hp-DGFEM for Partial Differential Equations with Nonnegative Characteristic Form

E. Süli[†], C. Schwab and P. Houston[†] [*]

hp-DGFEM for Partial Differential Equations with Nonnegative Characteristic Form

E. Süli[†], C. Schwab and P. Houston[†] [*]

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Abstract

We develop the error analysis for the hp-version of a discontinuous finite element approximation to second-order partial differential equations with nonnegative characteristic form. This class of equations includes classical examples of second-order elliptic and parabolic equations, first-order hyperbolic equations, as well as equations of mixed type. We establish an a priori error bound for the method which is of optimal order in the mesh size h and 1 order less than optimal in the polynomial degree p. In the particular case of a first-order hyperbolic equation the error bound is optimal in h and $1/2$ an order less than optimal in p.

1   Introduction

Discontinuous Galerkin Finite Element Methods (DGFEM) were introduced over quarter of a century ago for the numerical solution of first-order hyperbolic problems [14,11] and as nonstandard techniques for the approximation of second-order elliptic equations [12] (see also [13] for a historical survey). Although subsequently much of the research in the field of numerical analysis of partial differential equations has concentrated on the development and the analysis of conforming finite element methods for self-adjoint elliptic problems, stabilised continuous finite element methods for convection-diffusion equations, and finite difference and finite volume methods for hyperbolic problems, recent years have witnessed renewed interest in discontinuous schemes. This paradigm shift was stimulated by several factors: the desire to handle, within the finite element framework, nonlinear hyperbolic problems (see [6] and [7]) which are known to exhibit discontinuous solutions even when the data are perfectly smooth; the need to treat convection-dominated diffusion problems without excessive numerical stabilisation; the computational convenience of discontinuous finite element methods due to a large degree of locality; and the necessity to accommodate high-order hp-adaptive finite element discretisations in a flexible manner (see [5]).

The aim of this paper is to extend the error analysis of the hp-DGFEM, developed in our earlier work [8] for first-order hyperbolic equations, to general second-order partial differential equations with nonnegative characteristic form. In [8] an error bound, optimal both in terms of the local mesh size h and the local polynomial degree p, was derived for the hp-DGFEM supplemented by a streamline-diffusion type stabilisation involving a stabilisation parameter δ of size h/p. Here, we focus on the more subtle situation when $\delta = 0$, corresponding to no stabilisation. By exploiting theoretical tools similar to those in [8], we derive an error bound for the resulting scheme that is of optimal order in terms of the mesh size h and 1 order less than optimal in the polynomial degree p. For convection-dominated diffusion equations, suboptimality in p is compensated by the fact that the leading term in the error bound is multiplied by a small number, proportional to the square root of the norm of the diffusion matrix. Indeed, in the case of a first-order hyperbolic equation, our error bound collapses to one that is h-optimal, with a loss of only $1/2$ an order in p. The approximation technique adopted in the present paper involves a discontinuity-penalisation device based on the ideas of Nitsche [12], Wheeler [18] and Arnold [1], albeit with a slight (but important) modification which permits us to pass to the hyperbolic limit with inactive discontinuity-penalisation. The error analysis of the hp-DGFEM discretisation considered here can also be viewed as an extension of the work of Baumann [3], Oden, Babuška and Baumann [13], and Riviere and Wheeler [15] for a self-adjoint elliptic problem.

2 Model Problem and Discretisation

Given that Ω is a bounded Lipschitz domain in $\mathbb{R}^d$, $d \geq 2$, we consider the linear second-order partial differential equation

$$\mathcal{L}u \equiv -\sum_{i,j=1}^d \partial_j \left(a_{ij}(x)\,\partial_i u \right) + \sum_{i=1}^d b_i(x)\,\partial_i u + c(x)u = f(x) \ , \qquad (1)$$

where f is a real-valued function belonging to $L^2(\Omega)$, and the real-valued coefficients a, b, c have the following properties:

$$a(x) = \{a_{ij}(x)\}_{i,j=1}^d \in L^\infty(\Omega)_{\mathrm{sym}}^{d \times d} \ ,$$
$$\boldsymbol{b}(x) = \{b_i(x)\}_{i=1}^d \in W^{1,\infty}(\Omega)^d, \quad c(x) \in L^\infty(\Omega) \ . \qquad (2)$$

We shall suppose throughout that the characteristic form associated with the principal part of the differential operator $\mathcal{L}$ is nonnegative; namely,

$$\boldsymbol{\xi}^T a(x)\,\boldsymbol{\xi} \geq 0 \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d \text{ and a.e. } x \in \bar{\Omega} \ . \qquad (3)$$

For simplicity, we shall assume that the entries of the matrix a are piecewise continuous on $\bar{\Omega}$; this hypothesis is sufficiently general to cover most situations of practical relevance. Let $\boldsymbol{\mu}(x) = \{\mu_i(x)\}_{i=1}^d$ denote the unit outward normal vector to $\Gamma = \partial\Omega$ at $x \in \Gamma$ and define the following subsets of Γ:

$$\Gamma_0 = \left\{x \in \Gamma : \ \boldsymbol{\mu}^T a(x)\boldsymbol{\mu} > 0\right\} \ ,$$

$$\Gamma_- = \left\{x \in \Gamma \backslash \Gamma_0 : \ \boldsymbol{b} \cdot \boldsymbol{\mu} < 0\right\} \quad \text{and} \quad \Gamma_+ = \left\{x \in \Gamma \backslash \Gamma_0 : \ \boldsymbol{b} \cdot \boldsymbol{\mu} \geq 0\right\} \ .$$

The sets $\Gamma_\mp$ will be referred to as the inflow and outflow boundary, respectively. With these definitions we have that $\Gamma = \Gamma_0 \cup \Gamma_- \cup \Gamma_+$. We shall further decompose Γ_0 into two connected parts, Γ_D and Γ_N, and supplement the partial differential equation (1) with the following boundary conditions:

$$u = g_\mathrm{D} \quad \text{on } \Gamma_\mathrm{D} \cup \Gamma_- \quad \text{and} \quad \boldsymbol{\mu}^T a\nabla u = g_\mathrm{N} \text{ on } \Gamma_\mathrm{N} \ . \qquad (4)$$

We note that (1), (4) includes a range of physically relevant problems, such as the mixed boundary value problem for an elliptic equation corresponding to the case when (3) holds with strict inequality, as well as the case of a linear transport problem associated with the choice of $a \equiv 0$ on $\bar{\Omega}$. Our aim here is to develop, in a unified manner, the a priori error analysis of the hp-version of a discontinuous finite element approximation to (1), (4).

2.1 Finite element spaces

Let $\mathcal{T}$ be a subdivision of Ω into open element domains κ such that $\bar{\Omega} = \cup_{\kappa \in \mathcal{T}} \bar{\kappa}$. We shall assume that the family of subdivisions $\mathcal{T}$ is shape regular

and that each $\kappa \in \mathcal{T}$ is a smooth bijective image of a fixed master element $\hat{\kappa}$, that is, $\kappa = F_\kappa(\hat{\kappa})$ for all $\kappa \in \mathcal{T}$ where $\hat{\kappa}$ is either the open unit simplex or the open unit hypercube in $\mathbb{R}^d$. For an integer $r \geq 1$, we denote by $\mathcal{P}_r(\hat{\kappa})$ the set of polynomials of total degree $< r$ on $\hat{\kappa}$; when $\hat{\kappa}$ is the unit hypercube, we also consider $\mathcal{Q}_r(\hat{\kappa})$, the set of all tensor-product polynomials of degree $< r$ in each coordinate direction. Thus, to $\kappa \in \mathcal{T}$ we assign an integer $p_\kappa \geq 1$, collect the p_κ and F_κ in the vectors $\mathbf{p} = \{p_\kappa : \kappa \in \mathcal{T}\}$ and $\mathbf{F} = \{F_\kappa : \kappa \in \mathcal{T}\}$, respectively, and consider the finite element space

$$S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) = \{u \in L^2(\Omega) : \ u|_\kappa \circ F_\kappa \in \mathcal{R}_{p_\kappa}(\hat{\kappa}) \quad \forall \kappa \in \mathcal{T}\} \ ,$$

where $\mathcal{R}$ is either $\mathcal{P}$ or $\mathcal{Q}$. Given the subdivision $\mathcal{T}$ and $s > 0$, the associated broken Sobolev space $H^s(\Omega, \mathcal{T})$ is defined by

$$H^s(\Omega, \mathcal{T}) = \prod_{\kappa \in \mathcal{T}} H^s(\kappa) = \{u \in L^2(\Omega) : \ u|_\kappa \in H^s(\kappa) \quad \forall \kappa \in \mathcal{T}\} \ .$$

In the next section, we formulate the hp-DGFEM approximation of (1), (4).

2.2 The numerical method

Discretisation of the Low-Order Terms. Let us begin by considering the first-order partial differential operator $\mathcal{L}_b$ defined by

$$\mathcal{L}_b w = \boldsymbol{b} \cdot \nabla w + cw \ .$$

Given that κ is an element in the partition $\mathcal{T}$, we denote by $\partial\kappa$ the union of open faces of κ. This is non-standard notation in that $\partial\kappa$ is a subset of the boundary of κ. Let $x \in \partial\kappa$ and suppose that $\boldsymbol{\mu}(x)$ denotes the unit outward normal vector to $\partial\kappa$ at x. With these conventions, we define the inflow and outflow parts of $\partial\kappa$, respectively, by

$$\partial_-\kappa = \{x \in \partial\kappa : \ \boldsymbol{b}(x) \cdot \boldsymbol{\mu}(x) < 0\} \ , \qquad \partial_+\kappa = \{x \in \partial\kappa : \ \boldsymbol{b}(x) \cdot \boldsymbol{\mu}(x) \geq 0\} \ .$$

For each $v \in H^1(\Omega, \mathcal{T})$ and any $\kappa \in \mathcal{T}$, we denote by v^+ the interior trace of v on $\partial\kappa$ (the trace taken from within κ). Now consider an element κ such that the set $\partial_-\kappa \backslash \Gamma_-$ is nonempty; then for each $x \in \partial_-\kappa \backslash \Gamma_-$ (with the exception of a set of $(d-1)$-dimensional measure zero) there exists a unique element κ', depending on the choice of x, such that $x \in \partial_+\kappa'$. If $\partial_-\kappa \backslash \Gamma_-$ is nonempty for some $\kappa \in \mathcal{T}$, then we can also define the outer trace v^- of v on $\partial_-\kappa \backslash \Gamma_-$ relative to κ as the inner trace v^+ relative to those elements κ' for which $\partial_+\kappa'$ has intersection with $\partial_-\kappa \backslash \Gamma_-$ of positive $(d-1)$-dimensional measure. Further, we introduce the oriented jump of v across $\partial_-\kappa \backslash \Gamma_-$:

$$\lfloor v \rfloor = v^+ - v^- \ .$$

4

Supposing that $v, w \in H^1(\Omega, \mathcal{T})$, we define, as in [10], for example,

$$B_b(w,v) = \sum_{\kappa \in \mathcal{T}} \int_\kappa (\mathcal{L}_b w) v \ \mathrm{d}x \tag{5}$$

$$- \sum_{\kappa \in \mathcal{T}} \int_{\partial_- \kappa \setminus \Gamma_-} (\boldsymbol{b} \cdot \boldsymbol{\mu}) \lfloor w \rfloor v^+ \ \mathrm{d}s - \sum_{\kappa \in \mathcal{T}} \int_{\partial_- \kappa \cap \Gamma_-} (\boldsymbol{b} \cdot \boldsymbol{\mu}) w^+ v^+ \ \mathrm{d}s \ ,$$

and we put

$$\ell_b(v) = \sum_{\kappa \in \mathcal{T}} \int_\kappa f v \, \mathrm{d}x - \sum_{\kappa \in \mathcal{T}} \int_{\partial_- \kappa \cap \Gamma_-} (\boldsymbol{b} \cdot \boldsymbol{\mu}) \, g v^+ \ \mathrm{d}s \ .$$

Discretisation of the Leading Term. Let us suppose that the elements in the partition have been numbered in a certain way, regardless of the flow direction. We denote by $\mathcal{E}$ the set of element interfaces (edges for $d = 2$ or faces for $d = 3$) associated with the subdivision $\mathcal{T}$. Since hanging nodes are permitted in the DGFEM, $\mathcal{E}$ will be understood to consist of the smallest interfaces in $\partial \kappa$. With this notation, let Γ_{int} denote the union of all interfaces $e \in \mathcal{E}$. Given that $e \in \mathcal{E}$, there exist indices i and j such that $i > j$ and κ_i and κ_j share the interface e; we define the (numbering-dependent) jump of $v \in H^1(\Omega, \mathcal{T})$ across e and the mean value of v on e, respectively, by

$$[v] = v|_{\partial \kappa_i \cap e} - v|_{\partial \kappa_j \cap e} \quad \text{and} \quad \langle v \rangle = \frac{1}{2} \left(v|_{\partial \kappa_i \cap e} + v|_{\partial \kappa_j \cap e} \right) \ .$$

We note that, in general, $[v]$ is distinct from $\lfloor v \rfloor$ in that the latter depends on the sign of the normal flux on an element boundary, while the former is only dependent on the element numbering. With each face $e \in \mathcal{E}$ we associate the normal vector $\boldsymbol{\nu}$ which points from κ_i to κ_j; on boundary faces, we put $\boldsymbol{\nu} = \boldsymbol{\mu}$. Finally, we introduce, as in [13], the bilinear form

$$B_a(w,v) = \sum_{\kappa \in \mathcal{T}} \int_\kappa a(x) \nabla w \cdot \nabla v \, \mathrm{d}x + \int_{\Gamma_{\mathrm{D}}} \{ w((a\nabla v) \cdot \boldsymbol{\nu}) - ((a\nabla w) \cdot \boldsymbol{\nu}) v \} \, \mathrm{d}s$$

$$+ \int_{\Gamma_{\mathrm{int}}} \{ [w] \langle (a\nabla v) \cdot \boldsymbol{\nu} \rangle - \langle (a\nabla w) \cdot \boldsymbol{\nu} \rangle [v] \} \, \mathrm{d}s \ , \tag{6}$$

associated with the principal part of the partial differential operator $\mathcal{L}$, and the linear functional

$$\ell_a(v) = \int_{\Gamma_{\mathrm{D}}} g_{\mathrm{D}}((a\nabla v) \cdot \boldsymbol{\nu}) \, \mathrm{d}s + \int_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}} v \, \mathrm{d}s \ .$$

Discontinuity-Penalisation Term. Let $\bar{a} = ||a||_2$, with $|| \cdot ||_2$ denoting the matrix norm subordinate to the l^2 vector norm on $\mathbb{R}^d$, and let $\bar{a}_\kappa = \bar{a}|_\kappa$. To each e in $\mathcal{E}$ which is a common face of elements κ_i and κ_j in $\mathcal{T}$ we assign the nonnegative function $\langle \bar{a} p^2 \rangle_e = (p_{\kappa_i}^2 \bar{a}_{\kappa_i}|_e + p_{\kappa_j}^2 \bar{a}_{\kappa_j}|_e)/2$. Letting $\mathcal{E}_{\mathrm{D}}$ denote

the set of all faces contained in Γ_{D}, to each $e \in \Gamma_{\mathrm{D}}$ we assign the element $\kappa \in \mathcal{T}$ with that face and define $\langle \bar{a} p^2 \rangle_e = p_\kappa^2 \bar{a}_\kappa|_e$. Consider the function σ defined on $\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}$ by $\sigma(x) = K \langle \bar{a} p^2 \rangle_e / |e|$ for $x \in e$ and $e \in \mathcal{E} \cup \mathcal{E}_{\mathrm{D}}$, where $|e| = \mathrm{meas}_{d-1}(e)$ and K is a positive constant (whose value is irrelevant for the present analytical study, so we put $K = 1$), and introduce the bilinear form and the linear functional, respectively, by

$$B_s(w, v) = \int_{\Gamma_{\mathrm{D}}} \sigma w v \, \mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \sigma [w][v] \, \mathrm{d}s \ , \quad \ell_s(v) = \int_{\Gamma_{\mathrm{D}}} \sigma g_{\mathrm{D}} v \, \mathrm{d}s \ . \quad (7)$$

We highlight the fact that since the weight-function σ involves the norm of the matrix a, in the hyperbolic limit of $a \equiv 0$ the bilinear form $B_s(\cdot, \cdot)$ and the linear functional ℓ_s both vanish. This is a desirable property, since linear hyperbolic equations may possess solutions that are discontinuous across characteristic hypersurfaces, and penalising discontinuities across faces which belong to these would seem unnatural.

It is also worth noting here that, conceptually, the bilinear form $B_a(\cdot, \cdot) + B_s(\cdot, \cdot)$ should be regarded as a single entity, rather than a sum of two separate bilinear forms; the same comment applies to $\ell_a(\cdot) + \ell_s(\cdot)$. Although more convenient from the point of view of the presentation, separation into B_a, ℓ_a on the one hand and B_s, ℓ_s on the other is somewhat artificial and can only be justified on historical grounds (see [12,18,1]).

Definition of the Method. Finally, we define the bilinear form $B_{\mathrm{DG}}(\cdot, \cdot)$ and the linear functional $\ell_{\mathrm{DG}}(\cdot)$, respectively, by

$$B_{\mathrm{DG}}(w, v) = B_a(w, v) + B_b(w, v) + B_s(w, v) \ ,$$
$$\ell_{\mathrm{DG}}(v) = \ell_a(v) + \ell_b(v) + \ell_s(v) \ .$$

The *hp*-DGFEM approximation of (1), (4) is: find $u_{\mathrm{DG}} \in S^{\mathbf{P}}(\Omega, \mathcal{T}, \mathbf{F})$ such that

$$B_{\mathrm{DG}}(u_{\mathrm{DG}}, v) = \ell_{\mathrm{DG}}(v) \qquad \forall v \in S^{\mathbf{P}}(\Omega, \mathcal{T}, \mathbf{F}) \ . \quad (8)$$

In the next section we state the key properties of this method. Before we do so, however, we note that in the definitions of the bilinear forms and linear functionals above and in the arguments which follow it has been tacitly assumed that $a \in C(\kappa)$ for each $\kappa \in \mathcal{T}$, that the fluxes $(a\nabla u) \cdot \boldsymbol{\nu}$ and $(\boldsymbol{b} \cdot \boldsymbol{\mu})u$ are continuous across element interfaces, and that u is continuous in an (open) neighbourhood of the subset of Ω where a is not identically equal to zero. If the problem under consideration violates these properties, the scheme and the analysis have to be modified accordingly.

3   Analytical Results

Our first result concerns the positivity of the bilinear form $B_{\mathrm{DG}}(\cdot, \cdot)$ and the existence and uniqueness of a solution to (8).

Theorem 1. *Suppose that, in addition to the conditions (2) and (3), the function $\gamma \equiv c - \frac{1}{2}\nabla \cdot \boldsymbol{b}$ is nonnegative on $\bar{\Omega}$. Then,*

$$|||w|||^2_{\mathrm{DG}} \equiv B_{\mathrm{DG}}(w, w) = D + \sum_{\kappa \in \mathcal{T}} E_\kappa + \frac{1}{2}\sum_{\kappa \in \mathcal{T}} F_\kappa \ , \qquad (9)$$

where

$$D \equiv \int_{\Gamma_{\mathrm{D}}} \sigma w^2 \, ds + \int_{\Gamma_{\mathrm{int}}} \sigma[w]^2 \, ds \ , \quad E_\kappa \equiv \|\sqrt{a}\nabla w\|^2_{L^2(\kappa)} + \|\sqrt{\gamma}w\|^2_{L^2(\kappa)} \ ,$$

$$F_\kappa \equiv \int_{\partial_-\kappa \cap \Gamma_-} |\boldsymbol{b} \cdot \boldsymbol{\mu}| w_+^2 \, ds + \int_{\partial_+\kappa \cap \Gamma_+} |\boldsymbol{b} \cdot \boldsymbol{\mu}| w_+^2 \, ds + \int_{\partial_-\kappa \setminus \Gamma_-} |\boldsymbol{b} \cdot \boldsymbol{\mu}| \lfloor w \rfloor^2 \, ds \ ,$$

with $\sqrt{a}$ denoting the (nonnegative) square-root of the matrix a, and σ as in the definition of the discontinuity-penalisation. Furthermore, given that either a is positive definite or $\gamma > 0$ on each element κ in the partition $\mathcal{T}$, the hp-DGFEM (8) has a unique solution u_{DG} in $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$.

Proof. We begin by proving (9). First, we note that, trivially,

$$B_s(w, w) = \int_{\Gamma_{\mathrm{D}}} \sigma w^2 \, \mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \sigma[w]^2 \, \mathrm{d}s \ .$$

Further, as $(\boldsymbol{b} \cdot \nabla w)w = \frac{1}{2}\boldsymbol{b} \cdot \nabla(w^2)$, after integration by parts we have that

$$B_b(w, w) = \frac{1}{2}\sum_{\kappa \in \mathcal{T}} F_\kappa + \sum_{\kappa \in \mathcal{T}} \int_\kappa |\sqrt{\gamma(x)}w(x)|^2 \, \mathrm{d}x \ .$$

Finally, we observe that

$$B_a(w, w) = \sum_{\kappa \in \mathcal{T}} \int_\kappa |\sqrt{a(x)}\nabla w(x)|^2 \, \mathrm{d}x \ .$$

Upon adding these three identities, we arrive at (9).

To complete the proof of the lemma, we note that if either a is positive definite or $\gamma > 0$ on each element κ in the partition $\mathcal{T}$, then $B_{\mathrm{DG}}(w, w) > 0$ for all w in $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F}) \setminus \{0\}$, and hence we deduce the uniqueness of the solution u_{DG}. Further, since the linear space $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ is finite-dimensional, the existence of the solution to (8) follows from the fact that its homogeneous counterpart has the unique solution $u_{\mathrm{DG}} \equiv 0$. $\qquad\square$

Our second result provides a bound on the discretisation error. For simplicity, we shall assume that the entries of the matrix a are constant on each element $\kappa \in \mathcal{T}$ (with possible discontinuities across faces $e \in \mathcal{E}$) and $\boldsymbol{b}$ is a constant vector. We require the following approximation result [16].

Lemma 2. *Suppose that $u \in H^{k_\kappa}(\kappa)$, $k_\kappa \geq 0$, $\kappa \in \mathcal{T}$, and let $\Pi_{hp}u$ denote the orthogonal projection of u in $L^2(\Omega)$ onto the finite element space $S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$. Then, there exists a constant C dependent on k_κ and the angle condition of κ, but independent of u, $h_\kappa = diam(\kappa)$ and p_κ, such that*

$$\|u - \Pi_{hp}u\|_{L^2(\kappa)} \leq C \frac{h_\kappa^{\tau_\kappa}}{p_\kappa^{k_\kappa}} \|u\|_{H^{k_\kappa}(\kappa)} \quad , \tag{10}$$

where $\tau_\kappa = \min(p_\kappa, k_\kappa)$, $\kappa \in \mathcal{T}$.

Next, we state our main result, regarding the accuracy of the method (8).

Theorem 3. *Assume that there exists a positive constant γ_0 such that $\gamma \geq \gamma_0$ on each element κ in the partition $\mathcal{T}$. Then, assuming that $u \in H^{k_\kappa}(\kappa)$, $k_\kappa \geq 2$, for $\kappa \in \mathcal{T}$, the solution $u_{\mathrm{DG}} \in S^{\mathbf{p}}(\Omega, \mathcal{T}, \mathbf{F})$ of (8) obeys the error bound*

$$|||u - u_{\mathrm{DG}}|||_{\mathrm{DG}}^2 \leq C \sum_{\kappa \in \mathcal{T}} \left(\bar{a}_\kappa \frac{h_\kappa^{2(\tau_\kappa - 1)}}{p_\kappa^{2(k_\kappa - 2)}} + \bar{b} \frac{h_\kappa^{2(\tau_\kappa - 1/2)}}{p_\kappa^{2(k_\kappa - 1)}} \right) \|u\|_{H^{k_\kappa}(\kappa)}^2 \quad , \tag{11}$$

where $\tau_\kappa = \min(p_\kappa, k_\kappa)$ and $\bar{b}$ is the l^2 vector norm of $\mathbf{b}$.

Proof. Let us decompose $e = u - u_{\mathrm{DG}}$ as $e = \eta + \xi$ where $\eta = u - \Pi_{hp}u$, $\xi = \Pi_{hp}u - u_{\mathrm{DG}}$, and Π_{hp} is as in Lemma 2. Then, by virtue of Theorem 1,

$$|||\xi|||_{\mathrm{DG}}^2 = B_{\mathrm{DG}}(\xi, \xi) = B_{\mathrm{DG}}(e - \eta, \xi) = -B_{\mathrm{DG}}(\eta, \xi) \quad ,$$

where we have used the Galerkin orthogonality property: $B_{\mathrm{DG}}(u - u_{\mathrm{DG}}, \xi) = 0$ which follows from (8) with $v = \xi$ and the definition of the boundary value problem (1), (4), given the assumed smoothness of u. Thus, we deduce that

$$|||\xi|||_{\mathrm{DG}}^2 \leq |B_a(\eta, \xi)| + |B_b(\eta, \xi)| + |B_s(\eta, \xi)| \quad .$$

Now, from (7) we have that

$$|B_s(\eta, \xi)| \leq |||\xi|||_{\mathrm{DG}} \left(\int_{\Gamma_D} \sigma |\eta|^2 \, \mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \sigma [\eta]^2 \, \mathrm{d}s \right)^{1/2} \quad . \tag{12}$$

As $\nabla \cdot \mathbf{b} = 0$ on each $\kappa \in \mathcal{T}$, after integration by parts, we obtain

$$B_b(\eta, \xi) = \sum_\kappa \int_\kappa c\eta\xi \, \mathrm{d}x - \sum_{\kappa \in \mathcal{T}} \int_\kappa \eta(\mathbf{b} \cdot \nabla \xi) \, \mathrm{d}x + \sum_{\kappa \in \mathcal{T}} \int_{\partial_+\kappa \cap \Gamma_+} (\mathbf{b} \cdot \boldsymbol{\mu})\eta^+ \xi^+ \, \mathrm{d}s$$

$$+ \sum_{\kappa \in \mathcal{T}} \int_{\partial_+\kappa \backslash \Gamma_+} (\mathbf{b} \cdot \boldsymbol{\mu})\eta^+ \xi^+ \, \mathrm{d}s + \sum_{\kappa \in \mathcal{T}} \int_{\partial_-\kappa \backslash \Gamma_-} (\mathbf{b} \cdot \boldsymbol{\mu})\eta^- \xi^+ \, \mathrm{d}s \quad . \tag{13}$$

Denoting by $S_4 + S_5$ the sum of the last two (of the five) terms in (13), we find, after shifting the 'indices' in the summation in S_4, that

$$|S_4 + S_5| \leq \sum_{\kappa \in \mathcal{T}} \left(\int_{\partial_-\kappa \backslash \Gamma_-} |\mathbf{b} \cdot \boldsymbol{\mu}||\eta^-|^2 \, \mathrm{d}s \right)^{1/2} \left(\int_{\partial_-\kappa \backslash \Gamma_-} |\mathbf{b} \cdot \boldsymbol{\mu}|\lfloor\xi\rceil^2 \, \mathrm{d}s \right)^{1/2} \quad .$$

8

Also, since $\boldsymbol{b}$ is a constant vector, $\int_\kappa \eta(\boldsymbol{b} \cdot \nabla\xi)\,\mathrm{d}x = 0$. Thus, (13) yields

$$|B_b(\eta,\xi)| \leq C|||\xi|||_{\mathrm{DG}} \left(\|\eta\|_{L^2(\Omega)}^2 + \sum_{\kappa\in\mathcal{T}} \int_{\partial_+\kappa\cap\Gamma_+} |\boldsymbol{b}\cdot\boldsymbol{\mu}||\eta^+|^2\,\mathrm{d}s \right.$$
$$\left. + \sum_{\kappa\in\mathcal{T}} \int_{\partial_-\kappa\backslash\Gamma_-} |\boldsymbol{b}\cdot\boldsymbol{\mu}||\eta^-|^2\,\mathrm{d}s \right)^{1/2}, \qquad (14)$$

where C is a positive constant, as in the statement of the theorem.

Next,

$$|B_a(\eta,\xi)| \leq I + II + III ,$$

where

$$I \equiv \left| \sum_{\kappa\in\mathcal{T}} \int_\kappa a\nabla\eta\cdot\nabla\xi\,\mathrm{d}x \right| , \quad II \equiv \left| \int_{\Gamma_{\mathrm{D}}} \{\eta((a\nabla\xi)\cdot\boldsymbol{\nu}) - ((a\nabla\eta)\cdot\boldsymbol{\nu})\xi\}\,\mathrm{d}s \right| ,$$

$$III \equiv \left| \int_{\Gamma_{\mathrm{int}}} \{[\eta]\langle(a\nabla\xi)\cdot\boldsymbol{\nu}\rangle - \langle(a\nabla\eta)\cdot\boldsymbol{\nu}\rangle[\xi]\}\,\mathrm{d}s \right| .$$

Now,

$$I^2 \leq |||\xi|||_{\mathrm{DG}}^2 \sum_{\kappa\in\mathcal{T}} \|\sqrt{a}\nabla\eta\|_{L^2(\kappa)}^2 ,$$

$$II^2 \leq C|||\xi|||_{\mathrm{DG}}^2 \sum_{\kappa\,:\,\partial\kappa\cap\Gamma_{\mathrm{D}}\neq\emptyset} \left(\frac{\bar{a}_\kappa p_\kappa^2}{h_\kappa}\|\eta\|_{L^2(\partial\kappa\cap\Gamma_D)}^2 + \frac{\bar{a}_\kappa h_\kappa}{p_\kappa^2}\|\nabla\eta\|_{L^2(\partial\kappa\cap\Gamma_D)}^2 \right) ,$$

$$III^2 \leq C|||\xi|||_{\mathrm{DG}}^2 \sum_{\kappa\,:\,\partial\kappa\cap\Gamma=\emptyset} \left(\frac{\bar{a}_\kappa p_\kappa^2}{h_\kappa}\|[\eta]\|_{L^2(\partial\kappa)}^2 + \frac{\bar{a}_\kappa h_\kappa}{p_\kappa^2}\|\nabla\eta\|_{L^2(\partial\kappa)}^2 \right) .$$

Collecting the bounds on the terms I, II and III gives

$$|B_a(\eta,\xi)| \leq C|||\xi|||_{\mathrm{DG}} \left(\sum_{\kappa\in\mathcal{T}} \|\sqrt{a}\nabla\eta\|_{L^2(\kappa)}^2 \right.$$
$$+ \sum_{\kappa\,:\,\partial\kappa\cap\Gamma_{\mathrm{D}}\neq\emptyset} \left(\frac{\bar{a}_\kappa p_\kappa^2}{h_\kappa}\|\eta\|_{L^2(\partial\kappa\cap\Gamma_D)}^2 + \frac{\bar{a}_\kappa h_\kappa}{p_\kappa^2}\|\nabla\eta\|_{L^2(\partial\kappa\cap\Gamma_D)}^2 \right)$$
$$\left. + \sum_{\kappa\,:\,\partial\kappa\cap\Gamma=\emptyset} \left(\frac{\bar{a}_\kappa p_\kappa^2}{h_\kappa}\|[\eta]\|_{L^2(\partial\kappa)}^2 + \frac{\bar{a}_\kappa h_\kappa}{p_\kappa^2}\|\nabla\eta\|_{L^2(\partial\kappa)}^2 \right) \right)^{1/2}. \quad (15)$$

The required result now follows by noting that

$$|||u - u_{\mathrm{DG}}|||_{\mathrm{DG}} \leq |||\eta|||_{\mathrm{DG}} + |||\xi|||_{\mathrm{DG}} ,$$

using the estimates (12), (14) and (15) to bound $|||u - u_{\mathrm{DG}}|||_{\mathrm{DG}}$ in terms of $|||\eta|||_{\mathrm{DG}}$ and other norms of η, and applying Lemma 2, together with the Trace Inequality

$$\|v\|^2_{L^2(e)} \le C\left(\|v\|_{L^2(\kappa)}\|\nabla v\|_{L^2(\kappa)} + h_\kappa^{-1}\|v\|^2_{L^2(\kappa)}\right) \;, \quad v \in H^1(\kappa), \; e \subset \partial\kappa \;,$$

to estimate norms of η and $\nabla\eta$ over $\partial\kappa \cap \Gamma_{\mathrm{D}}$ and $\partial\kappa$ in terms of norms of over κ, $\kappa \in \mathcal{T}$. The argument is fairly standard, so we omit the details. □

We note that in the purely hyperbolic case of $a \equiv 0$ the error bound in Theorem 3 collapses to $O(h^{\tau-1/2}/p^{k-1})$; in the DG-norm, this is optimal with respect to h, while in p it is $1/2$ an order below the hp-optimal bound established in [8]. In fact, for $a \equiv 0$, the error bound of Theorem 3 is identical to the p-suboptimal hp error estimate of Bey and Oden [4], except that there a streamline-diffusion type stabilisation was included with stabilisation parameter $\delta = h/p^2$; Theorem 3 corresponds to $\delta = 0$.

In the case of non-constant $\boldsymbol{b}$, (11) should be supplemented with the term $|\boldsymbol{b}|^2_{W^{1,\infty}(\kappa)}(h^{2\tau_\kappa}/p^{2(k_\kappa-2)})\|u\|^2_{H^{k_\kappa}(\kappa)}$ under the summation sign on the right. When $\bar{a}_\kappa \ge c_0 > 0$ this additional term can be absorbed into the first term on the right; otherwise it degrades the error bound with respect to p. More generally, when streamline-diffusion stabilisation is added to (8), with stabilisation parameter $\delta = (h/p)\min\left(1, \bar{b}h/\bar{a}p^3\right)$, the bound (11) can be, simultaneously, extended to the case of non-constant $\boldsymbol{b}$ and sharpened to one that is still optimal in h, but now with only $1/2$ a power of p below the optimal rate in the diffusive part and of optimal order in p in the advective part. Specifically, when $\boldsymbol{b} \equiv 0$, we recover the bound $O(h^{\tau-1}/p^{k-3/2})$ of Riviere and Wheeler [15]; on the other hand, if $a = 0$, we arrive at the hp-optimal error bound $O(h^{\tau-1/2}/p^{k-1/2})$ of [8] in the DG-norm, proved with $\delta = h/p$, which represents the direct generalisation of the optimal h-version bound for the DGFEM (see [9] and [10]) to the hp-version. The proof of this is beyond the scope of the present paper and will be delivered in [17]. For further developments regarding these theoretical questions for hyperbolic and nearly-hyperbolic problems and numerical experiments, see [8,17].

References

1. Arnold, D.N.: An interior penalty finite element method with discontinuous elements. SIAM J. Numer. Anal. **19** (1982) 742–760
2. Babuška, I., Suri, M.: The hp-Version of the Finite Element Method with quasiuniform meshes. M^{2}AN Mathematical Modelling and Numerical Analysis **21** (1987) 199–238
3. Baumann, C.: An hp-Adaptive Discontinuous Galerkin FEM for Computational Fluid Dynamics. Doctoral Dissertation. TICAM, UT Austin, Texas, 1997
4. Bey, K.S., Oden, J.T.: hp-Version discontinuous Galerkin methods for hyperbolic conservation laws. Comput. Methods Appl. Mech. Engrg. **133** (1996) 259–286

10

5. Biswas, R., Devine, K., Flaherty, J.E.: Parallel adaptive finite element methods for conservation laws. App. Numer. Math. **14** (1994) 255–284
6. Cockburn, B., Hou, S., Shu, C.-W.: TVB Runge-Kutta local projection discontinuous Galerkin finite elements for hyperbolic conservation laws. Math. Comp. **54** (1990) 545–581
7. Cockburn, B., Shu, C.-W.: The local discontinuous Galerkin method for time-dependent reaction-diffusion systems. SIAM J. Numer. Anal. **35** (1998) 2440–2463
8. Houston, P., Schwab, Ch., Süli, E.: Stabilized hp-finite element methods for first–order hyperbolic problems. Oxford University Computing Laboratory Technical Report NA-98/14, 1998 (submitted for publication)
9. Johnson, C., Nävert, U., Pitkäranta, J.: Finite element methods for linear hyperbolic problems. Comp. Meth. Appl. Mech. Engrg. **45** (1984) 285–312
10. Johnson, C., Pitkäranta, J.: An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. Math. Comp. **46** (1986) 1–26
11. Lesaint, P., Raviart, P.-A.: On a finite element method for solving the neutron transport equation. In: Mathematical Aspects of Finite Elements in Partial Differential Equations. C.A. deBoor (Ed.), Academic Press, New York (1974) 89–145
12. Nitsche, J.: Über ein Variationsprinzip zur Lösung von Dirichlet Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. Abh. Math. Sem. Univ. Hamburg **36** (1971) 9-15
13. Oden, J.T., Babuška, I., Baumann, C.: A discontinuous hp-FEM for diffusion problems. J. Comp. Phys. **146** (1998) 491–519
14. Reed, W.H., Hill, T.R.: Triangular mesh methods for neutron transport equation. Los Alamos Scientific Laboratory report LA-UR-73-479, Los Alamos, NM, 1973
15. Riviere, B., Wheeler, M.F.: Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I. TICAM Technical Report, University of Texas at Austin, Texas, 1999
16. Schwab, Ch.: p- and hp-Finite Element Methods. Theory and Applications to Solid and Fluid Mechanics. Oxford University Press (1998)
17. Süli, E., Houston, P., Schwab, Ch.: hp-DGFEM for partial differential equations of mixed type. Oxford University Computing Laboratory Technical Report NA-99/05, 1999
18. Wheeler, M.F.: An elliptic collocation finite element method with interior penalties. SIAM J. Numer. Anal. **15** (1978) 152–161

Research Reports

No.	Authors	Title
99-13	E. Süli, C. Schwab, P. Houston	hp-DGFEM for Partial Differential Equations with Nonnegative Characteristic Form
99-12	K. Nipp	Numerical integration of differential algebraic systems and invariant manifolds
99-11	C. Lage, C. Schwab	Advanced boundary element algorithms
99-10	D. Schötzau, C. Schwab	Exponential Convergence in a Galerkin Least Squares hp-FEM for Stokes Flow
99-09	A.M. Matache, C. Schwab	Homogenization via p-FEM for Problems with Microstructure
99-08	D. Braess, C. Schwab	Approximation on Simplices with respect to Weighted Sobolev Norms
99-07	M. Feistauer, C. Schwab	Coupled Problems for Viscous Incompressible Flow in Exterior Domains
99-06	J. Maurer, M. Fey	A Scale-Residual Model for Large-Eddy Simulation
99-05	M.J. Grote	Am Rande des Unendlichen: Numerische Verfahren für unbegrenzte Gebiete
99-04	D. Schötzau, C. Schwab	Time Discretization of Parabolic Problems by the hp-Version of the Discontinuous Galerkin Finite Element Method
99-03	S.A. Zimmermann	The Method of Transport for the Euler Equations Written as a Kinetic Scheme
99-02	M.J. Grote, A.J. Majda	Crude Closure for Flow with Topography Through Large Scale Statistical Theory
99-01	A.M. Matache, I. Babuška, C. Schwab	Generalized p-FEM in Homogenization
98-10	J.M. Melenk, C. Schwab	The hp Streamline Diffusion Finite Element Method for Convection Dominated Problems in one Space Dimension
98-09	M.J. Grote	Nonreflecting Boundary Conditions For Electromagnetic Scattering
98-08	M.J. Grote, J.B. Keller	Exact Nonreflecting Boundary Condition For Elastic Waves
98-07	C. Lage	Concept Oriented Design of Numerical Software
98-06	N.P. Hancke, J.M. Melenk, C. Schwab	A Spectral Galerkin Method for Hydrodynamic Stability Problems
98-05	J. Waldvogel	Long-Term Evolution of Coorbital Motion
98-04	R. Sperb	An alternative to Ewald sums, Part 2: The Coulomb potential in a periodic system