Tenth MSU Conference on Differential Equations and Computational Simulations.
Electronic Journal of Differential Equations, Conference 23 (2016), pp. 59–76.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu

DISTRIBUTIONAL DERIVATIVES AND STABILITY OF DISCONTINUOUS GALERKIN FINITE ELEMENT APPROXIMATION METHODS

THOMAS LEWIS

Dedicated to Professor Ratnasingham Shivaji on his 60-th birthday

Abstract. The goal of this article is to explore and motivate stabilization requirements for various types of discontinuous Galerkin (DG) methods. A new approach for the understanding of DG approximation methods for second order elliptic partial differential equations is introduced. The approach explains the weaker stability requirements for local discontinuous Galerkin (LDG) methods when compared to interior-penalty discontinuous Galerkin methods while also motivating the existence of methods such as the minimal dissipation LDG method that are stable without the addition of interior penalization. The main idea is to relate the underlying DG gradient approximation to distributional derivatives instead of the traditional piecewise gradient operator associated with broken Sobolev spaces.

1. Introduction

In this article, a new approach for understanding discontinuous Galerkin (DG) finite element approximation methods for elliptic partial differential equations (PDEs) will be motivated and explored with regards to stability requirements. The main emphasis will be on the difficulties associated with extending weak solution theory for PDEs to broken Sobolev spaces. To this end, existence and uniqueness results, as well as Friedrichs inequalities, will be established for second order elliptic problems posed in broken Sobolev spaces. Once the extension has been made, we will be able to directly recover many existing discontinuous Galerkin approximation methods by using simple projection operators. We will gain new insight into the increased necessity for stabilization using penalization techniques for Interior-Penalty Discontinuous Galerkin (IPDG) methods, [12, 24, 2, 20, 21], versus the weaker penalization requirements for Local Discontinuous Galerkin (LDG) methods, [11, 7]. The approach will also motivate the existence of penalty-free methods such as the Minimal Dissipation Local Discontinuous Galerkin (MD-LDG) method, [9], and the Dual-Wind Discontinuous Galerkin (DWDG) method, [19, 15]. This

2010 *Mathematics Subject Classification.* 65J02, 65N02, 35C02.
Key words and phrases. Discontinuous Galerkin methods; broken Sobolev spaces; distributional derivatives; numerical derivatives.

Published March 21, 2016.

paper is meant as a survey paper that can shed new light on known properties of DG methods for researchers familiar with the topic while also motivating DG approximation methods for researchers familiar with elliptic PDE theory.

Let $H^m(\Omega)$ denote the set of all functions in $L^2(\Omega)$ whose distributional derivatives up to order m are in $L^2(\Omega)$. Define $L^2_g(\Omega)$ as the set of all functions in $L^2(\Omega)$ that have (a well-defined) trace value g for $g \in L^2(\partial\Omega)$ and let $H^m_0(\Omega)$ denote the set of all functions in $H^m(\Omega)$ whose traces vanish up to order $(m-1)$ on $\partial\Omega$. For transparency, we will refer to Poisson's equation with Dirichlet boundary conditions as our prototypical elliptic boundary value problem:

$$-\Delta u = f, \quad \text{in } \Omega, \tag{1.1a}$$

$$u = g, \quad \text{on } \partial\Omega, \tag{1.1b}$$

where $\Omega \subset \mathbb{R}^d$, $d \geq 2$, is an open, convex, polygonal domain; $g \in H^{1/2}(\partial\Omega)$; and $\Delta \equiv \sum_{j=1}^d \frac{\partial^2}{\partial x_j{}^2}$ denotes the Laplacian operator. In weak form, (1.1) is equivalent to finding the unique function $u \in H^1(\Omega) \cap L^2_g(\partial\Omega)$ such that

$$\big(\nabla u, \nabla \varphi\big)_\Omega = \big(f, \varphi\big)_\Omega \quad \forall \varphi \in H^1_0(\Omega), \tag{1.2}$$

where $(\cdot,\cdot)_S$ denotes the L^2 inner product on S.

To approximate the solution to (1.2) using discontinuous Galerkin methods, we first need to formalize the problem using broken Sobolev spaces. Traditionally, this can be done using a piecewise gradient operator and controlling the jumps/discontinuities along interior faces/edges. Such a framework naturally leads to IPDG methods, and, using the unified framework in [1], can be extended to LDG methods. In this paper, we consider a different approach that is based on using distributional derivatives instead of the piecewise gradient operator. Following the example of the unified framework, we will consider the flux-based formulation of LDG-based methods and the Bassi-Rebay method, [3], where the derivative approximation is based on a well-chosen single-valued trace. However, we will formally consider the consequences of such a flux-based definition to build intuition and then rigorously consider the relationship of such a flux-based approach within the setting of distributional derivatives. The approach in this paper complements the Hybrid Discontinuous Galerkin method, [10], where the interior fluxes are directly approximated and used to determine the approximate gradient. We will see that the new interpretation of differential operators for broken Sobolev functions inherently motivates how to control the jumps/discontinuities along interior faces/edges.

The remainder of this article is organized as follows. In Section 2 we will introduce broken Sobolev spaces and some known results. The IPDG method will be introduced in Section 3 as well as several applications to broken Sobolev spaces. The section will provide an introduction to DG methods and further develop the background for broken Sobolev spaces. A heuristic framework will be considered in Section 4 that will motivate the analysis in Section 5 where we consider distributional derivatives and how they relate to the LDG method. We will derive a stability result for distributional derivatives with regard to broken Sobolev spaces and then use projection operators to formalize DG methods. We end the paper with some concluding remarks in Section 6. Note that the ideas presented in this paper are readily extended to a general class of second order elliptic PDEs, and the ideas formally motivate both the Weak Galerkin Finite Element Method introduced in [23] and DG methods that can be expressed using the DG finite element differential calculus proposed in [16].

2. Broken Sobolev spaces

We now recall the traditional extension of (1.2) to broken Sobolev spaces. First, we must introduce some (standard) definitions and notation. Then we will pose (1.1) in the setting of broken Sobolev spaces and discuss the major obstacle that must be overcome. The fundamental result will be an extension of Friedrichs inequality to broken Sobolev spaces, a result that will be reinterpreted in the context of distributional derivatives in Section 5 where we successfully extend (1.2) to broken Sobolev spaces while preserving existence and uniqueness properties.

2.1. **Definitions and notation.**

Let $\mathcal{T}_h$ denote a locally quasi-uniform and shape-regular simplicial triangulation of Ω (see [8, 5]). Let $\mathcal{E}_h$ denote the set of all (d-1)-dimensional simplices in the triangulation, $\mathcal{E}_h^I \subset \mathcal{E}_h$ denote the set of all interior (d-1)-dimensional simplices in the triangulation, and $\mathcal{E}_h^B \subset \mathcal{E}_h$ denote the set of all boundary (d-1)-dimensional simplices in the triangulation. Notationally, we set h_K as the diameter of the simplex $K \in \mathcal{T}_h$, h_e the diameter of the simplex $e \in \mathcal{E}_h$, and $h = \max_{K \in \mathcal{T}_h} h_K$.

We define the piecewise L^2 inner product with respect to the triangulation by

$$\big(v, w\big)_{\mathcal{T}_h} \equiv \sum_{K \in \mathcal{T}_h} \int_K v\, w\, dx,$$

and the piecewise L^2 inner product with respect to the collection of faces/edges $\mathcal{E}_h$ by

$$\langle v, w\rangle_{\mathcal{S}_h} \equiv \sum_{e \in \mathcal{S}_h} \int_e v\, w\, ds$$

for all $\mathcal{S}_h \subseteq \mathcal{E}_h$. We use bold-faced formatting to denote a vector-valued space; for example, we have $\mathbf{L}^2(\mathcal{T}_h) \equiv \big[L^2(\mathcal{T}_h)\big]^d$.

We define our broken Sobolev spaces with respect to the triangulation as

$$W^{m,p}(\mathcal{T}_h) \equiv \prod_{K \in \mathcal{T}_h} W^{m,p}(K)$$

for all integers $m \geq 0$ and real numbers $p \in [1, \infty]$, where

$$W^{m,p}(K) \equiv \{v \in L^p(K) \mid D^\alpha v \in L^p(K) \text{ for all multi-indices } \alpha \text{ with } |\alpha| \leq m\}.$$

The broken Sobolev space $W^{m,p}(\mathcal{T}_h)$ is a Banach space with respect to the norm

$$\|v\|_{W^{m,p}(\mathcal{T}_h)} \equiv \sum_{\substack{k=1 \\ |\alpha| \leq k}}^{m} \|D^\alpha v\|_{L^p(\mathcal{T}_h)}.$$

We denote the Hilbert space $W^{m,2}(\mathcal{T}_h)$ by $H^m(\mathcal{T}_h)$, where the inner product is defined by

$$(v, w)_{H^m(\mathcal{T}_h)} \equiv \sum_{\substack{k=1 \\ |\alpha| \leq k}}^{m} \big(D^\alpha v, D^\alpha w\big)_{\mathcal{T}_h}.$$

The space $W_0^{m,p}(\mathcal{T}_h)$ denotes the set of all functions in $W^{m,p}(\mathcal{T}_h)$ whose trace on $\partial\Omega$ vanishes up to order (m-1). We also let $C_c^m(\Omega) \subset W_0^{m,p}(\mathcal{T}_h)$ denote the set of all functions in $C^m(\Omega)$ with compact support, where $C^m(\Omega)$ denotes the set of all continuously differentiable functions up to order m.

Broken Sobolev functions inherently are allowed to be multi-valued along interior faces/edges, provided trace values exists on each simplex in the triangulation. Let $K^+, K^- \in \mathcal{T}_h$ and $e = \partial K^- \cap \partial K^+$. Without a loss of generality, we assume the global labeling number of K^+ is smaller than that of K^-. We then define the sided-flux values for v as

$$v^+\big|_e \equiv v\big|_{e\cap\partial K^+}, \quad v^-\big|_e \equiv v\big|_{e\cap\partial K^-},$$

where $v\big|_{\partial K}$ is understood to be the trace of v defined on $\overline{K}$. Suppose K is a boundary simplex. We extend the sided-flux definitions to the boundary of Ω by

$$v^\pm\big|_{\partial K\cap\partial\Omega} \equiv v\big|_{\partial K\cap\partial\Omega}.$$

The standard jump and average operators on $\mathcal{E}_h$ are defined by

$$[v] \equiv \begin{cases} v^- - v^+ & \text{on } \mathcal{E}_h^I \\ v & \text{on } \mathcal{E}_h^B, \end{cases} \qquad \{v\} \equiv \frac{v^- + v^+}{2},$$

respectively. We impose the convention that the outward normal vector on e, denoted by $\mathbf{n}$, is always given by the outward normal vector for K^-.

Remark 2.1. We have $W^{m,p}(\Omega) \subset W^{m,p}(\mathcal{T}_h)$, where functions in $W^{m,p}(\mathcal{T}_h)$ are allowed to have discontinuities along interior faces/edges. Thus, the underlying gradient operator for broken Sobolev functions, denoted ∇_h, is understood piecewise with respect to the triangulation so that when acting on $W^{m,p}(\Omega)$ we have $\nabla_h = \nabla$.

2.2. **Friedrichs inequality and interior penalization.** Using the above notation, we are ready to derive a "weak formulation" for (1.1) in the broken Sobolev space $H^2(\mathcal{T}_h)$. Let $v \in H^2(\mathcal{T}_h) \cap L^2_g(\Omega)$ such that $-\big(\Delta v, \varphi\big)_{\mathcal{T}_h} = \big(f, \varphi\big)_{\mathcal{T}_h}$ for all $\varphi \in H^1_0(\mathcal{T}_h)$. Then, we have

$$\begin{aligned} \big(f, \varphi\big)_{\mathcal{T}_h} &= -\big(\Delta v, \varphi\big)_{\mathcal{T}_h} \\ &= \big(\nabla v, \nabla\varphi\big)_{\mathcal{T}_h} - \big\langle [\nabla v], \{\varphi\}\mathbf{n}\big\rangle_{\mathcal{E}_h^I} - \big\langle \{\nabla v\}, [\varphi]\mathbf{n}\big\rangle_{\mathcal{E}_h^I} \end{aligned} \tag{2.1}$$

for all $\varphi \in H^1_0(\mathcal{T}_h)$. Suppose $f = 0$ and $g = 0$. Then, $u = 0$ is the unique solution to (1.2). In contrast, there exists infinitely many piecewise constant functions $v \in H^2(\mathcal{T}_h) \cap L^2_0(\Omega)$ that solve (2.1). One such example is the function

$$v(x) = \begin{cases} 0 & \text{if } x \text{ is in a boundary simplex,} \\ 1 & \text{if } x \text{ is in an interior simplex.} \end{cases}$$

Remark 2.2.

(a) The formulation (2.1) is equivalent to the problem: find $v \in H^2(\mathcal{T}_h) \cap L^2_g(\Omega)$ such that

$$\big(f, \varphi\big)_K = -\big(\Delta v, \varphi\big)_K = \big(\nabla v, \nabla\varphi\big)_K - \big\langle \nabla v, \varphi\,\vec{n}_K\big\rangle_{\partial K} \tag{2.2}$$

for all $\varphi \in H^1(K)$ for all $K \in \mathcal{T}_h$, where $\vec{n}_K$ denotes the unit outward normal vector for $\overline{K}$. Problem (2.1) is obtained by summing (2.2) over all simplices K.

(b) The solution is assumed to be in the space $H^2(\mathcal{T}_h)$ instead of $H^1(\mathcal{T}_h)$ in the above weak formulation. The interior trace values of $\nabla v \cdot \mathbf{n}$ are not defined for all functions in $H^1(\mathcal{T}_h)$, see [18].

The main issue in the previous example is that (2.1) has no mechanism to address the potential for discontinuities that characterizes functions in $H^2(\mathcal{T}_h)$. Neighboring simplices in $\mathcal{T}_h$ have no communication across their shared boundary interface. As a consequence, when the gradient operator is utilized in an entirely local, piecewise fashion with respect to $\mathcal{T}_h$, Poisson's problem does not have a unique solution in the broken Sobolev space $H^2(\mathcal{T}_h) \cap L^2_g(\Omega)$ for any given combination of boundary data g and source data f. This observation can be explained by the extension of Friedrichs inequality to broken Sobolev spaces:

Lemma 2.3 ([1, Lemma 2.1]). *There exists a positive constant C independent of h such that*

$$\|v\|^2_{L^2(\Omega)} \le C\Big(\big\|\nabla v\big\|^2_{L^2(\mathcal{T}_h)} + \sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|v\|^2_{L^2(e)} + \sum_{e\in\mathcal{E}_h^I} h_e^{-1}\|[v]\|^2_{L^2(e)}\Big) \tag{2.3}$$

for all $v \in H^1(\mathcal{T}_h)$.

Thus, the L^2 norm of $\nabla_h v$ is not sufficient to control the L^2 norm of v for all functions $v \in H^1(\mathcal{T}_h)$. In order to control the L^2 norm of a broken Sobolev function using the L^2 norm of the piecewise gradient, we also need to control the L^2 norm of the interior jumps of the function. Such an idea is the basis of IPDG methods as will be seen in the following section. A more detailed account of Poincarè and Friedrichs inequalities for broken Sobolev spaces can be found in [4].

3. Interior-penalty discontinuous Galerkin methods

In this section, we introduce the (symmetric) IPDG method and consider its application to broken Sobolev spaces. The main idea for DG methods is to approximate the solution to (1.1) with a function defined in a finite dimensional subspace of $H^2(\mathcal{T}_h)$.

3.1. **Notation.** For a fixed integer $r \ge 0$, define the standard DG finite element space $V_{h,r} \subset C^\infty(\mathcal{T}_h) \subset H^1(\mathcal{T}_h)$ by

$$V_{h,r} \equiv \prod_{K\in\mathcal{T}_h} \mathbb{P}_r(K),$$

where $\mathbb{P}_r(K)$ denotes the set of all polynomials on K with degree not exceeding r. The analogous vector-valued DG space is given by $\mathbf{V}_{h,r} \equiv \big[V_{h,r}\big]^d$. Most of the approximation methods below assume $r \ge 1$, although many DG methods have been extended to the case $r = 0$.

The key component for extending continuous results that hold in broken Sobolev spaces to discrete spaces for DG methods is a projection operator from $L^2(\Omega)$ to $V_{h,r}$. A natural choice is L^2 projection. We denote by $\mathcal{P}_{h,r} : L^2(\Omega) \to V_{h,r}$ the L^2 projection operator onto $V_{h,r}$, which is defined by

$$\big(\mathcal{P}_{h,r} v, \varphi_h\big)_{\mathcal{T}_h} = \big(v, \varphi_h\big)_{\mathcal{T}_h} \quad \forall \varphi_h \in V_{h,r}$$

for all $v \in L^2(\Omega)$. We also let $\vec{\mathcal{P}}_{h,r} : \mathbf{L}^2(\Omega) \to \mathbf{V}_{h,r}$ denote the L^2 projection operator onto $\mathbf{V}_{h,r}$.

3.2. **Application to broken Sobolev spaces.** We now attempt to extend (1.1) to the broken Sobolev space $H^2(\mathcal{T}_h)$ using ideas from the (symmetric) IPDG method. While the extension may still lack uniqueness when posed in the full broken Sobolev space $H^2(\mathcal{T}_h)$, the extension does guarantee uniqueness when restricting the solution and test functions to (finite dimensional) DG finite element spaces. Thus, the formulation is rich enough to construct sequences of functions in broken Sobolev spaces such that the sequences converge to the solution of the original problem (1.1). The major difference in the interior-penalty based formulation and the first attempt in Section 2 is the introduction of interior jump stabilization terms in the "weak representation" of (1.1) which allows for the application of Lemma 2.3.

Define the (symmetric) bilinear form $B_h : H^2(\mathcal{T}_h) \times H^2(\mathcal{T}_h) \to \mathbb{R}$ by

$$\begin{aligned} B_h(v,w) \equiv \big(\nabla v, \nabla w\big)_{\mathcal{T}_h} - \big\langle \big\{\nabla v\big\}, [w]\mathbf{n}\big\rangle_{\mathcal{E}_h} - \big\langle [v], \{\nabla w\} \cdot \mathbf{n}\big\rangle_{\mathcal{E}_h} \\ + \gamma \sum_{e\in\mathcal{E}_h} h_e^{-1} \big\langle [v], [w]\big\rangle_e \end{aligned}$$

for all $v, w \in H^2(\mathcal{T}_h)$ for some constant $\gamma > 0$ called the penalty parameter. Define an auxiliary norm $|||\cdot||| : H^2(\mathcal{T}_h) \to [0, \infty)$ by

$$|||v|||^2 \equiv \sum_{K\in\mathcal{T}_h} \|\nabla v\|_K^2 + \sum_{e\in\mathcal{E}_h} \Big(2\frac{\gamma}{h_e} \|[v]\|_{L^2(e)}^2 + \frac{h_e}{\gamma} \|\big\{\nabla v \cdot \mathbf{n}\big\}\|_{L^2(e)}^2 \Big)$$

for all $v \in H^2(\mathcal{T}_h)$, and, lastly, define the linear functional $F_h : H^2(\mathcal{T}_h) \to \mathbb{R}$ by

$$F_h(w) \equiv \big(f, w\big)_{\mathcal{T}_h} - \big\langle g, \nabla w \cdot \mathbf{n}\big\rangle_{\mathcal{E}_h^B} + \gamma \sum_{e\in\mathcal{E}_h^B} h_e^{-1} \big\langle g, w\big\rangle_e$$

for all $w \in H^2(\mathcal{T}_h)$. Notice that the above traces involving both function values and gradient values are well-defined for all functions in $H^2(\mathcal{T}_h)$. The (symmetric) IPDG method is defined by finding $u_h \in V_{h,r}$ that solves the finite dimensional problem $B_h(u_h, \varphi_h) = F_h(\varphi_h)$ for all $\varphi_h \in V_{h,r}$. Then, we have u_h approximates the solution to (1.2).

We can extend (1.1) to broken Sobolev spaces by considering the (continuous) problem: Find $u \in H^2(\mathcal{T}_h)$ such that $B_h(u, \varphi) = F_h(\varphi)$ for all $\varphi \in H^2(\mathcal{T}_h)$. Notice that the Dirichlet boundary condition is enforced using a penalization along the boundary. We could also consider the problem of finding $u \in H^2(\mathcal{T}_h) \cap L_g^2(\Omega)$ with the test functions $\varphi \in H^2(\mathcal{T}_h) \cap L_0^2(\Omega)$. In this case, some of the boundary face/edge terms would disappear in the definitions of B_h, $|||\cdot|||$, and F_h. However, when implementing the IPDG method using piecewise polynomials, the boundary conditions are typically enforced weakly using boundary penalization due to the fact the Dirichlet data cannot be satisfied when g is not a polynomial with degree less than or equal to r.

The following facts are well-documented in the DG literature:

Lemma 3.1.

(i) $|||\cdot|||$ *is a norm on* $H^2(\mathcal{T}_h)$.

(ii) *Consistency: If* $u \in H^2(\Omega) \cap L_g^2(\Omega)$ *is a solution to* (1.1)*, then* $B_h(u, \varphi) = F_h(\varphi)$ *for all* $\varphi \in H^2(\mathcal{T}_h)$.

(iii) *Continuity: There exists a positive constant* $C_{B,1}$ *such that, for all* $v, w \in H^2(\mathcal{T}_h)$*, there holds* $|B_h(v,w)| \le C_{B,1} |||v|||\, |||w|||$.

(iv) *Restricted Coercivity: There exists constants* $\gamma_0 \geq 0$ *and* $C_{B,2} = C_{B,2}(\gamma_0) \geq 0$*, both independent of* h*, such that, for all* $\gamma \geq \gamma_0$*, there holds* $B_h(v_h, v_h) \geq C_{B,2} |||v_h|||^2$ *for all* $v_h \in V_{h,r} \subset H^2(\mathcal{T}_h)$*.*

(v) *Equivalency: If* $|||v|||_B \equiv \sqrt{B_h(v,v)}$*, then*

$$\sqrt{C_{B,2}}|||v_h||| \leq |||v_h|||_B \leq \sqrt{C_{B,1}}|||v_h|||$$

for all $v_h \in V_{h,r} \subset H^2(\mathcal{T}_h)$*.*

Remark 3.2.

(a) By the Lax-Milgram theorem, the (symmetric) IPDG method has a unique solution. Moreover, if $u \in H^{s+1}(\Omega)$ is the solution to (1.1) with $1 \leq s \leq r$, then there exists a positive constant C independent of h such that $\|u - u_h\|_{L^2(\Omega)} \leq Ch^{s+1}\|D^{s+1}u\|_{L^2(\Omega)}$.

(b) Consistency only holds for $u \in H^2(\Omega)$. The derivation of the formulation for B_h is based on the assumption $u \in H^2(\Omega)$ and the test functions $w \in H^2(\mathcal{T}_h)$. The inconsistency with (2.1) comes from the lack of the term $\langle [\nabla v], \{\varphi\}\mathbf{n}\rangle_{\mathcal{E}_h^I}$, a term that disappears if $v \in H^2(\Omega)$. We have also added the term $\langle [v], \{\nabla w\} \cdot \mathbf{n}\rangle_{\mathcal{E}_h^I}$ in the formulation of B_h where we have again used the ansatz that for $v \in H^2(\Omega)$, the term is zero-valued.

(c) Coercivity only holds for functions in the discrete space $V_{h,r}$. Since the formulation is consistent, we have a solution exists in the space $H^2(\mathcal{T}_h)$. Thus, the lack of coercivity for the entire broken Sobolev space $H^2(\mathcal{T}_h)$ implies a potential loss of uniqueness. Notice that the lack of coercivity arises from the fact we cannot control the norm associated with $H^2(\mathcal{T}_h)$ using the norm $|||\cdot|||$. We will use distributional derivatives as a means to preserve uniqueness.

(d) All of the results are dependent upon choosing $\gamma > 0$ sufficiently large. In contrast, LDG methods only require the penalty constant $\gamma > 0$ to preserve existence and uniqueness. We will explore this result further in the following sections.

4. Formal motivation

In this section, we heuristically motivate the analytic results in the following section. We will consider a formal definition of a "broken" derivative operator for broken Sobolev functions that acts as an alternative to the piecewise gradient considered above in Sections 2 and 3. Thus, we formally let $\overline{\nabla}_h : H^1(\mathcal{T}_h) \to \mathbf{L}^2(\mathcal{T}_h)$ be characterized locally by

$$\big(\overline{\nabla}_h v, \boldsymbol{\varphi}\big)_K = \big\langle \{v\}, \boldsymbol{\varphi} \cdot \vec{n}_K \big\rangle_{\partial K} - \big(v, \nabla \cdot \boldsymbol{\varphi}\big)_K \quad \forall \boldsymbol{\varphi} \in \mathbf{H}^1(K) \tag{4.1}$$

for all $K \in \mathcal{T}_h$ for $v \in H^1(\mathcal{T}_h)$ arbitrary. Observe, if $v \in H^1(\Omega)$, then we would have $\overline{\nabla}_h v = \nabla_h v = \nabla v$.

Remark 4.1.

(a) As seen in [16], the operator $\overline{\nabla}_h$ forms the basis for LDG methods when we restrict the test functions and the range of the operator to be in the finite dimensional spaces $V_{h,r} \subset H^1(\mathcal{T}_h)$ and $\mathbf{V}_{h,r} \subset \mathbf{H}^1(\mathcal{T}_h)$, respectively. Thus, the results obtained in this section are a prelude to the stability results for LDG methods that we present in Section 5.

(b) If $[v] \neq 0$, then $\overline{\nabla}_h v \notin \mathbf{L}^1_{\mathrm{loc}}(\mathcal{T}_h) \supset \mathbf{L}^2(\mathcal{T}_h)$. In general, $\overline{\nabla}_h v$ must be understood as a (potentially singular) distribution as will be considered in the following section. However, many of the formal results obtained below based on the assumption $\overline{\nabla}_h v \in \mathbf{L}^2(\mathcal{T}_h)$ will be seen to carry over into the distributional setting and provide insight into LDG methods that are based on the restriction of (4.1) to test and trial functions in the finite dimensional space $V_{h,r} \subset H^1(\mathcal{T}_h)$.

(c) Since each simplex K is convex and polygonal, the trace operator $T : H^1(K) \to H^{1/2}(\partial K)$ has a bounded right inverse, [18]. Thus, if $v \in H^1(\mathcal{T}_h)$ and $\{v\} \in L^2(\partial K) \setminus H^{1/2}(\partial K)$ for some simplex K, it follows that there does not exist an element $w \in H^1(\mathcal{T}_h)$ such that $\overline{\nabla}_h v = \nabla_h w$. If $v \in H^1(\mathcal{T}_h) \setminus C^0(\Omega)$, then we expect such discrepancies/jumps to occur on interior faces/edges. Therefore, the "broken" derivative operator $\overline{\nabla}_h$ has the potential to address issues related to the discontinuities of a general broken Sobolev function unlike the standard piecewise gradient.

(d) For comparison, the piecewise gradient operator $\nabla_h : H^1(\mathcal{T}_h) \to \mathbf{L}^2(\mathcal{T}_h)$ is fully characterized locally by

$$\big(\nabla_h v, \boldsymbol{\varphi}\big)_K = \big\langle v, \boldsymbol{\varphi} \cdot \vec{n}_K \big\rangle_{\partial K} - \big(v, \nabla \cdot \boldsymbol{\varphi}\big)_K \quad \forall \boldsymbol{\varphi} \in \mathbf{H}^1(K) \tag{4.2}$$

for all $K \in \mathcal{T}_h$ for $v \in H^1(\mathcal{T}_h)$ arbitrary.

Observe that (4.1) is equivalent to

$$\big(\overline{\nabla}_h v, \boldsymbol{\varphi}\big)_{\mathcal{T}_h} = -\big(v, \nabla \cdot \boldsymbol{\varphi}\big)_{\mathcal{T}_h} + \big\langle \{v\}, [\boldsymbol{\varphi}] \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} + \big\langle v, \boldsymbol{\varphi} \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^B} \tag{4.3}$$

for all $\boldsymbol{\varphi} \in \mathbf{H}^1(\mathcal{T}_h)$ while summing (4.2) over all simplices K yields

$$\begin{aligned}\big(\nabla_h v, \boldsymbol{\varphi}\big)_{\mathcal{T}_h} &= -\big(v, \nabla \cdot \boldsymbol{\varphi}\big)_{\mathcal{T}_h} + \big\langle [v], \{\boldsymbol{\varphi}\} \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} + \big\langle \{v\}, [\boldsymbol{\varphi}] \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} \\ &\quad + \big\langle v, \boldsymbol{\varphi} \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^B}\end{aligned} \tag{4.4}$$

for all $\boldsymbol{\varphi} \in \mathbf{H}^1(\mathcal{T}_h)$. Thus, $\overline{\nabla}_h$ is formally related to ∇_h by

$$\big(\overline{\nabla}_h v, \boldsymbol{\varphi}\big)_{\mathcal{T}_h} = \big(\nabla_h v, \boldsymbol{\varphi}\big)_{\mathcal{T}_h} - \big\langle [v], \{\boldsymbol{\varphi}\} \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} \tag{4.5}$$

for all $\boldsymbol{\varphi} \in \mathbf{H}^1(\mathcal{T}_h)$, a relationship that will be made precise in Section 5. In contrast to the standard piecewise gradient operator ∇_h, we see that $\overline{\nabla}_h v$ implicitly incorporates jump information of the function v into its definition. The incorporation of jump information directly into the value of $\overline{\nabla}_h v$ is the precise reason that $\overline{\nabla}_h$ must be treated as a distribution. The definition for $\overline{\nabla}_h$ is not entirely local; instead, it uses the trace information of its nearest neighbors. Therefore, we expect the "broken" derivative operator $\overline{\nabla}_h$ to be more faithful to the global gradient operator with respect to jumps in a function, as will be seen in the following subsections.

4.1. **Friedrichs inequality.** We now derive a result analogous to Lemma 2.3 that is based on the formal operator $\overline{\nabla}_h$. We will see that we no longer require control of interior jumps in order to maintain stability, even when considering functions in the broken Sobolev space $H^1(\mathcal{T}_h)$. We also note that the bound will trivially hold in the case $\overline{\nabla}_h v \notin \mathbf{L}^2(\mathcal{T}_h)$. Such a result is the first step in explaining the much weaker stability requirements for LDG methods when compared to IPDG methods.

Let $v \in H^1(\mathcal{T}_h)$ and suppose $\overline{\nabla}_h v \in \mathbf{L}^2(\Omega)$. Define $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ by $-\Delta \psi = v$. Then, since Ω is convex, there exists a constant C_1 depending only on Ω such that $\|\psi\|_{H^2(\Omega)} \le C_1 \|v\|_{L^2(\Omega)}$, see [17]. Thus, we have

$$\begin{aligned}\|v\|_{L^2(\Omega)}^2 &= \big(v, v\big)_{\mathcal{T}_h} = \big(v, -\Delta\psi\big)_{\mathcal{T}_h} = -\big(v, \nabla \cdot \nabla\psi\big)_{\mathcal{T}_h} \\ &= \big(\overline{\nabla}_h v, \nabla\psi\big)_{\mathcal{T}_h} - \big\langle \{v\}, [\nabla\psi] \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} - \big\langle v, \nabla\psi \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^B}\end{aligned}$$

$$= \big(\overline{\nabla}_h v, \nabla\psi\big)_{\mathcal{T}_h} - \big\langle v, \nabla\psi\cdot\mathbf{n}\big\rangle_{\mathcal{E}_h^B},$$

where we have used the fact $\big[\nabla\psi\big] = 0$ in the interior of the domain. Observe that, by Young's inequality and the trace inequality,

$$\begin{aligned}
\big(\overline{\nabla}_h v, \nabla\psi\big)_{\mathcal{T}_h} &\le \|\overline{\nabla}_h v\|_{L^2(\mathcal{T}_h)} \|\nabla\psi\|_{L^2(\Omega)} \\
&\le \|\overline{\nabla}_h v\|_{L^2(\mathcal{T}_h)} \|\psi\|_{H^2(\Omega)} \\
&\le C_1\|\overline{\nabla}_h v\|_{L^2(\mathcal{T}_h)} \|v\|_{L^2(\Omega)} \\
&\le \frac{1}{2}C_1^2\|\overline{\nabla}_h v\|_{L^2(\mathcal{T}_h)}^2 + \frac{1}{2}\|v\|_{L^2(\Omega)}^2
\end{aligned}$$

and

$$\begin{aligned}
\big\langle v, \nabla\psi\cdot\mathbf{n}\big\rangle_{\mathcal{E}_h^B} &\le \sum_{e\in\mathcal{E}_h^B} \|v\|_{L^2(e)} \|\nabla\psi\|_{L^2(e)} \\
&\le \frac{\theta}{2}\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|v\|_{L^2(e)}^2 + \frac{1}{2\theta}\sum_{e\in\mathcal{E}_h^B} h_e\|\nabla\psi\|_{L^2(e)}^2
\end{aligned}$$

with

$$h_e\|\nabla\psi\|_{L^2(e)}^2 \le \widehat{C}\left(\|\nabla\psi\|_{L^2(K)}^2 + h_e^2\|D^2\psi\|_{L^2(K)}^2\right) \le C'\|\psi\|_{H^2(K)}^2,$$

for C' a constant such that $C' \ge \max\{\widehat{C}, \widehat{C}h_e^2\}$ for all $e\in\mathcal{E}_h^B$. Hence,

$$\begin{aligned}
\|v\|_{L^2(\Omega)}^2 \le \frac{1}{2}C_1^2\|\overline{\nabla}_h v\|_{L^2(\mathcal{T}_h)}^2 + \frac{1}{2}\|v\|_{L^2(\Omega)}^2 + \frac{\theta}{2}\sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|v\|_{L^2(e)}^2 \\
+ \frac{N}{2\theta}C'C_1\|v\|_{L^2(\Omega)}^2
\end{aligned}$$

for N the maximum number of boundary faces/edges a boundary simplex can have. Choosing $\theta = 2NC'C_1$ and assuming $h_e \le 1$ for all faces/edges, we have

$$\|v\|_{L^2(\Omega)}^2 \le C\Big(\|\overline{\nabla}_h v\|_{L^2(\mathcal{T}_h)}^2 + \sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|v\|_{L^2(e)}^2\Big),$$

for C independent of h_e for all $e\in\mathcal{E}_h$. Thus, we have proved the following result.

Lemma 4.2. *There exists a constant C depending only on Ω and the lower bounds for the interior angles and adjacent-edge ratios of the triangulation such that*

$$\|v\|_{L^2(\Omega)} \le C\|\overline{\nabla}_h v\|_{L^2(\mathcal{T}_h)}$$

for all $v\in H^1(\mathcal{T}_h)\cap L_0^2(\Omega)$ and

$$\|v\|_{L^2(\Omega)}^2 \le C\Big(\|\overline{\nabla}_h v\|_{L^2(\mathcal{T}_h)}^2 + \sum_{e\in\mathcal{E}_h^B} h_e^{-1}\|v\|_{L^2(e)}^2\Big)$$

for all $v\in H^1(\mathcal{T}_h)$.

Remark 4.3. Observe that Lemma 4.2 does not require interior jump penalization unlike the standard result given by Lemma 2.3 that is based on the piecewise gradient operator ∇_h. Using Theorem 5.3 below, we can also conclude that for $v\in H^1(\mathcal{T}_h)$, there holds $\overline{\nabla}_h v = \mathbf{0}$ if and only if v is a continuous, constant-valued function. Thus, we would expect DG methods based on $\overline{\nabla}_h$ to not have the same uniqueness/stability issues observed in Sections 2 and 3.

4.2. **Formal existence and uniqueness results for Poisson's equation.** We now formally extend (1.2) to broken Sobolev functions in a way that preserves existence and uniqueness properties. The analytic extension will be given in Section 5 when we reformulate (1.2) using results based on distributional derivatives.

First, we define an appropriate norm so that we can apply the Lax-Milgram theorem. Define $|||\cdot|||_{\overline{1}} : H^1(\mathcal{T}_h) \to [0, \infty]$ by

$$|||v|||_{\overline{1}} \equiv \big(\overline{\nabla}_h v, \overline{\nabla}_h v\big)_{\mathcal{T}_h}$$

for all $v \in H^1(\mathcal{T}_h)$. Then, from the broken Poincarè/Friedrichs inequalities above, we immediately have $|||\cdot|||_{\overline{1}}$ defines an (extended) norm on $H^1(\mathcal{T}_h) \cap L^2_0(\Omega)$ and a semi-norm on $H^1(\mathcal{T}_h)$.

Using the "broken" derivative operator $\overline{\nabla}_h$, we formally recast (1.2) as: Find $u \in H^1(\mathcal{T}_h) \cap L^2_g(\Omega)$ such that

$$\big(\overline{\nabla}_h u, \overline{\nabla}_h \varphi\big)_{\mathcal{T}_h} = \big(f, \varphi\big)_{\mathcal{T}_h} \tag{4.6}$$

for all $\varphi \in H^1_0(\mathcal{T}_h)$. Since $H^1(\mathcal{T}_h)$ is a Hilbert space when equipped with the (extended) norm $\|\cdot\|^2 \equiv \|\cdot\|^2_{L^2(\mathcal{T}_h)} + |||\cdot|||_{\overline{1}}$ and coercivity follows directly from the Friedrichs' inequalities found in Lemma 4.2, the formal problem has a unique solution by the Lax-Milgram theorem.

Remark 4.4.

(a) The above formulation does not explicitly consider jumps in u. However, the formal operator $\overline{\nabla}_h$ does implicitly consider jumps when compared to ∇_h.

(b) The above formulation is consistent with the traditional formulation in $H^1(\Omega)$ whenever the solution $u \in H^1(\Omega) \cap L^2_g(\Omega)$. In other words, if the solution is in $H^1(\Omega)$, then it is also a formal solution to (1.2).

(c) By using the broken derivative operator $\overline{\nabla}_h$ instead of the global gradient operator ∇, the formal problem (4.6) *posed in the broken Sobolev space $H^1(\mathcal{T}_h)$ preserves the form of the original problem* (1.2) *posed in the Sobolev space $H^1(\Omega)$.*

5. Distributional derivatives and local discontinuous Galerkin methods

We now make the observations in Section 4 rigorous by extending the above results to the proper setting based on distributional derivatives. Distributions will allow us to extend (1.2) to the broken Sobolev space $H^1(\mathcal{T}_h)$ in such a way that consistency and uniqueness regarding the solution $u \in H^1(\Omega) \cap L^2_g(\Omega)$ for (1.2) are preserved when testing by the larger class of functions in $H^1_0(\mathcal{T}_h)$. The new formulation will immediately rule out any functions in the space $\big(H^1(\mathcal{T}_h) \cap L^2_g(\Omega)\big) \setminus \big(H^1(\Omega) \cap L^2(\Omega)\big)$ from being a solution, yielding a stability result that can be utilized by DG methods.

We first consider the distributional derivative of a broken Sobolev function. Next we will prove a stability result that will be the key observation motivating the use of approximate distributional derivatives when formulating DG methods. We will end the section by discussing the extension of our continuous results in $H^1(\mathcal{T}_h)$ to the finite dimensional setting of DG methods.

Definition 5.1. Let $v \in W^{1,p}(\mathcal{T}_h)$. The *distributional derivative* of v is defined as the distribution $\mathcal{D}(v) : \mathbf{C}^\infty_c(\Omega) \to \mathbb{R}$ such that

$$< \mathcal{D}(v), \boldsymbol{\varphi} > \equiv -\big(v, \nabla \cdot \boldsymbol{\varphi}\big)_\Omega$$

for all $\boldsymbol{\varphi} \in \mathbf{C}_c^\infty(\Omega)$.

Remark 5.2.
(a) For comparison, the distribution associated with $\nabla_h v$ is defined by

$$< \nabla_h(v), \boldsymbol{\varphi} > \equiv -\big(v, \nabla \cdot \boldsymbol{\varphi}\big)_\Omega + \big\langle [v], \boldsymbol{\varphi} \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} = \big(\nabla v, \boldsymbol{\varphi}\big)_{\mathcal{T}_h}$$

for all $\boldsymbol{\varphi} \in \mathbf{C}_c^\infty(\Omega)$. If $v \in W^{1,p}(\Omega)$, then we have

$$< \nabla_h(v), \boldsymbol{\varphi} > = < \mathcal{D}(v), \boldsymbol{\varphi} > = < \nabla(v), \boldsymbol{\varphi} > \equiv -\big(v, \nabla \cdot \boldsymbol{\varphi}\big)_\Omega$$

for all $\boldsymbol{\varphi} \in \mathbf{C}_c^\infty(\Omega)$. The distributional derivative will be shown to ensure stability through the ansatz that the definition of the derivative itself will naturally account for jumps.

(b) The distributional derivative can be associated with a function in $\mathbf{L}^2(\Omega)$ if and only if $v \in H^1(\Omega)$.

(c) If we let $\boldsymbol{\varphi} \in \mathbf{C}_c^\infty(\Omega)$ in (4.3), then we have $\overline{\nabla}_h v$ formally satisfies

$$\big(\overline{\nabla}_h v, \boldsymbol{\varphi}\big)_{\mathcal{T}_h} = -\big(v, \nabla \cdot \boldsymbol{\varphi}\big)_\Omega$$

for all $v \in W^{1,p}(\mathcal{T}_h)$. Thus, the formal operator considered in Section 4 behaves like a regularized/idealized version of the distribution $\mathcal{D}$.

5.1. **A stability result for distributional derivatives.** A property of functions in $H^1(\Omega)$ is that $\|\nabla v\|_{L^2(\Omega)} = 0$ if and only if v is constant valued for all $v \in H^1(\Omega)$. We now extend the result to distributional derivatives over the larger function space $H^1(\mathcal{T}_h)$. As noted above, such a result does not hold when considering ∇_h. Thus, ∇_h has a larger nullspace than $\mathcal{D}$ yielding a loss of uniqueness when deriving a weak formulation of (1.1).

Theorem 5.3. *Let $v \in H^1(\mathcal{T}_h)$. Then v is a continuous, constant-valued function over Ω if and only if $\mathcal{D}(v) = \vec{0}$, where $\vec{0}$ is understood as the zero-valued function in the set of all bounded, linear functionals acting on $\mathbf{C}_c^\infty(\Omega)$.*

Proof. The forward direction is immediate. Let w be a constant valued function over Ω. Since $w \in H^1(\Omega)$, we have $\mathcal{D}(w) = \nabla w = \vec{0}$.

We now prove the reverse direction. We use a variational proof that is based on choosing appropriate test functions since such a proof is commonplace in the DG literature. A much shorter proof is given in Remark 5.4 that uses results from distribution theory.

Let $v \in H^1(\mathcal{T}_h)$ and suppose $\mathcal{D}(v) = \vec{0}$. We first show v is piecewise constant with respect to the triangulation. To this end, we will choose suitable test functions that ensure $\nabla_h v = \mathbf{0}$. Pick $K \in \mathcal{T}_h$. Let $\eta_{K,\epsilon} \in C^\infty(\overline{K})$ denote a "cutoff" function on K, [17], such that

$$\eta_{K,\epsilon}(x) = \begin{cases} 1 & \text{if } \operatorname{dist}(x, \partial K) \geq 2\,\epsilon, \\ 0 & \text{if } \operatorname{dist}(x, \partial K) \leq \epsilon \end{cases}$$

for $\epsilon > 0$ sufficiently small. Let $\rho_\epsilon \in C^\infty\big(B(0,\epsilon)\big)$ denote the standard mollifier function on $B(0,\epsilon)$, see [14]. Then, we have $\operatorname{supp}(\rho_\epsilon) \subset B(0,\epsilon)$. Define $K_\epsilon = \{x \in K \mid \operatorname{dist}(x, \partial K) \geq \epsilon\}$. Then, $\rho_\epsilon * \varphi \in C^\infty(K_\epsilon)$ for all $\varphi \in L^2(\Omega)$, where $*$ denotes the convolution operator. Notationally, we let $\varphi^\epsilon \equiv \rho_\epsilon * \varphi$. Analogous results hold for $\boldsymbol{\varphi} \in \mathbf{L}^2(\Omega)$ and $\boldsymbol{\varphi}^\epsilon \equiv \rho_\epsilon \mathbf{1} * \boldsymbol{\varphi}$.

By the definition of $\mathcal{D}(v)$, we have

$$0 =< \mathcal{D}(v), \boldsymbol{\varphi} >= -\big(v, \nabla \cdot \boldsymbol{\varphi}\big)_\Omega$$

for all $\boldsymbol{\varphi} \in \mathbf{C}_c^\infty(\Omega)$. Letting $\boldsymbol{\varphi} = \eta_{K,\epsilon} \nabla v^\epsilon \in \mathbf{C}_c^\infty(K)$, where the ϵ superscript denotes a mollified version of the function, we have $\operatorname{supp}(\boldsymbol{\varphi}) \subset K_\epsilon$ and

$$\begin{aligned} 0 &= -\big(v, \nabla \cdot (\eta_{K,\epsilon} \nabla v^\epsilon)\big)_K \\ &= \big(\nabla v, \eta_{K,\epsilon} \nabla v^\epsilon\big)_K - \big\langle v, \eta_{K,\epsilon} \nabla v^\epsilon \cdot \vec{n}_K \big\rangle_{\partial K} \\ &= \big(\nabla v, \eta_{K,\epsilon} \nabla v^\epsilon\big)_K \\ &= \big(\nabla v, \nabla v^\epsilon\big)_{K_{2\epsilon}} + \big(\nabla v, \eta_{K,\epsilon} \nabla v^\epsilon\big)_{K_\epsilon \setminus K_{2\epsilon}}. \end{aligned}$$

Since $\nabla v^\epsilon \to \nabla v$ in $L^2(K_\epsilon)$, $\nabla v \in L^2(K)$, and $\mu\left(K_\epsilon \setminus K_{2\epsilon}\right) \to 0$, we have $\nabla v\big|_K = \mathbf{0}$. Therefore, v is constant on K.

We now show $v \in C^0(\Omega)$. Pick $e \in \mathcal{E}_h^I$. Let $K^+, K^- \in \mathcal{T}_h$ such that $\partial K^+ \cap \partial K^- = e$, and let x^+, x^- denote the barycentric centers of K^+, K^-, respectively. Define $\Omega_e \subset \mathbb{R}^d$ as the convex domain corresponding to the simplex formed by the vertices of e and the nodes x^+ and x^-.

By the definition of $\mathcal{D}(v)$ and the fact that v is piecewise constant with respect to the triangulation, there holds

$$0 = -\big(v, \nabla \cdot \boldsymbol{\varphi}\big)_\Omega = \big(\nabla_h v, \boldsymbol{\varphi}\big)_{\mathcal{T}_h} - \big\langle [v], \boldsymbol{\varphi} \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} = -\big\langle [v], \boldsymbol{\varphi} \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I}$$

for all $\boldsymbol{\varphi} \in C_c^\infty(\Omega)$. Let $j \in \{1, 2, \dots, d\}$ and let $\boldsymbol{\varphi} = \varphi\, \mathbf{e}_j$ for $\varphi \in C_c^\infty(\Omega)$ and $\mathbf{e}_j$ the jth Cartesian basis vector. Then,

$$0 = -\big\langle [v], \varphi\, n_j \big\rangle_{\mathcal{E}_h^I}.$$

Let $\eta_{e,\epsilon} \in C^\infty(\overline{e})$ denote a "cutoff" function on e such that

$$\eta_{e,\epsilon}(x) = \begin{cases} 1 & \text{if } \operatorname{dist}(x, \partial e) \geq 2\,\epsilon, \\ 0 & \text{if } \operatorname{dist}(x, \partial e) \leq \epsilon \end{cases}$$

for $\epsilon > 0$ sufficiently small. Define $e_\epsilon = \{x \in e \mid \operatorname{dist}(x, \partial e) \geq \epsilon\}$. Let $\rho_\epsilon' \in C^\infty\big(B(0,\epsilon)\big)$ denote the standard mollifier function on $B(0,\epsilon)$ in $\mathbb{R}^{d-1}$. Choose $\varphi \in C_c^\infty(\Omega)$ such that $\varphi\big|_e = \eta_{e,\epsilon}\, \rho_\epsilon' * [v] \operatorname{sign} n_j$ and $\operatorname{supp} \varphi \subset \Omega_e$, where we have used the fact that $\mathbf{n}$ is piecewise constant on $\mathcal{E}_h^I$ and $\overline{e} \cap \mathring{e}_1 = \emptyset$ (in the topology for $\mathbb{R}^{d-1}$) for all $e_1 \in \mathcal{E}_h$ such that $e \neq e_1$. Then, we have

$$\big\langle [v], \eta_{e,\epsilon}\, \rho_\epsilon' * [v]\, |n_j| \big\rangle_e = 0$$

for all $j = 1, 2, \dots, d$. Letting $\epsilon \to 0$ and using the fact $[v] \in L^2(e)$, we have v is continuous across e. Therefore, v is a continuous, constant-valued function over Ω since v is both piecewise constant and continuous. □

Remark 5.4. Using results from distribution theory, we could identify the zero distribution uniquely (up to sets of measure zero) as the zero function $\mathbf{0} \in \mathbf{L}^1_{\text{loc}}(\Omega)$. Since $\mathbf{0} \in \mathbf{L}^2(\Omega)$, we have $v \in H^1(\Omega)$ with $\nabla v = \mathbf{0}$. Thus, v is a continuous, constant-valued function over Ω. Note that the underlying results for distributions in this short argument are typically proved using density arguments combined with mollifiers, as was done in the proof of the theorem.

We can now precisely extend (1.2) to broken Sobolev spaces by exploiting properties of distributional derivatives. We will have that the solution u to problem (1.2) is also a solution to the new generalized problem. Furthermore, the solution to the generalized problem will be unique due to Theorem 5.3. Thus, the new generalized problem posed in broken Sobolev spaces extends the existence and uniqueness properties of the weak formulation given by (1.2).

In general, the distribution $\mathcal{D}(v)$ associated with $v \in W^{1,p}(\mathcal{T}_h)$ cannot be associated with a function $\mathcal{D}v \in \mathbf{L}^1_{\mathrm{loc}}(\Omega)$ defined by

$$\big(\mathcal{D}v, \boldsymbol{\varphi}\big)_\Omega = < \mathcal{D}(v), \boldsymbol{\varphi} >$$

for all $\boldsymbol{\varphi} \in \mathbf{C}^\infty_c(\Omega)$. The problem is that for discontinuous functions in $W^{1,p}(\mathcal{T}_h)$, the distributional derivative acts like a delta function/measure which has no representation in $\mathbf{L}^1_{\mathrm{loc}}(\Omega)$, [14]. Such a negative result is the key to successfully extending (1.2) to broken Sobolev spaces without a loss of uniqueness.

Let $\varphi \in H^1_0(\mathcal{T}_h)$ and $\{\varphi_k\}_{k=1}^\infty \subset C^\infty_c(\Omega)$ such that φ_k converges weakly to φ with respect to the L^2 inner product, i.e., $\varphi_k \to^* \varphi$ (cf. [22, Theorem 6.32]). Then, if $u \in H^1(\Omega) \cap L^2_g(\Omega)$, we have

$$\big(f, \varphi_k\big)_\Omega = \big(\mathcal{D}u, \mathcal{D}\varphi_k\big)_\Omega = \big(\nabla u, \mathcal{D}\varphi_k\big)_\Omega = \big(\nabla u, \nabla\varphi_k\big)_\Omega$$

for all $k \geq 1$. Using the fact $\varphi_k \to^* \varphi$, we can conclude

$$\big(f, \varphi\big)_\Omega = \lim_{k\to\infty} \big(\nabla u, \nabla\varphi_k\big)_\Omega,$$

consistent with the weak formulation (1.2).

Without a loss of generality, we can assume $g = 0$ and $f \neq 0$. Then, if $u \in H^1_0(\mathcal{T}_h) \setminus H^1(\Omega)$, we trivially have there exists a simplex K such that

$$\big(f, u\big)_K \neq \big(\mathcal{D}u, \mathcal{D}u\big)_K,$$

where we have let $\varphi = u\chi_K$ for χ_K the characteristic function supported on K. Thus, the problem: Find $u \in H^1(\mathcal{T}_h) \cap L^2_g(\Omega)$ such that

$$\big(\mathcal{D}u, \mathcal{D}\varphi\big)_\Omega = \big(f, \varphi\big)_\Omega \tag{5.1}$$

for all $\varphi \in H^1_0(\mathcal{T}_h)$ successfully captures the essentials of (1.2) without introducing new solutions in the larger discontinuous space $H^1(\mathcal{T}_h) \cap L^2_g(\Omega)$, where the left-hand side of the formulation must be interpreted in an approximate sense based on density arguments.

5.2. **Application to local discontinuous Galerkin methods.** We first introduce a simple version of the LDG method that will form as the basis for the application of results from Sections 4 and 5. To this end, we introduce the central DG derivative operator that will be used in the formulation of the LDG method. Define the spaces

$$C^m(\mathcal{T}_h) \equiv \prod_{K\in\mathcal{T}_h} C^m\left(\overline{K}\right), \qquad \mathcal{V}_h \equiv W^{1,1}(\mathcal{T}_h) \cap C^0(\mathcal{T}_h)$$

for $m \geq 0$ an integer. Notice $V_{h,r} \subset \mathcal{V}_h$.

Definition 5.5. The *central DG derivative operator* $\overline{\nabla}_{h,r} : \mathcal{V}_h \to \mathbf{V}_{h,r}$ is defined locally by

$$\big(\overline{\nabla}_{h,r} v, \boldsymbol{\varphi}_h\big)_K \equiv \big\langle \{v\}, \boldsymbol{\varphi}_h \cdot \vec{n}_K \big\rangle_{\partial K} - \big(v, \nabla\cdot\boldsymbol{\varphi}_{\boldsymbol{h}}\big)_K \quad \forall \boldsymbol{\varphi}_h \in \mathbf{V}_{h,r}$$

for all $K \in \mathcal{T}_h$ for all $v \in \mathcal{V}_h$. Then, summing over all simplexes $K \in \mathcal{T}_h$, we have

$$\big(\overline{\nabla}_{h,r} v, \boldsymbol{\varphi}_h\big)_{\mathcal{T}_h} = \big\langle \{v\}, [\boldsymbol{\varphi}_h] \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} + \big\langle v, \boldsymbol{\varphi}_h \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^B} - \big(v, \nabla \cdot \boldsymbol{\varphi}_{\boldsymbol{h}}\big)_{\mathcal{T}_h}$$

for all $\boldsymbol{\varphi}_h \in \mathbf{V}_{h,r}$.

Notice that every function $v \in \mathcal{V}_h$ has a well-defined trace in $L^1(\partial K)$ and every function $\boldsymbol{\varphi}_h \in \mathbf{L}^\infty(\partial K)$ for all $K \in \mathcal{T}_h$. Thus, by the auxiliary result in [19, Appendix A], we have the central DG derivative operator is well-defined for all functions in $\mathcal{V}_h$. If $v \in H^1(\Omega)$, then we immediately have the following approximation result.

Theorem 5.6 ([16]). *For any $v \in \mathcal{V}_h \cap H^1(\Omega)$, $\overline{\nabla}_h v$ coincides with the L^2 projection of ∇v onto $\mathbf{V}_{h,r}$. We write $\overline{\nabla}_h v = \vec{\mathcal{P}}_{h,r} \nabla v$, where $\vec{\mathcal{P}}_{h,r}$ denotes the L^2 projection operator onto $\mathbf{V}_{h,r}$.*

Remark 5.7.
(a) Formally, we have $\overline{\nabla}_{h,r}$ is the L^2 projection of the operator $\overline{\nabla}_h$ considered in Section 4.
(b) For $v_h \in V_{h,r}$, we have $\overline{\nabla}_{h,r} v_h \in \mathbf{V}_{h,r-1}$ if v_h is also in $H^1(\Omega)$. Furthermore, there exist functions $v_h \in V_{h,r}$ such that $\overline{\nabla}_{h,r} v_h \notin \mathbf{V}_{h,r-1}$. For comparison, we have the piecewise gradient operator satisfies $\nabla_h \left(V_{h,r}\right) \subset \mathbf{V}_{h,r-1}$.
(c) If the function $v \in \mathcal{V}_h$ has given Dirichlet boundary data, we define the *central DG derivative operator with given Dirichlet boundary data* $\overline{\nabla}_{h,r}^g : \mathcal{V}_h \to \mathbf{V}_{h,r}$ by

$$\big(\overline{\nabla}_{h,r}^g v, \boldsymbol{\varphi}_h\big)_{\mathcal{T}_h} = \big\langle \{v\}, [\boldsymbol{\varphi}_h] \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^I} + \big\langle g, \boldsymbol{\varphi}_h \cdot \mathbf{n} \big\rangle_{\mathcal{E}_h^B} - \big(v, \nabla \cdot \boldsymbol{\varphi}_{\boldsymbol{h}}\big)_{\mathcal{T}_h}$$

for all $\boldsymbol{\varphi}_h \in \mathbf{V}_{h,r}$.

Finally, using the results found in [16], we can define the LDG method (in primal form) by: find $u_h \in V_{h,r}$ such that

$$\big(\overline{\nabla}_{h,r}^g u_h, \overline{\nabla}_{h,r}^0 \varphi_h\big)_{\mathcal{T}_h} + \gamma \big\langle [u_h], [\varphi_h] \big\rangle_{\mathcal{E}_h^I} + \gamma \big\langle u_h - g, \varphi_h \big\rangle_{\mathcal{E}_h^B} = \big(f, \varphi_h\big)_{\mathcal{T}_h} \tag{5.2}$$

for all $\varphi_h \in V_{h,r}$. Then, for $r \geq 1$ and $\gamma > 0$, the LDG method defined by (5.2) has a unique solution (see [11, 1]). Thus, the proposed method based on $\overline{\nabla}_{h,r}$ is stable as long as any positive penalty parameter is used. Unlike the (symmetric) IPDG method defined in Section 3 based on the piecewise gradient operator ∇_h, we do not need to require the penalty parameter be sufficiently large in order to guarantee uniqueness. If we further assume $\gamma = \alpha h^{-1}$ for an appropriate positive constant α chosen independently of h and $u \in H^{r+2}(\Omega)$ for u the solution to (1.2), then there holds

$$\|u - u_h\|_{L^2(\Omega)} + h \|\nabla u - \overline{\nabla}_{h,r}^g u_h\|_{L^2(\Omega)} \leq C h^{r+1} \|u\|_{H^{r+2}(\Omega)}$$

for some positive constant C independent of h (cf. [7]).

Remark 5.8. If we choose $X_{h,r} \subset V_{h,r} \cap H^1(\Omega)$ and add the restriction $u_h, \varphi_h \in X_{h,r}$, then (5.2) is equivalent to the finite element method with weakly enforced boundary conditions.

The formal stability results found in Section 4 hint that only boundary penalization should be sufficient for the LDG method. However, as seen in Figure 5.2, for particular choices of r and h it is possible to construct non-constant, discontinuous functions in the nullspace of $\overline{\nabla}_{h,r}^0$ yielding a loss of uniqueness for the LDG method when choosing $\gamma = 0$. Thus, the LDG method given by (5.2) does require

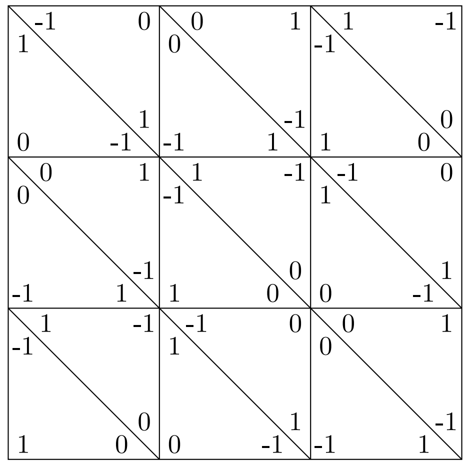

FIGURE 1. Example of a function in the nullspace of $\overline{\nabla}^0_{h,1}$, cf. [6]. The function is piecewise linear with the value of 1, -1, or 0 at each node in the mesh. Note that the function is not in the nullspace of $\overline{\nabla}_{h,1}$.

$\gamma > 0$ due to the loss of information when projecting onto the finite dimensional space. However, the MD-LDG method developed by Cockburn and Dong in [9] has extended the need for penalization to only a subset of the boundary faces/edges, consistent with the Friedrichs inequality given by Lemma 4.2. The main idea is to choose appropriate interior trace values in Definition 5.5 that eliminate the need to penalize along interior faces/edges. All of the interior trace values are in the set $\big\{\{v\}, v^+, v^-\big\}$. By incorporating two complementing DG derivative operators into the discretization of (1.1) and averaging, nonstandard LDG methods were developed in [19], called dual-wind DG methods, that eliminated the need for penalization along both interior and boundary faces/edges. Thus, the boundary conditions appear naturally in the method based on the choice of the underlying DG derivative operators. A short survey of when various DG methods do not require interior and/or boundary penalization can be found in Section 2.2 of [15].

We end this section by relating the LDG method to our results concerning distributional derivatives. We show that $\overline{\nabla}_{h,r}$ is in fact the L^2 projection of $\mathcal{D}(v)$ onto the finite dimensional space $\mathbf{V}_{h,r}$ for all $v \in \mathcal{V}_h$, when interpreted correctly. Proofs of the following results can all be found in [16].

Let $v \in \mathcal{V}_h$, $\mathcal{D}(v)$ denote the distributional derivative of v, and $\Xi \subset \Omega$ be a $(d$-$1)$-dimensional continuous and bounded surface. We define the *delta function* $\delta(\Xi, g, x)$ of variable strength supported on Ξ by (cf. [13])

$$< \delta(\Xi, g, x), \varphi > \equiv \int_{\Xi} g(s)\varphi(x(s))\, ds \tag{5.3}$$

for all $\varphi \in C^0(\Omega)$, where $x(s) \in \Xi$. We also extend the above definition to test functions from $V_{h,r}$ by

$$< \delta(\Xi, g, x), \varphi_h > \equiv \int_{\Xi} g(s)\big\{\varphi_h(x(s))\big\}\, ds \tag{5.4}$$

for all $\varphi_h \in V_{h,r}$. Using $\delta(\Xi, g, x)$, we have $\mathcal{D}(v)$ is characterized by

$$\mathcal{D}(v) = \sum_{K \in \mathcal{T}_h} \nabla v \, \chi_K - \sum_{e \in \mathcal{E}_h^I} \mathbf{n} \, \delta(e, [v], x) \quad \text{for a.e. } x \in \Omega, \tag{5.5}$$

where χ_K denotes the characteristic function supported on K. Thus, we are able to express the central DG derivative operator as the projection of the distributional derivative of v as follows:

Theorem 5.9. *For any $v \in \mathcal{V}_h$, $\overline{\nabla}_{h,r} v$ coincides with the L^2 projection of $\mathcal{D}(v)$ onto $\mathbf{V}_{h,r}$ in the sense that*

$$\left(\overline{\nabla}_{h,r} v, \boldsymbol{\varphi}_h\right)_{\mathcal{T}_h} = < \mathcal{D}(v), \boldsymbol{\varphi}_h > \tag{5.6}$$

for all $\boldsymbol{\varphi}_h \in \mathbf{V}_{h,r}$, where the right hand-side is understood according to (5.4)*. We write $\overline{\nabla}_{h,r} v = \vec{\mathcal{P}}_{h,r} \mathcal{D}(v)$.*

Using the above theorem, we can also formally compare the central DG derivative operator to the piecewise gradient operator using a lifting operator as follows:

Corollary 5.10. *Define the "lifting operator" $\mathcal{L}_{h,r} : L^1(\mathcal{E}_h) \to \mathbf{V}_{h,r}$ by*

$$\left(\mathcal{L}_{h,r} v, \boldsymbol{\varphi}_h\right)_{\mathcal{T}_h} \equiv \sum_{e \in \mathcal{E}_h^I} < \mathbf{n}(e) \delta(e, [v], \cdot), \boldsymbol{\varphi}_h > = \left\langle [v] \mathbf{n}, \left\{ \boldsymbol{\varphi}_h \right\} \right\rangle_{\mathcal{E}_h^I} \tag{5.7}$$

for all $\boldsymbol{\varphi}_h \in \mathbf{V}_{h,r}$. Then we have

$$\overline{\nabla}_{h,r} v = \nabla_h v - \mathcal{L}_{h,r} v \tag{5.8}$$

for all $v \in L^1(\mathcal{E}_h)$.

Thus, we see that jump stabilization is inherently enforced by the lifting operator that arises when comparing $\overline{\nabla}_{h,r}$ to the piecewise gradient operator ∇_h.

Finally, combining Theorems 5.6 and 5.9 and the well-known limiting characterization theorem of distributional derivatives (cf. [22, Theorem 6.32]), we obtain another characterization for our central DG derivative operator.

Theorem 5.11. *For any $v \in \mathcal{V}_h$, there exists a sequence of functions $\{v_j\}_{j \geq 1} \subset C_c^\infty(\Omega)$ such that*

(i) *$v_j \to v$ as $j \to \infty$ in $L^1(\Omega)$.*
(ii) *$\overline{\nabla}_{h,r} v_j \to \vec{\mathcal{P}}_{h,r} \mathcal{D}(v)$ as $j \to \infty$ weakly in $\mathbf{V}_{h,r}$ in the sense that*

$$\lim_{j \to \infty} \left(\overline{\nabla}_{h,r} v_j, \boldsymbol{\varphi}_h\right)_{\mathcal{T}_h} = \left(\vec{\mathcal{P}}_{h,r} \mathcal{D}(v), \boldsymbol{\varphi}_h\right)_{\mathcal{T}_h} \tag{5.9}$$

for all $\boldsymbol{\varphi}_h \in \mathbf{V}_{h,r}$.

Therefore, we have the LDG method is equivalent to a finite dimensional restriction of the distributional formulation of (1.2) over broken Sobolev spaces given by (5.1), where the Dirichlet boundary data is enforced weakly and interior penalization is added to ensure the preservation of uniqueness when projecting onto the finite dimensional solution space.

We make one more observation concerning the relationship between distributional derivatives and flux-based DG methods written in primal form. By adjusting the extension of the delta function of variable strength for test functions in $V_{h,r}$ given by (5.4), we are able to recover other DG derivative operators. The adjustment is simply choosing different flux values for φ_h such as φ_h^+ or φ_h^-. Thus, we can extend all of the above results and interpret a wide class of DG methods (including the MD-LDG method) as simply the L^2 projection of the broken Sobolev formulation (5.1) onto finite dimensional solution spaces where penalization is added on interior and/or boundary faces/edges only when necessary to ensure uniqueness and/or the enforcement of boundary conditions. We also see that DG methods that better approximate the distributional derivative instead of the piecewise gradient operator inherit better stability properties explaining the fundamental difference between the (symmetric) IPDG method and the LDG method with regards to stabilization.

6. Conclusion

The underlying goal of this paper was to extend uniqueness results for second order elliptic PDEs to broken Sobolev spaces as motivation for the design and understanding of various DG approximation methods. We showed that IPDG methods are based on a formulation that does not preserve uniqueness when the problem is posed in a broken Sobolev space using the natural choice of a piecewise gradient operator. In contrast, LDG methods are based on a formulation that guarantees uniqueness by considering distributional derivatives, derivative operators that naturally incorporate jump information. By combining uniqueness and consistency results for distributional derivatives, we see that not only does a solution exist in the broken Sobolev space setting, but the solution corresponds to the solution of the original PDE problem posed in (full) Sobolev spaces. Therefore, using distributional derivatives to extend PDE problems to broken Sobolev spaces provides the correct context for motivating a wide class of DG finite element approximation methods based on projections.

Another consequence of the above results is that flux-based DG methods can actually be considered conformal methods for the PDE problem (1.1) when the PDE problem is understood in terms of distributional derivatives. By posing the weak formulation (1.2) in the larger space of broken Sobolev functions, both continuous Galerkin methods (such as the finite element method) and DG methods can be discussed with respect to the same unifying PDE setting. The choice of (discrete) solution space $X_{h,r} \subset V_{h,r}$ and the choice of the DG derivative operator fully determine the approximation method and whether or not penalization is necessary for a given mesh $\mathcal{T}_h$. Therefore, the results found in this paper naturally motivate the use of the weak Galerkin finite element method ([23]) and the DG finite element differential calculus ([16]) for understanding and designing stable DG methods for approximating solutions to partial differential equations.

References

[1] D. Arnold, F. Brezzi, B. Cockburn, D. Marini; *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39:1749–1779, 2001.

[2] G. A. Baker; *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp., 31:45–59, 1977.

[3] F. Bassi, S.Rebay; *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations*, J. Comput. Phys., 131:267–279, 1997.

[4] S. C. Brenner; *Poincarè-Friedrichs inequalities for piecewise H^1 functions*, SIAM J. Numer. Anal., 41(1): 306–324, 2003.

[5] S. C. Brenner and L.R. Scott; *The Mathematical Theory of Finite Element Methods* (Third edition), Springer, 2008.

[6] F. Brezzi, M. Manzini, D. Marini, P. Pietra, A. Russo; *Discontinuous finite elements for Diffusion problems*, Francesco Brioschi (1824-1897) Convegno di Studi Matematici, October 22-23, 1997, Ist. Lomb. Acc. Sc. Lett., Incontro di studio N. 16, 197-217, 1999.
[7] P. Castillo, B. Cockburn, I. Perugia, D. Schötzau; *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems*, SIAM J. Numer. Anal., 38(5): 1676–1706, 2000.
[8] P. G. Ciarlet; *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
[9] B. Cockburn and B. Dong; *An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems*, J. Sci. Comput., 32(2): 233–262, 2007.
[10] B. Cockburn, J. Gopalakrishnan, R. Lazarov; *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal., 47: 1319–1365, 2009.
[11] B. Cockburn, C-W. Shu; *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal., 35(6): 2440–2463, 1998.
[12] J. Douglas Jr., T. Dupont; *Interior Penalty Procedures for Elliptic and Parabolic Galerkin Methods*, Lecture Notes in Phys. 58, Springer-Verlag, Berlin, 1976.
[13] B. Engquist, A. K. Tornberg, R. Tsai; *Discretization of dirac delta functions in level set methods*, J Comput. Phy., 207:28–51, 2005.
[14] L. C. Evans; *Partial Differential Equations, Graduate Studies in Mathematics*, 2nd Edition, American Mathematical Society, Providence, RI, 2010.
[15] W. Feng, T. Lewis, and S. Wise; *Discontinuous Galerkin derivative operators with applications to second order elliptic problems and stability*, Mathematical Meth. in App. Sciences, DOI: 10.1002/mma.3440, 2015.
[16] X. Feng, T. Lewis, M. Neilan; *Discontinuous Galerkin finite element differential calculus and applications to numerical solutions of linear and nonlinear partial differential equations*, accepted by J. Comput. Appl. Math., arXiv:1302.6984 [math.NA].
[17] D. Gilbarg, N. S. Trudinger; *Elliptic Partial Differential Equations of Second Order, Classics in Mathematics*, Springer-Verlag, Berlin, 2001, reprint of the 1998 edition.
[18] P. Grisvard; *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, MA, 1985.
[19] T. Lewis, M. Neilan; *Convergence analysis of a symmetric dual-wind discontinuous Galerkin method*, J. Sci. Comput., 59(3):602–625, 2014.
[20] B. Rivière, M.F. Wheeler, V. Girault; *Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. I.*, Comput. Geosci., 3(3–4): 337–360, 1999.
[21] B. Rivière, M. F. Wheeler, V. Girault; *A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems*, SIAM J. Numer. Anal. 39(3): 902–931, 2001.
[22] W. Rudin; *Functional Analysis*, Second Edition, McGraw-Hill, New York, 1991.
[23] J. Wang, X. Ye; *A weak Galerkin finite element method for second-order elliptic problems*, J. Comput. Appl. Math., 241:103–115, 2013.
[24] M. F. Wheeler; *An elliptic collocation-finite element method with interior penalties*, SIAM J. Numer. Anal., 15: 152–161, 1978.

Thomas Lewis
Department of Mathematics and Statistics, The University of North Carolina at Greensboro, Greensboro, NC 27412, USA
E-mail address: tllewis3@uncg.edu