

## Discontinuous Galerkin Methods for Ordinary Differential Equations\*

By M. Delfour, W. Hager and F. Trochu

**Abstract.** A class of Galerkin methods derived from discontinuous piecewise polynomial spaces is analyzed. For polynomials of degree  $k$ , these methods lead to a family of one-step schemes generating approximations up to order  $2k + 2$  for the solution of an ordinary differential equation.

**1. Introduction.** We study Galerkin approximations to ordinary differential equations using discontinuous, piecewise polynomial spaces. These schemes generalize a method proposed by Lesaint and Raviart [19]. The value of the approximation at  $t_j$ , a point of discontinuity in the approximating polynomial  $\mathbf{x}(\cdot)$ , is given by an average across the jump:  $\alpha_j \mathbf{x}(t_j^-) + (1 - \alpha_j) \mathbf{x}(t_j^+)$ . The case  $\alpha_j = 1$  is related to the scheme of Lesaint and Raviart while, for piecewise constant approximation, the values  $\alpha_j = 0, \frac{1}{2}$ , and 1 correspond, respectively, to Euler's explicit, improved, and implicit schemes.

For linear equations and piecewise polynomials of degree  $k$ , we prove superconvergence of order  $2k + 1$  when the  $\alpha_j$  lie in specified intervals. Experimentally, the same convergence rate is observed for nonlinear problems; and, moreover, for exceptional  $\alpha_j$ , the rate increases to  $2k + 2$ . The estimates of Lesaint and Raviart were based upon results of Butcher [6]–[9] and Crouzeix [11] for implicit Runge-Kutta methods. This strategy for estimating the error seems to fail when  $\alpha_j \neq 0$  or 1, and our approach is to view the bilinear form associated with the Galerkin problem as continuous with respect to *mesh-dependent norms* and then to study stability in these norms. This strategy is also used by Babuška and Osborn [3] for second order boundary value problems and Babuška, Osborn, and Pitkäranta [4] for some mixed methods applied to fourth order elliptic equations.

Galerkin approximations to differential equations using continuous piecewise polynomials are studied by Hulme [17], [18] and other references are found in his bibliography. Note, however, that these continuous approximations have order  $2k$  at the mesh points instead of  $2k + 1$ . Another discontinuous scheme with order  $2k + 2$  is studied by Delfour and Dubeau [13]. Applications of discontinuous methods are given by Lesaint and Raviart [19], Delfour and Trochu [14], Delfour [12], and Wellford and Oden [20], [22], [23], [24].

---

Received September 19, 1978; revised April 14, 1980.

1980 *Mathematics Subject Classification.* Primary 65L05.

\* This research was supported by National Sciences and Engineering Research Council Canada Grant A-8730, a Quebec Ministry of Education FCAC Grant, U. S. Office of Naval Research Grant N00014-76-C-0369, National Science Foundation Grant MCS-782556, and the Ford Foundation.

© 1981 American Mathematical Society  
0025-5718/81/0000-0056/\$05.75

**2. The Method.** Consider the initial value problem

$$(2.1) \quad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t), & t \in [0, T], \\ \mathbf{x}(0) = \mathbf{x}^0, \end{cases}$$

where  $\mathbf{x}: [0, T] \rightarrow R^n$  and  $\mathbf{f}: R^n \times [0, T] \rightarrow R^n$ . Assuming a solution exists, it is approximated using a discontinuous piecewise polynomial space. First, let us describe the space.

Given an integer  $N > 0$ , define  $h = T/N$ ,  $t_m = mh$ , and  $J_m = (t_{m-1}, t_m)$ . Let  $P^k[a, b]$  denote the set of polynomials defined on  $(a, b)$ , with degree at most  $k$ , and  $\Pi^h$  be the discontinuous piecewise polynomial space defined as follows:  $\mathbf{v} \in \Pi^h$  if and only if  $\mathbf{v}$  restricted to  $J_m$  lies in the  $n$ -fold Cartesian product

$$(P^k[t_{m-1}, t_m])^n = P^k[t_{m-1}, t_m] \times \cdots \times P^k[t_{m-1}, t_m]$$

for each  $m$ . The uniform mesh was introduced solely to simplify the exposition; the Galerkin scheme and analysis applies to grids for which the ratio between the biggest and the smallest mesh spacing is bounded uniformly in  $N$ .

Next, we formulate the Galerkin scheme. The dot product between vectors in  $R^n$  is denoted  $\mathbf{y} \cdot \mathbf{z}$  and the  $L^2[t_{m-1}, t_m]$  inner product is defined by

$$(\mathbf{v}, \mathbf{w})_m = \int_{t_{m-1}}^{t_m} \mathbf{v}(t) \cdot \mathbf{w}(t) dt.$$

Consider the problem: Find  $\mathbf{x} \in \Pi^h$  and  $\mathbf{x}(0^-)$  and  $\mathbf{x}(T^+) \in R^n$  such that

$$(2.2) \quad \begin{cases} \mathbf{x}_j^\alpha \cdot \mathbf{v}(t_j^-) = \mathbf{x}_{j-1}^\alpha \cdot \mathbf{v}(t_{j-1}^+) + (\mathbf{x}, \dot{\mathbf{v}})_j + (\mathbf{f}(\mathbf{x}), \mathbf{v})_j, \\ \mathbf{x}_0^\alpha = \mathbf{x}^0, \end{cases}$$

for all  $\mathbf{v} \in P^k[t_{j-1}, t_j]$  and  $j = 1, 2, \dots, N$ , where  $\mathbf{f}(\mathbf{x})$  denotes the map  $t \mapsto \mathbf{f}(\mathbf{x}(t), t)$ ,  $\alpha = \{\alpha_0, \alpha_1, \dots, \alpha_N\}$  is a given set of scalars, and  $\mathbf{x}^\alpha \in (R^n)^{N+1}$  is the vector with components

$$(2.3) \quad \mathbf{x}_j^\alpha = \alpha_j \mathbf{x}(t_j^-) + (1 - \alpha_j) \mathbf{x}(t_j^+).$$

Since the solution to (2.2) consists of both a piecewise polynomial and vectors  $\mathbf{x}(0^-)$  and  $\mathbf{x}(T^+) \in R^n$ , it is convenient to regard the elements of  $\Pi^h$  as defined on  $[0^-, T^+]$  with  $(\mathbf{x}(0^-), \mathbf{x}(T^+))$  and  $(\mathbf{x}(0^+), \mathbf{x}(T^-))$  possibly different.

In general, (2.2) is equivalent to  $n[N(k + 1) + 1]$  equations in  $n[N(k + 1) + 2]$  unknowns. To reduce the number of unknowns by  $n$ , we restrict ourselves to the following cases:

- (1)  $\alpha_0 = 0$  and  $\alpha_j \neq 1$  for  $j = 1, \dots, N$ , or
- (2)  $\alpha_N = 1$  and  $\alpha_j \neq 0$  for  $j = 0, \dots, N - 1$ .

In case (1), the unknown  $\mathbf{x}(0^-)$  disappears, giving us modified Runge-Kutta schemes (see Proposition 3.1); in case (2),  $\mathbf{x}(T^+)$  disappears and the schemes are generally implicit over the entire region  $[0^-, T]$  (see Appendix 2).

**3. Numerical Implementation.** To simplify the discussion in this section, we assume that  $n = 1$ . In solving (2.2), inner products are evaluated using  $(k + 1)$ -point quadrature. It is assumed that the abscissae  $\{\xi_0, \dots, \xi_k\} \subset [0, 1]$  and associated weights  $\{a_0, \dots, a_k\}$  are chosen so that

$$\int_0^1 p(t) dt = \sum_{i=0}^k a_i p(\xi_i)$$

for all  $p \in P^{2k-1}[0, 1]$ . Define the polynomials

$$\begin{aligned} \phi_m(\xi) &= \prod_{\substack{j=1 \\ j \neq m}}^k \frac{\xi - \xi_j}{\xi_m - \xi_j}, & 1 < m < k, \\ v_m(\xi) &= \int_{\xi}^1 \phi_m(t) dt, & 1 < m < k, \\ \psi_m(\xi) &= \prod_{\substack{j=0 \\ j \neq m}}^k \frac{\xi - \xi_j}{\xi_m - \xi_j}, & 0 < m < k, \end{aligned}$$

and

$$\psi_{jm}(t) = \begin{cases} \psi_m\left(\frac{t - t_{j-1}}{h}\right), & t \in J_j, \\ 0, & \text{elsewhere;} \end{cases}$$

$\phi_{jm}$  and  $v_{jm}$  are defined in a similar fashion. Expanding  $x \in \Pi^h$  in terms of the basis  $\{\psi_{jm}\}$ , we have

$$x(t) = \sum_{j=1}^N \sum_{m=0}^k x_{jm} \psi_{jm}(t),$$

where

$$x_{jm} = x(t_{jm}), \quad t_{jm} = t_{j-1} + h\xi_m.$$

**PROPOSITION 3.1.** Assume that  $\alpha_0 = 0$  and  $\alpha_j \neq 1$  for  $j = 1, \dots, N$ . If inner products are evaluated by quadrature, (2.2) is equivalent to the following system:

$$(3.1a) \quad x_{jl} = x_{j-1}^\alpha + h \sum_{m=0}^k a_{lm} f(x_{jm}, t_{jm}) + b_l z_{j-1}, \quad l = 0, \dots, k,$$

$$(3.1b) \quad x_j^\alpha = x_{j-1}^\alpha + h \sum_{m=0}^k a_m f(x_{jm}, t_{jm}),$$

where

$$(3.2) \quad \begin{cases} a_{lm} = a_m [\pi(\xi_l) \sigma_m + a_l^{-1} v_l(\xi_m)], & v_0 \equiv 0, \\ \sigma_m = - \sum_{i=1}^k \psi_i(0) v_i(\xi_m) a_i^{-1}, \\ b_l = \pi(\xi_l), \\ z_j = \frac{\alpha_j}{1 - \alpha_j} [x_j^\alpha - x(t_j^-)], \end{cases}$$

and  $\pi$  is the Legendre polynomial of degree  $k$ , translated and scaled to  $[0, 1]$ , with normalization  $\pi(0) = 1$ .

*Proof.* Inserting  $v = v_{jl}$  into (2.2), we get

$$(3.3) \quad 0 = v_l(0) x_{j-1}^\alpha + (x, \dot{v}_{jl})_j + (f(x), v_{jl})_j$$

for  $1 < l < k$ . Since the quadrature formula is exact for elements of  $P^{2k-1}$ ,

$$v_l(0) = \int_0^1 \phi_l(t) dt = a_l + a_0 \phi_l(\xi_0),$$

and

$$(x, \dot{v}_{jl})_j = -h^{-1}(x, \phi_{jl})_j = -[a_0\phi_l(\xi_0)x_{j0} + a_lx_{jl}].$$

But,  $a_l \neq 0$  so (3.3) becomes:

$$(3.4) \quad x_{jl} = x_{j-1}^\alpha - \frac{a_0}{a_l}\phi_l(\xi_0)[x_{j0} - x_{j-1}^\alpha] + \frac{1}{a_l}(f(x), v_{jl})_j.$$

Recalling that  $\pi$  is orthogonal to  $\phi_l$  and  $\pi\phi_l \in P^{2k-1}$ ,

$$0 = (\pi, \phi_l) = a_0\pi(\xi_0)\phi_l(\xi_0) + a_l\pi(\xi_l),$$

where  $(\cdot, \cdot)$  denotes the  $L^2[0, 1]$  inner product. Hence, (3.4) can be expressed as

$$(3.5) \quad x_{jl} = x_{j-1}^\alpha + \frac{\pi(\xi_l)}{\pi(\xi_0)} [x_{j0} - x_{j-1}^\alpha] + \frac{1}{a_l}(f(x), v_{jl})_j.$$

Now, by the definition of  $z_j$  and (2.3), we have

$$(3.6) \quad \begin{aligned} z_{j-1} &= \frac{\alpha_{j-1}}{1 - \alpha_{j-1}} [x_{j-1}^\alpha - x(t_{j-1}^-)] = x(t_{j-1}^+) - x_{j-1}^\alpha \\ &= \sum_{m=0}^k \psi_m(0) [x_{jm} - x_{j-1}^\alpha]. \end{aligned}$$

Substituting for  $x_{jm}$  using (3.5) gives us

$$(3.7) \quad \sum_{m=0}^k \psi_m(0) \frac{\pi(\xi_m)}{\pi(\xi_0)} (x_{j0} - x_{j-1}^\alpha) = - \sum_{m=1}^k \frac{\psi_m(0)}{a_m} (f(x), v_{jm})_j + z_{j-1}.$$

Since  $\pi \in P^k$  and  $\pi(0) = 1$ ,

$$1 = \pi(0) = \sum_{m=0}^k \pi(\xi_m)\psi_m(0).$$

Hence (3.7) reduces to the following:

$$(3.8) \quad x_{j0} = x_{j-1}^\alpha - \pi(\xi_0) \sum_{m=1}^k \frac{\psi_m(0)}{a_m} (f(x), v_{jm})_j + \pi(\xi_0)z_{j-1}.$$

Combining (3.5) and (3.8),

$$(3.9) \quad x_{jl} = x_{j-1}^\alpha + \pi(\xi_l) \left[ - \sum_{m=1}^k \frac{\psi_m(0)}{a_m} (f(x), v_{jm})_j + z_{j-1} \right] + \frac{1}{a_l}(f(x), v_{jl})_j$$

for  $1 < l < k$ . Finally, inserting  $v = 1$  in (2.2) yields

$$(3.10) \quad x_j^\alpha = x_{j-1}^\alpha + (f(x), 1)_j.$$

If inner products are evaluated by quadrature, (3.8)–(3.10) become (3.1).  $\square$

*Remark 3.2.* In general, (3.1a) is a nonlinear system of  $k + 1$  equations in the  $k + 1$  unknowns  $y_l = x_{jl}$ ,  $l = 0, \dots, k$ , which has the form

$$y = g(y).$$

If there exists a constant  $\lambda \geq 0$  such that

$$|f(x, t) - f(y, t)| \leq \lambda|x - y|$$

for all  $x, y \in R$  and  $t \in [0, T]$ , then, for  $h$  sufficiently small,  $g$  is a contraction on  $R^{k+1}$  and the successive substitutions

$$(3.11) \quad y^{l+1} = g(y^l)$$

converge to the solution of (3.1a). An initial guess is generated by extrapolating forward to the interval  $J_j$ , the polynomial computed on the previous mesh interval. This gives us

$$(3.12) \quad y_l^0 = \sum_{m=0}^k c_{lm} x_{j-1,m},$$

where  $c_{lm} = \psi_m(1 + \xi_j)$ .

If a column or a row of the matrix  $A = [a_{jm}]$  vanishes, one of the unknowns is computed explicitly, and we are left with  $k$  equations. The first row is zero if  $\xi_0 = 0$ , and we have

$$(3.13) \quad x_{j0} = x_{j-1}^\alpha + b_0 z_{j-1}.$$

The last column is zero when  $\xi_k = 1$ , and, moreover,

$$(3.14) \quad x_{jk} = x_{j-1}^\alpha + h \sum_{m=0}^{k-1} a_{km} f(x_{jm}, t_{jm}) + b_k z_{j-1}.$$

These values for  $\xi_0$  and  $\xi_k$  are produced by Gauss-Radau quadrature. For Gauss-Lobatto quadrature,  $\xi_0 = 0$  and  $\xi_k = 1$ ; hence, both  $x_{j0}$  and  $x_{jk}$  are explicitly determined, and we are left with  $k - 1$  equations. If  $\alpha_j = 0$  for all  $j$ , these Gauss-Radau and Gauss-Lobatto schemes reduce to those studied by Butcher [8].

When  $\xi_0 = 0$ , the computations associated with the interval  $J_j$  proceed as follows:

(1) Evaluate  $x_{j0}$  with (3.13).

(2) Compute  $\{x_{j1}, \dots, x_{jk}\}$  using the successive approximations (3.11) and the initial guess (3.12).

(3) Evaluate  $x_j^\alpha$  and  $z_j$  by (3.1b) and (3.2), respectively.

For  $\xi_k = 1$ , the computations become:

(1) Compute  $\{x_{j0}, \dots, x_{j,k-1}\}$  using successive substitutions on the first  $k$  equations in (3.1a).

(2) Evaluate  $x_{jk}$ ,  $x_j^\alpha$ , and  $z_j$  by (3.14), (3.1b), and (3.2), respectively.

*Remark 3.3.* If  $\alpha_j = 1$  for some  $j$ , the expression for  $z_j$  involves the factor  $1/0$ . However, making the new definition

$$z_j = \frac{1 - \alpha_j}{\alpha_j} [x_j^\alpha - x(t_j^+)],$$

we obtain a system similar to (3.1) that is valid when  $\alpha_j \neq 0, j = 0, \dots, N$ . In the special case  $\alpha_j = 1, j = 0, \dots, N$ , the new system is equivalent to (3.1), but with the  $b_j z_{j-1}$  term replaced by zero,  $\pi$  normalized so that  $\pi(1) = 1$ , and

$$\sigma_m = 1 - \sum_{i=1}^k \psi_i(1) v_i(\xi_m) a_i^{-1}.$$

*Remark 3.4.* Let  $\alpha^0 = 0$ . As  $\alpha_j$  for  $j > 1$  approaches  $-\infty$ , the piecewise polynomial satisfying (2.2) becomes *continuous*. Solving for  $x(t_j^+)$  from (2.3),

$$x(t_j^+) = [x_j^\alpha - \alpha_j x(t_j^-)] / (1 - \alpha_j).$$

Now take the limit as  $\alpha_j$  approaches  $-\infty$ . By (3.1), both  $x_j^\alpha$  and  $x(t_j^-)$  are independent of  $\alpha_j$ , and hence  $x(t_j^+) = x(t_j^-)$  in the limit. On the other hand,  $x_j^{-\infty} \neq x(t_j)$  in general.

For the limiting scheme, the computation of the approximating polynomial and  $\{x_j^{-\infty}\}$  essentially uncouples. We sum over  $j$  in (2.2), and integrate by parts to obtain

$$(3.15) \quad (\dot{x} - f(x), v) = 0$$

for all  $v \in \Pi^h$  with  $v$  continuous on  $[0, T]$  and  $v(T^-) = 0$ . When  $x \in \Pi^h$  is continuous and  $x(0^+) = x_0$ , (3.15) gives us  $kN$  equations in the same number of unknowns. After computing  $x$ , the  $\{x_j^{-\infty}\}$  is evaluated using (3.1b).

**4. Nodal Error.** The analysis of piecewise constant approximations, which follows conventional arguments, is given in Appendix 2. For  $k = 0$  and  $\alpha_0 = 0$ , the scheme is unstable unless  $\alpha_j < \frac{1}{2}$ ,  $1 \leq j \leq N$ . The value  $\alpha_j = \frac{1}{2}$  corresponds to the second order Euler improved scheme while the choice  $\alpha_j < \frac{1}{2}$  gives first order convergence. Henceforth, it is assumed that  $k > 1$ , and we make the simplifying assumption that

$$(4.1) \quad \mathbf{f}(\mathbf{x}, t) = A(t)\mathbf{x} + \mathbf{b}(t),$$

where  $A(t)$  is an  $n \times n$  matrix, and the elements of  $A$  and  $\mathbf{b}$  have square-integrable derivatives through order  $k$  on  $[0, T]$ . Our results, however, extend to the general nonlinear case. This section examines the error at the nodes.

*Notation.* The following spaces of functions  $\mathbf{f}: (a, b) \rightarrow R^n$  are utilized:

- $L^2[a, b]$ : Square-integrable functions.
- $L^\infty[a, b]$ : Essentially bounded functions.
- $H^k[a, b]$ : Functions with derivatives through order  $k$  in  $L^2[a, b]$ .
- $C[a, b]$ : Continuous functions.

For  $\mathbf{x} \in L^\infty[t_{j-1}, t_j]$ , let  $\|\mathbf{x}\|_{\infty, j}$  denote the essential supremum of  $|\mathbf{x}(t)|$  over  $t \in J_j$  where  $|\cdot|$  is the Euclidean norm, and for  $\mathbf{x} \in H^k[t_{j-1}, t_j]$ , define

$$\|\mathbf{x}\|_{k, j}^2 = \sum_{i=0}^k (\mathbf{x}^{(i)}, \mathbf{x}^{(i)})_j.$$

We also define the norms

$$\|\cdot\|_\infty = \max\{\|\cdot\|_{\infty, j} : j = 1, \dots, N\}$$

and

$$\|\cdot\|_k^2 = \sum_{j=1}^N \|\cdot\|_{k, j}^2.$$

The zero subscript on  $\|\cdot\|_0$  is often dropped; that is,  $\|\cdot\| \equiv \|\cdot\|_0$ . Finally, for any  $l > 0$ ,  $|\cdot|_\infty$  denotes the sup norm on  $R^l$ , and  $M^T$  is the transpose of the matrix  $M$ .

Letting  $\mathbf{x}^*$  denote the solution to (2.1) and assuming the existence of a solution  $\mathbf{x}^h$  to (2.2), we have

**THEOREM 4.1.** *There exists a constant  $c$ , independent of  $h$  and  $\alpha$ , such that*

$$(4.2) \quad |(\mathbf{x}^h - \mathbf{x}^*)^\alpha|_\infty \leq ch^k \|\mathbf{x}^h - \mathbf{x}^*\|.$$

*Proof.* Observe that  $\mathbf{x} = \mathbf{x}^*$  satisfies (2.2); subtracting relation (2.2) with  $\mathbf{x} = \mathbf{x}^h$  from the same relation with  $\mathbf{x} = \mathbf{x}^*$  gives us the following expression for the error

$\mathbf{e} = \mathbf{x}^* - \mathbf{x}^h$ :

$$(4.3) \quad \begin{cases} \mathbf{e}_0^\alpha = 0, \\ \mathbf{e}_j^\alpha \cdot \mathbf{v}(t_j^-) = \mathbf{e}_{j-1}^\alpha \cdot \mathbf{v}(t_{j-1}^+) + (\mathbf{e}, \dot{\mathbf{v}} + A^T \mathbf{v})_j, \end{cases}$$

for  $j = 1, 2, \dots, N$ , and  $\mathbf{v} \in \Pi^h$ . Summing (4.3) over  $j = 1, 2, \dots, m$ , yields

$$(4.4) \quad \mathbf{e}_m^\alpha \cdot \mathbf{v}(t_m^-) = \sum_{j=1}^m (\mathbf{e}, \dot{\mathbf{v}} + A^T \mathbf{v})_j$$

for all  $\mathbf{v} \in \Pi^h \cap C[0, T]$ . Let  $\mathbf{w} \in H^{k+1}[0, T]$  satisfy the relation

$$(4.5) \quad \begin{cases} \dot{\mathbf{w}} + A^T \mathbf{w} = 0, & \text{on } [0, T], \\ \mathbf{w}(t_m) = \mathbf{e}_m^\alpha. \end{cases}$$

By Lemma 4.2 below,

$$\|\mathbf{w}\|_{k+1} \leq c|\mathbf{e}_m^\alpha|,$$

where  $c$  denotes a generic constant throughout this paper.

Now let us recall, from Ciarlet [10] or Strang and Fix [21], that there exists an interpolation operator  $I: L^2[0, T] \rightarrow \Pi^h$  such that

$$(4.6) \quad \|\mathbf{w} - \mathbf{w}^I\|_s \leq ch^{m-s} \|\mathbf{w}\|_m,$$

whenever  $k + 1 \geq m \geq s$  and  $\|\mathbf{w}\|_m < \infty$ ; moreover, if  $k$  and  $m > 1$ ,  $\mathbf{w}^I$  can be chosen so that

$$\begin{cases} \mathbf{w}^I(t_j^+) = \mathbf{w}(t_j^+), \\ \mathbf{w}^I(t_j^-) = \mathbf{w}(t_j^-), \end{cases}$$

$j = 0, \dots, N$ . Hence, for the solution of (4.5),  $\mathbf{w}^I \in C[0, T]$ , and, by (4.6),

$$(4.7) \quad \|\mathbf{w} - \mathbf{w}^I\|_1 \leq ch^k \|\mathbf{w}\|_{k+1} \leq ch^k |\mathbf{e}_m^\alpha|.$$

Combining (4.5) and (4.7) leads to

$$(4.8) \quad \|\dot{\mathbf{w}}^I + A^T \mathbf{w}^I\| = \|\dot{\mathbf{w}}^I - \dot{\mathbf{w}} + A^T(\mathbf{w}^I - \mathbf{w})\| \leq ch^k |\mathbf{e}_m^\alpha|.$$

Finally, inserting  $\mathbf{v} = \mathbf{w}^I$  in (4.4), applying the Schwarz inequality, and utilizing (4.8) completes the proof.  $\square$

Although Theorem 4.1 applies to any  $\alpha$ , it does not guarantee convergence since the  $L^2$  error in  $\mathbf{x}^h$  may diverge. Also note that Theorem 4.1 holds for  $k = 0$ : just set  $\mathbf{v} = \mathbf{e}_m^\alpha$  in (4.4).

**LEMMA 4.2.** *Suppose that  $\mathbf{g}|_{J_m} \in H^k(J_m)$  for  $m = 1, 2, \dots, N$ , and  $v \in H^1[0, T]$  satisfies the equation*

$$(4.9) \quad \dot{\mathbf{v}} + A^T \mathbf{v} = \mathbf{g}$$

*almost everywhere on  $[0, T]$ . Then there exists a constant  $c$ , independent of  $\mathbf{g}$  and  $s \in [0, T]$ , such that*

$$\|\mathbf{v}\|_{k+1} \leq c\{|\mathbf{v}(s)| + \|\mathbf{g}\|_k\}.$$

*Proof.* Differentiate (4.9)  $k$  times to express  $\mathbf{v}^{(k+1)}$  in terms of  $\{\mathbf{v}, \mathbf{g}, \mathbf{g}^{(1)}, \dots, \mathbf{g}^{(k)}\}$ , then replace  $\mathbf{v}$  using the identity

$$\mathbf{v}(t) = \Phi(t, s)\mathbf{v}(s) + \int_s^t \Phi(t, \sigma)\mathbf{g}(\sigma) d\sigma,$$

where  $\Phi$  is the fundamental matrix associated with  $-A^T$ .  $\square$

5.  $L^2$  Error. In analyzing the  $L^2$  error, we utilize the following bilinear form

$$(5.1) \quad B(\mathbf{x}, \mathbf{v}) = (\mathbf{x}, \dot{\mathbf{v}} + A^T \mathbf{v}) + \sum_{j=0}^{N-1} \mathbf{x}_j^\alpha \cdot \delta \mathbf{v}_j - \mathbf{x}_N^\alpha \cdot \mathbf{v}(t_N^-),$$

where  $(\cdot, \cdot)$  denotes the  $L^2[0, 1]$  inner product and

$$\delta \mathbf{v}_j = \mathbf{v}(t_j^+) - \mathbf{v}(t_j^-).$$

Assuming  $f$  is given by (4.1), the Galerkin problem (2.2) is equivalent to the following: Find  $\mathbf{x} \in \Pi^h$  such that

$$(5.2) \quad B(\mathbf{x}, \mathbf{v}) + (\mathbf{b}, \mathbf{v}) + \mathbf{x}^0 \cdot \mathbf{v}(0^-) = 0$$

for all  $\mathbf{v} \in \Pi^h$ . Introducing the norms

$$\|\mathbf{x}\|_X^2 = \|\mathbf{x}\|^2 + \sum_{j=0}^N |\mathbf{x}_j^\alpha|^2$$

and

$$\|\mathbf{v}\|_V^2 = \|\dot{\mathbf{v}} + A^T \mathbf{v}\|^2 + |\mathbf{v}(t_N^-)|^2 + \sum_{j=0}^{N-1} |\delta \mathbf{v}_j|^2,$$

observe that  $|B(\mathbf{x}, \mathbf{v})| \leq \|\mathbf{x}\|_X \|\mathbf{v}\|_V$ .

The error  $\|\mathbf{x}^* - \mathbf{x}^h\|_X$  is now estimated. Define the balls

$$B_X = \{\mathbf{x} \in \Pi^h: \|\mathbf{x}\|_X = 1\} \quad \text{and} \quad B_V = \{\mathbf{v} \in \Pi^h: \|\mathbf{v}\|_V = 1\}$$

and the parameter

$$(5.3) \quad \gamma^h = \inf_{\mathbf{x} \in B_X} \sup_{\mathbf{v} \in B_V} B(\mathbf{x}, \mathbf{v}).$$

By Babuška [1], [2] or Brezzi [5], we have

**THEOREM 5.1.** *If  $\gamma^h > 0$ , there exists  $\mathbf{x}^h \in \Pi^h$  satisfying (2.2) and*

$$\|\mathbf{x}^* - \mathbf{x}^h\|_X \leq \left(1 + \frac{1}{\gamma^h}\right) \|\mathbf{x}^* - \mathbf{w}^h\|_X$$

for all  $\mathbf{w}^h \in \Pi^h$ .

Supposing that  $\gamma^h$  is bounded away from zero uniformly in  $h$ , let us substitute  $\mathbf{w}^h = \mathbf{x}^I$ , the interpolant of  $\mathbf{x}^*$ , to get

$$\|\mathbf{x}^* - \mathbf{x}^h\|_X \leq c \|\mathbf{x}^* - \mathbf{x}^I\|_X = c \|\mathbf{x}^* - \mathbf{x}^I\| \leq ch^{k+1} \|\mathbf{x}^*\|_{k+1}.$$

Since  $\|\mathbf{x}^* - \mathbf{x}^h\|_X \geq \|\mathbf{x}^* - \mathbf{x}^h\|$ , it follows that  $\|\mathbf{x}^* - \mathbf{x}^h\| = O(h^{k+1})$ . The remainder of this section addresses the problem of estimating  $\gamma^h$ .

Notice that  $\|\mathbf{x}\|_X$  consists of terms involving  $|\mathbf{x}^\alpha|$  and  $\|\mathbf{x}\|$ . As we shall see, the estimation of  $\gamma^h$  reduces to the companion cases:  $|\mathbf{x}^\alpha|/\|\mathbf{x}\|$  small or  $|\mathbf{x}^\alpha|/\|\mathbf{x}\|$  large. Let us begin with the first situation. Define the subspace

$$\Pi_0^h = \{\mathbf{x} \in \Pi^h: \mathbf{x}^\alpha = 0\},$$

let  $\Pi^{h,k-1} \subset \Pi^h$  be the subset of piecewise polynomials with degree at most  $k - 1$ , and let  $\mathcal{P}$  denote the  $L^2$  projection into  $\Pi^{h,k-1}$ .

LEMMA 5.2. *If there are  $\beta$  and  $\gamma$ , independent of  $h$ , such that one of the following conditions hold:*

(1)  $\alpha_0 = 0, -\infty < \beta < \alpha_j < \gamma < \frac{1}{2}$ , or

(2)  $\alpha_N = 1, \infty > \beta \geq \alpha_j \geq \gamma > \frac{1}{2}$ ,

for  $j = 1, 2, \dots, N - 1$ , then there exists  $\mu > 0$  (independent of  $h$ ) such that  $\mu\|\mathbf{x}\| < \|\mathcal{P}\mathbf{x}\| < \|\mathbf{x}\|$  for all  $\mathbf{x} \in \Pi_0^h$ .

*Proof.* Since  $\mathcal{P}$  is an  $L^2$  projection,  $\|\mathcal{P}\mathbf{x}\| < \|\mathbf{x}\|$ . Now let  $p$  denote the Legendre polynomial of degree  $k$ , translated and scaled to the interval  $[t_{j-1}, t_j]$  and normalized so that  $p(t_j) = 1$ . Hence, for any  $\mathbf{x} \in \Pi^h$ , there exists  $\mathbf{c}_j \in R^n$  such that

(5.4)  $\mathbf{x}(t) = (\mathcal{P}\mathbf{x})(t) + \mathbf{c}_j p(t), \quad t \in J_j.$

Recalling that  $p(t_{j-1}) = (-1)^k$ ,  $\mathbf{c}_j$  is evaluated:

$$\mathbf{c}_j = (-1)^k [\mathbf{x}(t_{j-1}^+) - \mathcal{P}\mathbf{x}(t_{j-1}^+)].$$

And substituting into (5.4) gives us

(5.5)  $\|\mathbf{x}\|_{\infty,j} \leq |\mathbf{x}(t_{j-1}^+)| + 2\|\mathcal{P}\mathbf{x}\|_{\infty,j},$

since  $\|p\|_{\infty,j} = 1$ . Squaring (5.5) and utilizing the relation

$$ab \leq (\eta a^2 + \eta^{-1} b^2)/2, \quad \eta > 0,$$

we have

$$\|\mathbf{x}\|_{\infty,j}^2 \leq (1 + \eta)|\mathbf{x}(t_{j-1}^+)|^2 + 4(1 + \eta^{-1})\|\mathcal{P}\mathbf{x}\|_{\infty,j}^2.$$

Summing over  $j = 1, \dots, N$  yields

(5.6)  $\sum_{j=1}^N \|\mathbf{x}\|_{\infty,j}^2 \leq (1 + \eta) \sum_{j=1}^N |\mathbf{x}(t_{j-1}^+)|^2 + 4(1 + \eta^{-1}) \sum_{j=1}^N \|\mathcal{P}\mathbf{x}\|_{\infty,j}^2.$

Now suppose that  $\mathbf{x} \in \Pi_0^h$ ; as noted earlier, the assumption  $\alpha_0 = 0$  implies that  $\mathbf{x}_0^+ = \mathbf{x}(0^+)$ . Since  $\mathbf{x}^- = 0$ , we see that  $\mathbf{x}(0^+) = 0$ , and, moreover, by (2.3),  $\mathbf{x}(t_j^+) = \rho_j \mathbf{x}(t_j^-)$ , where  $\rho_j = \alpha_j / (\alpha_j - 1)$  for  $j = 1, 2, \dots, N - 1$ . Since  $\lambda / (1 - \lambda)$  is a monotone function of  $\lambda < 1$ ,  $\lambda / (1 - \lambda) < 1$  for  $\lambda < \frac{1}{2}$  and, furthermore, both the parameter

$$\rho \equiv \text{maximum} \{ |\beta / (\beta - 1)|, |\gamma / (\gamma - 1)| \}$$

is  $< 1$ , and  $|\rho_j| < \rho$  for  $j = 1, \dots, N - 1$ . Combining these observations gives us

(5.7)  $\sum_{j=1}^N |\mathbf{x}(t_{j-1}^+)|^2 \leq \rho^2 \sum_{j=2}^N |\mathbf{x}(t_{j-1}^-)|^2 \leq \rho^2 \sum_{j=1}^{N-1} \|\mathbf{x}\|_{\infty,j}^2.$

Hence, by (5.6) and (5.7),

(5.8)  $[1 - (1 + \eta)\rho^2] \sum_{j=1}^N \|\mathbf{x}\|_{\infty,j}^2 \leq 4(1 + \eta^{-1}) \sum_{j=1}^N \|\mathcal{P}\mathbf{x}\|_{\infty,j}^2.$

Since  $\rho < 1$ , let us choose  $\eta$  so small that  $(1 + \eta)\rho^2 < 1$ . Recalling that there exist  $c_1, c_2 > 0$  (independent of  $h$ ) such that

$$c_1 \|\mathbf{x}\|^2 \leq h \sum_{j=1}^N \|\mathbf{x}\|_{\infty,j}^2 \leq c_2 \|\mathbf{x}\|^2$$

for all  $\mathbf{x} \in \Pi^h$ , (5.8) completes the proof when  $\alpha_0 = 0$ . The proof for  $\alpha_N = 1$  is almost identical.  $\square$

Applying this lemma gives us

**THEOREM 5.3.** *Under the hypotheses of Lemma 5.2, there exist positive constants  $c_1$ ,  $c_2$ , and  $\bar{h}$  such that*

$$\sup\{B(\mathbf{x}, \mathbf{v}) : \mathbf{v} \in B_\nu\} \geq c_1 \|\mathbf{x}\| - c_2 h^{1/2} |\mathbf{x}^\alpha|$$

for all  $\mathbf{x} \in \Pi^h$  and  $0 < h \leq \bar{h}$ .

*Proof.* Given  $\mathbf{x} \in \Pi^h$ , let  $\mathbf{z} \in \Pi^h$  be the continuous, piecewise linear polynomial satisfying  $\mathbf{z}(t_j) = \mathbf{x}_j^\alpha$  for  $0 < j \leq N$ , and define  $\mathbf{y} = (\mathbf{x} - \mathbf{z}) \in \Pi_0^h$ . By the triangle inequality,

$$(5.9) \quad |B(\mathbf{x}, \mathbf{v})| \geq |B(\mathbf{y}, \mathbf{v})| - |B(\mathbf{z}, \mathbf{v})|.$$

If  $\mathbf{v} \in H^1[0, T]$ ,  $\mathbf{v}(T^-) = 0$ , and  $\mathbf{v}(0^-) = \mathbf{v}(0^+)$ ,

$$|B(\mathbf{z}, \mathbf{v})| = |(\mathbf{z}, \dot{\mathbf{v}} + A^T \mathbf{v})| \leq \|\mathbf{z}\| \|\mathbf{v}\|_\nu.$$

Since  $\mathbf{z}$  is piecewise linear,

$$\|\mathbf{z}\|_{\infty, j} \leq \max\{|\mathbf{x}_j^\alpha|, |\mathbf{x}_{j-1}^\alpha|\}.$$

Therefore,  $\|\mathbf{z}\| \leq 2h^{1/2} |\mathbf{x}^\alpha|$ , and we have

$$(5.10) \quad |B(\mathbf{z}, \mathbf{v})| \leq 2h^{1/2} |\mathbf{x}^\alpha| \|\mathbf{v}\|_\nu.$$

Let  $\mathbf{w} \in H^1[0, T]$  satisfy the equation

$$\begin{cases} \dot{\mathbf{w}} + A^T \mathbf{w} = \mathcal{P} \mathbf{y}, & \text{a.e. on } [0, T], \\ \mathbf{w}(T^-) = 0, \end{cases}$$

where  $\mathcal{P}$  denotes the  $L^2$  projection defined above. By (4.6),

$$\|\mathbf{w} - \mathbf{w}'\|_1 \leq ch^k \|\mathbf{w}\|_{k+1},$$

and, by Lemma 4.2 and the fact that the polynomials composing  $\mathcal{P} \mathbf{y}$  have degree at most  $k - 1$ ,

$$\|\mathbf{w}\|_{k+1} \leq c \|\mathcal{P} \mathbf{y}\|_k = c \|\mathcal{P} \mathbf{y}\|_{k-1}.$$

Moreover, the mesh uniformity implies that

$$\|\mathcal{P} \mathbf{y}\|_{k-1} \leq ch^{1-k} \|\mathcal{P} \mathbf{y}\|,$$

the so-called inverse condition. Combining these relations, we get

$$(5.11) \quad \|\mathbf{w} - \mathbf{w}'\|_1 \leq ch \|\mathcal{P} \mathbf{y}\| \leq ch \|\mathbf{y}\|.$$

Now, by the definition of  $\mathbf{w}$  and Lemma 5.2,

$$\|\mathbf{w}\|_\nu = \|\mathcal{P} \mathbf{y}\| \begin{cases} \geq \mu \|\mathbf{y}\|, \\ \leq \|\mathbf{y}\|. \end{cases}$$

Hence, we have

$$(5.12) \quad \|\mathbf{w}'\|_\nu \begin{cases} \leq \|\mathbf{w}\|_\nu + \|\mathbf{w} - \mathbf{w}'\|_\nu \leq (1 + ch) \|\mathbf{y}\|, \\ \geq \|\mathbf{w}\|_\nu - \|\mathbf{w} - \mathbf{w}'\|_\nu \geq (\mu - ch) \|\mathbf{y}\|. \end{cases}$$

Again, by Lemma 5.2,

$$(5.13) \quad B(\mathbf{y}, \mathbf{w}) = (\mathbf{y}, \mathcal{P} \mathbf{y}) = (\mathcal{P} \mathbf{y}, \mathcal{P} \mathbf{y}) \geq \mu^2 \|\mathbf{y}\|^2,$$

and, from (5.11),

$$(5.14) \quad |B(\mathbf{y}, \mathbf{w}' - \mathbf{w})| \leq c \|\mathbf{w}' - \mathbf{w}\|_1 \|\mathbf{y}\| \leq ch \|\mathbf{y}\|^2.$$

Inserting  $v = w^l$  into (5.9), dividing by  $\|w^l\|_V$ , and utilizing (5.10) and (5.12)–(5.14) yields for  $h$  sufficiently small:

$$(5.15) \quad \sup_{v \in B_V} B(x, v) \geq (\mu^2 - ch)\|y\| - 2h^{1/2}|x^\alpha|.$$

On the other hand,

$$(5.16) \quad \|y\| \geq \|x\| - \|z\| \geq \|x\| - 2h^{1/2}|x^\alpha|.$$

Relations (5.15) and (5.16) complete the proof.  $\square$

Next, let us study the case where  $|x^\alpha|/\|x\|$  is large, and the following lemma is helpful.

**LEMMA 5.4.** *Given  $u \in (R^n)^N$ , suppose that  $v: [0, T] \rightarrow R^n$  satisfies the following relations:*

$$\begin{cases} \dot{v} + A^T v = 0 & \text{on } (t_{j-1}, t_j), \quad j = 1, 2, \dots, N, \\ \delta v_j = u_j, & j = 1, 2, \dots, N - 1, \\ v(T^-) = -u_N. \end{cases}$$

*Then there exists a constant  $c > 0$  (independent of  $h$  and  $u$ ) such that  $\|v\|_m < ch^{-1/2}|u|$ , whenever  $0 < m < k + 1$ .*

*Proof.* Define the function  $w: [0, T] \rightarrow R^n$  as follows:

$$(5.17) \quad w(t) = v(t) + \sum_{l=j}^N u_l, \quad t \in (t_{j-1}, t_j), \quad j = 1, \dots, N.$$

Observe that  $w$  belongs to  $H^1[0, T]$  and satisfies the identities

$$\begin{cases} w(T^-) = 0, \\ \dot{w}(t) + A(t)^T \left[ w(t) - \sum_{l=j}^N u_l \right] = 0, & t \in (t_{j-1}, t_j), j = 1, \dots, N. \end{cases}$$

Applying Lemma 4.2 with  $s = T$ , we get, for  $0 < m < k + 1$ ,

$$\|w\|_m^2 < c \sum_{j=1}^N h \left( \sum_{l=j}^N |u_l| \right)^2 < c \left\{ \sum_{j=1}^N |u_j| \right\}^2 < cN \sum_{j=1}^N |u_j|^2.$$

Combining this with (5.17) gives us

$$\|v\|_m < \|w\|_m + \left\{ \sum_{j=1}^N h \left( \sum_{l=j}^N |u_l| \right)^2 \right\}^{1/2} < cN^{1/2} \left\{ \sum_{j=1}^N |u_j|^2 \right\}^{1/2}. \quad \square$$

**THEOREM 5.5.** *There exist positive constants  $c_1, c_2$ , and  $\bar{h}$  such that, for any  $x \in \Pi^h$  and  $0 < h < \bar{h}$ ,*

$$\sup\{B(x, v): v \in B_V\} \geq c_1|x^\alpha| - c_2h^{1/2}\|x\|.$$

*Proof.* Given  $x \in \Pi^h$ , let  $w$  satisfy the following relations:

$$\begin{cases} \dot{w} + A^T w = 0 & \text{on } (t_{j-1}, t_j), \quad j = 1, \dots, N, \\ \delta w_j = x_j^\alpha, & j = 0, \dots, N - 1, \\ w(T^-) = -x_N^\alpha. \end{cases}$$

Since  $(\mathbf{w}' - \mathbf{w}) \in C[0, T]$ ,  $\mathbf{w}'(T^-) = \mathbf{w}(T^-)$ , and  $\mathbf{w}'(0^-) = \mathbf{w}(0^-)$ ,

$$(5.18) \quad \begin{aligned} \|\mathbf{w}' - \mathbf{w}\|_\nu &= \|(\dot{\mathbf{w}}' - \dot{\mathbf{w}}) + A^T(\mathbf{w}' - \mathbf{w})\| \\ &\leq c\|\mathbf{w}' - \mathbf{w}\|_1 \leq ch^k\|\mathbf{w}\|_{k+1} \leq ch^{k-1/2}|\mathbf{x}^\alpha|, \end{aligned}$$

where the last inequality comes from Lemma 5.4. Therefore, we have

$$(5.19) \quad \begin{aligned} |B(\mathbf{x}, \mathbf{w}')| &= |B(\mathbf{x}, \mathbf{w}) + B(\mathbf{x}, \mathbf{w}' - \mathbf{w})| \\ &\geq |\mathbf{x}^\alpha|^2 - \|\mathbf{x}\| \|\mathbf{w}' - \mathbf{w}\|_\nu \geq |\mathbf{x}^\alpha|^2 - ch^{k-1/2}\|\mathbf{x}\| |\mathbf{x}^\alpha| \end{aligned}$$

by (5.18). Similarly, (5.18) gives us

$$(5.20) \quad \|\mathbf{w}'\|_\nu \leq \|\mathbf{w}\|_\nu + \|\mathbf{w} - \mathbf{w}'\|_\nu = |\mathbf{x}^\alpha| + \|\mathbf{w} - \mathbf{w}'\|_\nu \leq (1 + ch^{k-1/2})|\mathbf{x}^\alpha|$$

and

$$(5.21) \quad \|\mathbf{w}'\|_\nu \geq (1 - ch^{k-1/2})|\mathbf{x}^\alpha|.$$

Dividing (5.19) by  $\|\mathbf{w}'\|_\nu$  and utilizing the inequalities (5.20) and (5.21), the proof is complete.  $\square$

Combining Theorems 5.3 and 5.5, we have

**COROLLARY 5.6.** *Under the hypotheses of Lemma 5.2,  $\gamma^h$  is uniformly bounded away from zero for  $h$  sufficiently small.*

*Proof.* Adding the inequalities in Theorems 5.3 and 5.5, we conclude that there exist positive scalars  $c_1, c_2$ , and  $\bar{h}$  such that

$$\sup_{\mathbf{v} \in B_\nu} B(\mathbf{x}, \mathbf{v}) \geq 2(c_1 - c_2h^{1/2})[\|\mathbf{x}\| + |\mathbf{x}^\alpha|]$$

for all  $0 < h \leq \bar{h}$  and  $\mathbf{x} \in B_X$ . Setting  $\eta = \text{minimum}\{\bar{h}, c_1^2/4c_2^2\}$  gives us

$$\sup_{\mathbf{v} \in B_\nu} B(\mathbf{x}, \mathbf{v}) \geq c_1[\|\mathbf{x}\| + |\mathbf{x}^\alpha|]$$

for all  $0 < h \leq \eta$ . But for  $\mathbf{x} \in B_X$ ,  $\|\mathbf{x}\| + |\mathbf{x}^\alpha| \geq 1$ ; therefore,  $\gamma^h > c_1$  for  $h$  sufficiently small.  $\square$

Combining Theorem 4.1, Corollary 5.6, and the remark below Theorem 5.1, we have

**THEOREM 5.7.** *If there are  $\beta$  and  $\gamma$  independent of  $h$  such that one of the following conditions hold:*

(1)  $\alpha_0 = 0, -\infty < \beta \leq \alpha_j \leq \gamma < \frac{1}{2}$ , or

(2)  $\alpha_N = 1, \infty > \beta \geq \alpha_j \geq \gamma > \frac{1}{2}$ ,

for  $j = 1, 2, \dots, N - 1$ , then

$$|(\mathbf{x}^h - \mathbf{x}^*)^\alpha|_\infty = O(h^{2k+1}) \quad \text{and} \quad \|\mathbf{x}^h - \mathbf{x}^*\| = O(h^{k+1}).$$

**Remark 5.8.** If  $\alpha_0 = 0$ , the assumption  $\alpha_j \leq \frac{1}{2}$  is necessary for the convergence of the Galerkin scheme. For  $\alpha_j > \frac{1}{2}$ , we detected numerical instability and divergence, while for the marginal cases  $\alpha_j = \frac{1}{2}$  or  $\alpha_j = -\infty$  and Gauss-Legendre quadrature, the convergence rates observed in numerical experiments with  $k \leq 3$  are reported in the following table.

TABLE A. Observed convergence rates for  $k \leq 3$  and  $\alpha_0 = 0$ .

k	even		odd	
	$\alpha_j$	1/2	$-\infty$	1/2
$\ \underline{x}^h - \underline{x}^*\ $	k+1	k	k	k+1
$ \underline{x}^h - \underline{x}^* _{\infty}^{\alpha}$	2k+2	2k	2k	2k+2

Similarly, if  $\alpha_N = 1$ , the scheme is unstable for  $\alpha_j < \frac{1}{2}$ ,  $j = 0, \dots, N - 1$ .

Remark 5.9. We have tested our scheme on the following problems with  $-\infty < \alpha_j < \frac{1}{2}$  and observed the same convergence rates established by Theorem 5.7:

- (i)  $\dot{x}(t) = x(t)$ ,  $t \in [0, 1], x(0) = 1$ .
- (ii)  $\dot{x}(t) = -10x(t)$ ,  $t \in [0, 1], x(0) = 1$ .
- (iii)  $\dot{x}(t) = x(t)^2$ ,  $t \in [0, .5], x(0) = 1$ .
- (iv)  $\dot{x}(t) = -2tx(t)^2$ ,  $t \in [0, 1], x(0) = 1$ .

Appendix 1. Examples.

TABLE B. Integration schemes for  $k = 1$ ,  $\alpha_0 = 0$  and  $\alpha_j \leq \frac{1}{2}$ .

	Legendre		Radau <sup>0</sup>		Radau <sup>1</sup>		Lobatto	
	m	0	1	0	1	0	1	0
$\xi_m$	$\frac{3-\sqrt{3}}{6}$	$\frac{3+\sqrt{3}}{6}$	0	$\frac{2}{3}$	$\frac{1}{3}$	1	0	1
$a_m$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$
$b_m$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	1	$-\frac{1}{3}$	$\frac{1}{3}$	-1	1	-1
$a_{0m}$	$\frac{1}{6}$	$\frac{2-\sqrt{3}}{6}$	0	0	$\frac{1}{3}$	0	0	0
$a_{1m}$	$\frac{2+\sqrt{3}}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	1	0	1	0
$\psi_m(1)$	$\frac{1-\sqrt{3}}{2}$	$\frac{1+\sqrt{3}}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	0	1	0	1
$c_{0m}$	$1-\sqrt{3}$	$\sqrt{3}$	x	x	$-\frac{1}{2}$	$\frac{3}{2}$	x	x
$c_{1m}$	$-\sqrt{3}$	$1+\sqrt{3}$	$-\frac{3}{2}$	$\frac{5}{2}$	x	x	x	x

An "x" is placed in row  $c_{jm}$  if  $x_{jl}$  is evaluated explicitly and the associated extrapolation coefficient is not needed.

TABLE C. *Integration schemes for  $k = 2$ ,  $\alpha_0 = 0$  and  $\alpha_j \leq \frac{1}{2}$ .*

m	Legendre			Radau <sup>9</sup>		
	0	1	2	0	1	2
$\xi_m$	$\frac{5-\sqrt{15}}{10}$	$\frac{1}{2}$	$\frac{5+\sqrt{15}}{10}$	0	$\frac{6-\sqrt{6}}{10}$	$\frac{6+\sqrt{6}}{10}$
$a_m$	$\frac{5}{18}$	$\frac{8}{18}$	$\frac{5}{18}$	$\frac{1}{9}$	$\frac{16+\sqrt{6}}{36}$	$\frac{16-\sqrt{6}}{36}$
$b_m$	$\frac{2}{5}$	$-\frac{1}{2}$	$\frac{2}{5}$	1	$-\frac{2+3\sqrt{6}}{25}$	$-\frac{2-3\sqrt{6}}{25}$
$a_{0m}$	$\frac{7}{60}$	$\frac{4-\sqrt{15}}{15}$	$\frac{7-2\sqrt{15}}{60}$	0	0	0
$a_{1m}$	$\frac{4+\sqrt{15}}{24}$	$\frac{1}{6}$	$\frac{4-\sqrt{15}}{24}$	$\frac{9+\sqrt{6}}{75}$	$\frac{24+\sqrt{6}}{120}$	$\frac{168-73\sqrt{6}}{600}$
$a_{2m}$	$\frac{7+2\sqrt{15}}{60}$	$\frac{4+\sqrt{15}}{15}$	$\frac{7}{60}$	$\frac{9-\sqrt{6}}{75}$	$\frac{168+73\sqrt{6}}{600}$	$\frac{24-\sqrt{6}}{120}$
$\psi_m(1)$	$\frac{5-\sqrt{15}}{6}$	$-\frac{2}{3}$	$\frac{5+\sqrt{15}}{6}$	$\frac{1}{3}$	$\frac{2-3\sqrt{6}}{6}$	$\frac{2+3\sqrt{6}}{6}$
$c_{0m}$	$\frac{13-3\sqrt{15}}{3}$	$\frac{4\sqrt{15}-20}{3}$	$\frac{10-\sqrt{15}}{3}$	x	x	x
$c_{1m}$	$\frac{10-\sqrt{15}}{3}$	$-\frac{17}{3}$	$\frac{10+\sqrt{15}}{3}$	$\frac{10-2\sqrt{6}}{3}$	$\frac{8-13\sqrt{6}}{6}$	$\frac{17\sqrt{6}-22}{6}$
$c_{2m}$	$\frac{10+\sqrt{15}}{3}$	$-\frac{20+4\sqrt{15}}{3}$	$\frac{13+3\sqrt{15}}{3}$	$\frac{10+2\sqrt{6}}{3}$	$-\frac{22+17\sqrt{6}}{6}$	$\frac{8+13\sqrt{6}}{6}$

TABLE D. *Integration schemes for  $k = 2$ ,  $\alpha_0 = 0$  and  $\alpha_j \leq \frac{1}{2}$ .*

m	Radau'			Lobatto		
	0	1	2	0	1	2
$\xi_m$	$\frac{4-\sqrt{6}}{10}$	$\frac{4+\sqrt{6}}{10}$	1	0	$\frac{1}{2}$	1
$a_m$	$\frac{16-\sqrt{6}}{36}$	$\frac{16+\sqrt{6}}{36}$	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$
$b_m$	$-\frac{2+3\sqrt{6}}{25}$	$-\frac{2+3\sqrt{6}}{25}$	1	1	$-\frac{1}{2}$	1
$a_{0m}$	$\frac{24-\sqrt{6}}{120}$	$\frac{24-11\sqrt{6}}{120}$	0	0	0	0
$a_{1m}$	$\frac{24+11\sqrt{6}}{120}$	$\frac{24+\sqrt{6}}{120}$	0	$\frac{1}{4}$	$\frac{1}{4}$	0
$a_{2m}$	$\frac{6-\sqrt{6}}{12}$	$\frac{6+\sqrt{6}}{12}$	0	0	1	0
$\psi_m(1)$	0	0	1	0	0	1
$c_{0m}$	$-\frac{16-7\sqrt{6}}{6}$	$\frac{2-3\sqrt{6}}{6}$	$\frac{10-2\sqrt{6}}{3}$	x	x	x
$c_{1m}$	$\frac{2+3\sqrt{6}}{6}$	$-\frac{16+7\sqrt{6}}{6}$	$\frac{10+2\sqrt{6}}{3}$	1	-3	3
$c_{2m}$	x	x	x	x	x	x

TABLE E. Convergence rates for  $k = 1$  and  $2$ ,  $\alpha_0 = 0$  and  $\alpha_j \leq \frac{1}{2}$ .

		$\ \underline{x}^h - \underline{x}^*\ $			$ (\underline{x}^h - \underline{x}^*)^\alpha _\infty$		
		$\frac{1}{2}$	$< \frac{1}{2}$	$-\infty$	$\frac{1}{2}$	$< \frac{1}{2}$	$-\infty$
$k = 1$	Gauss-Legendre	1	2	2	2	3	4
	Gauss-Radau	1	2	2	2	3	3
	Gauss-Lobatto	1	2	2	2	2	2
$k = 2$	Gauss-Legendre	3	3	2	6	5	4
	Gauss-Radau	3	3	2	5	5	4
	Gauss-Lobatto	3	3	2	4	4	4

Integration schemes derived from Gauss quadrature with  $k = 1$  and  $\alpha_j = 1$ ,  $j = 0, \dots, N$ , are listed below. In all cases, the convergence rates are 3 and 2 for  $|(\underline{x}^h - \underline{x}^*)^\alpha|_\infty$  and  $\|\underline{x}^h - \underline{x}^*\|$ , respectively, except that  $|(\underline{x}^h - \underline{x}^*)^\alpha|_\infty = O(h^2)$  for Gauss-Lobatto quadrature.

TABLE F. Integration schemes for  $k = 1$  and  $\alpha_j = 1$ .

		m	0	1
Gauss-Legendre	$a_{0m}$		$\frac{1}{3}$	$\frac{1-\sqrt{3}}{6}$
	$a_{1m}$		$\frac{1+\sqrt{3}}{6}$	$\frac{1}{3}$
Gauss-Radau <sup>0</sup>	$a_{0m}$		$\frac{1}{4}$	$-\frac{1}{4}$
	$a_{1m}$		$\frac{1}{4}$	$\frac{5}{12}$
Gauss-Radau <sup>1</sup>	$a_{0m}$		$\frac{5}{12}$	$-\frac{1}{12}$
	$a_{1m}$		$\frac{3}{4}$	$\frac{1}{4}$
Gauss-Lobatto	$a_{0m}$		$\frac{1}{2}$	$-\frac{1}{2}$
	$a_{1m}$		$\frac{1}{2}$	$\frac{1}{2}$

**Appendix 2. Piecewise Constant Approximation.** Let us consider the case  $k = 0$ . Given  $\mathbf{x} \in \Pi^h$ , let  $\mathbf{x}_m$  denote the value of  $\mathbf{x}$  on  $J_m$ . Hence, with  $\alpha_0 = 0$  and  $\alpha_j = \beta$  for  $j = 1, 2, \dots, N$ , (2.2) becomes the following:

$$(A1) \quad \begin{cases} \mathbf{x}_1 = \mathbf{x}^0, \\ (1 - \beta)(\mathbf{x}_2 - \mathbf{x}_1) = \mathbf{f}_1(\mathbf{x}_1), \\ (1 - \beta)(\mathbf{x}_{j+1} - \mathbf{x}_j) + \beta(\mathbf{x}_j - \mathbf{x}_{j-1}) = \mathbf{f}_j(\mathbf{x}_j), \quad j = 2, \dots, N, \end{cases}$$

where

$$(A2) \quad \mathbf{f}_j(\mathbf{x}_j) = \int_{t_{j-1}}^{t_j} \mathbf{f}(\mathbf{x}_j, t) dt.$$

The choices  $\beta = 0$  and  $1/2$  are related to Euler's explicit and improved methods, respectively.

Recall that (A1) is convergent if the zeros of the following characteristic equation are less than or equal to one in magnitude:

$$(1 - \beta)\lambda^2 + (2\beta - 1)\lambda - \beta = 0.$$

Since the zeros are  $1$  and  $\beta/\beta - 1$ , we see that (A1) is convergent for  $\beta \leq 1/2$ . Similarly, there is divergence for  $\beta > 1/2$ . Also recall that if (A1) is convergent, its rate is determined by the local truncation error. Expanding  $\mathbf{f}$  in a Taylor series about  $(\mathbf{x}_j, t_j)$ , it follows that the scheme is first order for  $\beta \leq 1/2$ .

Now let us consider the error in the  $\alpha$ -averaged variable

$$\mathbf{x}_j^\alpha = \beta\mathbf{x}_j + (1 - \beta)\mathbf{x}_{j+1}, \quad j = 1, \dots, N,$$

for  $\beta = 1/2$ . Adding (A1) to the same equation, but with  $j$  replaced by  $(j + 1)$ , and expanding  $\mathbf{f}$  in a Taylor series about  $(\mathbf{x}_j^\alpha, t_j)$ , we obtain

$$\mathbf{x}_{j+1}^\alpha = \mathbf{x}_{j-1}^\alpha + 2h\mathbf{f}(\mathbf{x}_j^\alpha, t_j) + O(h^3), \quad \beta = 1/2.$$

Deleting the  $O(h^3)$  term gives us Euler's improved method, a second order scheme. Hence  $\mathbf{x}_j^\alpha$  is correct to second order. As noted above, the Gauss-Legendre quadrature formula

$$\int_0^1 p(t) dt \approx p(1/2)$$

for (A2) preserves the second order convergence of  $\mathbf{x}_j^\alpha$ . For  $\beta < 1/2$ , the order drops to one.

If  $\alpha_N = 1$  and  $\alpha_j = \beta, j = 0, \dots, N - 1$ , (2.2) is equivalent to

$$(A3) \quad \begin{cases} \beta\mathbf{x}_0 + (1 - \beta)\mathbf{x}_1 = \mathbf{x}^0, \\ (1 - \beta)(\mathbf{x}_{j+1} - \mathbf{x}_j) + \beta(\mathbf{x}_j - \mathbf{x}_{j-1}) = \mathbf{f}_j(\mathbf{x}_j), \quad j = 1, \dots, N - 1, \\ \beta(\mathbf{x}_N - \mathbf{x}_{N-1}) = \mathbf{f}_N(\mathbf{x}_N). \end{cases}$$

For  $\beta = 1$ , this gives us Euler's implicit method, which is implicit over each mesh interval; but for  $\beta \neq 1$ , (A1) is implicit over  $[0, T]$ . Similar to the case  $\alpha_0 = 0$ , the

convergence rate for  $x^\alpha$  is 2 if  $\beta = 1/2$ , and 1 for  $\beta > 1/2$ ; the scheme diverges when  $\beta < 1/2$ .

Centre de Recherche de Mathématiques Appliquées  
Université de Montréal  
Montréal, Québec, Canada H3C 3J7

Department of Mathematics  
Pennsylvania State University  
University Park, Pennsylvania 16802

Centre de Recherche de Mathématiques Appliquées  
Université de Montréal  
Montréal, Québec, Canada H3C 3J7

1. I. BABUŠKA, "Error bounds for finite element method," *Numer. Math.*, v. 16, 1971, pp. 322–333.
2. I. BABUŠKA & A. K. AZIZ, "Survey lectures on the mathematical foundations of the finite element method," *The Mathematical Foundation of the Finite Element Method with Application to Partial Differential Equations* (A. K. Aziz, Ed.), Academic Press, New York, 1973.
3. I. BABUŠKA & J. OSBORN, "Analysis of finite element methods for second order boundary value problems using mesh dependent norms," *Numer. Math.*, v. 34, 1980, pp. 41–62.
4. I. BABUŠKA, J. OSBORN & J. PITKÄRANTA, "Analysis of mixed methods using mesh dependent norms," *Math. Comp.*, v. 35, 1980, pp. 1039–1062.
5. F. BREZZI, "On the existence, uniqueness, and approximation of saddle-point problems arising from Lagrangian multipliers," *R.A.I.R.O.*, v. 8, 1972, pp. 129–151.
6. J. C. BUTCHER, "Coefficients for the study of Runge-Kutta integration processes," *J. Austral. Math. Soc.*, v. 3, 1963, pp. 185–201.
7. J. C. BUTCHER, "Implicit Runge-Kutta processes," *Math. Comp.*, v. 18, 1964, pp. 50–64.
8. J. C. BUTCHER, "Integration processes based on Radau quadrature formulas," *Math. Comp.*, v. 26, 1964, pp. 233–244.
9. J. C. BUTCHER, "An algebraic theory of integration methods," *Math. Comp.*, v. 26, 1972, pp. 79–106.
10. P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
11. M. CROUZEIX, *Sur l'Approximation des Équations Différentielles Opérationnelles Linéaires par des Méthodes de Runge-Kutta*, Thèse de doctorat d'état es-sciences mathématiques, Université de Paris VI, mars, 1975.
12. M. C. DELFOUR, "The linear quadratic optimal control problem for hereditary differential systems: Theory and numerical solution," *J. Appl. Math. and Opt.*, v. 3, 1977, pp. 101–162.
13. M. C. DELFOUR & F. DUBEAU, *Piecewise Discontinuous Polynomial Approximation of Nonlinear Ordinary Differential Equations*, Centre de Recherches Mathématiques, Université de Montréal, Report #865, 1979.
14. M. C. DELFOUR & F. TROCHU, "Discontinuous finite element methods for the approximation of optimal control problems governed by hereditary differential systems," *Distributed Parameter Systems: Modelling and Identification* (A. Ruberti, Ed.), Springer-Verlag, New York, 1978, pp. 256–271.
15. M. C. DELFOUR & F. TROCHU, "Approximation des équations différentielles et problèmes de commande optimale," *Ann. Sci. Math. Québec*, v. 1, 1977, pp. 211–225.
16. M. C. DELFOUR & F. TROCHU, *Discontinuous Approximation of Ordinary Differential Equations and Application to Optimal Control Problems*, Centre de Recherches Mathématiques, Université de Montréal, Report #751, 1977.
17. B. L. HULME, "Discrete Galerkin and related one-step methods for ordinary differential equations," *Math. Comp.*, v. 26, 1972, pp. 881–891.
18. B. L. HULME, "One step piecewise polynomial Galerkin methods for initial value problems," *Math. Comp.*, v. 26, 1972, pp. 415–425.
19. P. LESANT & P. A. RAVIART, "On a finite element method for solving the neutron transport equation," *Mathematical Aspects of Finite Elements in Partial Differential Equations* (C. de Boor, Ed.), Academic Press, New York, 1974, pp. 89–123.

20. J. T. ODEN & L. C. WELLFORD, JR., "Discontinuous finite element approximations for the analysis of acceleration waves in elastic solids," *The Mathematics of Finite Elements and Applications II* (J. R. Whiteman, Ed.), Academic Press, London, 1976, pp. 269-284.

21. G. STRANG & G. FIX, *An Analysis of the Finite Element Method*, Prentice-Hall, New York, 1973.

22. L. C. WELLFORD, JR. & J. T. ODEN, "Discontinuous finite-element approximations for the analysis of shock waves in nonlinearly elastic materials," *J. Comput. Phys.*, v. 19, 1975, pp. 179-210.

23. L. C. WELLFORD, JR. & J. T. ODEN, "A theory of discontinuous finite element Galerkin approximations of shock waves in nonlinear elastic solids, Part I: Variational theory," *Comput. Methods Appl. Mech. Engrg.*, v. 8, 1976, pp. 1-16.

24. L. C. WELLFORD, JR. & J. T. ODEN, "A theory of discontinuous finite-element Galerkin approximations of shock waves in nonlinear elastic solids, Part II: Accuracy and convergence," *Comput. Methods Appl. Mech. Engrg.*, v. 8, 1976, pp. 17-36.