Adv Comput Math (2016) 42:155–174
DOI 10.1007/s10444-015-9415-2

# A weak Galerkin finite element method for the stokes equations

**Junping Wang**[1] **· Xiu Ye**[2]

Received: 16 April 2014 / Accepted: 10 April 2015 /
Published online: 23 May 2015


**Abstract** This paper introduces a weak Galerkin (WG) finite element method for the Stokes equations in the primal velocity-pressure formulation. This WG method is equipped with stable finite elements consisting of usual polynomials of degree $k \geq 1$ for the velocity and polynomials of degree $k-1$ for the pressure, both are discontinuous. The velocity element is enhanced by polynomials of degree $k-1$ on the interface of the finite element partition. All the finite element functions are discontinuous for which the usual gradient and divergence operators are implemented as distributions in properly-defined spaces. Optimal-order error estimates are established for the corresponding numerical approximation in various norms. It must be emphasized that the WG finite element method is designed on finite element partitions consisting of arbitrary shape of polygons or polyhedra which are shape regular.

**Keywords** Weak Galerkin · Finite element methods · The stokes equations · Polyhedral meshes

Communicated by: Jinchao Xu

This research was supported in part by National Science Foundation Grant DMS-1115097.

The research of Wang was supported by the NSF IR/D program, while working at the Foundation. However, any opinion, finding, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

✉ Xiu Ye
xxye@ualr.edu

Junping Wang
jwang@nsf.gov

1 Department of Mathematics, University of Arkansas at Little Rock, Little Rock, AR 72204, USA

2 Division of Mathematical Sciences, National Science Foundation, Arlington, VA 22230, USA

**Mathematics Subject Classifications (2010)** Primary · 65N15 · 65N30 · 76D07 · Secondary · 35B45 · 35J50

## 1 Introduction

In this paper, we are concerned with the development of weak Galerkin (WG) finite element methods for the Stokes problem which seeks unknown functions $\mathbf{u}$ and $p$ satisfying

$$-\nabla\cdot A\nabla\mathbf{u}+\nabla p=\mathbf{f}\quad\text{in }\Omega,\tag{1.1}$$
$$\nabla\cdot\mathbf{u}=0\quad\text{in }\Omega,\tag{1.2}$$
$$\mathbf{u}=\mathbf{g}\quad\text{on }\partial\Omega,\tag{1.3}$$

where $\Omega$ is a polygonal or polyhedral domain in $\mathbb{R}^d$ $(d=2,3)$. $A$ is a symmetric $d\times d$ matrix-valued function in $\Omega$. Assume that there exist two positive numbers $\lambda_1,\lambda_2>0$ such that

$$\lambda_1\xi^t\xi\le\xi^tA\xi\le\lambda_2\xi^t\xi,\qquad\forall\xi\in\mathbb{R}^d.$$

Here $\xi$ is understood as a column vector and $\xi^t$ is the transpose of $\xi$.

The weak form in the primal velocity-pressure formulation for the Stokes problem (1.1)–(1.3) seeks $\mathbf{u}\in[H^1(\Omega)]^d$ and $p\in L_0^2(\Omega)$ satisfying $\mathbf{u}=\mathbf{g}$ on $\partial\Omega$ and

$$(A\nabla\mathbf{u},\nabla\mathbf{v})-(\nabla\cdot\mathbf{v},p)=(\mathbf{f},\mathbf{v}),\tag{1.4}$$
$$(\nabla\cdot\mathbf{u},q)=0,\tag{1.5}$$

for all $\mathbf{v}\in[H_0^1(\Omega)]^d$ and $q\in L_0^2(\Omega)$. The conforming finite element method for Eqs. 1.1–1.3 developed over the last several decades is based on the weak formulation (1.4)–(1.5) by constructing a pair of finite element spaces satisfying the *inf-sup* condition of Babuška [2] and Brezzi [5]. Readers are referred to [11] for specific examples and details in the classical finite element methods for the Stokes equations.

Weak Galerkin refers to a general finite element technique for partial differential equations in which differential operators are approximated by weak forms as distributions for generalized functions. Two key features in weak Galerkin methods are (1) the approximating functions are discontinuous, and (2) the usual partial derivatives are taken as distributions or approximations of distributions. The idea of weak Galerkin method was first introduced by one of the authors in the *International Conference on Applied Mathematics and Interdisciplinary Research* in Chern Institute of Mathematics at Nankai University in June 2011. The method was successfully applied to the second order elliptic equations in the primal formulation [15], and then subsequently to the mixed formulation in [16] for general finite element partitions of arbitrary shape (see also [13, 14]). The goal of this paper is to develop a weak Galerkin finite element method for Eqs. 1.1–1.3 by combining the ideas presented in [16] and [14] over partitions of general polygonal/polyhedral elements. This new finite element scheme is efficient and robust in that (1) it can be easily hybridized for variable reduction purpose in implementation, and (2) it allows the use of discontinuous approximating functions on finite element partitions of arbitrary shape.

In general, weak Galerkin finite element formulations for partial differential equations can be derived naturally by replacing usual derivatives by weakly-defined derivatives in the corresponding variational forms, with the option of adding a stabilization term to enforce a weak continuity of the approximating functions. For the Stokes problem (1.1)–(1.3) interpreted by the variational formulation (1.4)–(1.5), the two principle differential operators are the gradient and the divergence operator defined in the Sobolev space $[H^1(\Omega)]^d$. Formally, our weak Galerkin method for the Stokes problem would take the following form: Find $\mathbf{u}_h$ and $p_h$ from properly-defined finite element spaces satisfying

$$(A\nabla_w \mathbf{u}_h, \nabla_w \mathbf{v}) + s(\mathbf{u}_h, \mathbf{v}) - (\nabla_w \cdot \mathbf{v}, p_h) = (\mathbf{f}, \mathbf{v}), \tag{1.6}$$

$$(\nabla_w \cdot \mathbf{u}_h, q) = 0 \tag{1.7}$$

for all test functions $\mathbf{v}$ and $q$ in the test spaces. Here $\nabla_w$ is a discrete weak gradient and $\nabla_w\cdot$ is a discrete weak divergence operator to be detailed in Section 2. The bilinear form $s(\cdot,\cdot)$ in Eq. 1.6 is a parameter-free stabilizer that shall enforce a certain weak continuity for the underlying approximating functions. The use of totally discontinuous functions and weak derivatives in the WG formulation provides the numerical scheme with many nice features. First, the construction of stable elements for the Stokes equations under WG formulation is straightforward with standard polynomials. Secondly, the WG method allows the use of finite element partitions with arbitrary shape of polygons in 2D or polyhedra in 3D with certain shape regularity. The later property provides a convenient and useful flexibility in both numerical approximation and mesh generation. Thirdly, our WG formulation is parameter-free and has competitive number of unknowns since lower degree of polynomials are used on element boundaries, and the unknowns corresponding to the interior of each element can be eliminated from the system.

The research on finite element methods with polytopal meshes has been an active topic in recent years. The discontinuous Galerkin methods (see, for example, [1] and [9] and the references cited therein) have the capability of dealing with polytopal partitions. The mimetic finite difference method [7] and the virtual element method [3] are two other representatives along this line. The central issue in this study is the cross-element continuity enforcement (strongly or weakly) for necessary variables. Discontinuous Galerkin achieves this goal mostly through a stabilization for the jump on each interface, while the virtual element method extends from the boundary to the interior for each element. Both WG and HDG use intermediate functions on the interface to weakly "glue" different pieces together. Consequently, our WG finite element scheme has structural similarity with the HDG scheme as presented in [9], but they make use of different polynomial approximating spaces and utilize different stabilization techniques.

Throughout the paper, we will follow the usual notation for Sobolev spaces and norms [8]. For any open bounded domain $D \subset \mathbb{R}^d$, $d = 2, 3$, with Lipschitz continuous boundary, we use $\|\cdot\|_{s,D}$ and $|\cdot|_{s,D}$ to denote the norm and seminorms in the Sobolev space $H^s(D)$ for any $s \geq 0$, respectively. The inner product in $H^s(D)$ is denoted by $(\cdot,\cdot)_{s,D}$. More precisely, for any integer $s \geq 0$, the seminorm $|\cdot|_{s,D}$ is given by

$$|v|_{s,D} = \left( \sum_{|\alpha|=s} \int_D |\partial^\alpha v|^2 dD \right)^{\frac{1}{2}}$$

with the usual notation

$$\alpha = (\alpha_1, \ldots, \alpha_d), \quad |\alpha| = \alpha_1 + \ldots + \alpha_d, \quad \partial^\alpha = \prod_{j=1}^{d} \partial_{x_j}^{\alpha_j}.$$

The Sobolev norm $\|\cdot\|_{m,D}$ is given by

$$\|v\|_{m,D} = \left( \sum_{j=0}^{m} |v|_{j,D}^2 \right)^{\frac{1}{2}}.$$

The space $H^0(D)$ coincides with $L^2(D)$, for which the norm and the inner product are denoted by $\|\cdot\|_D$ and $(\cdot,\cdot)_D$, respectively. When $D = \Omega$, we shall drop the subscript $D$ in the norm and inner product notation.

The paper is organized as follows. In Section 2, we introduce two weak differential operators, called weak gradient and weak divergence, and their discrete analogues. In Section 3, we develop a weak Galerkin finite element scheme for the Stokes problem (1.1)–(1.2). In Section 4, we shall study the stability and solvability of the WG scheme. In particular, the usual *inf-sup* condition is established for the WG scheme. In Section 5, we shall derive an error equation for the WG approximations. Optimal-order error estimates for the WG finite element approximations are derived in Section 6 in virtually an $H^1$ norm for the velocity, and $L^2$ norm for both the velocity and the pressure. In Section 7, we make some concluding remarks by mentioning some outstanding issues for future consideration. Finally, we present some technical estimates for quantities related to the local $L^2$ projections into various finite element spaces in Appendix A.

## 2 Weak differential operators and their approximations

The key to weak Galerkin methods is the use of weak derivatives in the place of strong derivatives that define the weak formulation for the underlying partial differential equations. The two differential operators used in the weak formulation (1.4) and (1.5) are gradient and divergence. Thus, it is essential to introduce a weak version for both the gradient and the divergence operator. In [16], a weak divergence operator has been introduced and employed to the mixed formulation of second order elliptic equations. In [15] and [13], a weak gradient operator was introduced for scalar functions. Those weakly defined differential operators shall be employed to the Stokes problem (1.4)–(1.5) in a weak Galerkin approximation. For convenience, the rest of the section will review the definition for the weak gradient and the weak divergence, respectively.

Note that the weak gradient shall be applied to each component when the underlying function is vector-valued, as is the case for the Stokes problem.

## 2.1 Weak gradient and discrete weak gradient

Let $K$ be any polygonal or polyhedral domain with boundary $\partial K$. A weak vector-valued function on the region $K$ refers to a vector-valued function $\mathbf{v} = \{\mathbf{v}_0, \mathbf{v}_b\}$ such that $\mathbf{v}_0 \in [L^2(K)]^d$ and $\mathbf{v}_b \in [L^2(\partial K)]^d$. The first component $\mathbf{v}_0$ can be understood as the value of $\mathbf{v}$ in $K$, and the second component $\mathbf{v}_b$ represents $\mathbf{v}$ on the boundary of $K$. Note that $\mathbf{v}_b$ may not necessarily be related to the trace of $\mathbf{v}_0$ on $\partial K$ should a trace be well-defined. Denote by $\mathcal{V}(K)$ the space of weak functions on $K$; i.e.,

$$\mathcal{V}(K) = \left\{ \mathbf{v} = \{\mathbf{v}_0, \mathbf{v}_b\} :\ \mathbf{v}_0 \in \left[L^2(K)\right]^d,\ \mathbf{v}_b \in \left[L^2(\partial K)\right]^d \right\}. \tag{2.1}$$

The weak gradient operator is defined as follows.

**Definition 2.1** For any $\mathbf{v} \in \mathcal{V}(K)$, the weak gradient of $\mathbf{v}$ is defined as a linear functional $\nabla_w \mathbf{v}$ in the dual space of $[H^1(K)]^{d\times d}$ whose action on each $q \in [H^1(K)]^{d\times d}$ is given by

$$\langle \nabla_w \mathbf{v}, q\rangle_K = -(\mathbf{v}_0, \nabla\cdot q)_K + \langle \mathbf{v}_b, q\cdot\mathbf{n}\rangle_{\partial K}, \tag{2.2}$$

where $\mathbf{n}$ is the outward normal direction to $\partial K$, $(\mathbf{v}_0, \nabla\cdot q)_K = \int_K \mathbf{v}_0(\nabla\cdot q)dK$ is the inner product of $\mathbf{v}_0$ and $\nabla\cdot q$ in $[L^2(K)]^d$, and $\langle \mathbf{v}_b, q\cdot\mathbf{n}\rangle_{\partial K} = \int_{\partial K} \mathbf{v}_b\, q\cdot\mathbf{n}ds$ is the inner product of $\mathbf{v}_b$ and $q\cdot\mathbf{n}$ in $[L^2(\partial K)]^d$.

The Sobolev space $[H^1(K)]^d$ can be embedded into the space $\mathcal{V}(K)$ by an inclusion map $i_{\mathcal{V}} :\ [H^1(K)]^d \to \mathcal{V}(K)$ defined as follows

$$i_{\mathcal{V}}(\phi) = \{\phi|_K, \phi|_{\partial K}\}, \qquad \phi \in \left[H^1(K)\right]^d.$$

With the help of the inclusion map $i_{\mathcal{V}}$, the Sobolev space $[H^1(K)]^d$ can be viewed as a subspace of $\mathcal{V}(K)$ by identifying each $\phi \in [H^1(K)]^d$ with $i_{\mathcal{V}}(\phi)$.

Let $P_r(K)$ be the set of polynomials on $K$ with degree no more than $r$.

**Definition 2.2** The discrete weak gradient operator, denoted by $\nabla_{w,r,K}$, is defined as the unique polynomial $(\nabla_{w,r,K}\mathbf{v}) \in [P_r(K)]^{d\times d}$ satisfying the following equation,

$$(\nabla_{w,r,K}\mathbf{v}, q)_K = -(\mathbf{v}_0, \nabla\cdot q)_K + \langle \mathbf{v}_b, q\cdot\mathbf{n}\rangle_{\partial K}, \qquad \forall q \in [P_r(K)]^{d\times d}. \tag{2.3}$$

## 2.2 Weak divergence and discrete weak divergence

To define weak divergence, we require weak function $\mathbf{v} = \{\mathbf{v}_0, \mathbf{v}_b\}$ such that $\mathbf{v}_0 \in [L^2(K)]^d$ and $\mathbf{v}_b\cdot\mathbf{n} \in L^2(\partial K)$. Denote by $V(K)$ the space of weak vector-valued functions on $K$; i.e.,

$$V(K) = \left\{ \mathbf{v} = \{\mathbf{v}_0, \mathbf{v}_b\} :\ \mathbf{v}_0 \in \left[L^2(K)\right]^d,\ \mathbf{v}_b\cdot\mathbf{n} \in L^2(\partial K) \right\}. \tag{2.4}$$

A weak divergence operator can be defined as follows.

**Definition 2.3** For any $\mathbf{v} \in V(K)$, the weak divergence of $\mathbf{v}$ is defined as a linear functional $\nabla_w \cdot \mathbf{v}$ in the dual space of $H^1(K)$ whose action on each $\varphi \in H^1(K)$ is given by

$$\langle \nabla_w \cdot \mathbf{v}, \varphi \rangle_K = -(\mathbf{v}_0, \nabla\varphi)_K + \langle \mathbf{v}_b \cdot \mathbf{n}, \varphi \rangle_{\partial K}, \tag{2.5}$$

where $\mathbf{n}$ is the outward normal direction to $\partial K$, $(\mathbf{v}_0, \nabla\varphi)_K$ is the inner product of $\mathbf{v}_0$ and $\nabla\varphi$ in $L^2(K)$, and $\langle \mathbf{v}_b \cdot \mathbf{n}, \varphi \rangle_{\partial K}$ is the inner product of $\mathbf{v}_b \cdot \mathbf{n}$ and $\varphi$ in $L^2(\partial K)$.

The Sobolev space $[H^1(K)]^d$ can be embedded into the space $V(K)$ by an inclusion map $i_V : [H^1(K)]^d \to V(K)$ defined as follows

$$i_V(\phi) = \{\phi|_K, \ \phi|_{\partial K}\}, \qquad \phi \in \left[H^1(K)\right]^d.$$

**Definition 2.4** A discrete weak divergence operator, denoted by $\nabla_{w,r,K} \cdot$, is defined as the unique polynomial $(\nabla_{w,r,K} \cdot \mathbf{v}) \in P_r(K)$ that satisfies the following equation

$$(\nabla_{w,r,K} \cdot \mathbf{v}, \varphi)_K = -(\mathbf{v}_0, \nabla\varphi)_K + \langle \mathbf{v}_b \cdot \mathbf{n}, \varphi \rangle_{\partial K}, \qquad \forall \varphi \in P_r(K). \tag{2.6}$$

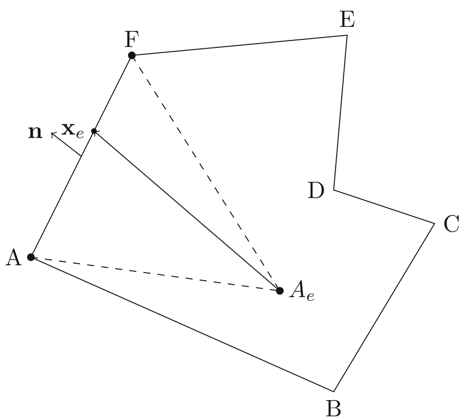
## 3 A weak Galerkin finite element scheme

Let $\mathcal{T}_h$ be a partition of the domain $\Omega$ with mesh size $h$ that consists of arbitrary polygons/polyhedra. Denote by $\mathcal{E}_h$ the set of all edges or flat faces in $\mathcal{T}_h$, and let $\mathcal{E}_h^0 = \mathcal{E}_h \backslash \partial\Omega$ be the set of all interior edges or flat faces. For every element $T \in \mathcal{T}_h$, we denote by $|T|$ the area or volume of $T$ and by $h_T$ its diameter. Similarly, we denote by $|e|$ the length or area of $e$ and by $h_e$ the diameter of edge or flat face $e \in \mathcal{E}_h$. We also set as usual the mesh size of $\mathcal{T}_h$ by

$$h = \max_{T \in \mathcal{T}_h} h_T.$$

All the elements in $\mathcal{T}_h$ are assumed to be closed and simply connected polygons or polyhedra, see Fig. 1.

**Fig. 1** Depiction of a shape-regular polygonal element $ABCDEFA$



Assume that the partition $\mathcal{T}_h$ is shape regular in the sense that the following conditions A1–A4 are satisfied [13, 16].

A1: Assume that there exist two positive constants $\varrho_v$ and $\varrho_e$ such that for every element $T \in \mathcal{T}_h$ we have

$$\varrho_v h_T^d \le |T|, \qquad \varrho_e h_e^{d-1} \le |e| \tag{3.1}$$

for all edges or flat faces of $T$.

A2: Assume that there exists a positive constant $\kappa$ such that for every element $T \in \mathcal{T}_h$ we have

$$\kappa h_T \le h_e \tag{3.2}$$

for all edges or flat faces $e$ of $T$.

A3: Assume that the mesh edges or faces are flat. We further assume that for every $T \in \mathcal{T}_h$, and for every edge/face $e \in \partial T$, there exists a pyramid $P(e, T, A_e)$ contained in $T$ such that its base is identical with $e$, its apex is $A_e \in T$, and its height is proportional to $h_T$ with a proportionality constant $\sigma_e$ bounded away from a fixed positive number $\sigma^*$ from below. In other words, the height of the pyramid is given by $\sigma_e h_T$ such that $\sigma_e \ge \sigma^* > 0$. The pyramid is also assumed to stand up above the base $e$ in the sense that the angle between the vector $\mathbf{x}_e - A_e$, for any $x_e \in e$, and the outward normal direction of $e$ (i.e., the vector $\mathbf{n}$ in Fig. 1) is strictly acute by falling into an interval $[0, \theta_0]$ with $\theta_0 < \frac{\pi}{2}$.

A4: Assume that each $T \in \mathcal{T}_h$ has a circumscribed simplex $S(T)$ that is shape regular and has a diameter $h_{S(T)}$ proportional to the diameter of $T$; i.e., $h_{S(T)} \le \gamma_* h_T$ with a constant $\gamma_*$ independent of $T$. Furthermore, assume that each circumscribed simplex $S(T)$ intersects with only a fixed and small number of such simplices for all other elements $T \in \mathcal{T}_h$.

Interested readers are referred to [7] for a similar shape regularity assumption for the mimetic finite difference method. In Fig. 1, we illustrate a polygonal element that is shape regular in the WG setting. The shape regularity assumption is essential for deriving error estimates for locally defined $L^2$ projection operators to be detailed in coming sections.

For any integer $k \ge 1$, we define a weak Galerkin finite element space for the velocity variable as follows

$$V_h = \left\{ \mathbf{v} = \{\mathbf{v}_0, \mathbf{v}_b\} : \{\mathbf{v}_0, \mathbf{v}_b\}|_T \in [P_k(T)]^d \times \left[P_{k-1}(e)\right]^d, \; e \subset \partial T \right\}.$$

We would like to emphasize that there is only a single value $\mathbf{v}_b$ defined on each edge $e \in \mathcal{E}_h$. For the pressure variable, we have the following finite element space

$$W_h = \left\{ q : q \in L_0^2(\Omega), \; q|_T \in P_{k-1}(T) \right\}.$$

Denote by $V_h^0$ the subspace of $V_h$ consisting of discrete weak functions with vanishing boundary value; i.e.,

$$V_h^0 = \{\mathbf{v} = \{\mathbf{v}_0, \mathbf{v}_b\} \in V_h, \mathbf{v}_b = 0 \text{ on } \partial\Omega\}.$$

The discrete weak gradient $\nabla_{w,k-1}$ and the discrete weak divergence $(\nabla_{w,k-1}\cdot)$ on the finite element space $V_h$ can be computed by using Eqs. 2.3 and 2.6 on each element

$T$, respectively. More precisely, they are given by

$$(\nabla_{w,k-1}\mathbf{v})|_T = \nabla_{w,k-1,T}(\mathbf{v}|_T), \qquad \forall\, \mathbf{v} \in V_h,$$
$$(\nabla_{w,k-1}\cdot\mathbf{v})|_T = \nabla_{w,k-1,T}\cdot(\mathbf{v}|_T), \qquad \forall\, \mathbf{v} \in V_h.$$

For simplicity of notation, from now on we shall drop the subscript $k-1$ in the notation $\nabla_{w,k-1}$ and $(\nabla_{w,k-1}\cdot)$ for the discrete weak gradient and the discrete weak divergence. The usual $L^2$ inner product can be written locally on each element as follows

$$(\nabla_w\mathbf{v},\ \nabla_w\mathbf{w}) = \sum_{T\in\mathcal{T}_h}(\nabla_w\mathbf{v},\ \nabla_w\mathbf{w})_T,$$
$$(\nabla_w\cdot\mathbf{v},\ q) = \sum_{T\in\mathcal{T}_h}(\nabla_w\cdot\mathbf{v},\ q)_T.$$

Denote by $Q_0$ the $L^2$ projection operator from $[L^2(T)]^d$ onto $[P_k(T)]^d$. For each edge/face $e\in\mathcal{E}_h$, denote by $Q_b$ the $L^2$ projection from $[L^2(e)]^d$ onto $[P_{k-1}(e)]^d$. We shall combine $Q_0$ with $Q_b$ by writing $Q_h=\{Q_0,Q_b\}$.

We are now in a position to describe a weak Galerkin finite element scheme for the Stokes Eqs. 1.1–1.3. To this end, we first introduce three bilinear forms as follows

$$s(\mathbf{v},\ \mathbf{w}) = \sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_b\mathbf{v}_0-\mathbf{v}_b,\ \ Q_b\mathbf{w}_0-\mathbf{w}_b\rangle_{\partial T},$$
$$a(\mathbf{v},\ \mathbf{w}) = (A\nabla_w\mathbf{v},\ \nabla_w\mathbf{w})+s(\mathbf{v},\mathbf{w}),$$
$$b(\mathbf{v},\ q) = (\nabla_w\cdot\mathbf{v},\ q).$$

**Weak Galerkin Algorithm 1** *A numerical approximation for Eqs. 1.1–1.3 can be obtained by seeking* $\boldsymbol{u}_h=\{\boldsymbol{u}_0,\boldsymbol{u}_b\}\in V_h$ *and* $p_h\in W_h$ *such that* $\boldsymbol{u}_b=Q_b\boldsymbol{g}$ *on* $\partial\Omega$ *and*

$$a(\boldsymbol{u}_h,\ \boldsymbol{v})-b(\boldsymbol{v},\ p_h) = (f,\ \boldsymbol{v}_0), \tag{3.3}$$
$$b(\boldsymbol{u}_h,\ q) = 0, \tag{3.4}$$

*for all* $\boldsymbol{v}=\{\boldsymbol{v}_0,\boldsymbol{v}_b\}\in V_h^0$ *and* $q\in W_h$.

## 4 Stability and solvability

The WG finite element scheme (3.3)–(3.4) is a typical saddle-point problem which can be analyzed by using the well known theory developed by Babuška [2] and Brezzi [5]. The core of the theory is to verify two properties: (1) boundedness and a certain coercivity for the bilinear form $a(\cdot,\cdot)$, and (2) boundedness and *inf-sup* condition for the bilinear form $b(\cdot,\cdot)$.

The finite element space $V_h^0$ is a normed linear space with a triple-bar norm given by

$$|\!|\!|\mathbf{v}|\!|\!|^2=\sum_{T\in\mathcal{T}_h}\|\nabla_w\mathbf{v}\|_T^2+\sum_{T\in\mathcal{T}_h}h_T^{-1}\|Q_b\mathbf{v}_0-\mathbf{v}_b\|_{\partial T}^2. \tag{4.1}$$

We claim that $|||\cdot|||$ indeed provides a norm in $V_h^0$. For simplicity, we shall only verify the positive length property for $|||\cdot|||$. Assume that $|||\mathbf{v}||| = 0$ for some $\mathbf{v} \in V_h^0$. It follows that

$$0 = (\nabla_w \mathbf{v}, \nabla_w \mathbf{v}) + \sum_{T \in \mathcal{T}_h} h_T^{-1} \langle Q_b \mathbf{v}_0 - \mathbf{v}_b,\ Q_b \mathbf{v}_0 - \mathbf{v}_b \rangle_{\partial T},$$

which implies that $\nabla_w \mathbf{v} = 0$ on each element $T$ and $Q_b \mathbf{v}_0 = \mathbf{v}_b$ on $\partial T$. Thus, we have from the definition (2.3) that for any $\tau \in [P_{k-1}(T)]^{d\times d}$

$$\begin{aligned}
0 &= (\nabla_w \mathbf{v}, \tau)_T \\
&= -(\mathbf{v}_0, \nabla \cdot \tau)_T + \langle \mathbf{v}_b, \tau \cdot \mathbf{n} \rangle_{\partial T} \\
&= (\nabla \mathbf{v}_0, \tau)_T - \langle \mathbf{v}_0 - \mathbf{v}_b, \tau \cdot \mathbf{n} \rangle_{\partial T} \\
&= (\nabla \mathbf{v}_0, \tau)_T - \langle Q_b \mathbf{v}_0 - \mathbf{v}_b, \tau \cdot \mathbf{n} \rangle_{\partial T} \\
&= (\nabla \mathbf{v}_0, \tau)_T.
\end{aligned}$$

Letting $\tau = \nabla \mathbf{v}_0$ in the equation above yields $\nabla \mathbf{v}_0 = 0$ on $T \in \mathcal{T}_h$. It follows that $\mathbf{v}_0 = const$ on every $T \in \mathcal{T}_h$. This, together with the fact that $Q_b \mathbf{v}_0 = \mathbf{v}_b$ on $\partial T$ and $\mathbf{v}_b = 0$ on $\partial \Omega$, implies that $\mathbf{v}_0 = 0$ and $\mathbf{v}_b = 0$.

Note that $|||\cdot|||$ defines only a semi-norm in $V_h$. It is not hard to see that $a(\mathbf{v},\ \mathbf{v}) = |||\mathbf{v}|||^2$ for any $\mathbf{v} \in V_h$. In fact, the trip-bar norm is equivalent to the standard $H^1$-norm, but was defined for weak finite element functions. It follows from the definition of $|||\cdot|||$ and the usual Cauchy-Schwarz inequality that the following boundedness and coercivity hold true for the bilinear form $a(\cdot,\cdot)$.

**Lemma 4.1** *For any $\boldsymbol{v}, \boldsymbol{w} \in V_h^0$, we have*

$$|a(\boldsymbol{v}, \boldsymbol{w})| \le |||\boldsymbol{v}|||\,|||\boldsymbol{w}|||, \tag{4.2}$$

$$a(\boldsymbol{v}, \boldsymbol{v}) = |||\boldsymbol{v}|||^2. \tag{4.3}$$

In addition to the projection $Q_h = \{Q_0, Q_b\}$ defined in the previous section, let $\mathbb{Q}_h$ and $\mathbf{Q}_h$ be two local $L^2$ projections onto $P_{k-1}(T)$ and $[P_{k-1}(T)]^{d\times d}$, respectively.

**Lemma 4.2** *The projection operators $Q_h$, $\boldsymbol{Q}_h$, and $\mathbb{Q}_h$ satisfy the following commutative properties*

$$\nabla_w (Q_h \boldsymbol{v}) = \boldsymbol{Q}_h(\nabla \boldsymbol{v}), \qquad \forall\, \boldsymbol{v} \in \left[H^1(\Omega)\right]^d, \tag{4.4}$$

$$\nabla_w \cdot (Q_h \boldsymbol{v}) = \mathbb{Q}_h(\nabla \cdot \boldsymbol{v}), \qquad \forall\, \boldsymbol{v} \in H(div, \Omega). \tag{4.5}$$

*Proof* Using Eq. 2.3, we have

$$(\nabla_w (Q_h \mathbf{v}),\ q)_T = -(Q_0 \mathbf{v},\ \nabla \cdot q)_T + \langle Q_b \mathbf{v},\ q \cdot \mathbf{n} \rangle_{\partial T}$$

for all $q \in [P_{k-1}(T)]^{d\times d}$. Next, we use the definition of $Q_h$ and $\mathbf{Q}_h$ and the usual integration by parts to obtain

$$\begin{aligned}-(Q_0\mathbf{v},\ \nabla\cdot q)_T + \langle Q_b\mathbf{v},\ q\cdot\mathbf{n}\rangle_{\partial T} &= -(\mathbf{v},\ \nabla\cdot q)_T + \langle \mathbf{v},\ q\cdot\mathbf{n}\rangle_{\partial T}\\ &= (\nabla\mathbf{v},\ q)\\ &= (\mathbf{Q}_h(\nabla\mathbf{v}),\ q).\end{aligned}$$

Thus,

$$(\nabla_w(Q_h\mathbf{v}),\ q)_T = (\mathbf{Q}_h(\nabla\mathbf{v}),\ q), \qquad \forall\, q\in[P_{k-1}(T)]^{d\times d},$$

which verifies the identity (4.4).

To verify (4.5), we use the discrete weak divergence (2.6) to obtain

$$(\nabla_w\cdot(Q_h\mathbf{v}),\ \varphi)_T = -(Q_0\mathbf{v},\ \nabla\varphi)_T + \langle Q_b\mathbf{v}\cdot\mathbf{n},\ \varphi\rangle_{\partial T}$$

for all $\varphi \in P_{k-1}(T)$. Next, we use the definition of $Q_h$ and $\mathbb{Q}_h$ and the usual integration by parts to arrive at

$$\begin{aligned}-(Q_0\mathbf{v},\ \nabla\varphi)_T + \langle Q_b\mathbf{v}\cdot\mathbf{n},\ \varphi\rangle_{\partial T} &= -(\mathbf{v},\ \nabla\varphi)_T + \langle \mathbf{v}\cdot\mathbf{n},\ \varphi\rangle_{\partial T}\\ &= (\nabla\cdot\mathbf{v},\ \varphi)_T\\ &= (\mathbb{Q}_h(\nabla\cdot\mathbf{v}),\ \varphi)_T.\end{aligned}$$

It follows that

$$(\nabla_w\cdot(Q_h\mathbf{v}),\ \varphi)_T = (\mathbb{Q}_h(\nabla\cdot\mathbf{v}),\ \varphi)_T, \qquad \forall\,\varphi\in P_{k-1}(T).$$

This completes the proof of Eq. 4.5, and hence the lemma. □

For the bilinear form $b(\cdot,\cdot)$, we have the following result on the *inf-sup* condition.

**Lemma 4.3** *There exists a positive constant $\beta$ independent of $h$ such that*

$$\sup_{\mathbf{v}\in V_h^0}\frac{b(\mathbf{v},\rho)}{|||\mathbf{v}|||} \ge \beta\|\rho\| \tag{4.6}$$

*for all $\rho\in W_h$.*

*Proof* For any given $\rho\in W_h\subset L_0^2(\Omega)$, it is well known [4, 6, 10–12] that there exists a vector-valued function $\tilde{\mathbf{v}}\in[H_0^1(\Omega)]^d$ such that

$$\frac{(\nabla\cdot\tilde{\mathbf{v}},\rho)}{\|\tilde{\mathbf{v}}\|_1} \ge C\|\rho\|, \tag{4.7}$$

where $C>0$ is a constant depending only on the domain $\Omega$. By setting $\mathbf{v}=Q_h\tilde{\mathbf{v}}\in V_h$, we claim that the following holds true

$$|||\mathbf{v}||| \le C_0\|\tilde{\mathbf{v}}\|_1 \tag{4.8}$$

for some constant $C_0$. To this end, we use Eq. 4.4 to obtain

$$\sum_{T\in\mathcal{T}_h}\|\nabla_w\mathbf{v}\|_T^2 = \sum_{T\in\mathcal{T}_h}\|\nabla_w(Q_h\tilde{\mathbf{v}})\|_T^2 = \sum_{T\in\mathcal{T}_h}\|\mathbf{Q}_h\nabla\tilde{\mathbf{v}}\|_T^2 \le \|\nabla\tilde{\mathbf{v}}\|^2. \tag{4.9}$$

Next, we use Eqs. A.4, A.1, and the definition of $Q_b$ to obtain

$$\begin{aligned}\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_b\mathbf{v}_0-\mathbf{v}_b\|_{\partial T}^2 &= \sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_b(Q_0\tilde{\mathbf{v}})-Q_b\tilde{\mathbf{v}}\|_{\partial T}^2\\ &= \sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_b(Q_0\tilde{\mathbf{v}}-\tilde{\mathbf{v}})\|_{\partial T}^2\\ &\le \sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0\tilde{\mathbf{v}}-\tilde{\mathbf{v}}\|_{\partial T}^2\\ &\le C\sum_{T\in\mathcal{T}_h}\left(h_T^{-2}\|Q_0\tilde{\mathbf{v}}-\tilde{\mathbf{v}}\|_T^2+\|\nabla(Q_0\tilde{\mathbf{v}}-\tilde{\mathbf{v}})\|_T^2\right)\\ &\le C\|\nabla\tilde{\mathbf{v}}\|^2. \end{aligned}\tag{4.10}$$

Combining the estimate (4.9) with (4.10) yields the desired inequality (4.8).

It follows from Eq. 4.5 and the definition of $\mathbb{Q}_h$ that

$$b(\mathbf{v},\ \rho)=(\nabla_w\cdot(Q_h\tilde{\mathbf{v}}),\ \rho)=(\mathbb{Q}_h(\nabla\cdot\tilde{\mathbf{v}}),\ \rho)=(\nabla\cdot\tilde{\mathbf{v}},\ \rho).$$

Using the above equation, (4.7) and (4.8), we have

$$\frac{|b(\mathbf{v},\rho)|}{|||\mathbf{v}|||}\ge\frac{|(\nabla\cdot\tilde{\mathbf{v}},\rho)|}{C_0\|\tilde{\mathbf{v}}\|_1}\ge\beta\|\rho\|$$

for a positive constant $\beta$. This completes the proof of the lemma. □

It follows from Lemma 4.1 and Lemma 4.3 that the following solvability holds true for the weak Galerkin finite element scheme (3.3)–(3.4).

**Lemma 4.4** *The weak Galerkin finite element scheme (3.3)–(3.4) has one and only one solution.*

## 5 Error equations

For simplicity of analysis, we assume the coefficient tensor $A=I$ in (1.1). The result can be extended to variable tensors without any difficulty, provided that the tensor $a$ is piecewise sufficiently smooth. Let $\mathbf{u}_h=\{\mathbf{u}_0,\mathbf{u}_b\}\in V_h$ and $p_h\in W_h$ be the weak Galerkin finite element solution arising from the numerical scheme (3.3)–(3.4). Denote by $\mathbf{u}$ and $p$ the exact solution of Eqs. 1.1–1.3. The $L^2$ projection of $\mathbf{u}$ in the finite element space $V_h$ is given by

$$Q_h\mathbf{u}=\{Q_0\mathbf{u},\ Q_b\mathbf{u}\}.$$

Similarly, the pressure $p$ is projected into $W_h$ as $\mathbb{Q}_hp$. Denote by $\mathbf{e}_h$ and $\varepsilon_h$ the corresponding error given by

$$\mathbf{e}_h=\{\mathbf{e}_0,\ \mathbf{e}_b\}=\{Q_0\mathbf{u}-\mathbf{u}_0,\ Q_b\mathbf{u}-\mathbf{u}_b\},\qquad \varepsilon_h=\mathbb{Q}_hp-p_h.\tag{5.1}$$

The goal of this section is to derive two equations for which the error $\mathbf{e}_h$ and $\varepsilon_h$ shall satisfy. The resulting equations are called *error equations*, which play a critical role in the convergence analysis for the weak Galerkin finite element method.

**Lemma 5.1** *Let $(\boldsymbol{w}; \rho) \in [H^2(\Omega)]^d \times H^1(\Omega)$ satisfy the following equation*

$$-\Delta \boldsymbol{w} + \nabla \rho = \eta \tag{5.2}$$

*in the domain $\Omega$. Let $Q_h \boldsymbol{w} = \{Q_0 \boldsymbol{w}, Q_b \boldsymbol{w}\}$ and $\mathbb{Q}_h \rho$ be the $L^2$ projection of $(\boldsymbol{w}; \rho)$ into the finite element space $V_h \times W_h$. Then, the following equation holds true*

$$(\nabla_w(Q_h \boldsymbol{w}), \nabla_w \boldsymbol{v}) - (\nabla_w \cdot \boldsymbol{v}, \mathbb{Q}_h \rho) = (\eta, \boldsymbol{v}_0) + \ell_{\boldsymbol{w}}(\boldsymbol{v}) - \theta_\rho(\boldsymbol{v}) \tag{5.3}$$

*for all $\boldsymbol{v} \in V_h^0$, where $\ell_{\boldsymbol{w}}$ and $\theta_\rho$ are two linear functionals on $V_h^0$ defined by*

$$\ell_{\boldsymbol{w}}(\boldsymbol{v}) = \sum_{T \in \mathcal{T}_h} \langle \boldsymbol{v}_0 - \boldsymbol{v}_b, \ \nabla \boldsymbol{w} \cdot \boldsymbol{n} - \boldsymbol{Q}_h(\nabla \boldsymbol{w}) \cdot \boldsymbol{n} \rangle_{\partial T},$$

$$\theta_\rho(\boldsymbol{v}) = \sum_{T \in \mathcal{T}_h} \langle \boldsymbol{v}_0 - \boldsymbol{v}_b, \ (\rho - \mathbb{Q}_h \rho) \boldsymbol{n} \rangle_{\partial T}.$$

*Proof* First, it follows from Eqs. 4.4, 2.3, and the integration by parts that

$$\begin{aligned}
(\nabla_w(Q_h \mathbf{w}), \nabla_w \mathbf{v})_T &= (\mathbf{Q}_h(\nabla \mathbf{w}), \nabla_w \mathbf{v})_T \\
&= -(\mathbf{v}_0, \nabla \cdot \mathbf{Q}_h(\nabla \mathbf{w}))_T + \langle \mathbf{v}_b, \mathbf{Q}_h(\nabla \mathbf{w}) \cdot \mathbf{n} \rangle_{\partial T} \\
&= (\nabla \mathbf{v}_0, \mathbf{Q}_h(\nabla \mathbf{w}))_T - \langle \mathbf{v}_0 - \mathbf{v}_b, \ \mathbf{Q}_h(\nabla \mathbf{w}) \cdot \mathbf{n} \rangle_{\partial T} \\
&= (\nabla \mathbf{w}, \nabla \mathbf{v}_0)_T - \langle \mathbf{v}_0 - \mathbf{v}_b, \mathbf{Q}_h(\nabla \mathbf{w}) \cdot \mathbf{n} \rangle_{\partial T}.
\end{aligned} \tag{5.4}$$

Next, by using Eqs. 4.5 and 2.6, the fact that $\sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_b, \rho \ \ \mathbf{n} \rangle_{\partial T} = 0$ and the integration by parts, we obtain

$$\begin{aligned}
(\nabla_w \cdot \mathbf{v}, \mathbb{Q}_h \rho) &= -\sum_{T \in \mathcal{T}_h} (\mathbf{v}_0, \nabla(\mathbb{Q}_h \rho))_T + \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_b, (\mathbb{Q}_h \rho) \mathbf{n} \rangle_{\partial T} \\
&= \sum_{T \in \mathcal{T}_h} (\nabla \cdot \mathbf{v}_0, \mathbb{Q}_h \rho)_T - \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b, (\mathbb{Q}_h \rho) \mathbf{n} \rangle_{\partial T} \\
&= \sum_{T \in \mathcal{T}_h} (\nabla \cdot \mathbf{v}_0, \rho)_T - \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b, (\mathbb{Q}_h \rho) \mathbf{n} \rangle_{\partial T} \\
&= -\sum_{T \in \mathcal{T}_h} (\mathbf{v}_0, \nabla \rho)_T + \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_0, \rho \mathbf{n} \rangle_{\partial T} - \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b, (\mathbb{Q}_h \rho) \mathbf{n} \rangle_{\partial T} \\
&= -\sum_{T \in \mathcal{T}_h} (\mathbf{v}_0, \nabla \rho)_T + \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b, \rho \mathbf{n} \rangle_{\partial T} - \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b, (\mathbb{Q}_h \rho) \mathbf{n} \rangle_{\partial T} \\
&= -(\mathbf{v}_0, \nabla \rho) + \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b, (\rho - \mathbb{Q}_h \rho) \mathbf{n} \rangle_{\partial T},
\end{aligned}$$

which leads to

$$(\mathbf{v}_0, \nabla \rho) = -(\nabla_w \cdot \mathbf{v}, \mathbb{Q}_h \rho) + \sum_{T \in \mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b, (\rho - \mathbb{Q}_h \rho) \mathbf{n} \rangle_{\partial T}. \tag{5.5}$$

Next we test (5.2) by using $\mathbf{v}_0$ in $\mathbf{v} = \{\mathbf{v}_0, \mathbf{v}_b\} \in V_h^0$ to obtain

$$-(\Delta \mathbf{w}, \mathbf{v}_0) + (\nabla \rho, \mathbf{v}_0) = (\eta, \mathbf{v}_0). \tag{5.6}$$

It follows from the usual integration by parts that

$$-(\Delta \mathbf{w}, \mathbf{v}_0) = \sum_{T\in\mathcal{T}_h} (\nabla \mathbf{w},\ \nabla \mathbf{v}_0)_T - \sum_{T\in\mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b,\ \nabla \mathbf{w}\cdot\mathbf{n}\rangle_{\partial T},$$

where we have used the fact that $\sum_{T\in\mathcal{T}_h} \langle \mathbf{v}_b, \nabla \mathbf{w}\cdot\mathbf{n}\rangle_{\partial T} = 0$. Using Eq. 5.4 and the equation above, we have

$$\begin{aligned} -(\Delta \mathbf{w}, \mathbf{v}_0) &= (\nabla_w(Q_h\mathbf{w}), \nabla_w \mathbf{v}) \\ &\quad - \sum_{T\in\mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b, \nabla \mathbf{w}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla \mathbf{w})\cdot\mathbf{n}\rangle_{\partial T}. \end{aligned} \tag{5.7}$$

Substituting Eqs. 5.5 and 5.7 into Eq. 5.6 yields

$$(\nabla_w(Q_h\mathbf{w}), \nabla_w \mathbf{v}) - (\nabla_w\cdot\mathbf{v},\ \mathbb{Q}_h\rho) = (\eta, \mathbf{v}_0) + \ell_{\mathbf{w}}(\mathbf{v}) - \theta_\rho(\mathbf{v}),$$

which completes the proof of the lemma. □

The following is a result on the error equation for the weak Galerkin finite element scheme (3.3)–(3.4).

**Lemma 5.2** *Let $\boldsymbol{e}_h$ and $\varepsilon_h$ be the error of the weak Galerkin finite element solution arising from Eqs. 3.3–3.4, as defined by Eq. 5.1. Then, we have*

$$a(\boldsymbol{e}_h,\ \boldsymbol{v}) - b(\boldsymbol{v},\ \varepsilon_h) = \varphi_{\boldsymbol{u},p}(\boldsymbol{v}), \tag{5.8}$$

$$b(\boldsymbol{e}_h,\ q) = 0, \tag{5.9}$$

*for all $\boldsymbol{v} \in V_h^0$ and $q \in W_h$, where $\varphi_{\boldsymbol{u},p}(\boldsymbol{v}) = \ell_{\boldsymbol{u}}(\boldsymbol{v}) - \theta_p(\boldsymbol{v}) + s(Q_h\boldsymbol{u}, \boldsymbol{v})$ is a linear functional defined on $V_h^0$.*

*Proof* Since $(\mathbf{u}; p)$ satisfies the Eq. 5.2 with $\eta = \mathbf{f}$, then from Lemma 5.1 we have

$$(\nabla_w(Q_h\mathbf{u}), \nabla_w \mathbf{v}) - (\nabla_w\cdot\mathbf{v}, \mathbb{Q}_h p) = (\mathbf{f}, \mathbf{v}_0) + \ell_{\mathbf{u}}(\mathbf{v}) - \theta_p(\mathbf{v}).$$

Adding $s(Q_h\mathbf{u}, \mathbf{v})$ to both side of the above equation gives

$$a(Q_h\mathbf{u}, \mathbf{v}) - b(\mathbf{v}, \mathbb{Q}_h p) = (\mathbf{f}, \mathbf{v}_0) + \ell_{\mathbf{u}}(\mathbf{v}) - \theta_p(\mathbf{v}) + s(Q_h\mathbf{u}, \mathbf{v}). \tag{5.10}$$

The difference of Eqs. 5.10 and 3.3 yields the following equation,

$$a(\mathbf{e}_h, \mathbf{v}) - b(\mathbf{v}, \varepsilon_h) = \ell_{\mathbf{u}}(\mathbf{v}) - \theta_p(\mathbf{v}) + s(Q_h\mathbf{u}, \mathbf{v})$$

for all $\mathbf{v} \in V_h^0$, where $\mathbf{e}_h = \{\mathbf{e}_0,\ \mathbf{e}_b\} = \{Q_0\mathbf{u} - \mathbf{u}_0,\ Q_b\mathbf{u} - \mathbf{u}_b\}$ and $\varepsilon_h = \mathbb{Q}_h p - p_h$. This completes the derivation of Eq. 5.8.

As to Eq. 5.9, we test Eq. 1.2 by $q \in W_h$ and use (4.5) to obtain

$$0 = (\nabla\cdot\mathbf{u}, q) = (\nabla_w\cdot Q_h\mathbf{u}, q). \tag{5.11}$$

The difference of Eqs. 5.11 and 3.4 yields the following equation

$$b(\mathbf{e}_h, q) = 0$$

for all $q \in W_h$. This completes the derivation of Eq. 5.9. □

## 6 Error estimates

In this section, we shall establish optimal order error estimates for the velocity approximation $\mathbf{u}_h$ in a norm that is equivalent to the usual $H^1$-norm, and for the pressure approximation $p_h$ in the standard $L^2$ norm. In addition, we shall derive an error estimate for $\mathbf{u}_h$ in the standard $L^2$ norm by applying the usual duality argument in finite element error analysis.

**Theorem 6.1** *Let* $(\boldsymbol{u}; p) \in [H_0^1(\Omega) \cap H^{k+1}(\Omega)]^d \times (L_0^2(\Omega) \cap H^k(\Omega))$ *with* $k \geq 1$ *and* $(\boldsymbol{u}_h; p_h) \in V_h \times W_h$ *be the solution of Eqs. 1.1–1.3 and Eqs. 3.3–3.4, respectively. Then, the following error estimate holds true*

$$|||Q_h\boldsymbol{u} - \boldsymbol{u}_h||| + \|\mathbb{Q}_h p - p_h\| \leq Ch^k(\|\boldsymbol{u}\|_{k+1} + \|p\|_k). \tag{6.1}$$

*Proof* By letting $\mathbf{v} = \mathbf{e}_h$ in Eq. 5.8 and $q = \varepsilon_h$ in Eq. 5.9 and adding the two resulting equations, we have

$$|||\mathbf{e}_h|||^2 = \varphi_{\mathbf{u},p}(\mathbf{e}_h). \tag{6.2}$$

It then follows from Eqs. A.6–A.8 (see Appendix A) that

$$|||\mathbf{e}_h|||^2 \leq Ch^k(\|\mathbf{u}\|_{k+1} + \|p\|_k)|||\mathbf{e}_h|||, \tag{6.3}$$

which implies the first part of Eq. 6.1. To estimate $\|\varepsilon_h\|$, we have from Eq. 5.8 that

$$b(\mathbf{v}, \varepsilon_h) = a(\mathbf{e}_h, \mathbf{v}) - \varphi_{\mathbf{u},p}(\mathbf{v}).$$

Using the equation above, (4.2), (6.3) and (A.6)–(A.8), we arrive at

$$|b(\mathbf{v}, \varepsilon_h)| \leq Ch^k(\|\mathbf{u}\|_{k+1} + \|p\|_k)|||\mathbf{v}|||.$$

Combining the above estimate with the *inf-sup* condition (4.6) gives

$$\|\varepsilon_h\| \leq Ch^k(\|\mathbf{u}\|_{k+1} + \|p\|_k),$$

which yields the desired estimate (6.1). □

In the rest of this section, we shall derive an $L^2$-error estimate for the velocity approximation through a duality argument. To this end, consider the problem of seeking $(\psi; \xi)$ such that

$$-\Delta\psi + \nabla\xi = \mathbf{e}_0 \quad \text{in } \Omega, \tag{6.4}$$

$$\nabla \cdot \psi = 0 \quad \text{in } \Omega, \tag{6.5}$$

$$\psi = 0 \quad \text{on } \partial\Omega. \tag{6.6}$$

Assume that the dual problem has the $[H^2(\Omega)]^d \times H^1(\Omega)$-regularity property in the sense that the solution $(\psi; \xi) \in [H^2(\Omega)]^d \times H^1(\Omega)$ and the following a priori estimate holds true:

$$\|\psi\|_2 + \|\xi\|_1 \leq C\|\mathbf{e}_0\|. \tag{6.7}$$

**Theorem 6.2** *Let* $(\boldsymbol{u}; p) \in [H_0^1(\Omega) \cap H^{k+1}(\Omega)]^d \times (L_0^2(\Omega) \cap H^k(\Omega))$ *with* $k \geq 1$ *and* $(\boldsymbol{u}_h; p_h) \in V_h \times W_h$ *be the solution of Eqs. 1.1–1.3 and Eqs. 3.3–3.4, respectively. Then, the following optimal order error estimate holds true*

$$\|Q_0\boldsymbol{u} - \boldsymbol{u}_0\| \leq Ch^{k+1}(\|\boldsymbol{u}\|_{k+1} + \|p\|_k). \tag{6.8}$$

*Proof* Since $(\psi; \xi)$ satisfies the Eq. 5.2 with $\eta = \mathbf{e}_0 = Q_0\mathbf{u} - \mathbf{u}_0$, then from Eq. 5.3 we have

$$(\nabla_w Q_h\psi, \nabla_w \mathbf{v}) - (\nabla_w \cdot \mathbf{v}, \mathbb{Q}_h\xi) = (\mathbf{e}_0, \mathbf{v}_0) + \ell_\psi(\mathbf{v}) - \theta_\xi(\mathbf{v}), \qquad \forall \mathbf{v} \in V_h^0.$$

In particular, by letting $\mathbf{v} = \mathbf{e}_h$ we obtain

$$\|\mathbf{e}_0\|^2 = (\nabla_w Q_h\psi, \nabla_w \mathbf{e}_h) - (\nabla_w \cdot \mathbf{e}_h, \mathbb{Q}_h\xi) - \ell_\psi(\mathbf{e}_h) + \theta_\xi(\mathbf{e}_h).$$

Adding and subtracting $s(Q_h\psi, \mathbf{e}_h)$ in the equation above yields

$$\|\mathbf{e}_0\|^2 = a(Q_h\psi, \mathbf{e}_h) - b(\mathbf{e}_h, \mathbb{Q}_h\xi) - \varphi_{\psi,\xi}(\mathbf{e}_h),$$

where $\varphi_{\psi,\xi}(\mathbf{v}) = \ell_\psi(\mathbf{e}_h) - \theta_\xi(\mathbf{e}_h) + s(Q_h\psi, \mathbf{e}_h)$. It follows from Eqs. 5.9, 6.5 and 5.11 that

$$b(\mathbf{e}_h, \mathbb{Q}_h\xi) = 0, \quad b(Q_h\psi, \varepsilon_h) = 0.$$

Combining the above two equations gives

$$\|\mathbf{e}_0\|^2 = a(\mathbf{e}_h, Q_h\psi) - b(Q_h\psi, \varepsilon_h) - \varphi_{\psi,\xi}(\mathbf{e}_h).$$

Using Eq. 5.8 and the equation above, we have

$$\|\mathbf{e}_0\|^2 = \varphi_{\mathbf{u},p}(Q_h\psi) - \varphi_{\psi,\xi}(\mathbf{e}_h). \tag{6.9}$$

To estimate the two terms on the right hand side of Eq. 6.9, we use the inequalities (A.6)–(A.8) with $(\mathbf{w}; \rho) = (\psi; \xi)$, $\mathbf{v} = \mathbf{e}_h$, and $r = 1$ to obtain

$$|\varphi_{\psi,\xi}(\mathbf{e}_h)| \leq Ch(\|\psi\|_2 + \|\xi\|_1)|\!|\!|\mathbf{e}_h|\!|\!| \leq Ch|\!|\!|\mathbf{e}_h|\!|\!| \, \|\mathbf{e}_0\|, \tag{6.10}$$

where we have used the regularity assumption (6.7). Each of the terms in $\varphi_{\mathbf{u},p}(Q_h\psi)$ can be handled as follows.

(i) For the stability term $s(Q_h\mathbf{u}, Q_h\psi)$, we use the definition of $Q_b$ and (A.4) to obtain

$$\begin{aligned} |s(Q_h\mathbf{u}, \ Q_h\psi)| &= \left| \sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_b(Q_0\mathbf{u} - \mathbf{u}), \ Q_b(Q_0\psi - \psi)\rangle_{\partial T} \right| \\ &\leq \left( \sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0\mathbf{u} - \mathbf{u}\|_{\partial T}^2 \right)^{1/2} \left( \sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0\psi - \psi\|_{\partial T}^2 \right)^{1/2} \\ &\leq Ch^{k+1}\|\mathbf{u}\|_{k+1}\|\psi\|_2. \end{aligned}$$

(ii) For the term $\ell_{\mathbf{u}}(Q_h\psi)$, we first use the definition of $Q_b$ and the fact that $\psi = 0$ on $\partial\Omega$ to obtain

$$\sum_{T\in\mathcal{T}_h} \langle \psi - Q_b\psi, \nabla\mathbf{u}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{u})\cdot\mathbf{n}\rangle_{\partial T} = \sum_{T\in\mathcal{T}_h} \langle \psi - Q_b\psi, \nabla\mathbf{u}\cdot\mathbf{n}\rangle_{\partial T} = 0.$$

Thus,

$$
\begin{aligned}
|\ell_{\mathbf{u}}(Q_h\psi)| &= \left|\sum_{T\in\mathcal{T}_h} \langle Q_0\psi - Q_b\psi, \nabla\mathbf{u}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{u})\cdot\mathbf{n}\rangle_{\partial T}\right| \\
&= \left|\sum_{T\in\mathcal{T}_h} \langle Q_0\psi - \psi, \nabla\mathbf{u}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{u})\cdot\mathbf{n}\rangle_{\partial T}\right| \\
&\le \left(\sum_{T\in\mathcal{T}_h} h_T\|\nabla\mathbf{u}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{u})\cdot\mathbf{n}\|_{\partial T}^2\right)^{1/2} \left(\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0\psi - \psi\|_{\partial T}^2\right)^{1/2} \\
&\le Ch^{k+1}\|\mathbf{u}\|_{k+1}\|\psi\|_2.
\end{aligned}
$$

(iii) For the term $\theta_p(Q_h\psi)$, we first use the definition of $Q_b$ and the fact that $\psi = 0$ on $\partial\Omega$ to obtain

$$
\sum_{T\in\mathcal{T}_h} \langle \psi - Q_b\psi, (p - \mathbb{Q}_h p)\mathbf{n}\rangle_{\partial T} = \sum_{T\in\mathcal{T}_h} \langle \psi - Q_b\psi, p\mathbf{n}\rangle_{\partial T} = 0.
$$

Thus, from Eqs. A.4 and A.3 we obtain

$$
\begin{aligned}
|\theta_p(Q_h\psi)| &= \left|\sum_{T\in\mathcal{T}_h} \langle Q_0\psi - Q_b\psi, (p - \mathbb{Q}_h p)\mathbf{n}\rangle_{\partial T}\right| \\
&= \left|\sum_{T\in\mathcal{T}_h} \langle Q_0\psi - \psi, (p - \mathbb{Q}_h p)\mathbf{n}\rangle_{\partial T}\right| \\
&\le \left(\sum_{T\in\mathcal{T}_h} h_T\|p - \mathbb{Q}_h p\|_{\partial T}^2\right)^{1/2} \left(\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0\psi - \psi\|_{\partial T}^2\right)^{1/2} \\
&\le Ch^{k+1}\|p\|_k\|\psi\|_2.
\end{aligned}
$$

The three estimates in (i), (ii), (iii), and the regularity (6.7) collectively yield

$$
\begin{aligned}
|\varphi_{\mathbf{u},p}(Q_h\psi)| &\le Ch^{k+1}(\|\mathbf{u}\|_{k+1} + \|p\|_k)\|\psi\|_2 \\
&\le Ch^{k+1}(\|\mathbf{u}\|_{k+1} + \|p\|_k)\|\mathbf{e}_0\|. \qquad (6.11)
\end{aligned}
$$

Finally, substituting Eqs. 6.10 and 6.11 into Eq. 6.9 gives

$$
\|\mathbf{e}_0\|^2 \le Ch^{k+1}(\|\mathbf{u}\|_{k+1} + \|p\|_k)\|\mathbf{e}_0\| + Ch|||\mathbf{e}_h|||\ \|\mathbf{e}_0\|.
$$

It follows that

$$
\|\mathbf{e}_0\| \le Ch^{k+1}(\|\mathbf{u}\|_{k+1} + \|p\|_k) + Ch|||\mathbf{e}_h|||,
$$

which, together with Theorem 6.1, completes the proof of the theorem. □

## 7 Concluding remarks

This paper introduced a new finite element method for the Stokes equations by using the general concept of weak Galerkin. The scheme is applicable to finite element partitions of arbitrary polygon or polyhedra. The paper has laid a solid theoretical foundation for the stability and convergence of the weak Galerkin method. There are, however, many open issues that need to be investigated in future work. Here we would like to list a few for interested readers to consider: (1) how the discretized linear systems can be solved efficiently by using techniques such as domain decomposition and multigrids? (2) can the weak Galerkin scheme for the Stokes equations be hybridized? If so, how such a hybridization may help in variable reduction and solution solving? and (3) what superconvergence can one develop for the weak Galerkin method? (4) is the weak Galerkin method more competitive than other existing finite element schemes in practical computation? (5) what stability do weak Galerkin methods have in other norms such as $L^p$, $p > 1$?

## Appendix A

In this Appendix, we shall provide some technical results regarding approximation properties for the $L^2$ projection operators $Q_h$, $\mathbf{Q}_h$, and $\mathbb{Q}_h$. These estimates have been employed in previous sections to yield various error estimates for the weak Galerkin finite element solution of the Stokes problem arising from the scheme (3.3)–(3.4).

**Lemma A.1** *Let $\mathcal{T}_h$ be a finite element partition of $\Omega$ satisfying the shape regularity assumption as specified in [16] and $\boldsymbol{w} \in [H^{r+1}(\Omega)]^d$ and $\rho \in H^r(\Omega)$ with $1 \le r \le k$. Then, for $0 \le s \le 1$ we have*

$$\sum_{T\in\mathcal{T}_h} h_T^{2s} \|\boldsymbol{w} - Q_0\boldsymbol{w}\|_{T,s}^2 \le h^{2(r+1)} \|\boldsymbol{w}\|_{r+1}^2, \tag{A.1}$$

$$\sum_{T\in\mathcal{T}_h} h_T^{2s} \|\nabla\boldsymbol{w} - \boldsymbol{Q}_h(\nabla\boldsymbol{w})\|_{T,s}^2 \le Ch^{2r} \|\boldsymbol{w}\|_{r+1}^2, \tag{A.2}$$

$$\sum_{T\in\mathcal{T}_h} h_T^{2s} \|\rho - \mathbb{Q}_h\rho\|_{T,s}^2 \le Ch^{2r} \|\rho\|_r^2. \tag{A.3}$$

*Here $C$ denotes a generic constant independent of the meshsize $h$ and the functions in the estimates.*

A proof of the lemma can be found in [16], which is based on some technical inequalities for functions defined on polygon/polyhedral elements with shape regularity. We emphasize that the approximation error estimates in Lemma A.1 hold true when the underlying mesh $\mathcal{T}_h$ consists of arbitrary polygons or polyhedra with shape regularity as detailed in [16] and [13].

Let $T$ be an element with $e$ as an edge/face. For any function $g \in H^1(T)$, the following trace inequality has been proved to be valid for general meshes satisfying the shape regular assumptions detailed in [16]:

$$\|g\|_e^2 \le C\left(h_T^{-1}\|g\|_T^2 + h_T\|\nabla g\|_T^2\right). \tag{A.4}$$

**Lemma A.2** *For any $\boldsymbol{v} = \{\boldsymbol{v}_0, \boldsymbol{v}_b\} \in V_h$, we have*

$$\sum_{T\in\mathcal{T}_h} \|\nabla \boldsymbol{v}_0\|_T^2 \le C|||v|||^2. \tag{A.5}$$

*Proof* For any $\mathbf{v} = \{\mathbf{v}_0, \mathbf{v}_b\} \in V_h$, it follows from the integration by parts and the definitions of weak gradient and $Q_b$,

$$\begin{aligned}(\nabla\mathbf{v}_0, \nabla\mathbf{v}_0)_T &= -(\mathbf{v}_0, \nabla\cdot\nabla\mathbf{v}_0)_T + \langle \mathbf{v}_0, \nabla\mathbf{v}_0\cdot\mathbf{n}\rangle_{\partial T}\\ &= -(\mathbf{v}_0, \nabla\cdot\nabla\mathbf{v}_0)_T + \langle \mathbf{v}_b, \nabla\mathbf{v}_0\cdot\mathbf{n}\rangle_{\partial T} + \langle \mathbf{v}_0-\mathbf{v}_b, \nabla\mathbf{v}_0\cdot\mathbf{n}\rangle_{\partial T}\\ &= (\nabla_w\mathbf{v}, \nabla\mathbf{v}_0)_T + \langle Q_b\mathbf{v}_0-\mathbf{v}_b, \nabla\mathbf{v}_0\cdot\mathbf{n}\rangle_{\partial T}.\end{aligned}$$

By applying the trace inequality (A.4) and the inverse inequality to the equation above, we obtain

$$\|\nabla\mathbf{v}_0\|_T^2 \le C\left(\|\nabla_w\mathbf{v}\|_T\|\nabla\mathbf{v}_0\|_T + h_T^{-\frac{1}{2}}\|Q_b\mathbf{v}_0-\mathbf{v}_b\|_{\partial T}\|\nabla\mathbf{v}_0\|_T\right).$$

Thus,

$$\|\nabla\mathbf{v}_0\|_T^2 \le C\left(\|\nabla_w\mathbf{v}\|_T^2 + h_T^{-1}\|Q_b\mathbf{v}_0-\mathbf{v}_b\|_{\partial T}^2\right),$$

which gives rise to Eq. A.5 after a summation over all $T \in \mathcal{T}_h$. □

**Lemma A.3** *Let $1 \le r \le k$ and $\boldsymbol{w} \in [H^{r+1}(\Omega)]^d$ and $\rho \in H^r(\Omega)$ and $\boldsymbol{v} \in V_h$. Assume that the finite element partition $\mathcal{T}_h$ is shape regular. Then, the following estimates hold true*

$$|s(Q_h\boldsymbol{w}, \boldsymbol{v})| \le Ch^r\|\boldsymbol{w}\|_{r+1}|||\boldsymbol{v}|||, \tag{A.6}$$
$$|\ell_{\boldsymbol{w}}(\boldsymbol{v})| \le Ch^r\|\boldsymbol{w}\|_{r+1}|||\boldsymbol{v}|||, \tag{A.7}$$
$$|\theta_\rho(\boldsymbol{v})| \le Ch^r\|\rho\|_r|||\boldsymbol{v}|||, \tag{A.8}$$

*where $\ell_{\boldsymbol{w}}(\cdot)$ and $\ell_\rho(\cdot)$ are two linear functionals on $V_h$ given by*

$$\ell_{\boldsymbol{w}}(\boldsymbol{v}) = \sum_{T\in\mathcal{T}_h} \langle \boldsymbol{v}_0-\boldsymbol{v}_b, \nabla\boldsymbol{w}\cdot\boldsymbol{n} - \mathbf{Q}_h(\nabla\boldsymbol{w})\cdot\boldsymbol{n}\rangle_{\partial T}, \tag{A.9}$$

$$\theta_\rho(\boldsymbol{v}) = \sum_{T\in\mathcal{T}_h} \langle \boldsymbol{v}_0-\boldsymbol{v}_b, (\rho-\mathbb{Q}_h\rho)\boldsymbol{n}\rangle_{\partial T}. \tag{A.10}$$

*Proof* Using the definition of $Q_b$, (A.4) and (A.1), we have

$$
\begin{aligned}
|s(Q_h\mathbf{w},\ \mathbf{v})| &= \left|\sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_b(Q_0\mathbf{w}) - Q_b\mathbf{w},\ Q_b\mathbf{v}_0 - \mathbf{v}_b\rangle_{\partial T}\right| \\
&= \left|\sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_b(Q_0\mathbf{w} - \mathbf{w}),\ Q_b\mathbf{v}_0 - \mathbf{v}_b\rangle_{\partial T}\right| \\
&= \left|\sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_0\mathbf{w} - \mathbf{w},\ Q_b\mathbf{v}_0 - \mathbf{v}_b\rangle_{\partial T}\right| \\
&\le \left(\sum_{T\in\mathcal{T}_h} (h_T^{-2}\|Q_0\mathbf{w} - \mathbf{w}\|_T^2 + \|\nabla(Q_0\mathbf{w} - \mathbf{w})\|_T^2)\right)^{1/2} \\
&\quad \left(\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_b\mathbf{v}_0 - \mathbf{v}_b\|_{\partial T}^2\right)^{1/2} \\
&\le Ch^r\|\mathbf{w}\|_{r+1}|\!|\!|\mathbf{v}|\!|\!|.
\end{aligned}
$$

It follows from Eqs. A.4 and A.2 that

$$
\begin{aligned}
|\ell_{\mathbf{w}}(\mathbf{v})| &= \left|\sum_{T\in\mathcal{T}_h} \langle \mathbf{v}_0 - \mathbf{v}_b,\ \nabla\mathbf{w}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{w})\cdot\mathbf{n}\rangle_{\partial T}\right| \\
&\le \left|\sum_{T\in\mathcal{T}_h} \langle \mathbf{v}_0 - Q_b\mathbf{v}_0,\ \nabla\mathbf{w}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{w})\cdot\mathbf{n}\rangle_{\partial T}\right| \\
&\quad + \left|\sum_{T\in\mathcal{T}_h} \langle Q_b\mathbf{v}_0 - \mathbf{v}_b,\ \nabla\mathbf{w}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{w})\cdot\mathbf{n}\rangle_{\partial T}\right|.
\end{aligned}
$$

To estimate the first term on the righ-hand side of the above inequality, we use Eqs. A.4, A.2, A.5 and the inverse inequality to obtain

$$
\begin{aligned}
&\left|\sum_{T\in\mathcal{T}_h} \langle \mathbf{v}_0 - Q_b\mathbf{v}_0,\ \nabla\mathbf{w}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{w})\cdot\mathbf{n}\rangle_{\partial T}\right| \\
\le\ & C\sum_{T\in\mathcal{T}_h} h_T\|\nabla\mathbf{w}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{w})\cdot\mathbf{n}\|_{\partial T}\|\nabla\mathbf{v}_0\|_{\partial T} \\
\le\ & C\left(\sum_{T\in\mathcal{T}_h} h_T\|\nabla\mathbf{w}\cdot\mathbf{n} - \mathbf{Q}_h(\nabla\mathbf{w})\cdot\mathbf{n}\|_{\partial T}^2\right)^{1/2}\left(\sum_{T\in\mathcal{T}_h}\|\nabla\mathbf{v}_0\|_T^2\right)^{1/2} \\
\le\ & Ch^r\|\mathbf{w}\|_{r+1}|\!|\!|\mathbf{v}|\!|\!|.
\end{aligned}
$$

Similarly, for the second term, we have

$$\begin{aligned}&\left|\sum_{T\in\mathcal{T}_h}\langle Q_b\mathbf{v}_0-\mathbf{v}_b,\ \nabla\mathbf{w}\cdot\mathbf{n}-\mathbf{Q}_h(\nabla\mathbf{w})\cdot\mathbf{n}\rangle_{\partial T}\right|\\ \le\ & C\left(\sum_{T\in\mathcal{T}_h}h_T\|\nabla\mathbf{w}\cdot\mathbf{n}-\mathbf{Q}_h(\nabla\mathbf{w})\cdot\mathbf{n}\|_{\partial T}^2\right)^{1/2}\left(\sum_{T\in\mathcal{T}_h}h_T^{-1}\|Q_b\mathbf{v}_0-\mathbf{v}_b\|_{\partial T}^2\right)^{1/2}\\ \le\ & Ch^r\|\mathbf{w}\|_{r+1}|\!|\!|\mathbf{v}|\!|\!|.\end{aligned}$$

The estimate (A.7) is verified by combining the above three estimates.

The same technique for proving (A.7) can be applied to yield the following estimate.

$$\begin{aligned}|\theta_\rho(\mathbf{v})| &= \left|\sum_{T\in\mathcal{T}_h}\langle\mathbf{v}_0-\mathbf{v}_b,\ (\rho-\mathbb{Q}_h\rho)\mathbf{n}\rangle_{\partial T}\right|\\ &\le Ch^r\|\rho\|_r|\!|\!|\mathbf{v}|\!|\!|.\end{aligned}$$

This completes the proof of the lemma. □

## References


1. Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. **39**, 1749–1779 (2002)
2. Babuška, I.: The finite element method with Lagrangian multiplier. Numer.Math. **20**, 179–192 (1973)
3. Beirao Da Veiga, L., Brezzi, F., Cangiani, A., Manzini, G., Marini, L.D., Russo, A.: Basic principles of virtual element methods. Math. Models Methods Appl. Sci. **23**(1), 199–214 (2013)
4. Brenner, S., Scott, R.: Mathematical theory of finite element methods. Springer (2002)
5. Brezzi, F.: On the existence, uniqueness, and approximation of saddle point problems arising from Lagrangian multipliers. RAIRO, Anal. Numér. **2**, 129–151 (1974)
6. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Elements. Springer-Verlag, New York (1991)
7. Brezzi, F., Lipnikov, K., Shashkov, M.: Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. SIAM J. Numer. Anal. **43**(5), 1872–1896 (2005)
8. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland (1978)
9. Cockburn, B., Gopalakrishnan, J., Nguyen, N.C., Peraire, J., Sayas, F.: Analysis of HDG methods for Stokes flow. Math. Comput. **80**(274), 723–760 (2011)
10. Crouzeix, M., Raviart, P.A.: Conforming and nonconforming finite element methods for solving the stationary Stokes equations. RAIRO Anal. Numer. **7**, 33–76 (1973)
11. Girault, V., Raviart, P.A.: Finite Element Methods for the Navier-Stokes Equations: Theory and Algorithms. Springer-Verlag, Berlin (1986)
12. Gunzburger, M.D.: Finite Element Methods for Viscous Incompressible Flows, A Guide to Theory, Practice and Algorithms. Academic, San Diego (1989)
13. Mu, L., Wang, J., Ye, X.: Weak Galerkin finite element methods on Polytopal Meshes. International J of Numerical Analysis and Modeling **12**, 31–53 (2015). arXiv:1204.3655v2
14. Mu, L., Wang, J., Ye, X.: A weak Galerkin finite element methods with polynomial reduction. J. Comp. and Appl. Math., in revision. arXiv:1304.6481
15. Wang, J., Ye, X.: A weak Galerkin finite element method for second-order elliptic problems. J. Comp. and Appl. Math. **241**, 103–115 (2013). arXiv:1104.2897
16. Wang, J., Ye, X.: A weak Galerkin mixed finite element method for second-order elliptic problems. Math. Comp. **83**(289), 2101–2126 (2014). S0025-5718(2014)02852-4. arXiv:1202.3655v2