+\frac{1}{2}} \left( \sum_{j=0}^n \left( \delta_j + \delta_{j-\frac{1}{2}} \right)^2 \right)^{\frac{1}{2}} \\ &\leq Ch^{m+\frac{\bar{\kappa}}{2}} + Ch^{m+\frac{1}{2}} E^{n+\frac{1}{2}}. \end{aligned}$$

Choosing  $h$  sufficiently small we can clearly eliminate the term involving  $E^{n+\frac{1}{2}}$  from the right-hand side. Noting that

$$\kappa = m + \frac{\bar{\kappa}}{2},$$

we have completed the proof of Theorem 4.1.  $\square$

## 5. INITIAL-BOUNDARY VALUE PROBLEMS

We now consider the initial-boundary value problem defined for  $x \in (0, 1)$ :

$$(68) \quad u_t = Au_x + Bu + F,$$

$$(69) \quad u^R(0, t) = M^L u^L(0, t) + \gamma^L(t),$$

$$(70) \quad u^L(1, t) = M^R u^R(1, t) + \gamma^R(t),$$

$$(71) \quad u(x, 0) = f(x).$$

We assume, for simplicity, that  $A$  is diagonal:

$$(72) \quad A = \begin{pmatrix} \Lambda^R & O \\ O & \Lambda^L \end{pmatrix},$$

where

$$\begin{aligned} \Lambda^R &\in \mathbb{R}^{d_R \times d_R}, \quad \Lambda^R < O, \\ \Lambda^L &\in \mathbb{R}^{d_L \times d_L}, \quad \Lambda^L > O. \end{aligned}$$

The partitioning

$$u = \begin{pmatrix} u^R \\ u^L \end{pmatrix}$$

thus corresponds to a splitting into right-going and left-going characteristic variables, and the boundary conditions (69), (70) explicitly express the incoming waves in terms of outgoing waves and inhomogeneous data. We note that any symmetric system can be diagonalized and the incoming and outgoing characteristic variables identified, which is all that is needed to construct the discrete approximations. For variable coefficient problems it is possible for the number of right-going and left-going waves to be different at each boundary, but for simplicity we exclude this possibility. The results extend to that case, however. It would also be possible for the number of ingoing and outgoing waves at a boundary to change as a function of time. Again we exclude this possibility, which would require further analysis.

We note that the problem (68)–(70) is strongly well-posed [16, 7.6]. Away from the boundaries the local estimates of Lemma 9.2 still hold. Near them, however, they must be modified to take account of the boundary data, so that additional inhomogeneous terms appear. Global estimates thus take the form [16, 7.6]:

$$(73) \quad \begin{aligned} &\sum_{j=0}^m \left( \|D^j u(\cdot, t)\|^2 + \int_0^t (|D^j u(0, \tau)|^2 + |D^j u(1, \tau)|^2) d\tau \right) \\ &\leq C e^{\alpha_m t} \sum_{j=0}^m \left( \|D^j f\|^2 + \int_0^t \left( \|D^j F(\cdot, \tau)\|^2 + \left| \frac{\partial^j \gamma^L}{\partial t^j}(\tau) \right|^2 + \left| \frac{\partial^j \gamma^R}{\partial t^j}(\tau) \right|^2 \right) d\tau \right). \end{aligned}$$

Time derivatives and mixed derivatives of the solution can be expressed by space derivatives; thus analogous estimates hold for them.

**5.1. Description of the method.** We introduce staggered grids satisfying (44) as before:

$$\begin{aligned} 0 &= x_0 < x_{\frac{1}{2}} < x_1 < x_{\frac{3}{2}} < x_2 < \cdots < x_{N-\frac{1}{2}} < x_N = 1, \\ \tilde{G}_h &= \{x_{j+\frac{1}{2}}, \quad 0 \leq j \leq N-1\} \cup \{x_0, x_N\}, \\ G_h &= \{x_j, \quad 0 \leq j \leq N\}. \end{aligned}$$

Note that the boundary points are included in both meshes.

Given Hermite data of order up through  $m$  on one of the grids, the  $(m, q)$  Hermite-Taylor update on the interior of the dual grid is carried out precisely as in the periodic cases described earlier. Thus we need only describe how to update the solution and its space derivatives at the points  $x_0$  and  $x_N$ .

The concept of our boundary update is as follows. Updated function and derivative values for outgoing components,  $u^L(0, t)$  and  $u^R(1, t)$ , are computed using one-sided formulas, that is, via (18) applied at the boundary point. We then update the incoming components by insisting that the boundary condition and time derivatives of the boundary condition are satisfied at the new time level. Time derivatives of the approximate solution appearing in the differentiated boundary condition are replaced by space derivatives using the equation (68).

For definiteness we look at an even half-step where we propagate data associated with the piecewise polynomial  $p^n$  from  $G_h$  to  $\tilde{G}_h$ . Applying the update directly at  $x = x_0$  and  $x = x_N$  we obtain values for the solution and its scaled space derivatives:

$$\mathcal{T}^{(q,l)} p^n(0, t_{n+\frac{1}{2}}), \quad \mathcal{T}^{(q,l)} p^n(1, t_{n+\frac{1}{2}}),$$

via (18). Note that these values only use one-sided spatial information and typically do not satisfy the boundary conditions. Thus we will use the boundary conditions to update the incoming characteristic variable and its space derivatives. Specializing to the left boundary,  $x = 0$ , we will derive recursion relations allowing us to effectively express  $D^l u^R$  in terms of  $D^l u^L$  alone. We use the one-sided Taylor updates to compute this outgoing data.

In what follows the vector space-time coefficients  $\tilde{c}_{ij}$  are understood to be approximations:

$$(74) \quad \tilde{c}_{ij} \approx u^{[i,j]}(0, t_{n+\frac{1}{2}}).$$

The tilde is to remind us that they are distinct from the coefficients  $c_{ij}$  used in the Taylor update (18). The coefficients  $c_{ij}$  approximate the space-time derivatives of  $u$  at the current time,  $t_n$ . Also it is to be understood that all updates are at the point  $(x, t) = (0, t_{n+\frac{1}{2}})$ .

We compute the approximation to  $u^L$  and its derivatives via (18):

$$(75) \quad \tilde{c}_{l0}^L = \sum_{s=0}^{q-l} c_{ls}^L \left(\frac{k}{2}\right)^s.$$

Note that the  $c_{ls}$  are obtained by a modified interpolation routine computing the expansion of the interpolating polynomial at an endpoint. The boundary condition (69) itself provides us with our approximation to  $u^R(0, t_{n+\frac{1}{2}})$  directly:

$$(76) \quad \tilde{c}_{00}^R = M^L \tilde{c}_{00}^L + \gamma^L.$$

Proceeding to first space derivatives we use  $\tilde{c}_{l0}^L$  and (68) to compute an approximation to  $u_t^L$ :

$$\tilde{c}_{01}^L = \Lambda^L \tilde{c}_{10}^L + (B\tilde{c}_{00})^L + F^L,$$

which can be substituted into the time derivative of (69) to produce an approximation to  $u_t^R$ :

$$\tilde{c}_{01}^R = M^L \tilde{c}_{01}^L + M^{L,[1]} \tilde{c}_{00}^L + \gamma^{L,[1]}.$$

(As the boundary coefficients are functions of time alone their scaled derivatives are expressed by a single superscript.) Finally, we compute an approximation to  $u^{R,[1,0]}$  using (68):

$$\tilde{c}_{10}^R = (\Lambda^R)^{-1} (\tilde{c}_{01}^R - (B\tilde{c}_{00})^R - F^R).$$

This process, though involved, can be repeated as desired. Proceeding recursively, suppose we know all space-time derivatives of order up through  $\nu - 1$ , as well as the outgoing data provided by (75). That is, we have values for  $\tilde{c}_{lj}$ ,  $l + j \leq \nu - 1$  and  $\tilde{c}_{l0}^L$ ,  $l \leq m$ . We compute the derivatives of order  $\nu$  as follows.

- i:** Approximate scaled order  $\nu$  space-time derivatives of  $u^L$  are computed using derivatives of (68). (See (19).) In particular, for  $j = 1, \dots, \nu$  we set

$$(77) \quad \begin{aligned} \tilde{c}_{ls}^L &= \frac{1}{s} \sum_{i=0}^l \sum_{j=0}^{s-1} (i+1) \Lambda^{L,[l-i,s-1-j]} \tilde{c}_{i+1,j}^L \\ &+ \frac{1}{s} \sum_{i=0}^l \sum_{j=0}^{s-1} \left( B^{[l-i,s-1-j]} \tilde{c}_{ij} \right)^L + \frac{1}{s} F^{L,[l,s-1]}. \end{aligned}$$

Note that if we proceed sequentially in  $j$  all terms on the right-hand side are known.

- ii:** Take the  $\nu$ th time derivative of (69) to compute  $\tilde{c}_{0\nu}^R$ :

$$(78) \quad \tilde{c}_{0\nu}^R = \sum_{l=0}^{\nu} M^{L,[\nu-l]} \tilde{c}_{0l}^L + \gamma^{L,[\nu]}.$$

- iii:** Finally, using (68) compute sequentially for  $j = 1, \dots, \nu$ :

$$(79) \quad \begin{aligned} j \Lambda^R \tilde{c}_{j,\nu-j}^R &= (\nu - j + 1) \tilde{c}_{j-1,\nu+1-j}^R - F^{R,[j-1,\nu-j]} \\ &- \sum_{i_1=0}^{j-1} \sum_{i_2=0}^{\nu-j} (i_1 + 1) \Lambda^{R,[j-i_1-1,\nu-j-i_2]} \tilde{c}_{i_1+1,i_2}^R \\ &- \sum_{i_1=0}^{j-1} \sum_{i_2=0}^{\nu-j} \left( B^{[j-1-i_1,\nu-j-i_2]} \tilde{c}_{i_1 i_2} \right)^R. \end{aligned}$$

(Here the prime denotes the exclusion of the  $i_1 = j - 1$ ,  $i_2 = \nu - j$  term from the double summation.)

At its termination, this process produces all the required derivatives of  $p^{n+\frac{1}{2}}$ , which are the coefficients  $\tilde{c}_{l0}$ . At  $x = 1$  we carry out the analogous process with the roles of  $\tilde{c}_{ij}^L$  and  $\tilde{c}_{ij}^R$  reversed and (69) replaced by (70). The update process for even time steps is identical. We note that this process defines an affine map between the outgoing Hermite data and the incoming Hermite data. We represent these maps via the following notation:

$$(80) \quad \tilde{c}_{\cdot,0}^R = P^{L,\tau} \tilde{c}_{\cdot,0}^L + \Gamma^{L,\tau}, \quad x = x_0,$$

$$(81) \quad \tilde{c}_{\cdot,0}^L = P^{R,\tau} \tilde{c}_{\cdot,0}^R + \Gamma^{R,\tau}, \quad x = x_N,$$

where  $\tau = n$  or  $\tau = n + \frac{1}{2}$ .

**5.2. Stability analysis.** We first note that the Orthogonality Lemma 3.2 still holds in this case, that is, for functions in  $H^{m+1}(0, 1)$ , so long as 0 and 1 are nodes. Also, we note that equations (77)–(79), with appropriate interpretation of the coefficients  $\tilde{c}_{i_s}^L$  and  $\tilde{c}_{i_s}^R$ , are satisfied by the exact solution  $u$ . A simplified stability and convergence analysis can then be constructed for problems where the Taylor time evolution is exact, that is, when

$$(82) \quad A(x, t) = \bar{A}, \quad M^L(t) = \bar{M}^L, \quad M^R(t) = \bar{M}^R, \quad B = 0, \quad F = 0, \quad \gamma = 0.$$

Precisely we have:

**Lemma 5.1.** *Suppose  $p^n$  is a  $C^m$  piecewise degree  $2m+1$  polynomial relative to the grid  $G_h$ , and  $\mathcal{S}p^n$  denotes the exact solution of (68)–(70) for  $t_n \leq t \leq t_{n+\frac{1}{2}}$  equal to  $p^n$  at  $t = t_n$ . Suppose further that (82) holds and that  $p^n$  is compatible; that is,  $p^n$  satisfies (80)–(81). Then if (17) holds,  $\mathcal{S}p^n$  satisfies (80)–(81) at  $t = t_{n+\frac{1}{2}}$ .*

*Proof.* For systems satisfying (82),  $u^L(0, t)$  is given by

$$(u^L)_i(0, t) = (u^L)_i(\Lambda_i^L \Delta t, t - \Delta t).$$

Thus if the data,  $p^n(x)$ , is a degree  $2m+1$  polynomial on  $x \in (0, \Lambda_i^L \Delta t + \epsilon)$ , then

$$\begin{aligned} \frac{1}{l!} D^l (\mathcal{S}p^n)_i^L(0, t_{n+\frac{1}{2}}) &= \frac{1}{l!} D^l (p^n)_i^L(\Lambda_i^L \Delta t) \\ &= \sum_{j=0}^{2m+1-l} \frac{(\Lambda_i^L \Delta t)^j}{l! j!} D^{j+l} (p^n)_i^L(0) \\ &= \sum_{j=0}^{2m+1-l} (c_{lj})_i^L (\Delta t)^j. \end{aligned}$$

Thus  $\mathcal{S}p^n$  satisfies the computational update formula (75). As our assumption that the compatibility conditions hold ensures a smooth solution, (77)–(79), which are consequences of the equation and boundary condition, are also satisfied.  $\square$

Now we must consider the semi-norm  $\|D^{m+1}p^n\|$ . We consider the simplest case when the reflection matrices are zero, that is,  $\bar{M}^L = O$ ,  $\bar{M}^R = O$ . Then the boundary conditions imply

$$(83) \quad D^j u^R(0, t) = 0, \quad D^j u^L(1, t) = 0.$$

Moreover, the system decouples into scalar equations:

$$(84) \quad u_{i,t}^L = \Lambda_i^L u_{i,x}^L, \quad u_{i,t}^R = \Lambda_i^R u_{i,x}^R.$$

Combining this with (83) we find

$$(85) \quad (\mathcal{S}p^n)_i^L = \begin{cases} p^n(x + \Lambda_i^L t), & 0 \leq x + \Lambda_i^L t \leq 1, \\ 0, & \text{otherwise,} \end{cases}$$

$$(86) \quad (\mathcal{S}p^n)_i^R = \begin{cases} p^n(x + \Lambda_i^R t), & 0 \leq x + \Lambda_i^R t \leq 1, \\ 0, & \text{otherwise,} \end{cases}$$

which implies

$$(87) \quad \|D^{m+1}\mathcal{S}p^n\| \leq \|D^{m+1}p^n\|.$$

Combining (87) with Lemma 3.2 we deduce the basic stability inequality:

$$\begin{aligned} \|D^{m+1}p^{n+\frac{1}{2}}\| &\leq \|D^{m+1}Sp^n\| \\ &\leq \|D^{m+1}p^n\|. \end{aligned}$$

In this special case we can now repeat the arguments leading to Theorem 3.1 to prove:

**Theorem 5.1.** *Let  $0 < \lambda = k/h < \rho(\bar{A})$  and  $T > 0$  be fixed. Let  $u$  solve (68)–(70) under the assumptions of Lemma 5.1 and suppose in addition that  $\bar{M}^L = O$ ,  $\bar{M}^R = O$ . Let  $f \in (H^{2m+2})^d$  be compatible with the boundary conditions and let  $p$  be the numerical approximation based on degree  $2m + 1$  Hermite interpolation and degree  $2m + 1$  Taylor evolution with time steps satisfying the CFL condition. Then there is a constant  $C$ , independent of  $h$ , so that*

$$\|u^n - p^n\| + \|u^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\| \leq Ch^{2m+1} \|D^{2m+2}f\| \quad \text{for } 0 \leq nk \leq T.$$

Theorem 5.1 can be generalized to problems with variable coefficients and nonzero reflections using the techniques of section 4 and the estimates of (73). The idea is again to estimate the error in truncating the Taylor series. Of course (88) is replaced in the argument by an inequality of the form (65). We have not carried out the details, but present some simple numerical experiments to illustrate the method’s performance.

6. CONVERGENCE IN MULTIPLE DIMENSIONS

We now extend the convergence result to multiple space dimensions, that is, to (21). For simplicity we restrict attention to the constant coefficient case:

$$(88) \quad A_j(x, t) = \bar{A}_j, \quad B = 0, \quad F = 0.$$

Generalizations follow using the arguments of section 4. We will also choose the degree,  $q$ , of the temporal Taylor series to be  $\mu(2m + 1)$  so that the evolution of the piecewise polynomial data is exact on the grid. In the experiments section we show that this isn’t necessary.

Now we may express the evolution of the interpolants as in section 3, that is, by (34)–(35). To generalize the proof of Theorem 3.1 we must generalize the semi-inner product and Orthogonality lemma to be compatible with piecewise Hermite tensor-product interpolation.

**6.1. Norms and interpolation.** We introduce

$$(89) \quad \|f\|_{\mu, [r]} = \sum_{\alpha \in \mathcal{Q}_r} \|D^\alpha f\|,$$

$$(90) \quad |f|_{\mu, [r]} = \|f\|_{\mu, [r]} - \|f\|,$$

where  $\mathcal{Q}_r$  is the set of  $\mu$ -dimensional multiindices defined by

$$(91) \quad \mathcal{Q}_r = \{\alpha | \alpha_j \in \{0, r\}\}.$$

In other words,  $\|f\|_{\mu, [r]}$  measures all partial derivatives of  $f$  involving derivatives in the coordinate directions of order  $r$  or 0. We also denote by  $\mathcal{H}_\mu^{[r]}$  the space of all functions for which  $\|\cdot\|_{\mu, [r]}$  is finite. Obviously,  $\mathcal{H}_\mu^{[r]}$  contains the Sobolev space

$H_{\mu, \text{per}}^{\mu}$ . However, the degree  $\mu(2m+1)$  tensor-product interpolants we are using are, in fact, elements of  $\mathcal{H}_{\mu}^{[m+1]}$  but are typically *not* elements of  $H_{\mu}^{(m+1)\mu}$  for  $\mu > 1$ .

We also inductively define a new semi-inner product for  $\mu > 1$  by

$$(92) \quad (f, g)_{\mu, [m+1]} = \left( \prod_{j=1}^{\mu} D_j^{m+1} f, \prod_{j=1}^{\mu} D_j^{m+1} g \right) + \sum_{i=1}^{\mu} \sum_{j_i=1}^{N_i} h_{i, j_i}(f, g)_{\mu-1, [m+1]}|_{x_i=x_{i, j_i}}.$$

(Clearly, the new inner product depends on the mesh, which is suppressed in our notation.) Note that if we expand this out we see that each term involves inner products of derivatives  $\prod_{j=i_1}^{i_r} D_j^{m+1}$  applied to  $f$  and  $g$  on  $r$ -dimensional faces. On one-dimensional faces we simply use the inner product from the previous sections. An inner product associated with the dual mesh is defined analogously.

We first note the following estimates of the interpolation error. For  $\mu = 2$  the estimates (93)–(95) follow from results in [2] and can be generalized to higher dimensions in a straightforward manner. Inequality (96) follows from the arguments leading to Corollary 3.1. A crucial point in (94) is that the convergence rate for approximations to  $g$  in  $\|\cdot\|_{\mu, [m+1]}$  is the same as for approximations to  $D_i^{m+1}g$  in  $L_2$  despite the fact that derivatives of higher order appear.

**Lemma 6.1.** *The Hermite interpolation operators,  $\mathcal{I}$  and  $\tilde{\mathcal{I}}$ , satisfy:*

$$(93) \quad \|g - \mathcal{I}g\| \leq Ch^{2m+2}|g|_{\mu, [2m+2]} \quad \text{for } g \in \mathcal{H}_{\mu}^{[2m+2]},$$

$$(94) \quad \|g - \mathcal{I}g\|_{\mu, [m+1]} \leq Ch^{m+1}|g|_{\mu, [2m+2]} \quad \text{for } g \in \mathcal{H}_{\mu}^{[2m+2]},$$

$$(95) \quad \|g - \mathcal{I}g\| \leq Ch^{m+1}|g|_{\mu, [m+1]} \quad \text{for } g \in \mathcal{H}_{\mu}^{[m+1]},$$

$$(96) \quad \|g - \mathcal{I}g\| \leq Ch^{m+1}|g - \mathcal{I}g|_{\mu, [m+1]} \quad \text{for } g \in \mathcal{H}_{\mu}^{[m+1]}.$$

We also have the generalization of Lemma 3.2,

**Lemma 6.2.** *For all  $f, g \in \mathcal{H}_{\mu}^{[m+1]}$ ,*

$$\left( \mathcal{I}f, g - \mathcal{I}g \right)_{\mu, [m+1]} = 0.$$

*Proof.* We proceed by induction on  $\mu$  and note that the lemma holds for  $\mu = 1$ . Also, restricted to the face  $x_i = x_{i, j_i}$ , the function  $\mathcal{I}f$  is a  $(\mu - 1)$ -dimensional interpolant and  $g - \mathcal{I}g$  is a  $(\mu - 1)$ -dimensional interpolation error. Thus each term in the double sum in (92) vanishes by the induction hypothesis. Hence we need only consider the first inner product on the right-hand side of (92). Looking at the integral on an arbitrary cell, integrating by parts the derivative  $D_{\mu}^{m+1}$  from  $g - \mathcal{I}g$  to  $\mathcal{I}f$  and noting that in the cell interior  $D_{\mu}^{2m+2}\mathcal{I}f = 0$ , we obtain for the integral:

$$\begin{aligned} & \sum_{i=0}^m (-1)^i \int_{F_+} D_{\mu}^{m+1+i} \prod_{j=1}^{\mu-1} D_j^{m+1} \mathcal{I}f \cdot D_{\mu}^{m-i} \prod_{j=1}^{\mu-1} D_j^{m+1} (g - \mathcal{I}g) \\ & + (-1)^{i+1} \int_{F_-} D_{\mu}^{m+1+i} \prod_{j=1}^{\mu-1} D_j^{m+1} \mathcal{I}f \cdot D_{\mu}^{m-i} \prod_{j=1}^{\mu-1} D_j^{m+1} (g - \mathcal{I}g), \end{aligned}$$

where  $F_{\pm}$  are the cell faces with normal in the  $x_{\mu}$ -direction. Let  $\mathcal{I}_{\mu}$  denote the interpolation operator in  $x_{\mu}$  and  $\mathcal{I}^{(\mu-1)}$  denote  $(\mu - 1)$ -dimensional interpolation

in the other variables. Then

$$(97) \quad \mathcal{I} = \mathcal{I}^{(\mu-1)} \circ \mathcal{I}_\mu.$$

Using the fact that  $D_\mu^{m-i} \mathcal{I}_\mu g = D_\mu^{m-i} g$  on  $F_\pm$ , we can rewrite the typical integral on a given face as

$$\int \left[ \prod_{j=1}^{\mu-1} D_j^{m+1} \mathcal{I}^{(\mu-1)} D_\mu^{m+1+i} \mathcal{I}_\mu f \right] \cdot \left[ \prod_{j=1}^{\mu-1} D_j^{m+1} \left( D_\mu^{m-i} g - \mathcal{I}^{(\mu-1)} D_\mu^{m-i} g \right) \right].$$

Summing over all faces with  $x_\mu = x_{\mu,i}$  and using the induction hypothesis we can write this as

$$\sum \left( \mathcal{I}^{(\mu-1)} \tilde{f}, \tilde{g} - \mathcal{I}^{(\mu-1)} \tilde{g} \right)_{\mu-1, [m+1]},$$

where

$$\tilde{f} = D_\mu^{m+1+i} \mathcal{I}_\mu f, \quad \tilde{g} = D_\mu^{m-i} g.$$

By the induction hypothesis this inner product is zero, completing the proof of the lemma.  $\square$

Secondly, we show that the new inner product can be used to control the semi-norm appearing in the error estimates:

**Lemma 6.3.** *There exist positive constants  $C_1, C_2$  independent of the grid  $G_h$  and the function  $g \in \mathcal{H}_\mu^{[m+1]}$  such that*

$$C_1 |g|_{\mu, [m+1]}^2 \leq (g, g)_{\mu, [m+1]} \leq C_2 \|g\|_{\mu, [m+1]}^2.$$

*Proof.* The second inequality follows directly from the standard Sobolev inequality in one dimension:

$$(98) \quad |f(x)| \leq C (\|f\| + \|D^{m+1} f\|),$$

followed by repeated integration. (Note that the scaling of the terms in (92) involving lower-dimensional inner products is necessary here for  $C_{1,2}$  to be grid-independent.) For the first inequality, we first prove by induction on  $\mu$  the following auxiliary inequality:

$$(99) \quad \|f\|^2 \leq Ch^\mu \left( (f, f)_{\mu, [m+1]} + \sum_{j_1=1}^{N_1} \cdots \sum_{j_\mu=1}^{N_\mu} f^2(x_{1,j_1}, \dots, x_{\mu,j_\mu}) \right).$$

We assume  $h < \frac{1}{2}$ . Taking  $\mu = 1$  and considering an interval  $(x_j, x_{j+1})$  we have

$$\begin{aligned} \max_{x_j \leq x \leq x_{j+1}} f^2(x) &\leq f^2(x_j) + 2 \int_{x_j}^{x_{j+1}} |f| \cdot |Df| dx \\ &\leq f^2(x_j) + \|f\|_{(x_j, x_{j+1})}^2 + \|Df\|_{(x_j, x_{j+1})}^2 \\ &\leq f^2(x_j) + h_{j+\frac{1}{2}} \max_{x_j \leq x \leq x_{j+1}} f^2(x) + \|Df\|_{(x_j, x_{j+1})}^2, \end{aligned}$$

which implies

$$(100) \quad \max_{x_j \leq x \leq x_{j+1}} f^2(x) \leq 2f^2(x_j) + 2\|Df\|_{(x_j, x_{j+1})}^2.$$

Integrating (100) yields

$$(101) \quad \|f\|_{(x_j, x_{j+1})}^2 \leq 2h_{j+\frac{1}{2}} \left( f^2(x_j) + \|Df\|_{(x_j, x_{j+1})}^2 \right).$$

Summing over  $j$  and estimating  $\|Df\|$  by  $\|D^{m+1}f\|$  yields (99) for  $\mu = 1$ .

Now assume that the analogue of (101) holds in dimensions up through  $\mu - 1$ . That is, for a cell

$$(102) \quad \mathcal{C}_{j_1, \dots, j_{\mu-1}} = \bigotimes_{i=1}^{\mu-1} (x_{i, j_i}, x_{i, j_i+1}),$$

$$(103) \quad \|f\|_{\mathcal{C}_{j_1, \dots, j_{\mu-1}}}^2 \leq Ch^{\mu-1} \left( f^2(x_{1, j_1}, \dots, x_{\mu-1, j_{\mu-1}}) + (f, f)_{\mu-1, [1], \mathcal{C}_{j_1, \dots, j_{\mu-1}}} \right),$$

where the inner product is restricted to the cell. Now consider a  $\mu$ -dimensional cell,  $\mathcal{C}_{j_1, \dots, j_\mu}$ . For any fixed  $x_\mu \in (x_{\mu, j_\mu}, x_{\mu, j_\mu+1})$ , (103) holds. Integrating in  $x_\mu$  then yields

$$(104) \quad \|f\|_{\mathcal{C}_{j_1, \dots, j_\mu}}^2 \leq Ch^{\mu-1} (I_1 + I_2),$$

where

$$(105) \quad I_1 = \int_{x_{\mu, j_\mu}}^{x_{\mu, j_\mu+1}} f^2(x_{1, j_1}, \dots, x_{\mu-1, j_{\mu-1}}, x_\mu) dx_\mu,$$

$$(106) \quad I_2 = \int_{x_{\mu, j_\mu}}^{x_{\mu, j_\mu+1}} (f(\cdot, x_\mu), f(\cdot, x_\mu))_{\mu-1, [1], \mathcal{C}_{j_1, \dots, j_{\mu-1}}} dx_\mu.$$

Now we apply (101) to estimate  $I_1$  and  $I_2$ :

$$(107) \quad I_1 \leq 2h_{\mu, j_\mu + \frac{1}{2}} \left( f^2(x_{1, j_1}, \dots, x_{\mu, j_\mu}) + \int_{x_{\mu, j_\mu}}^{x_{\mu, j_\mu+1}} (D_\mu f)^2(x_{1, j_1}, \dots, x_{\mu-1, j_{\mu-1}}, x_\mu) dx_\mu \right),$$

$$(108) \quad I_2 \leq 2h_{\mu, j_\mu + \frac{1}{2}} \left( (f(\cdot, x_{\mu, j_\mu}), f(\cdot, x_{\mu, j_\mu}))_{\mu-1, [1], \mathcal{C}_{j_1, \dots, j_{\mu-1}}} + 2h_{\mu, j_\mu + \frac{1}{2}} \int_{x_{\mu, j_\mu}}^{x_{\mu, j_\mu+1}} (D_\mu f(\cdot, x_\mu), D_\mu f(\cdot, x_\mu))_{\mu-1, [1], \mathcal{C}_{j_1, \dots, j_{\mu-1}}} dx_\mu \right).$$

Noting that all of the integral terms are bounded by  $(f, f)_{\mu, [1], \mathcal{C}_{j_1, \dots, j_\mu}}$  we conclude that (103) holds in dimension  $\mu$ . Equation (99) follows from summing over the cells and bounding  $(f, f)_{\mu, [1]}$  by  $(f, f)_{\mu, [m+1]}$ .

Now we prove the lemma by induction on  $\mu$ , again noting that we have already proven it for  $\mu = 1$ . Note that

$$(109) \quad |g|_{\mu, [m+1]}^2 = \left( \sum_i \int |D_i^{m+1} g|_{\mu-1, [m+1]}^2 dx_i + \sum_i \|D_i^{m+1} g\|^2 \right).$$

Using the induction hypothesis we have

$$(110) \quad \begin{aligned} \int |D_i^{m+1} g|_{\mu-1, [m+1]}^2 dx_i &\leq C \int (D_i^{m+1} g, D_i^{m+1} g)_{\mu-1, [m+1]} dx_i \\ &\leq C(g, g)_{\mu, [m+1]}, \end{aligned}$$

and by (99) with  $\mu$  replaced by  $\mu - 1$  and  $f$  replaced by  $D_i^{m+1}g$ :

$$\begin{aligned}
 \|D_i^{m+1}g\|^2 &= \int \|D_i^{m+1}g(\dots, x_i, \dots)\|^2 dx_i \\
 &\leq Ch^{\mu-1} \int (D_i^{m+1}g, D_i^{m+1}g)_{\mu-1, [m+1]} dx_i \\
 (111) \quad &+ Ch^{\mu-1} \sum \dots \sum \int (D_i^{m+1}g)^2(x_{1,j_1}, \dots, x_i, \dots, x_{\mu,j_\mu}) dx_i \\
 &\leq C(g, g)_{\mu, [m+1]}.
 \end{aligned}$$

This completes the proof of Lemma 6.3.  $\square$

**6.2. Convergence theorem.** We can now use this semi-inner product to generalize Theorem 3.1 to the multidimensional case.

**Theorem 6.1.** *Let  $u$  solve (21) with  $B = 0$ ,  $F = 0$  and  $A_j = \bar{A}_j$ , constant. Let  $f \in \left(\mathcal{H}_{\text{per}}^{[2m+2]}\right)^d$  and let  $p$  be the numerical approximation based on degree  $\mu(2m+1)$  Hermite tensor-product interpolation and degree  $\mu(2m+1)$  Taylor evolution with time step  $k$  satisfying (33). Finally, let  $T > 0$  be fixed. Then there is a constant  $C$ , independent of the grids, so that*

$$\|u^n - p^n\| + \|u^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\| \leq Ch^{2m+1} |f|_{\mu, [2m+2]} \quad \text{for } 0 \leq nk \leq T.$$

*Proof.* As in the proof of Theorem 3.1 we use the fact that the Taylor formula to advance  $p^n$  in time is exact at the nodes. Thus we repeat the arguments of Theorem 3.1 using the analogues of Lemmas 3.1–3.2 proven above.

Set

$$(112) \quad \varepsilon_n^2 = (e^n, e^n)_{\mu, [m+1]}, \quad \delta_n^2 = (\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n, \mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n)_{\mu, [m+1]},$$

and, by Lemmas 6.1 and 6.3,

$$(113) \quad \eta^n = u^{n+1} - \mathcal{I}u^{n+1}, \quad (\eta^n, \eta^n)_{\mu, [m+1]} = O(h^{2m+2}).$$

Then using Lemma 6.2 we derive (56) and (57). To estimate the error in the  $L_2$ -norm we again follow the arguments from the one-dimensional case and use (96) in conjunction with Lemma 6.3 to conclude

$$(114) \quad \|\mathcal{S}e^n - \tilde{\mathcal{I}}\mathcal{S}e^n\| \leq Ch^{m+1} \delta_n,$$

which leads to the final result.  $\square$

## 7. NUMERICAL EXPERIMENTS

In this section we present the results of a series of simple numerical experiments which demonstrate the resolution capabilities of the Hermite methods and illustrate the theoretical results.

An accepted framework within which to compare the efficiency of methods of different orders is to find the minimal number of points per wavelength required to propagate a wave a given number of periods with a fixed accuracy (e.g. [9, Ch. 3]). For standard central difference methods and periodic problems, this number is always greater than two, approaching two points per wavelength in the limit of infinite order, which yields the Fourier pseudospectral method [5]. As Hermite methods require more work and storage per point, to be competitive they must have superior resolution; that is, they must be capable of using fewer than two

points per wavelength. This will be demonstrated below. Note that more extensive experiments can be found in [7]. Also, the methods are used to discretize high-order radiation boundary conditions for the linearized Euler equations in [10].

Consider the initial value problem (1) for  $a = 1$  with

$$(115) \quad f(x) = e^{i\omega x},$$

and where  $\omega$  is an integer. We assume a uniform mesh of width  $h$  and evolve the discrete solution with the  $m$ -derivative Hermite method. To study subgrid resolution we allow

$$(116) \quad |\omega|h > \pi,$$

i.e., the function  $f(x) = e^{i\omega x}$  may oscillate between consecutive grid points  $x_j, x_{j+1}$ . We set

$$(117) \quad \xi = \omega h,$$

and note that the wavelength of the wave  $u = e^{i\omega(x+t)}$  is

$$(118) \quad \Lambda = \frac{2\pi}{|\omega|}.$$

Therefore,

$$(119) \quad \frac{\Lambda}{h} = \frac{2\pi}{|\omega h|} = \frac{2\pi}{|\xi|}$$

where  $\Lambda/h$  is the number of mesh points per wavelength  $\Lambda$ . For example,  $\xi = \pi$  corresponds to two mesh points per wavelength,  $\xi = 2\pi$  corresponds to one mesh point per wavelength, and  $\xi = \pi/10$  corresponds to 20 mesh points per wavelength.

Define the error function (the maximum is taken over all  $x_{j+\frac{1}{2}} \in \tilde{G}_h$ ),

$$(120) \quad \text{error}(\xi, m, \lambda, n) = \max \left\{ |u(x_{j+\frac{1}{2}}, nk) - p^n(x_{j+\frac{1}{2}})| \right\}.$$

We show some representative results for the error function (120). We have fixed the CFL number as  $\lambda = .8$  and the number of half time steps as  $n = 200$ . For these values the functions

$$(121) \quad \xi \rightarrow \eta(\xi, m) = \text{error}(\xi, m, \lambda = .8, n = 200)$$

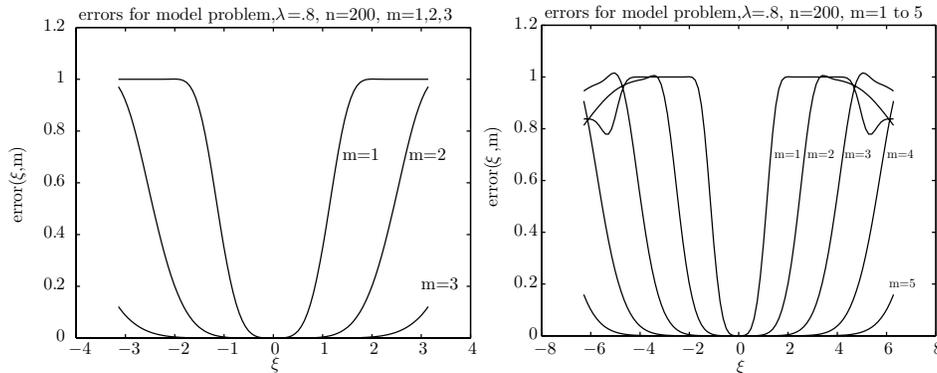


FIGURE 2. Error functions  $\eta(\xi, m)$  for  $|\xi| \leq \pi$  and  $|\xi| \leq 2\pi$

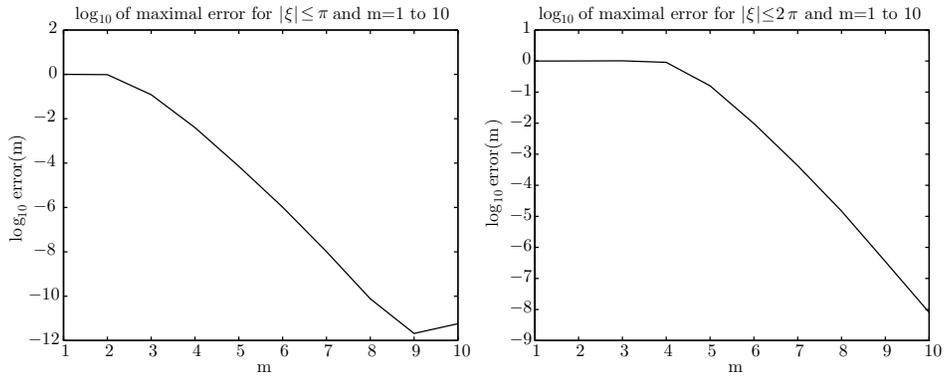


FIGURE 3. Logarithm of error functions maximized over  $|\xi| \leq \pi$  and  $|\xi| \leq 2\pi$

are plotted in Figure 2 for  $|\xi| \leq \pi$  and  $m = 1, 2, 3$  and for  $|\xi| \leq 2\pi$  and  $m = 1$  to 5. Note that the functions vanish to high order at  $\xi = 0$ , which is to be expected for any high-order method.

Again fixing  $\lambda = 0.8$  and  $n = 200$ , we have maximized these functions over  $\xi$  and have varied  $m$  from 1 to 10. In Figure 3 we plot the functions

$$(122) \quad m \rightarrow \log_{10} \max\{\eta(\xi, m) : |\xi| \leq \pi\}$$

and

$$(123) \quad m \rightarrow \log_{10} \max\{\eta(\xi, m) : |\xi| \leq 2\pi\} .$$

From these figures it is clear that, for  $m$  large, subgrid resolution is achieved. Thus the number of degrees of freedom required is comparable to that required by high-order difference methods, while we may take significantly larger time steps.

It is also of interest to study the convergence of the method in the limit  $m \rightarrow \infty$  when the number of wavelengths per cell is large. Ainsworth [1] has considered this problem for finite-element methods applied to the Helmholtz equation.

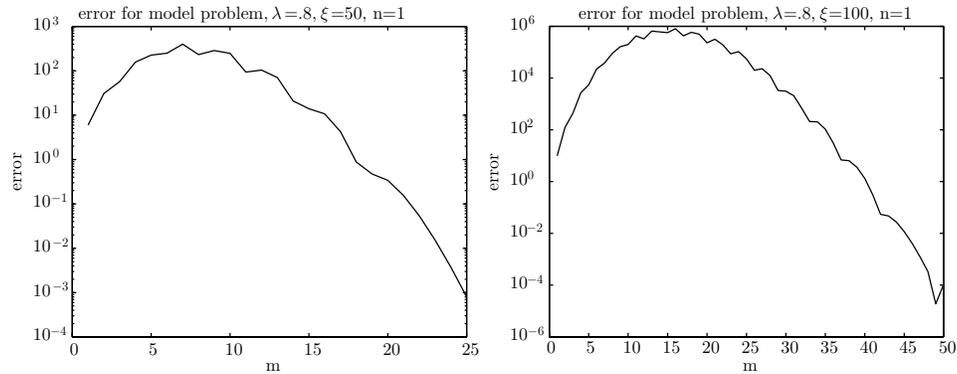


FIGURE 4. Logarithm of error functions with increasing  $m$

TABLE 1. Errors at  $t = 100$  in 1d with a uniform mesh

$m$	$q$	$N$	CFL	error	$p_{\text{obs}}$
2	5	40	0.8	$4.36E - 2$	
2	5	80	0.8	$1.64E - 3$	4.7
2	5	120	0.8	$2.90E - 4$	4.3
2	6	40	0.8	$4.52E - 2$	
2	6	80	0.8	$1.63E - 3$	4.8
2	6	120	0.8	$3.00E - 4$	4.2
4	9	25	0.8	$1.36E - 3$	
4	9	50	0.8	$3.98E - 6$	8.4
4	9	75	0.8	$1.16E - 7$	8.7
4	10	25	0.8	$1.11E - 3$	
4	10	50	0.8	$4.20E - 6$	8.0
4	10	75	0.8	$1.22E - 7$	8.7
7	15	10	0.6	$2.37E - 2$	
7	15	20	0.6	$5.10E - 7$	15.5
7	15	30	0.6	$9.73E - 10$	15.4
7	16	10	0.6	$8.92E - 3$	
7	16	20	0.6	$4.68E - 7$	14.2
7	16	30	0.6	$9.21E - 10$	15.4
7	17	10	0.8	$2.32E - 1$	
7	17	20	0.8	$7.77E - 7$	18.2
7	17	30	0.8	$9.78E - 10$	16.5
10	21	8	0.4	$2.37E - 3$	
10	21	16	0.4	$1.35E - 9$	20.7
10	21	24	0.4	$5.31E - 12$	13.7
10	22	8	0.4	$2.34E - 3$	
10	22	16	0.4	$1.38E - 9$	20.7
10	22	24	0.4	$5.23E - 12$	13.7
10	25	10	0.8	$3.05E - 4$	
10	25	15	0.8	$9.13E - 8$	20.0
10	25	20	0.8	$9.78E - 10$	26.3

Following that work, we fix  $N = 1$  and plot the error after two full time steps for  $\xi = 50$  and  $\xi = 100$ . (See Figure 4.) The graphs are qualitatively similar to those shown in [1], exhibiting, in particular, nonmonotonic behavior in increasing  $m$  followed by a superexponential convergence. Note that, in contrast to [1], our calculations include the effects of time-stepping.

We now consider full-blown implementations of the method in one and two space dimensions and study the order of convergence under mesh refinement. In all cases we take  $m = 2, 4, 7, 10$ . In our first experiment we solve

$$(124) \quad u_t = \left(1 + \frac{\sin x}{2}\right) u_x + f(x, t),$$

with periodic boundary conditions on the domain

$$(125) \quad (x, t) \in [0, 2\pi] \times [0, 100],$$

where  $f$  is chosen so that the solution is

$$(126) \quad u(x, t) = \cos(16(x + t)).$$

TABLE 2. Errors at  $t = 100$  in 1d with a random mesh

$m$	$q$	$N$	CFL	error	$p_{\text{obs}}$
2	5	55	0.8	$5.34E-2$	
2	5	110	0.8	$1.95E-3$	4.8
2	5	165	0.8	$3.02E-4$	4.6
2	6	50	0.8	$9.28E-2$	
2	6	100	0.8	$3.33E-3$	4.8
2	6	150	0.8	$5.08E-4$	4.6
4	9	25	0.8	$1.26E-2$	
4	9	50	0.8	$5.81E-5$	7.8
4	9	75	0.8	$1.23E-6$	9.5
4	10	25	0.8	$1.26E-2$	
4	10	50	0.8	$5.82E-5$	7.8
4	10	75	0.8	$1.23E-6$	9.5
7	15	12	0.8	$9.50E-3$	
7	15	24	0.8	$1.35E-6$	12.8
7	15	36	0.8	$5.46E-9$	13.6
7	16	12	0.8	$9.47E-3$	
7	16	24	0.8	$1.35E-6$	12.8
7	16	36	0.8	$5.46E-9$	13.6
10	21	10	0.8	$1.31E-3$	
10	21	20	0.8	$3.47E-10$	21.8
10	22	10	0.8	$1.34E-3$	
10	22	20	0.8	$3.48E-10$	21.9

We varied both the number of mesh points,  $N$ , and the number of terms,  $q$ , in the temporal Taylor series. In most cases we chose a CFL number of 0.8 and various numbers of mesh points. For the larger values of  $m$  and minimal  $q$  this was sometimes unstable. Then we found that the method could be stabilized either by decreasing the CFL number or by increasing  $q$ . (The latter method is typically more efficient.) We compute an estimated convergence rate by comparing errors for different meshes and computing

$$(127) \quad p_{\text{obs}} = \frac{\log\left(\frac{\text{err}_2}{\text{err}_1}\right)}{\log\left(\frac{h_2}{h_1}\right)}.$$

In our first experiment we use a uniform mesh, with results shown in Table 1. (The errors are relative  $l_2$  errors at  $t = 100$ .) Clearly, the observed convergence rates are consistent with the theory. Lower rates are observed only when errors are at a level where roundoff plays a role. We see that increasing  $q$  had little or no effect on the accuracy, but for  $m$  large had significant effects on the stability.

In our second experiment we solve the same problem, but introduce random perturbations to the mesh. Precisely we set

$$(128) \quad x_j = (j + \tau_j)h, \quad x_{j+\frac{1}{2}} = \left(j + \frac{1}{2} + \tau_{j+\frac{1}{2}}\right)h,$$

with  $\tau$  chosen randomly from a uniform distribution on  $(-.2, .2)$  and  $h = \frac{2\pi}{N}$ . Thus the mesh is not smooth. Note that the CFL constraint in this case is based on the

TABLE 3. Errors in 1d at  $t = 100$  with a reflective boundary condition

$m$	$q$	$N$	CFL	error	$p_{\text{obs}}$
2	5	40	0.7	$5.04E - 2$	
2	5	80	0.7	$1.85E - 3$	4.8
2	5	120	0.7	$4.51E - 4$	3.5
2	6	40	0.7	$4.32E - 2$	
2	6	80	0.7	$1.26E - 3$	5.1
2	6	120	0.7	$2.02E - 4$	4.5
4	9	25	0.6	$4.74E - 3$	
4	9	50	0.6	$5.41E - 6$	9.8
4	9	75	0.6	$1.23E - 7$	9.3
4	10	25	0.6	$1.12E - 3$	
4	10	50	0.6	$4.05E - 6$	8.1
4	10	75	0.6	$1.07E - 7$	9.0
7	15	10	0.4	$1.82E - 2$	
7	15	20	0.4	$4.74E - 7$	15.2
7	15	30	0.4	$9.24E - 10$	15.3
7	16	10	0.4	$9.37E - 3$	
7	16	20	0.4	$4.18E - 7$	14.5
7	16	30	0.4	$7.48E - 10$	15.5
7	17	10	0.5	$5.26E - 2$	
7	17	20	0.5	$3.36E - 7$	17.3
7	17	30	0.5	$8.66E - 10$	14.7
10	21	7	0.3	$8.38E - 3$	
10	21	14	0.3	$1.14E - 8$	19.5
10	22	7	0.3	$1.15E - 2$	
10	22	14	0.3	$1.31E - 8$	19.7
10	24	7	0.4	$8.17E - 2$	
10	24	14	0.4	$1.36E - 7$	19.2

smallest cell. Here it was never necessary to decrease the CFL number to achieve stability. The results are given in Table 2. We use average mesh width to compute  $p_{\text{obs}}$ . Of course this is somewhat imprecise. Given this imprecision, the agreement of the observed and theoretical convergence rates is reasonable. The results are essentially independent of  $q$ .

In our third set of experiments we test our implementation of nonperiodic boundary conditions as described in section 5. Here we solve the system:

$$(129) \quad u_{1,t} = \frac{\sin x}{2} u_{1,x} + u_{2,x} + f_1,$$

$$(130) \quad u_{2,t} = u_{1,x} + \frac{\sin x}{2} u_{2,x} + f_2,$$

where  $x \in [0, 2\pi]$  and  $f_1, f_2$  are chosen so that

$$(131) \quad u_1 = \sin(16x) \cos(16t), \quad u_2 = \cos(16x) \sin(16t).$$

TABLE 4. Errors in 2d at  $t = 100$ 

$m$	$q$	$N$	CFL	error	$p_{\text{obs}}$
2	5	50	0.8	$5.60E-2$	
2	5	100	0.8	$1.10E-3$	5.7
2	7	55	0.8	$4.09E-2$	
2	7	110	0.8	$1.32E-3$	5.0
2	10	60	0.8	$2.66E-2$	
2	10	120	0.8	$8.53E-4$	5.0
4	9	20	0.7	$4.08E-3$	
4	9	40	0.7	$5.77E-6$	9.5
4	13	20	0.7	$3.65E-3$	
4	13	40	0.7	$7.72E-6$	8.9
4	18	20	0.7	$3.65E-3$	
4	18	40	0.7	$7.72E-6$	8.9
7	15	6	0.6	$9.90E-2$	
7	15	12	0.6	$5.27E-6$	14.2
7	22	6	0.6	$7.15E-2$	
7	22	12	0.6	$6.99E-6$	13.3
7	30	6	0.6	$7.15E-2$	
7	30	12	0.6	$6.99E-6$	13.3
10	21	4	0.6	$4.02E-2$	
10	21	8	0.6	$9.46E-8$	18.7
10	31	4	0.6	$2.92E-2$	
10	31	8	0.6	$6.69E-8$	18.7
10	42	4	0.6	$2.92E-2$	
10	42	8	0.6	$6.69E-8$	18.7

We impose the boundary conditions

$$(132) \quad u_1 = 0, \quad x = 0, 2\pi.$$

Note that at the boundaries the local characteristic variables are

$$(133) \quad u^R = u_1 - u_2, \quad u^L = u_1 + u_2,$$

so that the boundary condition takes the form  $u^R = -u^L$ . We use a uniform mesh.

The results, shown in Table 3, are similar to the periodic case, except that we needed to decrease the CFL number somewhat. We believe that this is due to the particular implementation of the boundary condition. We plan to carefully examine the stability characteristics of different implementations in the future. Note that again increasing  $q$  sometimes allowed us to compute with larger time steps.

In our final experiment we consider a periodic problem in two space dimensions:

$$(134) \quad u_t = .6u_{x_1} + .8u_{x_2}, \quad (x_1, x_2) \in [0, 2\pi] \times [0, 2\pi],$$

$$(135) \quad u = \cos(7x_1 + 9x_2 + 11.4t).$$

The results are shown in Table 4. Note that the theory we developed assumes  $q = 4m + 2$ . The experiments show that  $q$  can in practice be taken much smaller.

## 8. CONCLUSIONS

In conclusion, we have demonstrated that minimal stencil methods using derivatives of order  $m$  in each coordinate direction combined with Taylor series in time yield approximations to hyperbolic initial-boundary value problems which

- i:** are of order  $2m + 1$  in space and time;
- ii:** are stable under an order-independent stability constraint;
- iii:** allow a natural implementation of boundary conditions to the order of the method;
- iv:** achieve subgrid resolution for  $m$  large.

In future work we will extend our analysis to hyperbolic-parabolic systems and discuss the efficient implementation of the method for problems with simple nonlinearities. We believe the methods have great potential for applications to problems in scientific computing, and we are currently developing codes which utilize them to solve the compressible Navier-Stokes equations.

## 9. APPENDIX: AUXILIARY RESULTS

In this section we give, for completeness, basic lemmas estimating Hermite interpolants as well as the solution of (14). We begin with pointwise estimates of Hermite interpolants.

**Lemma 9.1.** *The Hermite interpolation operators,  $\mathcal{I}$  and  $\tilde{\mathcal{I}}$ , satisfy for  $0 \leq q \leq 2m + 1$ :*

$$(136) \quad \max_{x_k \leq x \leq x_{k+1}} |D^q \mathcal{I}g(x)| \leq C \|g\|_{C^{2m+2}(x_k, x_{k+1})},$$

$$(137) \quad \max_{x_k \leq x \leq x_{k+1}} |D^q \mathcal{I}g(x)| \leq Ch_{k+\frac{1}{2}}^\kappa \|\mathcal{I}g\|_{H^{m+1}(x_k, x_{k+1})},$$

$$\kappa = \min\left(-\frac{1}{2}, m + \frac{1}{2} - q\right),$$

$$(138) \quad \|D^q \mathcal{I}g\|_{L^2(x_k, x_{k+1})}^2 \leq C \sum_{j=0}^m h_{k+\frac{1}{2}}^{2j-2q+1} ((D^j g(x_k))^2 + (D^j g(x_{k+1}))^2).$$

*Proof.* We define the Hermite-Lagrange basis functions,  $Q_j(z)$ ,  $j = 0, \dots, m$ , on the standard interval  $z \in [0, 1]$  by:

- (1)  $Q_j$  is a polynomial of degree  $2m + 1$ ;
- (2)  $Q_j$  satisfies for  $0 \leq i, j \leq m$ :

$$\frac{d^i Q_j}{dz^i}(1) = \delta_{ij}, \quad \frac{d^i Q_j}{dz^i}(0) = 0.$$

Setting  $G(z) = g(x_i + h_{i+\frac{1}{2}}z)$  and constructing the Hermite interpolant,  $\mathcal{I}_S G$ , of  $G$  on the standard interval we find that  $\mathcal{I}g$  may be written:

$$(139) \quad \mathcal{I}g(x) = \sum_{j=0}^m \left( (-h_{i+\frac{1}{2}})^j D^j g(x_i) Q_j(1-z) + h_{i+\frac{1}{2}}^j D^j g(x_{i+1}) Q_j(z) \right),$$

where

$$z = \frac{x - x_i}{h_{i+\frac{1}{2}}}.$$

The interpolation error is expressed using the well-known Peano kernel representation [11, Ch. 5]. For  $g \in H^p$  and integers  $0 \leq q < p \leq 2m + 2$  we have

$$\begin{aligned}
 D^q(g - \mathcal{I}g) &= h_{i+\frac{1}{2}}^{-q} \frac{d^q}{dz^q} (G - \mathcal{I}_S G) \\
 (140) \quad &= h_{i+\frac{1}{2}}^{-q} \int_0^1 K_{q,m,p}(z, s) \frac{d^p G}{dz^p}(s) ds \\
 &= h_{i+\frac{1}{2}}^{p-1-q} \int_{x_i}^{x_{i+1}} K_{q,m,p}(h_{i+\frac{1}{2}}^{-1}(x - x_i), h_{i+\frac{1}{2}}^{-1}(y - x_i)) D^p g(y) dy.
 \end{aligned}$$

Note that the kernel,  $K_{q,m,p}(z, s)$ , depends only on  $q, m, p$ , and its arguments,  $z$  and  $s$ , but not on  $g$ . It is also bounded.

To prove the first statement in Lemma 9.1 we write

$$|D^q \mathcal{I}g| \leq |D^q(g - \mathcal{I}g)| + |D^q g|.$$

By (140) the first term is bounded by  $C h_{i+\frac{1}{2}} \max_x |D^{2m+2}g(x)|$ . Thus (136) is immediate. To prove the remaining results we use norm equivalence arguments on the vector space of degree  $2m + 1$  polynomials of  $z$ . Suppose that  $0 \leq q \leq 2m + 1$  and  $0 \leq p \leq q$ . Then

$$\max_{0 \leq z \leq 1} \left| \frac{d^q \mathcal{I}_S G}{dz^q}(z) \right|^2 \leq C(q, m, p) \left\| \frac{d^p \mathcal{I}_S G}{dz^p} \right\|^2.$$

Thus

$$\max_{x_i \leq x \leq x_{i+1}} |D^q \mathcal{I}g(x)|^2 \leq C(q, m, p) h_{i+\frac{1}{2}}^{2p-1-2q} \int_{x_i}^{x_{i+1}} (D^p \mathcal{I}g(x))^2 dx.$$

Now (137) follows from choosing  $p = q$  when  $q \leq m + 1$  and  $p = m + 1$  for  $q > m + 1$ .

Finally we also have for any  $q$ :

$$\int_0^1 \left( \frac{d^q \mathcal{I}_S G}{dz^q}(z) \right)^2 dz \leq C(q, m) \sum_{j=0}^m \left( \left( \frac{d^j \mathcal{I}_S G}{dz^j}(0) \right)^2 + \left( \frac{d^j \mathcal{I}_S G}{dz^j}(1) \right)^2 \right).$$

Transforming back to  $x$  coordinates as before yields (138). □

We now consider solutions of (14). For simplicity we set

$$(141) \quad c = \max_{x,t} |A(x, t)|.$$

**Lemma 9.2.** *Suppose*

$$A, B \in C^r([0, T]; (C_{\text{per}}^r)^{d \times d}), \quad F \in C^r([0, T]; (C_{\text{per}}^r)^d), \quad f \in (C_{\text{per}}^r)^d.$$

*Then equation (14) has a solution,  $u \in C^r([0, T]; (C_{\text{per}}^r)^d)$ . Moreover, for any  $0 \leq q \leq r$  there exists  $\alpha(q), C(q)$  independent of the initial data and  $F$  such that for any  $x_0 \leq x_1, 0 \leq t_0 \leq t_1 \leq T$ :*

$$\begin{aligned}
 \left\| \frac{\partial^q u}{\partial x^q}(\cdot, t_1) \right\|_{(x_0, x_1)} &\leq e^{\alpha \Delta t} \left\| \frac{\partial^q u}{\partial x^q}(\cdot, t_0) \right\|_{(x_0 - c \Delta t, x_1 + c \Delta t)} \\
 &\quad + C \Delta t \sum_{j=0}^{q-1} \left\| \frac{\partial^j u}{\partial x^j}(\cdot, t_0) \right\|_{(x_0 - c \Delta t, x_1 + c \Delta t)} \\
 &\quad + C \Delta t \max_{t_0 \leq t \leq t_1} \sum_{j=0}^q \left\| \frac{\partial^j F}{\partial x^j}(\cdot, t) \right\|_{(x_0 - c \Delta t, x_1 + c \Delta t)},
 \end{aligned}$$

$$\begin{aligned}
\left| \frac{\partial^q u}{\partial x^q}(x, t_1) \right| &\leq e^{\alpha \Delta t} \max_{|z-x| \leq c \Delta t} \left| \frac{\partial^q u}{\partial x^q}(z, t_0) \right| \\
&+ C \Delta t \sum_{j=0}^{q-1} \max_{|z-x| \leq c \Delta t} \left| \frac{\partial^j u}{\partial x^j}(z, t_0) \right| \\
&+ C \Delta t \max_{t_0 \leq t \leq t_1} \sum_{j=0}^q \max_{|z-x| \leq c \Delta t} \left| \frac{\partial^j F}{\partial x^j}(z, t) \right|,
\end{aligned}$$

and for  $q_1 \geq 1$ :

$$\begin{aligned}
\left\| \frac{\partial^q u}{\partial t^{q_1} \partial x^{q-q_1}}(\cdot, t_1) \right\|_{(x_0, x_1)} &\leq C \sum_{j=0}^q \left\| \frac{\partial^j u}{\partial x^j}(\cdot, t_0) \right\|_{(x_0 - c \Delta t, x_1 + c \Delta t)} \\
&+ C \sum_{j=0}^q \sum_{l=0}^{\min(q_1-1, j)} \max_{t_0 \leq t \leq t_1} \left\| \frac{\partial^j F}{\partial t^l \partial x^{(j-l)}}(\cdot, t) \right\|_{(x_0 - c \Delta t, x_1 + c \Delta t)}, \\
\left| \frac{\partial^q u}{\partial t^{q_1} \partial x^{q-q_1}}(x, t_1) \right| &\leq C \sum_{j=0}^q \max_{|z-x| \leq c \Delta t} \left\| \frac{\partial^j u}{\partial x^j}(z, t_0) \right\| \\
&+ C \sum_{j=0}^q \sum_{l=0}^{\min(q_1-1, j)} \max_{t_0 \leq t \leq t_1} \max_{|z-x| \leq c \Delta t} \left| \frac{\partial^j F}{\partial t^l \partial x^{(j-l)}}(z, t) \right|.
\end{aligned}$$

Here  $\Delta t = t_1 - t_0$ .

*Proof.* The proof of the first two statements is based on standard energy arguments and domain of dependence considerations, e.g. [16, Ch. 3,6]. In particular, consider first

$$w = \frac{\partial^q u}{\partial x^q},$$

which satisfies an equation of the form

$$(142) \quad \frac{\partial w}{\partial t} = A \frac{\partial w}{\partial x} + B_q w + F_q,$$

where  $F_q$  depends linearly on space derivatives of  $u$  of order 1 through  $q-1$  as well as on the  $q$ th space derivative of  $F$ . An energy inequality for the entire domain takes the form

$$\frac{d}{dt} \|w\| \leq \alpha(q) \|w\| + \beta \|F_q\|.$$

This inequality implies

$$\|w(\cdot, t_1)\| \leq e^{\alpha(t_1-t_0)} \|w(\cdot, t_0)\| + \beta \frac{e^{\alpha(t_1-t_0)} - 1}{\alpha} \max_{t_0 \leq t \leq t_1} \|F_q(\cdot, t)\|.$$

Combined with a simple induction argument on  $q$  this inequality yields the desired result when the integrals are taken over the entire domain. The second inequality follows from using derivatives of the equation to express time derivatives in terms of space derivatives. Our final form follows from domain of dependence considerations [16, Ch. 3,6].

To derive maximum norm estimates we again look at (142), but now use the method of characteristics. Diagonalizing the system by a unitary transformation,

$Q(x, t)$ , we obtain for  $v = Q^T w$ :

$$\frac{\partial v}{\partial t} = \Lambda \frac{\partial v}{\partial x} + \tilde{B}_q v + \tilde{F},$$

where

$$\Lambda = \text{diag}(\lambda_i(x, t)).$$

Thus along the characteristic  $X_i(\zeta, t)$  defined by

$$\frac{dX_i}{dt} = -\lambda_i(X_i, t), \quad X_i(t_0) = \zeta,$$

we have

$$\frac{dv_i}{dt} = \left( \tilde{B}_q v + \tilde{F} \right)_i.$$

From these we easily derive the inequality

$$\max_x |v(x, t)| \leq \max_x |v(x, t_0)| + C \int_{t_0}^t \left( \max_x |v(x, s)| + \max_x |\tilde{F}(x, s)| \right) ds.$$

Application of Gronwall's lemma implies

$$\max_x |v(x, t_1)| \leq e^{\alpha \Delta t} \max_x |v(x, t_0)| + C \Delta t \max_{t_0 \leq t \leq t_1} \max_x |\tilde{F}(x, t)|.$$

The desired inequalities for all space derivatives then follow by induction and domain of dependence considerations. The inequalities for the mixed space-time derivatives follow from expressing time derivatives in terms of space derivatives.  $\square$

#### REFERENCES

1. M. Ainsworth, *Dispersive properties of high order finite elements*, Mathematical and Numerical Aspects of Wave Propagation Phenomena (G. Cohen, E. Heikkola, P. Joly, and P. Neittaanmäki, eds.), Springer, 2003, pp. 3–10. MR2077971
2. G. Birkhoff, M. Schultz, and R. Varga, *Piecewise Hermite interpolation in one and two variables with applications to partial differential equations*, Numer. Math. **11** (1968), 232–256. MR0226817 (37:2404)
3. B. Cockburn, G. E. Karniadakis, and C.-W. Shu (eds.), *Discontinuous Galerkin methods: Theory, computation and applications*, Lecture Notes in Computational Science and Engineering, vol. 11, New York, Springer-Verlag, 2000. MR1842160 (2002b:65004)
4. G. Cohen, *Higher-order numerical methods for transient wave equations*, Springer-Verlag, New York, 2002. MR1870851 (2002m:65069)
5. B. Fornberg and D. Sloan, *A review of pseudospectral methods for solving partial differential equations*, Acta Numerica **3** (1994), 203–267. MR1288098 (95j:65122)
6. P. Garabedian, *Partial differential equations*, Wiley, New York, 1964. MR0162045 (28:5247)
7. J. Goodrich, *Accurate finite difference algorithms*, Barriers and Challenges in Computational Fluid Dynamics (V. Venkatakrisnan, M. Salas, and S. Chakravarthy, eds.), Kluwer Academic Publishers, 1998, pp. 43–61. MR1607381
8. A. Griewank, *Evaluating derivatives: Principles and techniques of algorithmic differentiation*, SIAM, Philadelphia, 2000. MR1753583 (2001b:65003)
9. B. Gustafsson, H.-O. Kreiss, and J. Olinger, *Time-dependent problems and difference methods*, John Wiley, New York, 1995. MR1377057 (97c:65145)
10. T. Hagstrom and J. Goodrich, *Accurate radiation boundary conditions for the linearized Euler equations in Cartesian domains*, SIAM J. Sci. Comput. **24** (2002), 770–795. MR1950511 (2003k:35194)
11. G. Hämmerlin and K.-H. Hoffmann, *Numerical mathematics*, Springer-Verlag, New York, 1991. MR1088482 (92d:65001)
12. A. Harten, B. Engquist, S. Osher, and S. Chakravarthy, *Uniformly high-order accurate essentially nonoscillatory schemes, III*, J. Comput. Phys. **71** (1987), 231–303. MR0897244 (90a:65199)

13. J. Hesthaven and T. Warburton, *Nodal high-order methods on unstructured grids: I. Time-domain solution of Maxwell's equations*, J. Comput. Phys. **181** (2002), 186–221. MR1925981 (2003f:78034)
14. G.-S. Jiang and E. Tadmor, *Nonoscillatory central schemes for multidimensional hyperbolic conservation laws*, SIAM J. Sci. Comput. **19** (1998), 1892–1917. MR1638064 (99f:65128)
15. A. Jorba and M. Zou, *A software package for the numerical integration of ODEs by means of high-order Taylor methods*, Preprint, 2001.
16. H.-O. Kreiss and J. Lorenz, *Initial-Boundary Value Problems and the Navier–Stokes Equations*, Academic Press, New York, 1989. MR0998379 (91a:35138)
17. J.-L. Lions and O. Pironneau, *High precision with low order finite elements*, Preprint, 2000.
18. P. Olsson, *Summation by parts, projections and stability I*, Math. Comp. **64** (1995), 1035–1065. MR1297474 (96a:65130)
19. H. Takewaki, A. Nishiguchi, and T. Yabe, *Cubic interpolated pseudo-particle (CIP) method for solving hyperbolic-type equations*, J. Comput. Phys. **61** (1985), 261–268. MR0814444 (86m:65124)
20. H. Takewaki and T. Yabe, *The cubic interpolated pseudo-particle (CIP) method: Application to nonlinear and multi-dimensional hyperbolic equations*, J. Comput. Phys. **70** (1987), 355–372.
21. T. Yabe, T. Ishikawa, P. Wang, T. Aoki, Y. Kadota, and F. Ikeda, *A universal solver for hyperbolic equations by cubic-polynomial interpolation. II. Two- and three-dimensional solvers*, Comput. Phys. Comm. **66** (1991), 233–242. MR1125405 (92j:65132)

ACOUSTICS BRANCH, NASA GLENN RESEARCH CENTER, CLEVELAND, OHIO 44135

*E-mail address:* `John.Goodrich@grc.nasa.gov`

DEPARTMENT OF MATHEMATICS AND STATISTICS, THE UNIVERSITY OF NEW MEXICO, ALBUQUERQUE, NEW MEXICO 87131

*E-mail address:* `hagstrom@math.unm.edu`

DEPARTMENT OF MATHEMATICS AND STATISTICS, THE UNIVERSITY OF NEW MEXICO, ALBUQUERQUE, NEW MEXICO 87131

*E-mail address:* `lorenz@math.unm.edu`