COMMUNICATIONS ON
PURE AND APPLIED ANALYSIS
Volume 6, Number 3, September 2007

Website: http://AIMsciences.org

pp. 719–740

ERROR ANALYSIS OF THE P-VERSION DISCONTINUOUS GALERKIN METHOD FOR HEAT TRANSFER IN BUILT-UP STRUCTURES

Hideaki Kaneko*

Department of Mathematics and Statistics
Old Dominion University
Norfolk, Virginia 23529-0077, USA

Kim S. Bey

Thermal Structures Branch, Structure Division
NASA Langley Research Center
Hampton, VA 23681, USA

Peter Z. Daffer

Northrup Grumman Space Technology
293 Highway 247 South, WRG4
Warner Robins, GA 31088, USA

Puntip Toghaw

Department of Mathematics
Mahidol University, Bangkok, Thailand

Abstract. In this paper, we provide an error analysis for the p-version of the discontinuous Galerkin finite element method for a class of heat transfer problems in built-up structures. Also, a general form of the matrix associated with the discretization of time variable using the p-finite element basis functions is established. Many interesting properties of this matrix are obtained. Numerical examples are provided in the last section.

1. **Introduction.** The purposes of this paper are to report the state of the art information on the time discretization techniques in the discontinuous Galerkin method for parabolic problems (this section) and to establish an error analysis for the p-version of the finite element method for such problems (Section 2). One of the applications in which we are particularly interested is a class of heat conduction problems over a thin body in R^3. We list all known facts and theorems in this section. Using the theorems in this section, we may obtain immediately error analysis for the current method. They will be presented in Section 2. In Section 3, various time discretization techniques will be discussed. A general form of the matrix which results from the discretization of time variable using p-finite element basis functions is presented. A computational method that uses the real Schur

2000 *Mathematics Subject Classification.* Primary: 65M60, 65N30; Secondary: 35K05.
Key words and phrases. Discontinuous Galerkin method, heat equations.
*This author is supported by NASA- Grant NAG-1-01092.

decomposition is presented. The real Schur decomposition along with a block backward substitution technique is used in examples in the last section, demonstrating the effectiveness of the method. For a more computational efficiency, it is desirable that the matrices involved in the discontinuous Galerkin method are diagonalizable as, in that instance, the system of linear equations arising from the discontinuous Galerkin method can be decoupled into a cluster of linear systems each of a smaller size. An extensive investigation concerning the issue of diagonalization of one of the matrices involved was done and many interesting properties of this matrix were obtained. They will be presented in Section 4. Despite these results, the issue of diagonalization is unsolved. However, we decided to include the results in Section 4 since they are interesting in themselves and they may serve to promote the reader of his interest to obtain a complete resolution of this problem. The reader is referred to the recent papers of Schötzau and Schwab [11] and Werber, Gerdes, Schötzau and Schwab [12] for related topics. After briefly describing a problem of start-up singularity related to parabolic problems in Section 5, we close this paper by presenting numerical examples in Section 6.

The discontinuous Galerkin method is applied to the following standard model problem of parabolic type:

Find u such that

$$\begin{aligned} &u_t(x,t) - \Delta u(x,t) = f(x,t), && x \in \Omega, t \in R^+,\\ &u(x,t) = 0, && x \in \partial\Omega, t \in R^+,\\ &u(x,0) = u_0(x), && x \in \Omega, \end{aligned} \tag{1.1}$$

where Ω is a closed and bounded set in R^2 or in R^3 with boundary $\partial\Omega$, $R^+ = (0,\infty)$, $\Delta u = \partial^2 u/\partial x^2 + \partial^2/\partial y^2$ or $\Delta u = \partial^2 u/\partial x^2 + \partial^2/\partial y^2 + \partial^2 u/\partial z^2$, $u_t = \partial u/\partial t$, and the functions f and u_0 are given data. In this paper, the region Ω, when it is located in R^3, is assumed to be a thin body in R^3, such as a panel on the wing or fuselage of an aerospace vehicle. A p-version of the finite element method is considered in all directions including time variable. Because of the special characteristics of the region, it is assumed that, through the thickness, only one element is taken. This allows us not only to avoid a construction of finite elements that are too thin to violate the quasiuniform condition (see (1.11) below) but also to treat (1.1) proposed in R^3 as a heat equation in R^2. The recent paper [6] addresses the similar issue under the framework of the 'modified' hp-finite element scheme. In the following, we use the notations used in [11]. For a Banach space X and $I = (0,T)$ indicating a time interval, we denote by $L^p(I;X)$, $1 \le p \le \infty$, and $H^k(I;X)$, $0 \le k \in R$, the Lebesgue and Sobolev spaces. Also $P^p(I;X)$ denotes the set of all polynomials of degree $\le p$ with coefficients in X, i.e., $q(t) \in P^p(I;X)$ if and only if $q(t) = \sum_{j=0}^{p} x_j t^j$ for some $x_j \in X$ and $t \in I$. Let T_I be a partition of I into $M(I)$ subintervals $\{I_n = (t_{n-1}, t_n)\}_{n=1}^{M(I)}$. We set $k_n = t_n - t_{n-1}$. Denote by u_n^+ and u_n^- the right and the left limits of u at t_n respectively. We set $[u]_n = u_n^+ - u_n^-$, $n = 1, \ldots, M(I) - 1$. For each time interval I_n, an approximation order $p_n \ge 0$ is assigned and they are stored in the vector $\bar{p} = \{p_n\}_{n=1}^{M(I)}$. The semidiscrete space is then given by

$$V^{\bar{p}}(T_I;X) = \{u\colon I \to X \colon u|_{I_n} \in P^{p_n}(I_n;X),\ 1 \le n \le M(I)\}.$$

If $\bar{p}$ is a constant vector, i.e., $p_n = p$ for all $1 \le n \le M(I)$, then we write $V^p(T_I;X)$. The number of degrees of freedom of the time discretization will be denoted by $\mathrm{NRDOF}(V^{\bar{p}}(T_I;X))$ and it is, of course, equal to $\sum_{n=1}^{M(I)}(p_n+1)$. The semidiscrete

solution $U \in P^{p_n}(I_n; X)$ of the problem (1.1), if U is already determined on I_k, $1 \le k \le n-1$, is found by solving the following problem:
Find $U \in P^{p_n}(I_n; X)$ such that

$$\int_{I_n} \{(U_t, V)_2 + (\bigtriangledown U, \bigtriangledown V)_2\} dt + (U_{n-1}^+, V_{n-1}^+)_2 = \int_{I_n} (g, V)_{X^* \times X} dt + (U_{n-1}^-, V_{n-1}^+)_2 \tag{1.2}$$

for all $V \in P^{p_n}(I_n; X)$ and $U_0^- = u_0$.

Here we assume that $L^2(\Omega)$ is densely embedded in a Banach space X. The following theorem and its subsequent corollary are reported in [11]. Theorem describes the error estimate for the semidiscrete solution explicitly in terms of time steps, the orders of spatial discretization and the local regularities of the solution.

Theorem 1. [11] *Let u be the solution of (1.1) and U the semidiscrete solution in $V^{\bar{p}}(T_I; X)$. Assume that $u|_{I_n} \in H^{s_n+1}(I_n; X)$ for $1 \le n \le M(I)$ and s_n, $1 \le n \le M(I)$, nonnegative integers. Then*

$$\|u - U\|^2_{L^2(I;X)} \le C \sum_{n=1}^{M(I)} (\frac{k_n}{2})^{2(s_n+1)} \max(1, p_n)^{-2} \frac{\Gamma(p_n + 1 - s_n^*)}{\Gamma(p_n + 1 + s_n^*)} \|u\|^2_{H^{s_n^*+1}(I_n;X)}$$

for any $0 \le s_n^* \le \min(p_n, s_n)$.

In the p-version of the discontinuous Galerkin finite element method, a temporal partition T_I is fixed and convergence is obtained by $p_n \to \infty$. For $p_n = p$ for each $n = 1, \ldots, M(I)$, and for a smooth solution u, we obtain the following:

Corollary 1. [11] *Let $p_n = p$ for each $n = 1, \ldots M(I)$ and $k_{max} = \max\{k_n\}$. Let $u \in H^{s_0+1}(I; X)$, for a nonnegative integer s_0, be the exact solution of (1.1) and $U \in V^p(T_I; X)$ the semidiscrete solution. Then*

$$\|u - U\|^2_{L^2(I;X)} \le C \frac{k_{max}^{2\{\min(s_0,p)+1\}}}{\max(1,p)^{2(s_0+1)}} \|u\|^2_{H^{s_0+1}(I;X)},$$

where C depends on s_0, but is independent of k_{max} and p.

Remark 1. As pointed out in [11], Corollary 1.2 shows that for smooth solutions where s_0 is large, it is better to increase p than to reduce k_{max} at a fixed, often low p. Since $N \equiv \mathrm{NRDOF}(V^{\bar{p}}(T_I; X)) \sim p$, we see that for p-version of the finite element method,

$$\|u - U\|_{L^2(I;X)} \le C N^{-s_0-1}. \tag{1.3}$$

Using the standard approximation theory for analytic functions, (1.3) reduces to

$$\|u - U\|_{L^2(I;X)} \le C e^{-bp} \tag{1.4}$$

for some $b > 0$ independent of p. If the solution is not smooth in time, it is still possible to approximate it in exponential orders by a hp-finite element method which combines a certain geometric partition with the semidiscrete space $V^{\bar{p}}(T_I; X)$ where $\bar{p}$ is linearly distributed, (see [11] for detail). Using the h-finite element method with non-uniform graded time partitions, such non-smooth solutions can be approximated in the order that is algebraically optimal, but not exponentially, using different approaches (see, e.g., [6], [11]). The standard p-finite element method does not perform well in this context. Therefore, since we aim to establish the p-version of the finite element method for (1.1) in this paper, we assume for the remainder of this paper that solutions are smooth in time. This assumption allows us to establish the order of approximation in time variable that is compatible with the approximation orders in the spatial variables.

Now we consider the problem of discretizing the space Ω. An extensive treatment of the subject can be found, -e.g., in [9]. In order to make this paper self contained, included below is a brief introduction to the p-finite element method applied to the problem of current interest. For simplicity, we assume that

$$\Omega = \omega \times [-\frac{d}{2}, \frac{d}{2}] = \sum_{l=1}^{M(h)} K^l \times [-\frac{d}{2}, \frac{d}{2}]$$

where ω is divided into $M(h)$ number of triangular elements, each denoted by K^l. Let $K_{\xi,\eta}$ denote the master triangular element defined by

$$K_{\xi,\eta} = \{(\xi,\eta) \in R^2 : \begin{array}{ll} 0 \le \eta \le (1+\xi)\sqrt{3} & -1 \le \xi \le 0 \text{ or} \\ 0 \le \eta \le (1-\xi)\sqrt{3} & 0 \le \xi \le 1\}. \end{array}$$

Let $S^p(K_{\xi,\eta})$ denote the space of polynomials of degree $\le p$ on $K_{\xi,\eta}$, -i.e.,

$$S^p(K_{\xi,\eta}) = span\{\xi^i \eta^j : i,j = 0,1,\dots,p\,; i+j \le p\}.$$

First, the shape functions for the master element $K_{\xi,\eta}$ are formed. To accomplish this, the barycentric coordinates are introduced via

$$\lambda_1 = (1-\xi-\eta/\sqrt{3})/2, \quad \lambda_2 = (1+\xi-\eta/\sqrt{3})/2, \quad \lambda_3 = \eta/\sqrt{3}.$$

λ_i's form a partition of unity and λ_i is identically equal to one at a vertex of $K_{\xi,\eta}$ and vanishes on the opposite side of $K_{\xi,\eta}$. The hierarchical shape functions on $K_{\xi,\eta}$ consists of internal as well as external functions. The normalized antiderivatives of the Legendre polynomials are defined by

$$\bar{\psi}_i(\zeta) = \sqrt{\frac{2i+1}{2}} \int_{-1}^{\zeta} P_i(t)dt \qquad i = 1,2,3,\dots$$

Now, the external shape functions consist of 3 nodal shape functions

$$N_i(\xi,\eta) = \lambda_i \qquad i = 1,2,3,$$

and $3(p-1)$ side shape functions $N_i^{[j]}(\xi,\eta)$, $i = 1,\dots,p-1$, $j = 1,2,3$. The index j indicates one of three sides of $K_{\xi,\eta}$. Noting that $\bar{\psi}_i(\pm 1) = 0$,

$$\bar{\psi}_i(\eta) = \frac{1}{4}(1-\eta^2)\varphi_i(\eta) \qquad i = 1,2,3,\dots \tag{1.5}$$

where $\varphi_i(\eta)$ is a polynomial of degree $i-1$. For instance, $\varphi_1(\eta) = -\sqrt{6}$, $\varphi_2(\eta) = -\sqrt{10}\eta$ and $\varphi_3(\eta) = \frac{\sqrt{14}}{4}(1-5\eta^2)$, etc. The side shape functions are constructed as follows:

$$\begin{array}{ll} N_i^{[1]}(\xi,\eta) & = \lambda_2\lambda_3\varphi_i(\lambda_3-\lambda_2) \\ N_i^{[2]}(\xi,\eta) & = \lambda_3\lambda_1\varphi_i(\lambda_1-\lambda_3) \qquad i = 1,\dots,p-1 \\ N_i^{[3]}(\xi,\eta) & = \lambda_1\lambda_2\varphi_i(\lambda_2-\lambda_1) \end{array} \tag{1.6}$$

From (1.5) and (1.6), there are $3+3(p-1)$ shape functions. As $dim(S^p(K_{\xi,\eta})) = \frac{(p+1)(p+2)}{2}$, the remaining $\frac{(p-1)(p-2)}{2}$ basis elements are constructed in terms of internal shape functions. Clearly, nontrivial internal shape functions on $K_{\xi,\eta}$ exists only if $p \ge 3$. For $p = 3$, the bubble function on $K_{\xi,\eta}$ defined below serves as the internal function;

$$b_{K_{\xi,\eta}}(\xi,\eta) = \lambda_1\lambda_2\lambda_3 = \frac{\eta}{4\sqrt{3}}((1-\frac{\eta}{\sqrt{3}})^2 - \xi^2).$$

Moreover, the collection $I^p(K_{\xi,\eta})$ of higher-order internal shape functions can be constructed from

$$I^p(K_{\xi,\eta}) = \{b_{K_{\xi,\eta}}v\colon v \in S^{p-3}(K_{\xi,\eta})\} = \{b_{K_{\xi,\eta}}\} \otimes S^{p-3}(K_{\xi,\eta}) \qquad p \geq 3.$$

Let T_h, $h > 0$, be a triangulation of ω. let $x = Q_x^l(L_1, L_2, L_3)$ and $y = Q_y^l(L_1, L_2, L_3)$ be the element mappings of the standard triangle $K_{\xi,\eta}$ to the lth triangular element $K^l \in T_h$, -e.g., the linear mapping onto K^l with vertices $\{(x_i^l, y_i^l)\}_{i=1}^3$,

$$Q_x^l(L_1, L_2, L_3) = \sum_{i=1}^3 x_i^l L_i, \qquad Q_y^l(L_1, L_2, L_3) = \sum_{i=1}^3 y_i^l L_i.$$

The space of all polynomials of degree $\leq p$ on K^l is denoted by $S^p(K^l)$ and its basis can be formed from the shape functions of $S^p(K_{\xi,\eta})$ described above by transforming them under Q_x^l and Q_y^l. The finite element space $S^{p,l}(\omega, T_h)$ is now defined. For ω, $p \geq 0$ and $k \geq 0$,

$$S^{p,k}(\omega, T_h) = \{u \in H^k(\omega)\colon u|_K \in S^p(K),\ K \in T_h\}.$$

Assume that a triangulation $\{T_h\}$, $h > 0$, of ω consists of $\{K_h^l\}_{l=1}^{M(h)}$ and that $h_l = diam(K_h^l)$, for $l = 1, \ldots, M(h)$.

In the z-variable for through the thickness approximation, the local variable τ is defined in the reference element $[-1, 1]$ and Γ is mapped onto the reference element by Q_z, i.e.,

$$\Gamma = Q_z([-1, 1]) \qquad z = Q_z(\tau).$$

Clearly, Q_z is a linear function defined by

$$z = Q_z(\tau) = \frac{1}{2}(1-\tau)(-\frac{d}{2}) + \frac{1}{2}(1+\tau)\frac{d}{2} \qquad \tau \in [-1, 1].$$

The Jacobian of Q_z is constant

$$\frac{dz}{d\tau} = \frac{d}{2}.$$

In this paper, the basis functions of $P^p([-1, 1])$ are taken to be the one-dimensional hierarchical shape functions, see [9].

For example, in approximating an element in $H^l[-1, 1]$, with $l = 0$, $\psi_i(\tau) = P_{i-1}(\tau)$, $1 \leq i \leq p+1$, where P_{i-1} is the Legendre polynomial of degree $i-1$, form the hierarchical basis functions. With $l = 1$, the external (ψ_1 and ψ_2) and internal (ψ_i, $i \geq 3$) shape functions are defined by

$$\begin{aligned} \psi_1(\tau) &= \frac{1-\tau}{2}, \quad \psi_2(\tau) = \frac{1+\tau}{2} \\ \psi_i(\tau) &= (\frac{2i-3}{2})^{1/2} \int_{-1}^{\tau} P_{i-2}(t)dt \qquad 3 \leq i \leq p+1. \end{aligned} \tag{1.7}$$

Note that ψ_i's form an orthogonal family with respect to the energy inner product $(\cdot, \cdot)_E$,

$$(\psi_i, \psi_j)_E \equiv \int_{-1}^1 \psi_i'(t)\psi_j'(t)dt = \int_{-1}^1 P_i(t)P_j(t)dt = \delta_{ij}.$$

Also note that the internal shape functions satisfy

$$\psi_i(\pm 1) = 0 \qquad \text{for } 3 \leq i \leq p+1.$$

For the case $l = 2$ and $p \geq 3$, the four nodal shape functions and the remaining $p-3$ internal shape functions given by

$$\psi_1(\tau) = \frac{1}{4}(1-\tau)^2(1+\tau), \quad \psi_2(\tau) = \frac{1}{4}(1-\tau)^2(2+\tau)$$
$$\psi_3(\tau) = -\frac{1}{4}(1+\tau)^2(1-\tau), \quad \psi_4(\tau) = \frac{1}{4}(1+\tau)^2(2-\tau)$$
$$\psi_i(\tau) = (\frac{2i-5}{2})^{1/2} \int_{-1}^{\tau}(\tau-\eta)P_{i-3}(\eta)d\eta \qquad i = 5, \dots, p+1.$$

In this case, the internal shape functions satisfy

$$\frac{d^j\psi_i}{d\tau^j}(\pm 1) = 0 \qquad \text{for } 5 \le i \le p+1 \text{ and } j = 0, 1.$$

For example, using the shape functions in (1.7), any element $u \in H^1[-1,1]$ can be approximated by $u_p \in P^p([-1,1])$, in the form

$$u_p(\tau) = \frac{1-\tau}{2}u(-1) + \frac{1+\tau}{2}u(1) + \sum_{i=3}^{p+1} a_i\psi_i(\tau).$$

For approximating a solution of parabolic problem with the homogeneous Dirichlet boundary condition, the first two terms will be dropped, as $u(-1) = u(1) = 0$. A sequence of triangulations $\{T_h\}_{h>0}$ is called the quasiuniform mesh if

$$\frac{h}{diam(K)} \le \gamma \qquad \text{for all } h > 0,$$

with $h = \max_{K \in T_h} diam(K)$, and for some $\gamma > 0$.

The following is proved by Babuška, Szabo and Katz in [1]. See also [2] by Babuška and Suri on a related discussion. Here Ω_0 denotes a bounded polygonal domain in R^2.

Theorem 2. [1] *Let $u \in H^k(\Omega_0)$. Then there exists a sequence $z_p \in P^p(\Omega_0)$, $p = 1, 2, \dots,$ such that, for any $0 \le l \le k$,*

$$\|u - z_p\|_{l,\Omega_0} \le Cp^{-(k-l)}\|u\|_{k,\Omega_0},$$

where C is independent of u and p.

The parameters k and l are not necessarily integral. Their proof relies heavily on the approximation power of the trigonometric polynomials.

With $l = 0$ in Theorem 1.3 and using the usual duality argument, the results in Theorem 1.3 are further improved by Babuška and Suri in [2] (Theorem 2.9), (see also a series of papers by Gui and Babuška [10]), to the *hp*-finite element setting as follows:

Theorem 3. [2] *Let T_h be a quasiuniform partition of Ω_0. Then for $k \ge 1$, $u \in H^k(\Omega_0)$,*

$$\inf_{v \in S^{p,k}(\omega,T_h)} \|u - v\|_{L_2(\Omega_0)} \le Ch^\nu p^{-k}\|u\|_{H^k(\Omega_0)}$$

where $\nu = \min(k, p+1)$.

Note that Corollary 1.2 can be derived also from Theorem 1.4, and they establish the same result in terms of time variable approximation and spatial variable approximation, respectively. The corresponding error estimate in the $\|\cdot\|_{H^k(\Omega_0)}$ is also available in [2].

Now assume that Ω_0 in Theorems 1.3 and 1.4 is $\Omega = \omega \times [-\frac{d}{2}, \frac{d}{2}]$ and consider the problem of approximating elements in $H^k(\omega \times [-\frac{d}{2}, \frac{d}{2}])$ by the tensor product space $S^{p_1,k}(\omega, T_h) \otimes P^{p_2}[-\frac{d}{2}, \frac{d}{2}]$. Using Theorems 1.3 and 1.4, the following is proved in [6]:

Theorem 4. [6] *Let* $u \in H^k(\omega \times [-\frac{d}{2}, \frac{d}{2}])$. *Then there exists* $u^* \in S^{p_1,k}(\omega, T_h) \otimes P^{p_2}([-\frac{d}{2}, \frac{d}{2}])$,

$$\|u - u^*\|_{L_2(\omega\times[-\frac{d}{2},\frac{d}{2}])} = O(h^{\nu_1} p_1^{-k} + d^{\nu_2} p_2^{-k})$$

where $\nu_i = \min(k, p_i + 1)$ *for* $i = 1, 2$ *and* $h = \max_{K \in T_h} diam(K)$, *with* T_h *a triangulation of* ω.

2. p-Version of discontinuous Galerkin finite element method.

In this section, a p-version of the discontinuous Galerkin finite element method for the parabolic problem (1.1) is described. The main goal here is to provide an error analysis for the p-finite element method using the results from Section 1. The semidiscrete approximation equation (1.2) is now upgraded to a fully discretized equation below. First, we modify Theorem 1.5 by recording below the approximation power of an element in $H^k(\omega \times [-\frac{d}{2}, \frac{d}{2}] \times I)$ by an element from its subspace $S^{p_1,k}(\omega, T_h) \otimes P^{p_2}[-\frac{d}{2}, \frac{d}{2}] \otimes P^{p_3}(I)$.

Theorem 5. *Let* $u \in H^k(\omega \times [-\frac{d}{2}, \frac{d}{2}] \times I)$. *Then there exists* $\bar{u} \in S^{p_1,k}(\omega, T_h) \otimes P^{p_2}([-\frac{d}{2}, \frac{d}{2}]) \otimes P^{p_3}(I)$,

$$\|u - \bar{u}\|_{L_2(\omega\times[-\frac{d}{2},\frac{d}{2}]\times I)} = O(h^{\nu} p_1^{-k} + d^{\nu} p_2^{-k} + k_{max}^{\nu} p_3^{-k})$$

where $\nu_i = \min(k, p_i + 1)$, $i = 1, 2, 3$ *and* $h = \max_{K \in T_h} diam(K)$, *with* T_h *a triangulation of* ω *and* $k_{max} = \max_{1 \le n \le N} k_n$.

Proof. Take X in $V^p(I; X)$ used in Corollary 1.2 to be $H^k(\omega \times [-\frac{d}{2}, \frac{d}{2}])$. Let $u^* \in V^{p_3}(I; H^k(\omega \times [-\frac{d}{2}, \frac{d}{2}]))$ be the element approximating u. From Corollary 1.2,

$$\|u - u^*\|_{L_2(I;H^k(\omega\times[-\frac{d}{2},\frac{d}{2}])} = O(k_{max}^{\nu_3} p_3^{-k}).$$

Approximate $u^*|_{\omega\times[-\frac{d}{2},\frac{d}{2}]}$ by $\bar{u}$ by Theorem 1.5. Then,

$$\begin{aligned}\|u - \bar{u}\|_{L_2(\omega\times[-\frac{d}{2},\frac{d}{2}]\times I)} &= \|u - u^*\|_{L_2(\omega\times[-\frac{d}{2},\frac{d}{2}]\times I)} + \|u^* - \bar{u}\|_{L_2(\omega\times[-\frac{d}{2},\frac{d}{2}]\times I)} \\ &= C_1 h^{\nu_1} p_1^{-k} + C_2 d^{\nu_2} p_2^{-k} + C_3 k_{max}^{\nu_3} p_3^{-k}\end{aligned}$$

where C_1, C_2 and C_3 are constants independent of h, k_{max}, d, p_1, p_2 and p_3. □

Using (1.4), we obtain

Corollary 2. *Let* $u \in H^\infty(\omega \times [-\frac{d}{2}, \frac{d}{2}] \times I)$. *Then there exists* $u^* \in S^{p_1,k}(\omega, T_h) \otimes P^{p_2}([-\frac{d}{2}, \frac{d}{2}]) \otimes P^{p_3}(I)$,

$$\|u - u^*\|_{L_2(\omega\times[-\frac{d}{2},\frac{d}{2}]\times I)} \le C e^{-b\min(p_1,p_2,p_3)}$$

where C *and* b *are independent of* p_1,p_2 *and* p_3.

Define

$$W^{(p_1,p_2,p_3)} = \{V : R^+ \to S^{p_1,k} \otimes P^{p_2}[-\frac{d}{2}, \frac{d}{2}] : V|_{I_n} \in P^{p_3}(I_n), n = 1, 2, \ldots, N\} \tag{2.1}$$

where

$$P^{p_3}(I_n) = \{V(t) = \sum_{i=0}^{p_3} V_i t^i : V_i \in S^{p_1,k} \otimes P^{p_2}[-\frac{d}{2}, \frac{d}{2}]\}.$$

Then the fully discretized discontinuous Galerkin method can be described as follows:

Find $U \in W^{(p_1,p_2,p_3)}$ such that for $n = 1, 2, \cdots$,

$$\int_{I_n} \{(U_t, V) + (\bigtriangledown U, \bigtriangledown V)\} dt + ([U]_{n-1}, V_{n-1}^+) = \int_{I_n} (f, v) dt \quad \text{for all } V \in W^{(p_1,p_2,p_3)}, \tag{2.2}$$

We consider in (1.1) only the case of isotropic materials along with Dirichlet boundary conditions. However, extensions to anisotropic materials as well as mixed boundary conditions where Neumann boundary conditions are incorporated are possible and the present analysis carries over to these cases. Numerical experiments recently done in [8] treat transversely anisotropic materials as well as isotropic materials along with a mixed boundary condition.

The solution u is approximated over $K^l \times (-\frac{d}{2}, \frac{d}{2}) \times I_n$ using the outer tensor products:

$$u|_{K^l \times (-\frac{d}{2}, \frac{d}{2}) \times I_n} \simeq (\phi \otimes \psi \otimes \theta)^T \mathbf{a}^n \equiv \chi^T \mathbf{a}^n \tag{2.3}$$

which is

$$u|_{K^l \times (-\frac{d}{2}, \frac{d}{2}) \times I_n} \simeq \sum_{\alpha=0}^{p_1} \sum_{\beta=0}^{p_2} \sum_{\gamma=0}^{p_3} \varphi_\alpha^l(x, y) \psi_\beta(z) \theta_\gamma(t) a_{\alpha\beta\gamma}^n.$$

Equation (2.2) becomes

$$\begin{aligned} \textstyle\sum_{K^l \in T_h} \quad & \textstyle\int_{K^l} \int_\Gamma [\int_{I_n} \chi (\frac{\partial \chi}{\partial t})^T + \bigtriangledown\chi(\bigtriangledown\chi)^T dt] + \chi(t_{n-1}^+)\chi(t_{n-1}^+)^T \, dz d\omega \cdot \mathbf{a}^n \\ & = \textstyle\sum_{k^l \in T_h} \int_{K^l} \int_\Gamma \int_{I_n} f\chi \, dt dz d\omega + \int_{K^l} \int_\Gamma \chi(t_{n-1}^-)\chi(t_{n-1}^-)^T dz d\omega \cdot \mathbf{a}^{n-1}. \end{aligned} \tag{2.4}$$

Equation (2.4) can be written in the following matrix form:

$$\sum_{K^l \in T_h} (\mathbf{C}_{K^l} + \mathbf{K}_{K^l} + \mathbf{M}_{K^l}) \cdot \mathbf{a}^n = \sum_{K^l \in T_h} (\mathbf{H}_{K^l} + \mathbf{M}_{K^l} \cdot \mathbf{a}^{n-1}) \tag{2.5}$$

where $\mathbf{C}_{K^l}$, $\mathbf{K}_{K^l}$ and $\mathbf{M}_{K^l}$ are, respectively, the element capacitance, conductance (stiffness) and mass matrices, whereas $\mathbf{H}_{K^l}$ is the element load vector. In the next section, we consider different bases in time variable, delineating an advantage of each choice of basis elements. The following theorem can be proved by minor modifications to the proof of Theorem 1.1, [3] and by making use of Theorem 2.1.

Theorem 6. *Let $u \in H^k(\omega \times [-\frac{d}{2}, \frac{d}{2}] \times I)$. Suppose that there is a constant γ such that the time steps k_n satisfy $k_n \leq \gamma k_{n+1}$, $n = 1, \ldots, N-1$ and let U^n denote the solution of (2.2) approximating u at t_n. Then there is a constant C depending only on γ and a constant β, where $\rho_K \geq \beta h_K$ and ρ_K is the diameter of the circle inscribed in K for all $K \in T_h$, such that for $n = 1, 2, \ldots, N$,*

$$\|u - U^n\|_{L_2(\omega \times [-\frac{d}{2}, \frac{d}{2}] \times I)} \leq C(1 + \log \frac{t_n}{k_n})^{1/2} \{C_1 h^{\nu_1} p_1^{-k} + C_2 d^{\nu_2} p_2^{-k} + C_3 k_{max}^{\nu_3} p_3^{-k}\}$$

where $\nu_i = \min(k, p_i + 1)$, $i = 1, 2, 3$; C_1, C_2 and C_3 are constants independent of h, k_{max}, d, p_1, p_2 and p_3.

Similarly, Corollary 2.2 implies the following

Corollary 3. *Let $u \in H^\infty(\omega \times [-\frac{d}{2}, \frac{d}{2}] \times I)$. Suppose that there is a constant γ such that the time steps k_n satisfy $k_n \leq \gamma k_{n+1}$, $n = 1, \ldots, N-1$ and let U^n denote the solution of (2.2) approximating u at t_n. Then there is a constant C depending only on γ and a constant β, where $\rho_K \geq \beta h_K$ and ρ_K is the diameter of the circle inscribed in K for all $K \in T_h$, such that for $n = 1, 2, \ldots, N$,*

$$\|u - U^n\|_{L_2(\omega \times [-\frac{d}{2}, \frac{d}{2}] \times I)} \leq C e^{-b \min(p_1, p_2, p_3)}$$

where C and b are independent of p_1,p_2 and p_3.

Remark 2. Maintaining throughout computation a certain accuracy of numerical solution obtained from (2.2) is always desirable and Theorem 2.3 gives an insight to the following adaptive scheme. Suppose that a tolerance of δ is required. Then, as h, d and k_{max} are known *a priori*, the minimum approximation degrees' requirement in spatial, through the thickness and time variable are obtained from

$$h^{\nu_1} p_1^{-k} < \delta, \quad d^{\nu_2} p_2^{-k} < \delta, \quad k_{max}^{\nu_3} p_3^{-k} < \delta. \tag{2.6}$$

Moreover, for the case of infinitely differentiable u, with $p = \min(p_1, p_2, p_3)$ where p_i's satisfy inequalities in (2.6), Corollary 2.2 implies that

$$\|u - U^n\|_{L_2(\omega\times[-\frac{d}{2},\frac{d}{2}]\times I)} \leq Ce^{-bp} < \delta.$$

Remark 3. Thus far, the case for the constant degree vectors was considered. Nonconstant degree vectors can also be incorporated easily from Theorem 1.1. For instance, nonconstant degree vector $\bar{p} = (p_n)$ can be derived from the inequality in Theorem 1.1. Note that the bound given in Theorem 1.1 combines all errors in time approximation over $M(I)$ intervals. Thus, for each $n = 1, 2, \ldots, M(I)$, the error in disrectization in time variable over I_n is given by

$$\|u-U\|^2_{L^2(I_n;X)} \leq C\,(\frac{k_n}{2})^{2(s_n+1)} \max(1,p_n)^{-2} \frac{\Gamma(p_n+1-s_n^*)}{\Gamma(p_n+1+s_n^*)} \|u\|^2_{H^{s_n^*+1}(I_n;X)} \tag{2.7}$$

for any $0 \leq s_n^* \leq \min(p_n, s_n)$. Since $\frac{\Gamma(p_n+1-s_n^*)}{\Gamma(p_n+1+s_n^*)} \sim p_n^{-2s_n^*}$ as $p_n \to \infty$, (2.7) becomes

$$\|u - U\|_{L^2(I_n;X)} \leq Ck_n^{s_n+1} p_n^{-(s_n+1)}.$$

As k_n's are known, construct the degree vector $\bar{p}$ by requiring each component p_n to satisfy

$$k_n^{s_n+1} p_n^{-(s_n+1)} < \delta.$$

A construction of nonconstant degree vector corresponding to a triangulation T_h for the region ω is similar.

3. **Discretizations in time variable.** In this section, effects of the use of different basis elements to approximate the solution in time variable are considered. Recall from (2.3), the solution u is approximated over $K^l \times (-\frac{d}{2}, \frac{d}{2}) \times I_n$ using the outer tensor products:

$$u|_{K^l\times(-\frac{d}{2},\frac{d}{2})\times I_n} \simeq (\phi \otimes \psi \otimes \theta)^T \mathbf{a}^n \equiv \chi^T \mathbf{a}^n.$$

To better illustrate the choice of basis elements for time variable, we write a semidiscrete solution $U_n \in P^{p_n}(I_n, X)$ by

$$U_n = \sum_{j=0}^{p_n} u_{j,n}\theta_{j,n}. \tag{3.1}$$

Here $u_{j,n} \in X$ are unknown coefficients to be determined. V_n is defined similarly. For convenience, we let $I = I_n$ and drop the time step index n. Substituting U_n of (3.1) and the corresponding V_n into (1.2), we obtain the following:
Find $\{u_j\}_{j=0}^p \subset X$ such that

$$\sum_{i,j=0}^{p} \{[\int_I \theta_j' \theta_i dt + \theta_j^+(t_{n-1})\theta_i^+(t_{n-1})](u_j, v_i)_X + \int_I \theta_j\theta_i dt (\nabla u_j, \nabla v_i)_X\}$$

$$= \sum_{i=0}^{p} \{ \int_I (g, v_i\theta_i)_{X^*\times X} dt + (U_{n-1}^-, v_i\theta_i^+(t_{n-1})\} \qquad \text{for all } \{v_i\}_{i=0}^p \subset X. \qquad (3.2)$$

Transforming the integrals into the reference interval $[-1,1]$ under F^{-1} and letting (see [12])

$$\hat{A}_{ij} \equiv \int_{-1}^{1} \hat{\theta}'_j \hat{\theta}_i d\hat{t} + \hat{\theta}_j^+(-1)\hat{\theta}_i^+(-1) \qquad \hat{B}_{ij} \equiv \int_{-1}^{1} \hat{\theta}_j\hat{\theta}_i d\hat{t}$$

$$\hat{f}_i^1(v_i) \equiv (\int_{-1}^{1} (g\circ F)\hat{\theta}_i d\hat{t}, v_i)_X \qquad \hat{f}_i^2(v_i) \equiv (U_{n-1}^-\hat{\theta}_i(-1), v_i)_X$$

equation (3.2) becomes:
Find the coefficient $\{u_j\}_{j=0}^p \subset X$ such that for all $\{v_j\}_{i=0}^p \subset X$

$$\sum_{i,j=0}^{p} \hat{A}_{ij}(u_j, v_i)_X + \frac{k}{2}\hat{B}_{ij}(\nabla u_j, \nabla v_i)_X = \sum_{i=0}^{p} \frac{k}{2}\hat{f}_i^1(v_i) + \hat{f}_i^2(v_i). \qquad (3.3)$$

Here $k = k_n$. The strong form of equation (3.3) is

$$\sum_{j=0}^{p} \hat{A}_{ij}u_j + \frac{k}{2}\hat{B}_{ij}\Delta u_j = \frac{k}{2}\hat{l}_i^1 + \hat{l}_i^2, \qquad i = 0,1,\dots p, \qquad (3.4)$$

where $\hat{l}_i^1 = \int_{-1}^1 (g\circ F)\hat{\theta}_i d\hat{t}$ and $\hat{l}_i^2 = U_{n-1}^-\hat{\theta}_i(-1)$. Hence, in order to execute the p-version of the finite element method, the following integrals must be computed for assembly of the global matrix.

$$\int_{t_{n-1}}^{t_n} \theta(\frac{d\theta}{dt})^T dt, \quad \int_{t_{n-1}}^{t_n} \theta\theta^T dt, \quad \int_{t_{n-1}}^{t_n} \theta dt.$$

The set of the canonical polynomials $\theta_{\nu+1}(t) = t^\nu$ were used in [8] and the components of the matrices which represent $\hat{A}_{ij}$ and $\hat{B}_{ij}$ were computed exactly. We consider two other alternatives for θ.

First, we consider $\theta_\nu(\mathbf{t}) = (\frac{\mathbf{t}-\mathbf{t_{n-1}}}{\mathbf{k_n}})^\nu$: Let $U|_{I_n} = \sum_{\nu=0}^p (\frac{t-t_{n-1}}{k_n})^\nu \Phi_\nu^n(\bar{x})$. Since $U^{n-1,+} = \Phi_0^n(\bar{x})$, the advantage of this basis is that the term $(U^{n-1,+}, v^{n-1,+})$ in (2.1) is simplified. Also,

$$\int_{t_{n-1}}^{t_n} \theta(\frac{d\theta}{dt})^T dt = \begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 0 & \frac{1}{2} & \frac{2}{3} & \dots & \frac{p}{p+1} \\ 0 & \frac{1}{3} & \frac{2}{4} & \dots & \frac{p}{p+2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & \frac{1}{p+1} & \frac{2}{p+2} & \dots & \frac{p}{2p} \end{bmatrix}$$

$$\int_{t_{n-1}}^{t_n} \theta\theta^T dt = k_n \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{p+1} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots & \frac{1}{p+2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \frac{1}{p+1} & \frac{1}{p+2} & \frac{1}{p+3} & \dots & \frac{1}{2p+1} \end{bmatrix}$$

and

$$\int_{t_{n-1}}^{t_n} \theta dt = k_n \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{p+1} \end{bmatrix}^T.$$

As stated in [12], the ideal choice for $\hat{\theta}_i$ would be the one under which the matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ diagonalize simultaneously. If the diagonalizations of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are possible, then equation (3.4) decouples into $p+1$ systems of linear equations, reducing the size of computation. The canonical basis and the basis just considered generate the full matrices for $\hat{\mathbf{A}}$ and for $\hat{\mathbf{B}}$. In order to select basis functions in time variable which takes into account of the structures of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$, we now turn to the Legendre polynomials for θ.

Second: Here, we consider the translated Legendre polynomials for $\theta_\nu(t)$. This is essentially the same as the normalized Legendre polynomials used in [12]. We extend the discussion in [12] by exhibiting the general form for $\hat{\mathbf{A}}$ and state its characteristics. The orthogonal nature of the Legendre polynomials guarantees the matrix $\hat{\mathbf{B}}$ to be diagonal. Hence it remains to analyze the matrix $\hat{\mathbf{A}}$. In [12], it is stated that '..,this (diagonalization) seems not to be possible with time shape functions in $\mathbf{R}$, but numerical experiments show that $\hat{\mathbf{A}}$ ... is diagonalizable in $\mathbf{C}$ at least for $0 \le p \le 100$'. We study this issue of diagonalization of $\hat{\mathbf{A}}$ in the next section. In this section, we are interested in presenting a general form of the matrix $\hat{\mathbf{A}}$ which can be used to generate it quite simply for any order n. A discussion is also included for an application of the real Schur decomposition of the matrix $\hat{\mathbf{A}}$ to establish a solution scheme for (3.4). This parallels the use of the complex Schur decomposition in [12]. The real Schur decomposition leads to a modified backward substitution scheme. The method is mathematically justifiable for any degree p. Moreover, the current method avoids the complex number arithmetic, and since we are dealing with real data in applications, it may be more appropriate for numerical computations.

The advantage of this choice as basis elements in time variable lies in the formations of $\int_{t_{n-1}}^{t_n} \theta\theta^T dt$, $\int_{t_{n-1}}^{t_n} \theta dt$ as well as $\int_{t_{n-1}}^{t_n} \theta(\frac{d\theta}{dt})^T dt$, all of which are of banded structures as seen below. Recall that $P_i(x)$ denote the Legendre polynomial of degree i defined over $[-1,1]$ (see (1.7)). Let $x^n(t) = \frac{2}{t_n - t_{n-1}} t - \frac{t_n + t_{n-1}}{t_n - t_{n-1}}$ and $L_i^n(t) = P_i(x^n(t))$ for each $n = 1, 2, \ldots, M(I)$ and $i = 0, 1, \ldots$. Obviously,

$$\frac{2}{t_n - t_{n-1}} \int_{t_{n-1}}^{t_n} L_i^n(t) L_j^n(t) dt = \delta_{ij} \qquad \text{for each } i, j = 0, 1, \ldots, .$$

Thus,

$$\int_{t_{n-1}}^{t_n} \theta\theta^T dt = diag[\frac{k_n}{2} \quad \cdots \quad \frac{k_n}{2}]$$

and

$$\int_{t_{n-1}}^{t_n} \theta dt = [k_n \quad 0 \quad \cdots \quad 0]^T$$

Also

$$\int_{t_{n-1}}^{t_n} \theta(\frac{d\theta}{dt})^T dt = \begin{bmatrix} 0 & k_n & 0 & k_n & 0 & k_n & \cdots \\ 0 & 0 & k_n & 0 & k_n & 0 & \cdots \\ 0 & 0 & 0 & k_n & 0 & k_n & \cdots \\ 0 & 0 & 0 & 0 & k_n & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

The formation of the last matrix $\int_{t_{n-1}}^{t_n} \theta(\frac{d\theta}{dt})^T dt$ requires some tedious but straightforward calculations which are not so obvious. Thus, we include them below.

Clearly, it is sufficient to derive the following:

$$\int_{-1}^{1} \theta(\frac{d\theta}{dt})^T dt = \begin{bmatrix} 0 & 2 & 0 & 2 & 0 & 2 & \cdots \\ 0 & 0 & 2 & 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & 2 & 0 & 2 & \cdots \\ 0 & 0 & 0 & 0 & 2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{3.5}$$

For the Legendre polynomials of the first kind, $P_i(x)$, $i \geq 0$, $x \in [-1, 1]$, first note that

$$P_{2i}(-1) = P_{2i}(1) = 1, \quad P_{2i+1}(-1) = -1 \text{ and } P_{2i+1}(1) = 1 \text{ for } i \geq 0.$$

To derive (3.5), we must show that

$$\int_{-1}^{1} P_i(t)P_j'(t)dt = 0 \qquad \text{for all } i \text{ and } j \text{ with } i \geq j \geq 0 \tag{3.6}$$

and

$$\int_{-1}^{1} P_i(t)P_{i+2l}'(t)dt = 0 \text{ and } \int_{-1}^{1} P_i(t)P_{i+2l+1}'(t)dt = 2 \qquad \text{for all } i \geq 0 \text{ and } l \geq 0. \tag{3.7}$$

Equations (3.6) and (3.7) are verified by induction. As $P_0(t) = 1$, (3.6) is verified with $i = 0$. Also

$$\int_{-1}^{1} P_0(t)P_{2l}'(t)dt = P_0(t)P_{2l}(t)|_{-1}^{1} - \int_{-1}^{1} P_0'(t)P_{2l}(t)dt = 0 \qquad \text{for all } l \geq 0.$$

and similarly

$$\int_{-1}^{1} P_0(t)P_{2l+1}'(t)dt = 2 \qquad \text{for all } l \geq 0.$$

Now assume that (3.6) and (3.7) are satisfied for some $i = i^* - 1$ and for all j such that $i^* - 1 \geq j \geq 0$ and for all $l \geq 0$. First, for the diagonal element,

$$\int_{-1}^{1} P_{i^*}(t)P_{i^*}'(t)dt = P_{i^*}(t)P_{i^*}(t)|_{-1}^{1} - \int_{-1}^{1} P_{i^*}(t)P_{i^*}'(t)dt$$

implies that $\int_{-1}^{1} P_{i^*}(t)P_{i^*}'(t)dt = 0$. Also, for $i^* > j \geq 0$,

$$\int_{-1}^{1} P_{i^*}(t)P_j'(t)dt = P_{i^*}(t)P_j(t)|_{-1}^{1} - \int_{-1}^{1} P_{i^*}'(t)P_j(t)dt = \begin{cases} 2-2=0 & \text{for } i^*-j \text{ odd} \\ 0-0=0 & \text{for } i^*-j \text{ even} \end{cases}$$

where the inductive assumption was used in computing the second integral. This shows (3.6) for all i and $i \geq j \geq 0$. It remains to prove that $\int_{-1}^{1} P_{i^*}(t)P_{i^*+2l}'(t)dt = 0$ and $\int_{-1}^{1} P_{i^*}(t)P_{i^*+2l+1}'(t)dt = 2$ for all $l \geq 0$. The case for $l = 0$ is done. For $l > 0$,

$$\int_{-1}^{1} P_{i^*}(t)P_{i^*+2l}'(t)dt = P_{i^*}(t)P_{i^*+2l}(t)|_{-1}^{1} - \int_{-1}^{1} P_{i^*}'(t)P_{i^*+2l}(t)dt.$$

The last integral is 0 by (3.6) and the term $P_{i^*}(t)P_{i^*+2l}(t)|_{-1}^{1}$ is 0 regardless of i^* even or odd. Hence $\int_{-1}^{1} P_{i^*}(t)P_{i^*+2l}'(t)dt = 0$. $\int_{-1}^{1} P_{i^*}(t)P_{i^*+2l+1}'(t)dt = 2$ is similar. Thus, (3.5) is verified.

Now, from (3.5), it is clear that

$$\hat{\mathbf{A}}_n \equiv \int_{-1}^{1} \theta(\frac{d\theta}{dt})^T dt + \theta_{n-1}^{+}\theta_{n-1}^{+} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & \cdots \\ -1 & 1 & 1 & 1 & 1 & 1 & \cdots \\ 1 & -1 & 1 & 1 & 1 & 1 & \cdots \\ -1 & 1 & -1 & 1 & 1 & 1 & \cdots \\ 1 & -1 & 1 & -1 & 1 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}. \tag{3.8}$$

Formula (3.8) provides a general construction method for the assembly of the p-finite element matrix with any order p when the Legendre polynomials are used in time variable. In particular, for $p = 5$, with normalizing factors incorporated, (3.8) can be used to derive the matrix $\hat{\mathbf{A}}$ in [12], (eq. (4.11)). As in (2.3), denote the semidiscrete solution U as well as V in (1.2) as

$$U = \sum_{j=0}^{p} u_j\theta_j, \qquad V = \sum_{j=0}^{p} v_j\theta_j \tag{3.9}$$

where the subscript n is dropped. Let $\theta_i = L_i$, the transformed Legendre polynomial of degree i and $\hat{\theta}_i = P_i$, the Legendre polynomial of degree i defined over $[-1, 1]$ in equations (3.2) and (3.3). Also, denote a basis for the finite element space $S^{p_1,k}(\omega, T_h) \otimes P^{p_2}[-\frac{k}{2}, \frac{k}{2}]$ for X by $\{s_j\}_{j=1}^{D}$ with $D = \dim(S^{p_1,k}(\omega, T_h) \otimes P^{p_2}[-\frac{k}{2}, \frac{k}{2}])$. The trial and test functions u_j and v_i in the fully-discrete system (3.2) above are further approximated by

$$u_j \sim \sum_{l=1}^{D} u_j^l s_l(x), \qquad v_i \sim \sum_{k=1}^{D} v_i^k s_k(x). \tag{3.10}$$

Substituting (3.10) into (3.3), the fully discrete p-finite element system can be written as (ref. [12]), for the unknown coefficient vectors $\vec{u}_j = (u_j^1, u_j^2, \ldots, u_j^D)^T \in R^D$,

$$\begin{bmatrix} \hat{A}_{00}M + \frac{k}{2}S & \cdots & \hat{A}_{0p}M \\ \vdots & \ddots & \vdots \\ \hat{A}_{p0}M & \cdots & \hat{A}_{pp}M + \frac{k}{2}S \end{bmatrix} \begin{bmatrix} \vec{u}_0 \\ \vdots \\ \vec{u}_p \end{bmatrix} = \frac{k}{2} \begin{bmatrix} \vec{f}_0^1 \\ \vdots \\ \vec{f}_p^1 \end{bmatrix} + \begin{bmatrix} \vec{f}_0^2 \\ \vdots \\ \vec{f}_p^2 \end{bmatrix} \tag{3.11}$$

where

$$M = \{(s_l, s_k)_X\}_{l,k=1}^{D}, \qquad S = \{(\nabla s_l, \nabla s_k)_X\}_{l,k=1}^{D}$$

and

$$\vec{f}_j^1 = (\hat{f}_j^1(s_1), \hat{f}_j^1(s_2), \ldots, \hat{f}_j^1(s_D))^T$$
$$\vec{f}_j^2 = (\hat{f}_j^2(s_1), \hat{f}_j^2(s_2), \ldots, \hat{f}_j^2(s_D))^T.$$

Note that the use of the translated Legendre polynomial served well because $\hat{\mathbf{B}}_{ij} = \delta_{ij}$ in (3.4). Equation (3.4) can be written in the vector form as

$$\hat{\mathbf{A}}_{p+1}\vec{u} + \frac{k}{2}[\delta_{ij}]\triangle\vec{u} = \vec{F} \tag{3.12}$$

where $\vec{F} = (\frac{k}{2}\int_{-1}^{1}(g \circ x^n)\hat{\theta}_0 d\hat{t} + U_{n-1}^{-}\hat{\theta}_0(-1), \ldots, \frac{k}{2}\int_{-1}^{1}(g \circ x^n)\hat{\theta}_p d\hat{t} + U_{n-1}^{-}\hat{\theta}_p(-1))^T$. In [12], $\hat{\mathbf{A}}_{p+1}$ is diagonalized numerically for $0 \le p \le 99$ and equation (3.12) is solved for $\vec{w} = Q^T\vec{u}$ from

$$Q^T\hat{\mathbf{A}}_{p+1}Q\vec{w} + \frac{k}{2}\triangle\vec{w} = Q^T\vec{F}. \tag{3.13}$$

It is reported in [12] that, as the result of their numerical experiments, the matrix $\hat{A}_{p+1}$ is diagonalizable up to its order $p = 99$. Subsequently, equation (3.12) is decoupled via (3.13) into $p+1$ systems of linear equations, each system having the size $D \times D$ and requires complex arithmetic to solve. As an alternative approach, the complex form of Schur decomposition is used in [12] to transform $\hat{A}_{p+1}$ into an upper triangular matrix and subsequently (3.13) was solved by the standard backward substitution. Since one would more likely to deal with real data in applications, the real form of Schur decomposition may perhaps be more useful in this connection.

Theorem 7. *(Real Schur Decomposition,* [4] *p.219) If $A \in R^{n\times n}$, then there exists an orthogonal matrix $Q \in R^{n\times n}$ such that*

$$Q^T AQ = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ 0 & R_{22} & \cdots & R_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_{mm} \end{bmatrix} \tag{3.14}$$

where each R_{ii} is either a 1×1 matrix or a 2×2 matrix having complex conjugate eigenvalues of A.

By Theorem 4.1 in the next section, it is guaranteed that no 1×1 matrix in Schur decomposition for $\hat{\mathbf{A}}_{p+1}$ is 0. Also, it is worth noting that the Schur decomposition is an orthogonal similarity transformation and thus avoid the computation of Q^{-1}. Equation (3.13) is then solved by the backward substitution using block matrices. More specifically, $\vec{w}_{p-1}$ and $\vec{w}_p$ are found from

$$(R_{mm}M + \frac{k}{2}S) \begin{bmatrix} \vec{w}_{p-1} \\ \vec{w}_p \end{bmatrix} = \begin{bmatrix} \vec{F}_{p-1} \\ \vec{F}_p \end{bmatrix} \tag{3.15}$$

or from

$$(R_{mm}M + \frac{k}{2}S)\vec{w}_p = \vec{F}_p \tag{3.16}$$

according to R_{mm} being 2×2 or 1×1 matrix respectively. Here, we used the standard convention that if $R_{mm} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $R_{mm}M = \begin{bmatrix} aM & bM \\ cM & dM \end{bmatrix}$. Computation proceeds to find $\vec{w}_{p-2}$ as well as $\vec{w}_{p-3}$ if $R_{m-1\,m-1}$ is 2×2. Namely, assuming that R_{mm} is 2×2, if $R_{m-1\,m-1}$ is a scalar, then $\vec{w}_{p-2}$ is found by solving

$$\left[R_{m-1\,m-1}M + \frac{k}{2}S\right] \vec{w}_{p-2} = \vec{F}_{p-2} - [R_{m-1\,m}M] \begin{bmatrix} \vec{w}_{p-1} \\ \vec{w}_p \end{bmatrix}.$$

If $R_{m-1\,m-1}$ is 2×2, then $\vec{w}_{p-3}$ and $\vec{w}_{p-2}$ are found from

$$\left[R_{m-1\,m-1}M + \frac{k}{2}S\right] \begin{bmatrix} \vec{w}_{p-3} \\ \vec{w}_{p-2} \end{bmatrix} = \begin{bmatrix} \vec{F}_{p-3} \\ \vec{F}_{p-2} \end{bmatrix} - [R_{m-1\,m}M] \begin{bmatrix} \vec{w}_{p-1} \\ \vec{w}_p \end{bmatrix}.$$

The use of the real Schur decomposition along with the block backward substitution technique just described is demonstrated in examples in Section 6.

4. **The Matrix $\hat{\mathbf{A}}$:** As stated earlier, a diagonalization of the matrix $\hat{\mathbf{A}}$ results in saving in computational cost. The reports in [12] show that in practice we can go ahead and proceed with this diagonalization. Also, we observed that the real Schur decomposition approach works well in our numerical experiments. Therefore, it is more of a theoretical interest that we present in this section some observations we

obtained concerning the characteristics of the matrix $\hat{\mathbf{A}}$ and the possibility of it being diagonalized. These observations are just one avenue of solution method and the reader may take a completely different approach to a solution. First, we have

Theorem 8. *Let $\hat{\mathbf{A}}_n$ denote the $n \times n$ leading principle matrix of (3.8) for each $n = 1, 2, \ldots$. Then $\hat{\mathbf{A}}_n$ is invertible for each n.*

Proof. This follows immediately by noting that $det(A_n) = 2^{n-1}$. □

The matrix $\hat{\mathbf{A}}_n$ is not positive definite for all n. However:

Theorem 9. *The matrix $\hat{\mathbf{A}}_n$ is positive semi-definite for all $n = 1, 2, \ldots$.*

Proof. This can be verified as follows: For n even, $n = 2k$, for some k, with

$$x^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_{2k-1} & x_{2k} \end{bmatrix}$$

$$x^T \hat{\mathbf{A}}_n x = (\sum_{i=1}^{k} x_{2k-1})^2 + (\sum_{i=1}^{k} x_{2i})^2$$

Similarly, for n odd, $n = 2k + 1$, for some k, with

$$x^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_{2k} & x_{2k+1} \end{bmatrix}$$

$$x^T \hat{\mathbf{A}}_n x = (\sum_{i=1}^{k+1} x_{2k-1})^2 + (\sum_{i=1}^{k} x_{2k})^2$$

Proceed with induction to complete the proof. □

The characteristic polynomial of $\hat{\mathbf{A}}_n$ can be derived as follows: First, for the characteristic polynomial

$$f_n(\lambda) = |\hat{\mathbf{A}} - \lambda I| \tag{4.3}$$

a sequence of row operations (starting with the second row, change the sign of a row and add it to the row above it; after completing this operation to the last row, multiply the next-to-last row by $\frac{1}{2}$ and add it to the row below) can be applied to obtain the following,

$$f_n(\lambda) = \begin{vmatrix} 2-\lambda & \lambda & 0 & 0 & \cdots & 0 & 0 \\ -\lambda & 2 & \lambda & 0 & \cdots & 0 & 0 \\ 0 & -\lambda & 2 & \lambda & \cdots & 0 & 0 \\ 0 & 0 & -\lambda & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 2 & \lambda \\ 0 & 0 & 0 & 0 & \cdots & -\frac{\lambda}{2} & 1-\frac{\lambda}{2} \end{vmatrix}. \tag{4.4}$$

For convenience, we introduce $Q_n(\lambda)$ which is defined by

$$Q_n(\lambda) = \begin{vmatrix} 2 & \lambda & 0 & 0 & \cdots & 0 & 0 \\ -\lambda & 2 & \lambda & 0 & \cdots & 0 & 0 \\ 0 & -\lambda & 2 & \lambda & \cdots & 0 & 0 \\ 0 & 0 & -\lambda & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 2 & \lambda \\ 0 & 0 & 0 & 0 & \cdots & -\frac{\lambda}{2} & 1-\frac{\lambda}{2} \end{vmatrix}. \tag{4.5}$$

Developing $f_n(\lambda)$ along the first row in (4.4), we get the following relation:

$$f_n(\lambda) = (2-\lambda)Q_{n-1}(\lambda) + \lambda^2 Q_{n-2}(\lambda) \tag{4.6}$$

For the Q's themselves, we have the following recursion relation:

$$Q_n - 2Q_{n-1} - \lambda^2 Q_{n-2} = 0 \tag{4.7}$$

Combining (4.6) and (4.7), the following expression for the characteristic polynomial is obtained,

$$f_n(\lambda) = Q_n(\lambda) - \lambda Q_{n-1}(\lambda). \tag{4.8}$$

Lemma 1. *The solution of the recursion relation (4.7) is given by*

$$\begin{aligned} Q_n(\lambda) &= \frac{\lambda^2-\lambda+2+(2-\lambda)\sqrt{1+\lambda^2}}{4(1+\lambda^2+\sqrt{1+\lambda^2})}(1+\sqrt{1+\lambda^2})^n \\ &\quad+\frac{\lambda^2-\lambda+2-(2-\lambda)\sqrt{1+\lambda^2}}{4(1+\lambda^2-\sqrt{1+\lambda^2})}(1-\sqrt{1+\lambda^2})^n \end{aligned}$$

for each $n \geq 0$.

Proof. From (4.7),

$$Q_n(\lambda) = C_0(1+\sqrt{1+\lambda^2})^n + C_1(1-\sqrt{1+\lambda^2})^n$$

for some constants C_0 and C_1. We get $C_0 = \frac{\lambda^2-\lambda+2+(2-\lambda)\sqrt{1+\lambda^2}}{4(1+\lambda^2+\sqrt{1+\lambda^2})}$ by multiplying Q_1 by $1-\sqrt{1+\lambda^2}$ and subtracting from Q_2. Also, $C_1 = \frac{\lambda^2-\lambda+2-(2-\lambda)\sqrt{1+\lambda^2}}{4(1+\lambda^2-\sqrt{1+\lambda^2})}$ can be obtained by multiplying Q_1 by $1+\sqrt{1+\lambda^2}$ and subtracting from Q_2. □

Using Lemma 4.3 along with (4.2), we can construct $f_n(\lambda)$. We leave the straightforward calculation to the reader.

$$\begin{aligned} f_n(\lambda) &= \frac{(\lambda^2-\lambda+2+(2-\lambda)\sqrt{1+\lambda^2})^2}{4(1+\lambda^2+\sqrt{1+\lambda^2})}(1+\sqrt{1+\lambda^2})^{n-2} \\ &\quad+\frac{(\lambda^2-\lambda+2-(2-\lambda)\sqrt{1+\lambda^2})^2}{4(1+\lambda^2-\sqrt{1+\lambda^2})}(1-\sqrt{1+\lambda^2})^{n-2} \end{aligned} \tag{4.9}$$

Let $[x]$ denote the greatest integer less than or equal to x. The expression of the characteristic polynomial $f_n(\lambda)$ in (4.9) can be further reduced to the following:

$$f_n(\lambda) = \begin{cases} (\lambda^2-2\lambda+2)\sum_{k=0}^{[\frac{n}{2}]-1}\binom{n-2}{2k}(1+\lambda^2)^k \\ \qquad -(\lambda^3-2\lambda^2+2\lambda-2)\sum_{k=0}^{[\frac{n}{2}]-1}\binom{n-2}{2k+1}(1+\lambda^2)^k, & n \text{ odd;} \\ (\lambda^2-2\lambda+2)\sum_{k=0}^{\frac{n}{2}-1}\binom{n-2}{2k}(1+\lambda^2)^k \\ \qquad -(\lambda^3-2\lambda^2+2\lambda-2)\sum_{k=0}^{\frac{n}{2}-2}\binom{n-2}{2k+1}(1+\lambda^2)^k, & n \text{ even.} \end{cases} \tag{4.10}$$

To obtain (4.10), first put $f_n(\lambda)$ over a common denominator to get

$$\begin{aligned} f_n(\lambda) &= \frac{(\lambda^2-\lambda+2+(2-\lambda)\sqrt{1+\lambda^2})^2(1+\lambda^2-\sqrt{1+\lambda^2})(1+\sqrt{1+\lambda^2})^{n-2}}{4(\lambda^4+\lambda^2)} \\ &\quad+\frac{(\lambda^2-\lambda+2-(2-\lambda)\sqrt{1+\lambda^2})^2(1+\lambda^2+\sqrt{1+\lambda^2})(1-\sqrt{1+\lambda^2})^{n-2}}{4(\lambda^4+\lambda^2)} \end{aligned} \tag{4.11}$$

Rewriting the numerators in (4.11) as

$$\begin{aligned}&(\lambda^2-\lambda+2+(2-\lambda)\sqrt{1+\lambda^2})^2(1+\lambda^2-\sqrt{1+\lambda^2})(1+\sqrt{1+\lambda^2})^{n-2}\\ &=2\lambda^2(1+\lambda^2)(\lambda^2-2\lambda+2)-2\lambda^2(\lambda^3-2\lambda^2+2\lambda-2)\sqrt{1+\lambda^2}\\ &(\lambda^2-\lambda-2+(2-\lambda)\sqrt{1+\lambda^2})^2(1+\lambda^2+\sqrt{1+\lambda^2})(1-\sqrt{1+\lambda^2})^{n-2}\\ &=2\lambda^2(1+\lambda^2)(\lambda^2-2\lambda+2)+2\lambda^2(\lambda^3-2\lambda^2+2\lambda-2)\sqrt{1+\lambda^2}\end{aligned}$$

and substituting these into (4.11), we obtain

$$\begin{aligned}f_n(\lambda)\quad&=\frac{\lambda^2-2\lambda+2}{2}[(1+\sqrt{1+\lambda^2})^{n-2}+(1-\sqrt{1+\lambda^2})^{n-2}]\\ &-\frac{\lambda^3-2\lambda^2+2\lambda-2}{2\sqrt{1+\lambda^2}}[(1+\sqrt{1+\lambda^2})^{n-2}-(1-\sqrt{1+\lambda^2})^{n-2}]\end{aligned}\tag{4.12}$$

Finally, noting that

$$\begin{aligned}(1+\sqrt{1+\lambda^2})^{n-2}+(1-\sqrt{1+\lambda^2})^{n-2}&=2\binom{n-2}{0}+2\binom{n-2}{2}(1+\lambda^2)\\ +2\binom{n-2}{4}(1+\lambda^2)^2+\cdots\end{aligned}$$

and that

$$\begin{aligned}(1+\sqrt{1+\lambda^2})^{n-2}-(1-\sqrt{1+\lambda^2})^{n-2}&=2\binom{n-2}{1}(1+\lambda^2)\\ +2\binom{n-2}{3}(1+\lambda^2)^3+2\binom{n-2}{5}(1+\lambda^2)^5+\cdots\end{aligned}$$

and (4.12) can be written as (4.10). The first few of the characteristic polynomials are listed below.

$$\begin{aligned}f_1(\lambda)\quad&=1-\lambda\\ f_2(\lambda)\quad&=\lambda^2-2\lambda+2\\ f_3(\lambda)\quad&=-\lambda^3+3\lambda^2-4\lambda+4\\ f_4(\lambda)\quad&=\lambda^4-4\lambda^3+8\lambda^2-8\lambda+8\\ f_5(\lambda)\quad&=-\lambda^5+5\lambda^4-12\lambda^3+20\lambda^2-16\lambda+16\end{aligned}$$

Examining the roots of these characteristic polynomials, it appears that the following is true.

Conjecture: All eigenvalues of $\hat{\mathbf{A}}_n$ are complex as well as distinct for even n, and for odd n, there is one positive real eigenvalue and the remaining eigenvalues are distinct complex values.

Of course, a proof of this conjecture would settle the question of the diagonalization of $\hat{\mathbf{A}}_n$.

5. **Start-up singularities:** Here, we make some comments on the start-up singularities normally associated with the parabolic problems. The regularity assumptions in Theorem 2.1 and Corollary 2.2 were taken so that the current fully p-finite element method could provide numerical solutions where the discretization error associated in time can be made consistent with the discretization error associated in space. However, in many occasions, singularities of the solution of the parabolic problem arise due to various types of incompatible initial data. To capture properly such singularities in numerical computation, an hp-version of the finite element method must be considered. In [11], a nonuniform time discretization is determined by considering the conditions on f as well as the initial function u_0 in (1.1). More

specifically, the function f is assumed to be piecewise analytic as a function on $[0,T]$ with values in H, i.e.,

$$\|f^{(l)}\|_H \le Cd^l\Gamma(l+1), \qquad t\in[0,T],\ l\in N_0$$

with constants C and d independent of l and t. Also, u_0 is assumed to be in $H_\theta = (H,X)_{\theta,2}$, $0\le\theta\le 1$, where $(H,X)_{\theta,2}$ is a space between H and X determined by the K-method of interpolation, (see [9]). An h-version of the discontinuous Galerkin finite element method developed in [5] establishes a similar nonuniform time discretization scheme which is based upon the regularity of $\|u^{(1)}(t)\|_X$. The direct inspection into the regularity of $\|u^{(1)}(t)\|_X$ in determining time partition takes into account of various possibilities under which the start-up singularities associated with parabolic problems arise. It is possible from computed solution to estimate the regularity of $\|u^{(1)}(t)\|_X$. Analysis used in determining the nonuniform graded partition points in [5] is distinct from the one used in [11] and an example is provided in [5], which demonstrates that, in some instances, the time discretization technique of Kaneko, Bey and Hou gives more relaxed time partition points than the ones given in [11].

6. **Numerical examples:** In this section, we present two numerical examples of heat conduction problems. The use of the real Schur decomposition and the backward substitution technique described in Section 3 are employed in both examples.

Example 1: Now we consider (1.1) over three-dimensional plate $\Omega = [-2,2]\times[-2,2]\times[-\frac{1}{10},\frac{1}{10}]$. A total of 256 elements over the surface $[-2,2]\times[-2,2]$ are used. Linear splines are employed for approximating u in all variables x,y,z including t. The solution is taken to be

$$u(x,y,z,t) = (1-0.25x^2)(1-0.25y^2)(1-100z^2)^2t$$

and the time interval of $[0,5]$ is partitioned uniformly into 10 intervals. The initial condition as well as the boundary condition are taken exactly from the analytical solution. Errors in temperature at the cross sections of $z1=0$, $z2=1/30$ and $z3=2/30$ are given below. Because of the symmetry of this problem, the temperatures at $z=-1/30$ and $z=-2/30$ follow from these data.

	Example 1		
t	error at z1	error at z2	error at z3
t_1	0.0016	0.0012	0.0004
t_2	0.0032	0.0024	0.0006
t_3	0.0048	0.0036	0.0008
t_4	0.0056	0.0042	0.0011
t_5	0.0069	0.0051	0.0012
t_6	0.0081	0.0060	0.0013
t_7	0.0093	0.0069	0.0013
t_8	0.0106	0.0078	0.0014
t_9	0.0118	0.0087	0.0015
t_{10}	0.0130	0.0095	0.0016

Temperature profiles over the three cross sections $z_0=0, z_1=1/3$ and $z_2=2/3$ at three different time levels, t_1, t_5 and t_9 are drawn below.

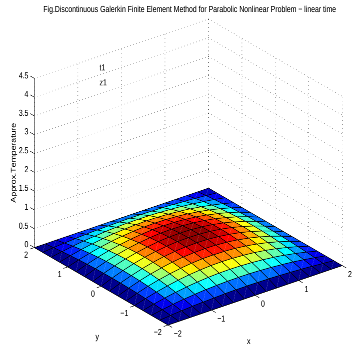


Fig.Discontinuous Galerkin Finite Element Method for Parabolic Nonlinear Problem − linear time
t1
z1
Approx.Temperature
y
x

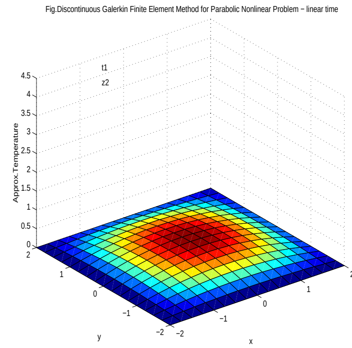


Fig.Discontinuous Galerkin Finite Element Method for Parabolic Nonlinear Problem − linear time
t1
z2
Approx.Temperature
y
x

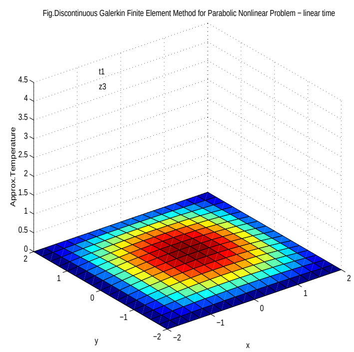


Fig.Discontinuous Galerkin Finite Element Method for Parabolic Nonlinear Problem − linear time
t1
z3
Approx.Temperature
y
x

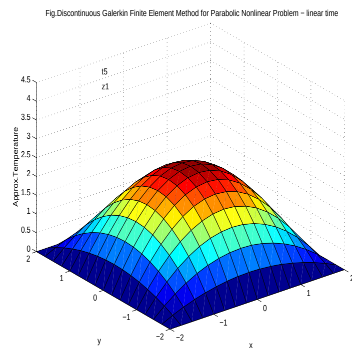


Fig.Discontinuous Galerkin Finite Element Method for Parabolic Nonlinear Problem − linear time
t5
z1
Approx.Temperature
y
x

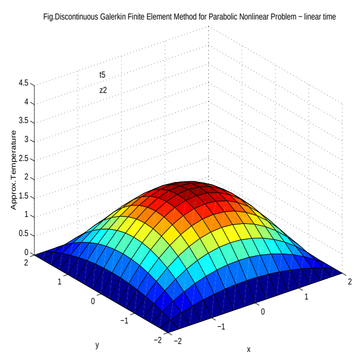


Fig.Discontinuous Galerkin Finite Element Method for Parabolic Nonlinear Problem − linear time
t5
z2
Approx.Temperature
y
x

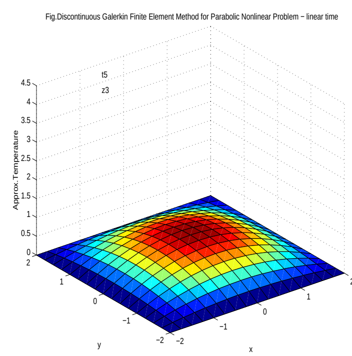


Fig.Discontinuous Galerkin Finite Element Method for Parabolic Nonlinear Problem − linear time
t5
z3
Approx.Temperature
y
x

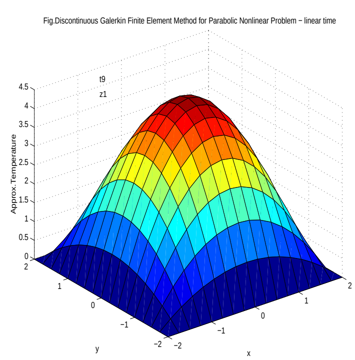


Fig.Discontinuous Galerkin Finite Element Method for Parabolic Nonlinear Problem − linear time
t9
z1
Approx.Temperature
y
x

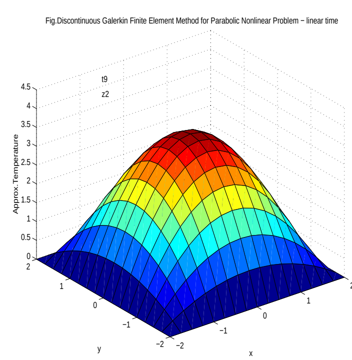
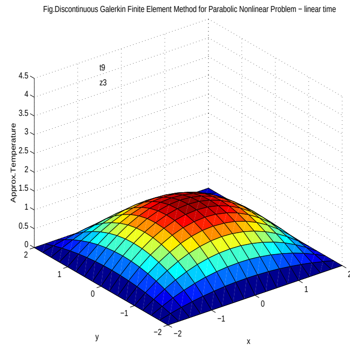


Fig.Discontinuous Galerkin Finite Element Method for Parabolic Nonlinear Problem − linear time
t9
z2
Approx.Temperature
y
x



Example 2: We consider (1.1) over two-dimensional plate $\Omega = [0, \pi] \times [0, \pi]$ with boundary and initial conditions are given respectively by

$$u = 0 \quad \text{on } \partial\Omega \qquad \text{and} \qquad u(x, y, 0) = (\pi - x)(\pi - y).$$

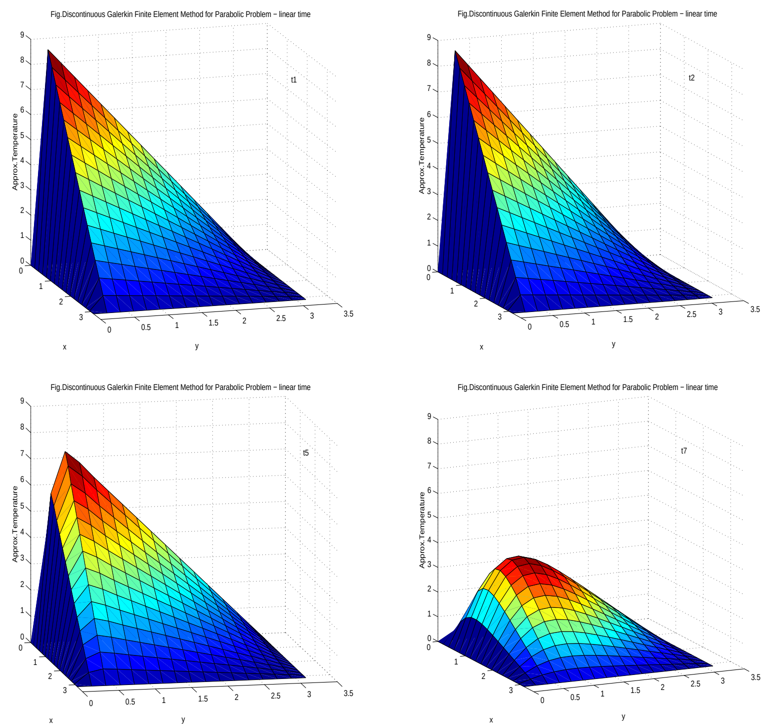
The exact solution is given by

$$u = \sum_{j,k=1}^{\infty} A_{jk} e^{-(j^2+k^2)t} \sin(jx) \sin(ky),$$

where $A_{jk} = \frac{4}{jk}$. Time interval of $[0, 5]$ is taken. The solution contains a transient phase due to the incompatibility between the boundary condition and the initial condition as was briefly described in the previous section. The time discretization technique which was explored in [5] is used to capture accurately this phase of the solution. The real Schur decomposition is employed along with the block backward substitution technique described in Section 3. In Matlab program, we used $u = \sum_{j,k=1}^{200} A_{jk} e^{-(j^2+k^2)t} \sin(jx) \sin(ky)$, as a substitute for the exact solution. Also, it is noted that during the transient phase and when time is fixed, a boundary layer along $x = 0$ and $y = 0$ develops (see, for example, the temperature profiles below at the first two time frames). In this case, it is important to refine finite elements in the region so that the solution can be approximated accurately. The linear splines are used to discretize both in time and spatial variables. Time interval is partitioned into 10 nonuniform subintervals and there are a total of 961 $(= (2^5 - 1)^2)$ elements are used in computation. A table describing approximation errors at each time step as well as the temperature profiles at four different time levels, t_1, t_2, t_5 and t_8 are given below.

	Example 2					
t	t_1	t_2	t_3	t_4	t_5	t_6
Error	0.0016	0.0029	0.0042	0.0049	0.0058	0.0066

Example 2			
t_7	t_8	t_9	t_{10}
0.0074	0.0083	0.0091	0.0010



Acknowledgement. The authors thank the referees who made several constructive comments which improved this paper.

REFERENCES

[1] I. Babuška, B.A. Szabo and I.N. Katz, *The p-version of the finite element method,* SIAM J. Num. Anal., **18** (1981), 515–545.

[2] I. Babuška and M. Suri, *The p and hp versions of the finite element method, basic principles and properties,* SIAM Review, **36** (1994), 578–632.

[3] Kenneth Eriksson and Claes Johnson, *Adaptive finite element methods for parabolic problems I: a linear model problem,* SIAM J. Num. Anal., **28** (1991), 43–77.

[4] Gene H. Golub and Charles F. Van Loan, *Matrix Computations,* Johns Hopkins Unv. Press, (1985).

[5] Hideaki Kaneko, Kim S. Bey and Gene J.W. Hou, *Discontinuous Galerkin finite element method for parabolic problems,* Applied Mathematics and Computation, **182** (2006), 388–402.

[6] Hideaki Kaneko and Gene J.W. Hou, *Discontinuous Galerkin finite element method for parabolic problems with non-smmoth data,* Far East Jl. of Applied Mathematics, an invited paper in a special issue, **14** (2004), 137–155.

[7] Hideaki Kaneko, Kim S. Bey and Gene J.W. Hou, *Discontinuous Galerkin finite element method for parabolic problems with modified hp finite element approximation technique,* Applied Mathematics and Computation, **182** (2006), 1405–1417.

[8] J.P. Tomey, “p-version Discontinuous Galerkin Method for Heat Transfer in Built-up Structures,” M.S. thesis, George Washington University, May, 2001, (Research conducted at NASA Langley).

[9] Ch. Schwab, *p- and hp- Finite Element Methods,* Numerical Mathematics and Scientific Computation, Oxford Science Publication, Oxford University Press, (1998).

[10] W. Gui and I. Babuška, *The h, p and hp versions of the finite element method in 1 dimension, Part I. The error analysis of the p-version,* Numer. Math., **49** (1986), 577–612. *Part II. The error analysis of the h and hp versions,* Numer. Math., **49** (1986), 613–657. *Part III. The adaptive hp version,* Numer. Math., **49** (1986), 659–683.

[11] D. Schötzau and C. Schwab, *Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method,* SIAM J. Num. Anal., **38** (2000), 837–875.

[12] T. Werber, K. Gerdes, D. Schötzau and C. Schwab, *hp discontinuous Galerkin time stepping for parabolic problems,* Research Report No. 2000-01, April 2000, Seminar für Angewandte Mathematik, ETH, Zürich.

Received April 2006; revised January 2007.

E-mail address: hkaneko@odu.edu
E-mail address: kim.s.bey@nasa.gov
E-mail address: Peter.Daffer.Ctr@ROBINS.AF.MIL
E-mail address: g4737882@student.mahidol.ac.th