IMA Journal of Numerical Analysis (2009) **29**, 827–855
doi:10.1093/imanum/drn038
Advance Access publication on July 4, 2008

Compact embeddings of broken Sobolev spaces and applications

ANNALISA BUFFA
Istituto di Matematica Applicata e Tecnologie Informatiche del Consiglio Nazionale delle Ricerche, Via Ferrata 1, 27100 Pavia, Italy

AND

CHRISTOPH ORTNER†
Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, UK

[Received on 19 April 2007; revised on 18 May 2008]

In this paper, we present several extensions of theoretical tools for the analysis of discontinuous Galerkin (DG) method beyond the linear case. We define broken Sobolev spaces for Sobolev indices in $[1, \infty)$, and we prove generalizations of many techniques of classical analysis in Sobolev spaces. Our targeted application is the convergence analysis for DG discretizations of energy minimization problems of the calculus of variations. Our main tool in this analysis is a theorem which permits the extraction of a 'weakly' converging subsequence of a family of discrete solutions and which shows that any 'weak limit' is a Sobolev function. As a second application, we compute the optimal embedding constants in broken Sobolev–Poincaré inequalities.

Keywords: discontinuous Galerkin method; broken Sobolev spaces; embedding theorems; compactness; Γ-convergence.

1. Introduction

In this article, we develop several tools for the analysis of the discontinuous Galerkin finite-element method (DGFEM) which, in this generality, have only been available in classical Sobolev spaces. We define broken Sobolev norms for Sobolev indices $p \in [1, \infty)$ and prove several embedding theorems such as broken Poincaré–Sobolev inequalities (see also Lasis & Süli, 2003; Brenner, 2003, 2004) and trace theorems; see Section 4. These broken embedding theorems are based on combining the known results in classical Sobolev spaces and the space of functions of bounded variation with a *continuous reconstruction operator* which maps any discontinuous Galerkin finite-element (DGFE) function to a Lipschitz function. This operator is analysed in detail in Section 3.

These results are then used to prove a compactness theorem for broken Sobolev spaces on succesively refined meshes when endowed with suitable mesh-dependent topologies. In our opinion, this compactness theorem is the most important result of the present work.

Our original motivation to prove these results was to understand how one could use a DGFEM to discretize energy minimization problems of the calculus of variations which occur in many areas of applied mathematics. A possible idea was provided by Ten Eyck & Lew (2006) which we briefly motivate in Section 1.1 and analyse in detail in Section 6. The tools which we develop in Sections 3–5 allow us to give a rigorous convergence analysis for a general class of energy minimization problems.

†Email: christoph.ortner@comlab.ox.ac.uk

© The author 2008. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.

As a second application, we present a technique to prove that the constant in a broken embedding inequality is the same as in its classical version, provided that the continuous version of the embedding is compact. We demonstrate the technique at the example of the Poincaré–Sobolev inequality.

We anticipate that the tools and techniques which we develop in this paper will have numerous applications in the analysis of DGFEMs. For example, the embedding results can be useful for any nonlinear problem where bounds on lower-order nonlinear terms are required. The compactness results may be useful for any problem where no 'classical' analysis based on coercivity or an inf–sup condition is possible (for example, in the presence of multiplicity of solutions) and where only weak convergence can be expected.

In Sections 1.1 and 1.2, we provide an introduction to our two targeted applications. We will use notation which is not introduced until Section 2 but which is standard in the literature on DGFEMs. Furthermore, we would like to stress that these sections are intended as an informal introduction and therefore some statements are intentionally not made fully precise.

1.1 *The variational DGFEM*

Let $S^k(\mathscr{T}_h)$ denote the space of possibly discontinuous, piecewise polynomial functions of degree k with respect to a partition $\mathscr{T}_h$ of a domain $\Omega \subset \mathbb{R}^n$ with boundary $\partial\Omega = \Gamma_\mathrm{D} \cup \Gamma_\mathrm{N}$. Let Γ_int denote the interior skeleton of the partition and let h denote the global and $\mathsf{h}(x)$ the local mesh size.

The basic problem of the calculus of variations is to minimize the functional

$$\mathscr{I}(u) = \int_\Omega f(x, u, \nabla u)\mathrm{d}x + \int_{\Gamma_\mathrm{N}} g(x, u)\mathrm{d}s \tag{1.1}$$

over a set of admissible functions, say,

$$\mathscr{A} = \left\{u \in W^{1,p}(\Omega)^m : u|_{\Gamma_\mathrm{D}} = u_\mathrm{D}\right\}, \tag{1.2}$$

where $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ and $g : \Gamma_\mathrm{N} \times \mathbb{R}^m \to \mathbb{R}$. Under suitable conditions on f and g, the existence of minimizers follows from the direct method of the calculus of variations (Dacorogna, 1989).

To discretize (1.1) by a conforming finite-element method, one would construct a finite-dimensional subspace $\mathscr{A}_h$ of $\mathscr{A}$ (by means of the finite-element method) and aim to minimize $\mathscr{I}$ over $\mathscr{A}_h$ instead. When f and g satisfy suitable conditions, one can then modify the direct method to prove the convergence of *discrete minimizers* to a minimizer of the original problem. Such a technique completely avoids the use of the Euler–Lagrange equations and is therefore particularly useful when they are not available or when it is known that the minimizers sought are singular and therefore may not satisfy these equations (Ball, 2001).

The question which we wish to adress here, and in more detail in Section 6, is whether a similar technique can be applied for the DGFEM. Naively, one might try to define a discrete functional as follows:

$$\mathscr{I}_h(u_h) = \int_\Omega f(x, u_h, \nabla u_h)\mathrm{d}x + \int_{\Gamma_\mathrm{N}} g(x, u_h)\mathrm{d}s + \int_{\Gamma_\mathrm{int}} \mathsf{h}^{-1}|[\![u_h]\!]|^2\,\mathrm{d}s + \int_{\Gamma_\mathrm{D}} \mathsf{h}^{-1}|u_h - u_\mathrm{D}|^2\,\mathrm{d}s, \tag{1.3}$$

where ∇u_h denotes the elementwise gradient of u_h, $[\![u_h]\!]$ denotes the jump of u_h between two elements and h is the local mesh size (see Section 2 for the precise definitions). The two latter terms would, respectively, impose weak continuity across element interfaces and the Dirichlet boundary condition. However, it turns out that this discretization is *not* convergent, which is due to fact that we used an *inconsistent* discretization for the gradient. Since DGFE functions u_h are not continuous, their distributional gradient has a contribution from the jumps; more precisely,

$$\langle Du_h, \varphi \rangle = \int_\Omega \nabla u \cdot \varphi \, \mathrm{d}x - \int_{\Gamma_{\mathrm{int}}} [\![u_h]\!] \cdot \varphi \, \mathrm{d}s \quad \forall \varphi \in C_c^\infty(\Omega)^{m \times n}, \tag{1.4}$$

where $[\![u_h]\!]$ is the jump of u_h across the faces of Γ_{int}, which should be taken into account. Ten Eyck & Lew (2006) used a *lifting operator* defined by

$$\int_\Omega R(u_h) \cdot \varphi_h \, \mathrm{d}x = - \int_{\Gamma_{\mathrm{int}} \cup \partial\Omega} [\![u_h]\!] \cdot \{\varphi_h\} \mathrm{d}s, \tag{1.5}$$

where $\{\varphi_h\}$ is a suitable average (*flux*) of the bi-valued function φ_h on the skeleton, to define

$$\begin{aligned} \mathscr{I}_h(u_h) = \int_\Omega f(x, u_h, \nabla u_h + R(u_h)) \mathrm{d}x + \int_{\Gamma_{\mathrm{N}}} g(x, u_h) \mathrm{d}s \\ + \int_{\Gamma_{\mathrm{int}}} \mathsf{h}^{-1} |[\![u_h]\!]|^2 \, \mathrm{d}s + \int_{\Gamma_{\mathrm{D}}} \mathsf{h}^{-1} |u_h - u_{\mathrm{D}}|^2 \, \mathrm{d}s. \end{aligned} \tag{1.6}$$

Using our compactness result, Theorem 5.2, for motivation it was natural to arrive at the same discretization. In fact, our theoretical results in Sections 4 and 5 make it straightforward to prove convergence of minimizers of $\mathscr{I}_h$ in $S^k(\mathscr{T}_h)^m$ to a minimizer of $\mathscr{I}$ in $\mathscr{A}$; see Theorem 6.1. The proof of this theorem mimics the direct method (or rather a closely related technique known as Γ-convergence; Braides, 2002; Dal Maso, 1993) where our compactness results feature prominently. In addition, we do not restrict ourselves to the case $p = 2$ but will use more general Sobolev indices in our discretization. It will become clear that the appropriate choice strongly depends on the properties of f and g.

We conclude this dicussion with a remark on the minimization problem (1.1). Depending on the particular properties of f, the computation of minimizers to (1.1) is a largely unsolved problem. For example, for typical stored energy densities of finite elasticity, it is unknown whether a conforming Galerkin finite-element discretization of (1.1) converges (Ball, 2001; Le Tallec, 1994). Our own analysis in the present work covers only the case where f is convex in the third argument, and satisfies certain growth conditions, which are insufficient to cover physically realistic stored energies (where f is at best polyconvex and is infinite for certain gradients) and it can therefore only be considered as an exploratory first step towards the solution of the general model problem (1.1) by the DGFEM. However, we hope that the flexibility of the discontinuous Galerkin (DG) method will allow us in the future to tackle some of the more difficult problems in this class.

1.2 *Optimal embedding constants*

In Section 4, we prove several broken embedding theorems, such as the broken Sobolev–Poincaré inequality

$$\|u_h - (u_h)_\Omega\|_{L^q(\Omega)} \leqslant C_h |u_h|_{W^{1,p}(\mathscr{T}_h)} \quad \forall u_h \in S^k(\mathscr{T}_h), \tag{1.7}$$

where $(u_h)_\Omega = |\Omega|^{-1} \int_\Omega u_h \, \mathrm{d}x$ and where $p \in [1, n)$ and $q \in [1, np/(n-p)]$; see Lemma 4.1. The proofs of these embedding inequalities are not sharp and do not give optimal constants, even if one would make the effort to compute them explicitly.

Thus, in Section 7, we demonstrate a technique which allows us to determine the asymptotic behaviour of the constant C_h as $h \to 0$ by comparing it to its classical counterpart

$$\|u - (u)_\Omega\|_{L^q(\Omega)} \leqslant C\|\nabla u\|_{L^p(\Omega)} \quad \forall\, u \in W^{1,p}(\Omega). \tag{1.8}$$

For example, if we define the broken Sobolev norm as

$$|u_h|_{W^{1,p}(\mathscr{T}_h)} = \|\nabla u_h\|_{L^p(\Omega)} + \alpha \left(\int_{\Gamma_{\text{int}}} \mathsf{h}^{1-p} |[\![u_h]\!]|^p \, \mathrm{d}s \right)^{1/p}$$

(see also Lemma 2), then we can prove that if α is *small*, then $\liminf_{h\to 0} C_h > C$, whereas if α is *large*, then $\lim_{h\to 0} C_h = C$. We obtain this result by rewriting the embedding inequalities as minimization problems and then using techniques similar to those of Section 6.

2. Discontinuous finite-element spaces

Let $\mathscr{H}^{n-1}$ denote the $(n-1)$-dimensional Hausdorff measure and, for a set $A \subset \mathbb{R}^n$, let $\dim_{\mathrm{H}} A$ denote the Hausdorff dimension of A.

Let $\Omega \subset \mathbb{R}^n$ be a polyhedral Lipschitz domain. We divide the boundary $\partial\Omega$ into a Dirichlet boundary Γ_{D} and a Neumann boundary Γ_{N} such that $\Gamma_{\mathrm{N}} \cap \Gamma_{\mathrm{D}} = \emptyset$ and $\mathscr{H}^{n-1}(\partial\Omega \setminus (\Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}})) = 0$. Let $(\mathscr{T}_h)_{h\in(0,1]}$ be a family of partitions of $\bar{\Omega}$ into convex polyhedral elements which are affine images of a set of reference polyhedra. More precisely, we assume that there exists a finite number of convex reference polyedra $\hat{\kappa}_1, \ldots, \hat{\kappa}_r$ such that $|\hat{\kappa}_i| = 1$ for $i = 1, \ldots, r$ and that for each $\kappa \in \mathscr{T}_h$, there exist an invertible affine map F_κ and a reference element $\hat{\kappa}_i$ such that $\kappa = F_\kappa(\hat{\kappa}_i)$. The symbol h denotes the global mesh size, i.e. $h = \max_{\kappa\in\mathscr{T}_h} \operatorname{diam}(\kappa)$. Without loss of generality, we assume that $h \in (0, 1]$. We will provide further assumptions on the mesh regularity in Section 2.1.

Throughout, we shall use the symbols $\approx$, $\lesssim$ and $\gtrsim$ to compare quantities which differ only up to positive constants that do not depend on the local or global mesh size or on any function which appears in the estimate.

2.1 *Mesh regularity*

In this section, we propose a set of assumptions on the family of partitions $(\mathscr{T}_h)_{h\in(0,1]}$ which are required in order to apply the theory developed in this paper. As it is standard in the finite-element literature, we define the set of $(n-1)$-dimensional faces $\mathscr{E}_h$ of the partition as follows:

$$\begin{aligned} \mathscr{E}_h = \{&\kappa \cap \kappa' : \kappa, \kappa' \in \mathscr{T}_h, \dim_{\mathrm{H}}(\kappa \cap \kappa') = n-1\} \\ &\cup \{\kappa \cap \partial\Omega : \kappa \in \mathscr{T}_h, \dim_{\mathrm{H}}(\kappa \cap \partial\Omega) = n-1\}. \end{aligned}$$

Furthermore, we use Γ_{int} to denote the union of all faces $e \in \mathscr{E}_h$ such that $\dim_{\mathrm{H}}(e \cap \partial\Omega) < n-1$.

Let $h_\kappa = \operatorname{diam}(\kappa)$ for all $\kappa \in \mathscr{T}_h$ and $h_e = \operatorname{diam}(e)$ for all $e \in \mathscr{E}_h$.We denote by $\mathsf{h}(x)$ the local mesh size defined as a piecewise constant function defined as $\mathsf{h}(x) = h_\kappa$, $x \in \operatorname{int}(\kappa)$, and $\mathsf{h}(x) = h_e$, $x \in e$.

ASSUMPTION 2.1 (Mesh quality) We assume throughout that the family $(\mathscr{T}_h)_{h\in(0,1]}$ satisfies the following conditions:

(a) Shape regularity: There exist $C_1, C_2 > 0$ such that

$$C_1 h_\kappa^n \leqslant |\kappa| \leqslant C_2 h_\kappa^n \quad \forall\, \kappa \in \mathscr{T}_h, \ \ \forall\, h \in (0, 1].$$

(b) Contact regularity: There exists a constant $C_1 > 0$ such that

$$C_1 h_\kappa^{n-1} \leqslant \mathscr{H}^{n-1}(e) \quad \forall\, e \in \mathscr{E}_h, \kappa \in \mathscr{T}_h \text{ s.t. } e \subset \bar{\kappa}, \quad \forall\, h \in (0, 1].$$

In particular, we have $h_e \approx h_\kappa$ under the above condition.

(c) Submesh condition: There exists a regular, conforming, simplicial submesh $\widetilde{\mathscr{T}}_h$ (without hanging nodes, edges, etc.) such that

1. for each $\tilde{\kappa} \in \widetilde{\mathscr{T}}_h$, there exists a $\kappa \in \mathscr{T}_h$ such that $\tilde{\kappa} \subset \kappa$;
2. the family $(\widetilde{\mathscr{T}}_h)_{h \in (0,1]}$ satifies (a) and (b) and
3. there exists a constant $\tilde{c}$ such that whenever $\tilde{\kappa} \subset \kappa$, $h_\kappa \leqslant \tilde{c} h_{\tilde{\kappa}}$.

REMARK 2.2 The existence of a simplicial submesh is an entirely technical assumption which may be tedious to verify in practice. We have included it since it seemed a fairly general assumption under which we were able to prove the required results. We note also that in dimension $n = 2, 3$, such a submesh can be constructed under fairly mild assumptions on the partition $\mathscr{T}_h$ (Brenner, 2003, Corollary 7.3). In fact, it seems straightforward to generalize this proof to arbitrary dimensions.

LEMMA 1 There exists a constant C, independent of h, such that

$$\sharp\{e \in \mathscr{E}_h : e \subset \kappa\} \leqslant C \quad \forall\, \kappa \in \mathscr{T}_h, \quad \forall\, h \in (0, 1].$$

Proof. Let $\kappa \in \mathscr{T}_h$ and let $E \subset \mathscr{E}_h$ be the set of faces contained in κ. Using Assumptions 2.1(a) and 2.1(b), we have

$$\sharp E h_\kappa^{n-1} \approx \sum_{e \in E} h_e^{n-1} \approx \sum_{e \in E} \mathscr{H}^{n-1}(e) = \mathscr{H}^{n-1}(\partial\kappa) \approx h_\kappa^{n-1}.$$

Upon dividing by h_κ^{n-1}, we obtain $\sharp E \approx 1$. □

2.2 *Broken Sobolev spaces and DGFE spaces*

Let $p \in [1, \infty)$. We will use standard Sobolev spaces $W^{1,p}(\Omega)$ and L^p spaces $L^p(\Omega)$ with their corresponding norms, with a self-evident notation. The broken Sobolev space $W^{1,p}(\mathscr{T}_h)$ is defined by

$$W^{1,p}(\mathscr{T}_h) = \{u \in L^1(\Omega) : u|_\kappa \in W^{1,p}(\kappa) \text{ for all } \kappa \in \mathscr{T}_h\}.$$

The dual index is denoted by $p' = p/(p-1)$. The Sobolev index appearing in the Sobolev embedding theorems (see Adams & Fournier, 2003) is denoted by $p^* = np/(n-p)$ if $p < n$ and $p^* = \infty$ if $p \geqslant n$. We recall that $W^{1,p}(\Omega) \subset L^q(\Omega)$, $q \in [1, p^*] \setminus \{+\infty\}$, and this embedding is compact for all $q < p^*$ (Adams & Fournier, 2003).

The subspace of discontinuous finite-element functions of polynomial degree no higher than k is defined as

$$S^k(\mathscr{T}_h) = \{u \in L^1(\Omega) : u|_\kappa \in P^k \text{ for all } \kappa \in \mathscr{T}_h\},$$

where P^k denotes the space of polynomials of degree k in $\mathbb{R}^n$. For each face $e \in \mathscr{E}_h$, $e \subset \Gamma_{\text{int}}$, we denote by κ^+ and κ^- its neighbouring elements. We write ν^+, ν^- to denote the outward normal unit vectors to the boundaries $\partial\kappa^\pm$, respectively. The jump of a vector-valued function $\varphi \in W^{1,1}(\mathscr{T}_h)^m$ and the average of a matrix-valued function $\varphi \in W^{1,1}(\mathscr{T}_h)^{m\times n}$ with traces $\varphi = \varphi^{\pm}$ from $\kappa^{\pm}$ are, respectively, defined as

$$[\![\varphi]\!] = \varphi^{+} \otimes \nu^{+} + \varphi^{-} \otimes \nu^{-} \quad \text{and}$$

$$\{\varphi\} = \tfrac{1}{2}(\varphi^{+} + \varphi^{-}).$$

For $u \in W^{1,p}(\mathscr{T}_h)^m$, we define the broken Sobolev seminorms:

$$|u|^p_{W^{1,p}(\mathscr{T}_h)} = \|\nabla u\|^p_{L^p(\Omega)} + \int_{\Gamma_{\mathrm{int}}} \mathsf{h}^{1-p} |[\![u]\!]|^p \,\mathrm{d}s,$$

$$|u|^p_{W^{1,p}_{\mathrm{D}}(\mathscr{T}_h)} = |u|^p_{W^{1,p}(\mathscr{T}_h)} + \int_{\Gamma_{\mathrm{D}}} \mathsf{h}^{1-p} |u|^p \,\mathrm{d}s.$$

Next, we recall some important facts about the Banach space $\mathrm{BV}(\Omega)^m$ of functions of bounded variation which contains the spaces $W^{1,p}(\mathscr{T}_h)^m$. The space is equipped with the norm

$$\|u\|_{\mathrm{BV}} = \|u\|_{L^1(\Omega)} + |Du|(\Omega),$$

where Du is the measure representing the distributional derivative of u and $|Du|(\Omega)$ is its total variation, defined by

$$|Du|(\Omega) = \sup_{\substack{\varphi \in C^1_{\mathrm{c}}(\Omega)^{m\times n}, \\ \|\varphi\|_{L^\infty} \leqslant 1}} \int_{\Omega} u \cdot \mathrm{div}\varphi \,\mathrm{d}x.$$

The symbol $C^1_{\mathrm{c}}(\Omega)$ denotes the space of continuously differential functions with compact support in Ω. Here and throughout, we use $a \cdot b$ to denote the usual euclidean inner product of either vectors or matrices a, b of the same dimensions. Weak-$*$ compactness of bounded sets and many other properties of the space $\mathrm{BV}(\Omega)$ will play an important role in our analysis.

The variation (distributional derivative) of a broken Sobolev function $u \in W^{1,p}(\mathscr{T}_h)^m$ is given by the following formula, which can be easily verified using integration by parts on every element of the mesh:

$$-\int_{\Omega} u \cdot \mathrm{div}\varphi \,\mathrm{d}x = \int_{\Omega} \nabla u \cdot \varphi \,\mathrm{d}x - \int_{\Gamma_{\mathrm{int}}} [\![u]\!] \cdot \varphi \,\mathrm{d}s \quad \forall \varphi \in C^1_{\mathrm{c}}(\Omega)^{m\times n}. \tag{2.1}$$

The following result is the starting point to lift results for the space BV to DGFE spaces.

LEMMA 2 There exists a constant C, independent of h and p, such that for all $p \in [1, \infty)$,

$$|Du|(\Omega) \leqslant C|u|_{W^{1,p}(\mathscr{T}_h)} \quad \forall\, u \in W^{1,p}(\mathscr{T}_h)^m, \ \ \forall\, h \in (0, 1].$$

Proof. The proof is a straightforward generalization of Lew *et al.* (2004, Theorem 3.26) to the case $p \neq 2$. For the sake of completeness, we include a brief sketch. The variation is bounded by

$$|Du|(\Omega) \leqslant \|\nabla u\|_{L^1(\Omega)} + \int_{\Gamma_{\mathrm{int}}} |[\![u]\!]| \mathrm{d}s.$$

Since $|\Omega| < +\infty$, we have $\|\nabla u\|_{L^1(\Omega)} \leqslant |\Omega|^{1-1/p}\|\nabla u\|_{L^p(\Omega)}$. We can use Hölder's inequality and Assumption 2.1 to estimate

$$\begin{aligned}\int_{\Gamma_{\mathrm{int}}} |[\![u]\!]| \mathrm{d}s &= \int_{\Gamma_{\mathrm{int}}} \mathsf{h}^{1/p'}\, \mathsf{h}^{(1-p)/p} |[\![u]\!]| \mathrm{d}s \\ &\leqslant \left(\int_{\Gamma_{\mathrm{int}}} \mathsf{h}\, \mathrm{d}s\right)^{1/p'} \left(\int_{\Gamma_{\mathrm{int}}} \mathsf{h}^{1-p} |[\![u]\!]|^p\, \mathrm{d}s\right)^{1/p} \\ &\lesssim \left(\sum_{e\subset\Gamma_{\mathrm{int}}} h_e^n\right)^{1/p'} \left(\int_{\Gamma_{\mathrm{int}}} \mathsf{h}^{1-p} |[\![u]\!]|^p\, \mathrm{d}s\right)^{1/p}.\end{aligned}$$

By Assumption 2.1 as well as Lemma 1, we have

$$\sum_{e\subset\Gamma_{\mathrm{int}}} h_e^n \lesssim \sum_{e\subset\Gamma_{\mathrm{int}}} \sum_{\substack{\kappa\in\mathscr{T}_h,\\ e\subset\kappa}} h_\kappa^n \lesssim \sum_{\kappa\in\mathscr{T}_h} h_\kappa^n \approx |\Omega|,$$

which gives the result. □

We conclude this section with an approximation result.

LEMMA 3 Suppose $u \in W^{1,p}(\Omega)^m$ for some $p \in [1,\infty)$; then for each $h \in (0,1]$, there exists $u_h \in S^1(\mathscr{T}_h)^m$ such that

$$\|u - u_h\|_{L^p(\Omega)} + |u - u_h|_{W^{1,p}(\mathscr{T}_h)} \to 0 \quad \text{as } h \to 0.$$

Proof. Since Ω is assumed to be a Lipschitz domain, it follows that $C^\infty(\bar{\Omega})^m$ is dense in $W^{1,p}(\Omega)^m$ and hence we may assume without loss of generality that $u \in C^\infty(\bar{\Omega})^m$. For such a smooth function, this result follows from standard polynomial approximation theory (Ciarlet, 1978). □

3. Reconstruction operator

As is the case in many works on DG methods, ranging from *a posteriori* error estimation (Karakashian & Pascal, 2003) to the proof of broken Poincaré-type inequalities (Brenner, 2003, 2004; Lasis & Süli, 2003; Ortner & Süli, 2007), we require at several points a continuous reconstruction operator. In this section, we will make use of the assumption that there exists a regular simplicial submesh of $\mathscr{T}_h$ (see Assumption 2.1(c)).

Our goal is to define a family of quasi-interpolation operators $Q_h : S^k(\mathscr{T}_h)^m \to W^{1,\infty}(\Omega)^m$ and provide localized error estimates for $Q_h u - u$ in L^q-norms, $q \in [1,\infty)$. Our results are more general than previous ones in that we consider arbitrary Sobolev indices, but weaker than those in Brenner (2003), for example, since we restrict ourselves to a fixed polynomial degree. In fact, our proofs do not carry over to arbitrary $W^{1,p}(\mathscr{T}_h)$ functions in an obvious way since we make use of local inverse inequalities. The idea of using quasi-interpolation operators was inspired by Lasser & Toselli (2003).

In order to simplify the notation, our discussion in this section is for scalar functions only. The corresponding results for vector-valued functions follow trivially.

3.1 *Local projection operators*

Let us first introduce some notation for the submesh $\widetilde{\mathscr{T}}_h$ (see Assumption 2.1(c)). We denote by $\widetilde{\mathscr{N}_h}$ the set of nodes of $\widetilde{\mathscr{T}}_h$ and by $\widetilde{\mathscr{N}_h^0}$ the subset of internal nodes. For every $z \in \widetilde{\mathscr{N}_h}$, we define the star-shaped patch

$$\widetilde{T}_z = \bigcup \{\widetilde{\kappa} \in \widetilde{\mathscr{T}}_h : z \in \widetilde{\kappa}\}, \tag{3.1}$$

and we set $h_z = \mathrm{diam}(\widetilde{T}_z)$. Due to the assumptions on the submesh $\widetilde{\mathscr{T}}_h$, it is clear that $\widetilde{T}_z$ contains a finite number of elements which is independent of the mesh size.

Next, we establish the existence of linear maps $\pi_z : \mathrm{BV}(\Omega) \to \mathbb{R},\ z \in \widetilde{\mathscr{N}_h}$, such that

$$\|u - \pi_z(u)\|_{L^1(\widetilde{T}_z)} \leqslant C h_z |Du|(\widetilde{T}_z) \quad \forall z \in \widetilde{\mathscr{N}_h}, \ \ \forall u \in \mathrm{BV}(\Omega), \tag{3.2}$$

where C is independent of h and z. To achieve this, we have to distinguish between the cases when z lies on the boundary $\partial\Omega$ and in the interior of the domain Ω. If $z \in \widetilde{\mathscr{N}_h^0}$, i.e. $z \in \mathrm{int}(\Omega)$, let $B_z = B(z, \rho_z)$, where $\rho_z = \min_{x \in \partial \widetilde{T}_z} |x - z|_2$ such that $B_z \subset \widetilde{T}_z$. From Assumption 2.1(c), it follows that $\rho_z \approx h_z$. Setting $\pi_z(u) = (u)_{B_z}$ (the mean value over the ball B_z), we obtain the following result. We note that our construction as well as the proofs of the estimates are only minor modifications of the L^2 case treated by Verfürth (1999, Lemma 4.1).

LEMMA 4 Let $K \subset \mathbb{R}^n$ be star shaped with respect to the point $x_0 \in K$ and define

$$\rho_1 = \inf_{x \in \partial K} |x - x_0|_2 \quad \text{and} \quad \rho_2 = \sup_{x \in \partial K} |x - x_0|_2.$$

There exists a constant C depending only on ρ_2/ρ_1 and n such that

$$\|u\|_{L^1(K)} \leqslant C(\rho_2/\rho_1)(\|u\|_{L^1(B)} + \rho_1 |Du|(K)) \quad \forall\, u \in \mathrm{BV}(K), \tag{3.3}$$

where $B = B(x_0, \rho_1)$, and

$$\|u - (u)_B\|_{L^1(K)} \leqslant C(\rho_2/\rho_1)\rho_1 |Du|(K) \quad \forall\, u \in \mathrm{BV}(K). \tag{3.4}$$

Since the proof of this lemma is technical, we postpone it to the appendix.

We note that Lemma 4 together with Assumption 2.1(c) (shape regularity of the submesh $\widetilde{\mathscr{T}}_h$) immediately implies (3.2) for interior nodes.

If z lies at the boundary, we define h_z as before but we now set

$$\rho_z = \inf_{x \in \partial \widetilde{T}_z \setminus \partial\Omega} |z - x|_2.$$

Let $\widetilde{B}_z = B(z, \rho_z) \cap \widetilde{T}_z = B(z, \rho_z) \cap \bar{\Omega}$. Repeating the proof of Lemma 4 verbatim, we obtain

$$\|v\|_{L^1(\widetilde{T}_z)} \leqslant C \left(\|v\|_{L^1(\widetilde{B}_z)} + h_z |Dv|(\widetilde{T}_z) \right) \quad \forall\, v \in \mathrm{BV}(\widetilde{T}_z). \tag{3.5}$$

Since $\widetilde{B}_z$ is not necessarily convex, we apply a further reduction to the first term on the right-hand side of (3.5). Since $\partial\Omega$ is Lipschitz continuous, there exists a cone $\mathscr{C}$ with positive opening angle α, which can be chosen independently of z, and apex 0 such that $(z + \mathscr{C}) \cap B(z, \varepsilon) \subset \mathbb{R}^n \setminus \widetilde{T}_z$ for some $\varepsilon > 0$. Let $a \in \mathbb{R}^n$, $|a|_2 = \rho_z/2$, be the direction of the axis of the cone C pointing into $\widetilde{T}_z$ and define $z' = z + a$. It can be easily seen that $\widetilde{B}_z$ is star shaped with respect to z' and that there exists a value $r_0 \in (0, 1/2]$ which depends only on α such that $B_z := B(z', r_0\rho_z) \subset \widetilde{B}_z \subset \widetilde{T}_z$. Hence, we may define $\pi_z(u) = (u)_{B_z}$ again (but note that B_z is defined differently now) to obtain the following result.

LEMMA 5 For $z \in \widetilde{\mathscr{N}_h}$ and $u \in \mathrm{BV}(\Omega)$, let $\pi_z(u) = (u)_{B_z}$, where B_z is defined as in the above discussion. Then, (3.2) holds with a constant C independent of the mesh size.

Proof. For interior vertices, we have already shown that (3.2) holds with a constant depending only on h_z/ρ_z, which measures mesh quality, and it remains to prove a similar bound for boundary vertices.

Using (3.5) with $v = u - \pi_z(u)$, we have

$$\|u - \pi_z(u)\|_{L^1(\widetilde{T}_z)} \lesssim \|u - \pi_z(u)\|_{L^1(\widetilde{B}_z)} + h_z|Du|(\widetilde{T}_z).$$

We now apply Lemma 4 with $K = \widetilde{B}_z$, $B = B_z$, $h = \rho_z$ and $\rho = r_0\rho_z$ to obtain

$$\|u - \pi_z(u)\|_{L^1(\widetilde{B}_z)} \lesssim h_z|Du|(\widetilde{B}_z).$$

Combining this estimate with the previous formula, we obtain

$$\|u - \pi_z(u)\|_{L^1(\widetilde{T}_z)} \lesssim h_z|Du|(\widetilde{T}_z).$$ □

3.2 *Construction and analysis of Q_h*

Finally, we are in a position to define and analyse the reconstruction operator. For each $h \in (0, 1]$, let $Q_h : S^k(\mathscr{T}_h) \to W^{1,\infty}(\Omega)$ be the linear operator defined by

$$Q_h u = \sum_{z \in \widetilde{\mathscr{N}_h}} \pi_z(u)\lambda_z, \tag{3.6}$$

where λ_z is the standard P^1 nodal basis function on the mesh $\widetilde{\mathscr{T}}_h$ associated with the vertex z.

For later use, we define for each $z \in \widetilde{\mathscr{N}_h}$, $\kappa \in \mathscr{T}_h$ and $e \in \mathscr{E}_h$:

$$T_z = \bigcup\{\kappa \in \mathscr{T}_h : z \subset \kappa\}, \quad T_\kappa = \bigcup\{T_z : z \subset \kappa\} \quad \text{and} \quad T_e = \bigcup\{T_\kappa : e \subset \kappa\}.$$

Furthermore, for $A \subset \Omega$, we define the notation

$$\mathscr{T}_h \cap A = \{\kappa \in \mathscr{T}_h : \kappa \subset A\}.$$

Since $\widetilde{\mathscr{T}}_h$ is a submesh of $\mathscr{T}_h$, we have that $T_z \supset \widetilde{T}_z$, where $\widetilde{T}_z$ is as defined in (3.1). If we denote by $\mathscr{K}_\kappa$ the number of elements $\kappa' \in \mathscr{T}_h \cap T_\kappa$, due to Assumption 2.1(b) (contact regularity), it follows that $\mathscr{K}_\kappa$ is bounded independent of h and κ. Together with Assumption 2.1(c), this implies that

$$h_z = \mathrm{diam}(\widetilde{T}_z) \approx \mathrm{diam}(T_z) \approx \max_{\kappa, z \subset \kappa} \mathrm{diam}(T_\kappa)$$

and also

$$\mathrm{diam}(T_\kappa) \approx \min_{\kappa' \subset T_\kappa} h_{\kappa'} \approx h_\kappa.$$

THEOREM 3.1 Fix $p, q \in [1, \infty)$. The reconstruction operator Q_h defined in (3.6) satisfies the local estimates for all $u \in S^k(\mathscr{T}_h)$,

$$\|u - Q_h u\|_{L^q(\kappa)} \lesssim h_\kappa^{\frac{n}{q}-\frac{n}{p}+1} |u|_{W^{1,p}(\mathscr{T}_h \cap T_\kappa)} \quad \forall \kappa \in \mathscr{T}_h, \tag{3.7}$$

$$\|u - Q_h u\|_{L^q(e)} \lesssim h_e^{\frac{(n-1)}{q}-\frac{n}{p}+1} |u|_{W^{1,p}(\mathscr{T}_h \cap T_e)} \quad \forall e \in \mathscr{E}_h \setminus \Gamma_{\text{int}}, \tag{3.8}$$

$$\|\nabla Q_h u\|_{L^p(\kappa)} \lesssim |u|_{W^{1,p}(\mathscr{T}_h \cap T_\kappa)} \quad \forall \kappa \in \mathscr{T}_h. \tag{3.9}$$

Furthermore, for $q \in [p, p^*] \setminus \{\infty\}$, we have the global estimates

$$\|u - Q_h u\|_{L^q(\Omega)} \lesssim h^{\frac{n}{q}-\frac{n}{p}+1} |u|_{W^{1,p}(\mathscr{T}_h)} \quad \text{and} \tag{3.10}$$

$$\|\nabla Q_h u\|_{L^p(\Omega)} \lesssim |u|_{W^{1,p}(\mathscr{T}_h)}, \tag{3.11}$$

where h denotes the global mesh size.

Proof. Fix $q \in [1, \infty)$. For each $z \in \widetilde{\mathscr{N}_h}$, we use Lemma A1 to obtain

$$\|u - \pi_z(u)\|_{L^q(\widetilde{T}_z)} \approx h_z^{\frac{n}{q}-n} \|u - \pi_z(u)\|_{L^1(\widetilde{T}_z)}.$$

Our local projection result Lemma 5 gives

$$\begin{aligned}\|u - \pi_z(u)\|_{L^q(\widetilde{T}_z)} &\lesssim h_z^{\frac{n}{q}-n+1} |Du|(\widetilde{T}_z) \\ &\lesssim h_z^{\frac{n}{q}-n+1} \|\nabla u\|_{L^1(T_z)} + h_z^{\frac{n}{q}-n+1} \sum_{e \in \mathscr{E}_h \cap T_z} \int_e |[\![u]\!]| \mathrm{d}s.\end{aligned}$$

For the bulk term $\|\nabla u\|_{L^1(T_z)}$, we use Lemma A1 and for the surface term we use Hölder's inequality (as in the proof of Lemma 2) to deduce

$$\begin{aligned}\|u - \pi_z(u)\|_{L^q(\widetilde{T}_z)} &\lesssim h_z^{\frac{n}{q}-\frac{n}{p}+1} \|\nabla u\|_{L^p(T_z)} + h_z^{\frac{n}{q}-\frac{n}{p}+1} \left(\sum_{e \in \mathscr{E}_h \cap T_z} h_e^{1-p} \int_e |[\![u]\!]|^p \,\mathrm{d}s \right)^{1/p} \\ &\lesssim h_z^{\frac{n}{q}-\frac{n}{p}+1} |u|_{W^{1,p}(\mathscr{T}_h \cap T_z)}.\end{aligned} \tag{3.12}$$

We now prove the local estimate (3.7). Using the fact that the hat functions $\{\lambda_z\}_{z \in \widetilde{\mathscr{N}_h}}$ form a partition of unity, we have

$$\|u - Q_h u\|_{L^q(\kappa)}^q = \left\| \sum_{z \in \widetilde{\mathscr{N}_h} \cap \kappa} (u - \pi_z(u)) \lambda_z \right\|_{L^q(\kappa)}^q.$$

Rearranging terms and recalling that $\|\lambda_z\|_{L^\infty(\Omega)} = 1$ and $\lambda_z = 0$ outside $\widetilde{T}_z$, we compute

$$\|u - Q_h u\|_{L^q(\kappa)}^q \lesssim \sum_{z \in \widetilde{\mathscr{N}_h} \cap \kappa} \|u - \pi_z(u)\|_{L^q(\kappa \cap \widetilde{T}_z)}^q \lesssim \sum_{z \in \widetilde{\mathscr{N}_h} \cap \kappa} \|u - \pi_z(u)\|_{L^q(\widetilde{T}_z)}^q.$$

Using (3.12), we obtain

$$\|u - Q_h u\|_{L^q(\kappa)}^q \lesssim \sum_{z \in \widetilde{\mathscr{N}_h} \cap \kappa} h_z^{q\left(\frac{n}{q} - \frac{n}{p} + 1\right)} |u|_{W^{1,p}(\mathscr{T}_h \cap T_z)}^q.$$

Rearranging terms, using the definition of T_κ and recalling that the cardinality of $\widetilde{\mathscr{N}_h} \cap \kappa$ is uniformly bounded,

$$\|u - Q_h u\|_{L^q(\kappa)} \lesssim h_\kappa^{\frac{n}{q} - \frac{n}{p} + 1} \left(\sum_{z \in \widetilde{\mathscr{N}_h} \cap \kappa} |u|_{W^{1,p}(\mathscr{T}_h \cap T_z)}^q \right)^{1/q} \lesssim h_\kappa^{\frac{n}{q} - \frac{n}{p} + 1} |u|_{W^{1,p}(\mathscr{T}_h \cap T_\kappa)},$$

which concludes the proof of (3.7).

If $e \in \mathscr{E}_h \cap \partial\Omega$, then

$$\|u - Q_h u\|_{L^q(e)} \leqslant \sum_{z \in \widetilde{\mathscr{N}_h} \cap e} \|u - \pi_z(u)\|_{L^q(e \cap \widetilde{T}_z)}.$$

The set $e \cap \widetilde{T}_z$ is a union of faces of elements in $\widetilde{\mathscr{T}}_h$. We can therefore use the local inverse estimate

$$\|u - \pi_z(u)\|_{L^q(e \cap \widetilde{T}_z)}^q \lesssim h_z^{-1} \|u - \pi_z(u)\|_{L^q(\widetilde{T}_z)}^q,$$

after which proceed as above to obtain (3.8). The third local estimate (3.9) follows along the same lines.

To prove the first global estimate (3.10), we assume $q \in [p, p^*]$, $q \neq \infty$. It then holds that $\frac{n}{q} - \frac{n}{p} + 1 \geqslant 0$, and we set $h^* = h^{\frac{n}{q} - \frac{n}{p} + 1}$ (recall that h is the global mesh size). We sum (3.7) (to power q) over $\kappa \in \mathscr{T}_h$ to obtain

$$\begin{aligned}\|u - Q_h u\|_{L^q(\Omega)}^q &\lesssim (h^*)^q \sum_{\kappa \in \mathscr{T}_h} \left(\|\nabla u\|_{L^p(T_\kappa)}^p + \int_{\Gamma_{\text{int}} \cap T_\kappa} \mathsf{h}^{1-p} |[\![u]\!]|^p \, \mathrm{d}s \right)^{q/p} \\ &\lesssim (h^*)^q \left(\sum_{\kappa \in \mathscr{T}_h} \left[\|\nabla u\|_{L^p(T_\kappa)}^p + \int_{\Gamma_{\text{int}} \cap T_\kappa} \mathsf{h}^{1-p} |[\![u]\!]|^p \, \mathrm{d}s \right] \right)^{q/p},\end{aligned}$$

where we used the fact $\sum |a_i|^\alpha \leqslant \left(\sum |a_i|\right)^\alpha$ for $\alpha \geqslant 1$. Finally, we note that due to Lemma 1, each element κ appears only in finitely many sets $T_{\kappa'}$ and thus, taking the qth root, we obtain the result.

The second global estimate can be proved in the same way. □

4. Broken embedding theorems

4.1 *Poincaré inequalities*

In this section, we prove broken Sobolev–Poincaré inequalities for any $p \in [1, n)$. Similar results were previously derived by Lasis & Süli (2003) for $p = 2$. The idea in our proof is the same as in the proof of Theorem 3.1 to use the known results in $\mathrm{BV}(\Omega)$ and the Sobolev spaces $W^{1,p}(\Omega)$ together with local norm equivalence and the reconstruction operator.

THEOREM 4.1 (Sobolev–Poincaré inequalities) Let $p < n$ and let $p^* = np/(n-p)$. There exists a constant C_S such that

$$\|u - (u)_\Omega\|_{L^{p^*}(\Omega)} \leqslant C_S |u|_{W^{1,p}(\mathscr{T}_h)} \quad \forall\, u \in S^k(\mathscr{T}_h)^m, \quad \forall\, h \in (0,1]. \tag{4.1}$$

In particular, it holds that

$$\|u\|_{L^{p^*}(\Omega)} \leqslant C_S \left(\|u\|_{L^1(\Omega)} + |u|_{W^{1,p}(\mathscr{T}_h)}\right) \quad \forall\, u \in S^k(\mathscr{T}_h)^m, \quad \forall\, h \in (0,1]. \tag{4.2}$$

Proof. Let $v = u - (u)_\Omega$. It is easy to see that $Q_h w = w$ if w is a constant function. Hence, it follows that $Q_h v = Q_h u - (u)_\Omega$ and

$$\|v\|_{L^{p^*}(\Omega)} \leqslant \|v - Q_h v\|_{L^{p^*}(\Omega)} + \|Q_h v - (Q_h v)_\Omega\|_{L^{p^*}(\Omega)} + \|(Q_h v)_\Omega\|_{L^{p^*}(\Omega)}. \tag{4.3}$$

For the first term on the right-hand side of (4.3), we use Theorem 3.1 to estimate

$$\|v - Q_h v\|_{L^{p^*}(\Omega)} \lesssim |v|_{W^{1,p}(\mathscr{T}_h)}.$$

For the second term on the right-hand side of (4.3), we employ the Poincaré–Sobolev inequality for $W^{1,p}(\Omega)^m$ and (3.11) to obtain

$$\|Q_h v - (Q_h v)_\Omega\|_{L^{p^*}(\Omega)} \lesssim \|\nabla Q_h v\|_{L^p(\Omega)} \lesssim |v|_{W^{1,p}(\mathscr{T}_h)}.$$

For the last term, we note that $\|(Q_h v)_\Omega\|_{L^{p^*}(\Omega)} \lesssim \|Q_h v\|_{L^1(\Omega)}$ and

$$\begin{aligned}\|Q_h v\|_{L^1(\Omega)} &\leqslant \|Q_h v - v\|_{L^1(\Omega)} + \|v\|_{L^1(\Omega)} \\ &\lesssim h|v|_{W^{1,1}(\mathscr{T}_h)} + |Dv|(\Omega),\end{aligned}$$

where we used Theorem 3.1 on the first term and the Poincaré inequality for $\mathrm{BV}(\Omega)$ on the second term on the right-hand side.

Using our estimate in Lemma 2, we deduce that $|Dv|(\Omega) = |Du|(\Omega) \lesssim |u|_{W^{1,p}(\mathscr{T}_h)}$, and we can combine our estimates to give the first result.

The second result follows immediately from $\|(u)_\Omega\|_{L^{p^*}(\Omega)} \lesssim \|u\|_{L^1(\Omega)}$. □

4.2 *Trace theorem*

We first recall some facts about traces of functions of bounded variation. The following result summarizes Theorems 1 and 2 in Evans & Gariepy (1992, Section 5.3).

THEOREM 4.2 Let Ω be a Lipschitz domain in $\mathbb{R}^n$. There exists a bounded, linear operator $T : \mathrm{BV}(\Omega)^m \to L^1(\partial\Omega)^m$ (we write $Tu = u$) such that

$$\int_\Omega u \cdot \mathrm{div}\varphi \,\mathrm{d}x = -\int_\Omega \varphi \cdot \mathrm{d}Du + \int_{\partial\Omega} (u \otimes \nu) \cdot \varphi \,\mathrm{d}s \quad \forall u \in \mathrm{BV}(\Omega)^m, \quad \forall\, \varphi \in C^1(\mathbb{R}^n)^{m\times n},$$

where ν is the unit outward normal to $\partial\Omega$.

If $u \in \mathrm{BV}(\Omega)$, then for $\mathscr{H}^{n-1}$ almost every $x \in \partial\Omega$, the identity

$$Tu(x) = \lim_{r\to 0} ⨍_{B(x,r)\cap\Omega} u \,\mathrm{d}x \tag{4.4}$$

holds.

First, we note that identity (4.4) immediately implies a Friedrichs inequality for $\mathrm{BV}(\Omega)$ and therefore, by Theorem 4.1, a broken Sobolev–Poincaré inequality with respect to a broken norm which penalizes boundary values.

LEMMA 6 (Friedrichs inequality for BV) Let $u \in \mathrm{BV}(\Omega)$ and let Γ_{D} be a subset of $\partial\Omega$ with positive surface measure. Then, there exists a constant C_F such that

$$\|u\|_{L^1(\Omega)} \leqslant C_F \left(|Du|(\Omega) + \int_{\Gamma_{\mathrm{D}}} |u| \mathrm{d}s \right) \quad \forall u \in \mathrm{BV}(\Omega).$$

Proof. We use the standard compactness technique to prove this result. For contradiction, suppose that no such constant C_F exists. Then, there exists a sequence $u_j \in \mathrm{BV}(\Omega)$ such that $\|u_j\|_{L^1(\Omega)} = 1$ and $|Du_j|(\Omega) + \|u_j\|_{L^1(\Gamma_{\mathrm{D}})} \to 0$ as $j \to \infty$. Since $\|u_j\|_{\mathrm{BV}}$ is bounded, there exists a subsequence (not relabelled) and $u \in \mathrm{BV}(\Omega)$ such that $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega)$. Since this implies $u_j \to u$ strongly in $L^1(\Omega)$, it follows that $\|u\|_{L^1(\Omega)} = 1$. Since the functional $v \mapsto |Dv|(\Omega) + \|v\|_{L^1(\Gamma_{\mathrm{D}})}$ is convex and strongly continuous, it is also lower semicontinuous with respect to weak-$*$ convergence. Therefore, $|Du|(\Omega) = 0$, which implies that u is constant in Ω. Since $\|u\|_{L^1(\Gamma_{\mathrm{D}})} = 0$, the trace of u at Γ_{D} vanishes which means that $u = 0$ and contradicts the assumption that $\|u\|_{L^1(\Omega)} = 1$. □

COROLLARY 4.3 (Broken Friedrichs-type inequality) Let $p \in [1, n)$ and suppose that $\Gamma_{\mathrm{D}} \subset \partial\Omega$ has positive surface measure. Then, there exists a constant C_{BF}, independent of h, such that

$$\|u\|_{L^{p^*}(\Omega)} \leqslant C_{\mathrm{BF}} \left(\|u\|_{L^p(\Gamma_{\mathrm{D}})} + |u|_{W^{1,p}(\mathscr{T}_h)} \right) \quad \forall u \in S^k(\mathscr{T}_h)^m, \ \forall h \in (0, 1].$$

Proof. Using Theorem 4.1 and Lemmas 2 and 6, we obtain

$$\begin{aligned} \|u\|_{L^{p^*}(\Omega)} &\lesssim \|u\|_{L^1(\Omega)} + |u|_{W^{1,p}(\mathscr{T}_h)} \\ &\lesssim \|u\|_{L^1(\Gamma_{\mathrm{D}})} + |Du|(\Omega) + |u|_{W^{1,p}(\mathscr{T}_h)} \\ &\lesssim \|u\|_{L^p(\Gamma_{\mathrm{D}})} + |u|_{W^{1,p}(\mathscr{T}_h)}. \end{aligned}$$

□

One may argue that, strictly speaking, Lemma 4.3 is a Poincaré-type inequality. However, we chose to label it a Friedrichs-type inequality since it trivially implies

$$\|u\|_{L^{p^*}(\Omega)} \leqslant C'_{\mathrm{BF}} |u|_{W^{1,p}_{\mathrm{D}}(\mathscr{T}_h)}. \tag{4.5}$$

THEOREM 4.4 (Broken trace theorem) Let $p \in (1, n]$ and set $q = p(n-1)/(n-p)$ $\left(\text{i.e. } q \text{ satifies } \frac{(n-1)}{p} - \frac{(n-1)}{q} = 1 - \frac{1}{p}\right)$. There exists a constant C_{BT}, independent of h, such that

$$\|u\|_{L^q(\partial\Omega)} \leqslant C_{\mathrm{BT}} \left(\|u\|_{L^1(\Omega)} + |u|_{W^{1,p}(\mathscr{T}_h)} \right) \quad \forall u \in S^k(\mathscr{T}_h)^m, \ \forall h \in (0, 1]. \tag{4.6}$$

Proof. Summing qth powers of (3.8) over the faces on $\partial\Omega$, we obtain the following:

$$\|u\|^q_{L^q(\partial\Omega)} \lesssim \|Q_h u\|^q_{L^q(\partial\Omega)} + \sum_{e\in\mathscr{E}_h, e\subset\partial\Omega} h_\kappa^{n-1-\frac{nq}{p}+q} |u|^q_{W^{1,p}(\mathscr{T}_h\cap T_e)}.$$

For the choice $q = p(n-1)/(n-p)$, we have $n - 1 - nq/p + q = 0$ and furthermore, $q/p \geqslant 1$. The latter property can be used to estimate

$$\sum_{i=1}^J |a_j|^{q/p} \leqslant \left(\sum_{i=1}^J |a_j|\right)^{q/p}.$$

Hence, we can estimate further

$$\begin{aligned}\|u\|^q_{L^q(\partial\Omega)} &\lesssim \|Q_h u\|^q_{L^q(\partial\Omega)} + \sum_{e\in\mathscr{E}_h, e\subset\partial\Omega} |u|^q_{W^{1,p}(\mathscr{T}_h\cap T_e)} \\ &\lesssim \|Q_h u\|^q_{L^q(\partial\Omega)} + \left(\sum_{e\in\mathscr{E}_h, e\subset\partial\Omega} |u|^p_{W^{1,p}(\mathscr{T}_h\cap T_e)}\right)^{q/p} \\ &\lesssim \|Q_h u\|^q_{L^q(\partial\Omega)} + |u|^q_{W^{1,p}(\mathscr{T}_h)}.\end{aligned}$$

The trace inequality (4.6) is obtained by employing the trace theorem (see for instance Theorem 6.4.1 in Kufner *et al.*, 1977) for $Q_h u$, the continuity property of Q_h and the estimate (3.11) of Theorem 3.1. □

5. Compactness in $W^{1,p}(\mathscr{T}_h)$

In this section, we will generalize the compactness properties of classical Sobolev spaces to broken Sobolev spaces. This requires a *consistent* discretization of the gradient.

Using integration by parts on each element, it can be easily seen that the distributional derivative Du of a broken Sobolev function is given by

$$\langle Du, \varphi\rangle = \int_\Omega \nabla u \cdot \varphi \,\mathrm{d}x - \int_{\Gamma_{\mathrm{int}}} [\![u]\!] \cdot \varphi \,\mathrm{d}s \quad \forall \varphi \in C^\infty_{\mathrm{c}}(\Omega)^{m\times n}.$$

In order to use compactness properties of Lebesgue spaces, we construct a bulk representation of the jump contribution. To this end, we choose a polynomial degree $l \geqslant 0$ and then define the lifting operator $R\colon W^{1,p}(\mathscr{T}_h)^m \to S^l(\mathscr{T}_h)^{m\times n}$ via

$$\int_\Omega R(u)\cdot\varphi \,\mathrm{d}x = -\int_{\Gamma_{\mathrm{int}}} [\![u]\!] \cdot \{\varphi\} \mathrm{d}s \quad \forall \varphi \in S^l(\mathscr{T}_h)^{m\times n}. \tag{5.1}$$

The polynomial degree l will later become a discretization parameter and can be chosen arbitrarily.

REMARK 5.1 We note that for the sake of the theory developed in this paper, the averages $\{\varphi\}$ in the right-hand side of the definition (5.1) can be replaced by any linear flux $\hat\varphi$ such that $\hat\varphi = \varphi$ whenever φ is continuous across all inter-element boundaries.

We first analyse the main features of the lifting operator. The left-hand side in (5.1) is an inner product on a finite-dimensional space (cf. also Lemma A2) while the right-hand side, for $u \in W^{1,p}(\mathscr{T}_h)^m$ fixed, is a linear functional on $S^l(\mathscr{T}_h)^{m\times n}$ and hence R is well defined. Next, we prove the boundedness of R in different broken Sobolev spaces.

LEMMA 7 Let $p \in [1, \infty)$. There exists a constant C_R such that

$$\|R(u)\|_{L^p(\Omega)} \leqslant C_R \left(\int_{\Gamma_{\rm int}} \mathsf{h}^{1-p} |[\![u]\!]|^p \,\mathrm{d}s\right)^{1/p} \quad \forall\, u \in W^{1,p}(\mathscr{T}_h)^m, \;\; \forall\, h \in (0,1].$$

Proof. For each $u \in W^{1,p}(\mathscr{T}_h)^m$ and for each $\varphi \in S^l(\mathscr{T}_h)^{m\times n}$, we have

$$\begin{aligned}\int_{\Gamma_{\rm int}} [\![u]\!] \cdot \{\varphi\} \mathrm{d}s &\leqslant \int_{\Gamma_{\rm int}} |\mathsf{h}^{-1/p'} [\![u]\!]|\, |h^{1/p'}\{\varphi\}| \mathrm{d}s \\ &\leqslant \left(\int_{\Gamma_{\rm int}} \mathsf{h}^{1-p} |[\![u]\!]|^p \,\mathrm{d}s\right)^{1/p} \left(\frac{1}{2^{p'}} \int_{\Gamma_{\rm int}} \mathsf{h}(|\varphi^+| + |\varphi^-|)^{p'} \,\mathrm{d}s\right)^{1/p'}.\end{aligned}$$

We can further bound the second term in the last estimate by

$$\begin{aligned}\int_{\Gamma_{\rm int}} \mathsf{h}(|\varphi^+| + |\varphi^-|)^{p'} \,\mathrm{d}s &\leqslant 2^{p'-1} \int_{\Gamma_{\rm int}} \mathsf{h}(|\varphi^+|^{p'} + |\varphi^-|^{p'}) \mathrm{d}s \\ &\lesssim \sum_{\kappa \in \mathscr{T}_h} \int_{\partial\kappa} \mathsf{h} |\varphi|^{p'} \,\mathrm{d}s \\ &\lesssim \sum_{\kappa \in \mathscr{T}_h} \int_{\kappa} |\varphi|^{p'} \,\mathrm{d}x.\end{aligned}$$

Thus, we have shown that

$$\int_{\Gamma_{\rm int}} [\![u]\!] \cdot \{\varphi\} \mathrm{d}s \leqslant C \left(\int_{\Gamma_{\rm int}} \mathsf{h}^{1-p} |[\![u]\!]|^p \,\mathrm{d}s\right)^{1/p} \|\varphi\|_{L^{p'}(\Omega)} \tag{5.2}$$

$$\forall\, u \in W^{1,p}(\mathscr{T}_h)^m, \;\; \forall\, \varphi \in S^l(\mathscr{T}_h)^{m\times n},$$

where C depends only on the mesh quality and on p. Using the inf–sup condition of Lemma A2, we obtain the result. □

THEOREM 5.2 (Compactness in $W^{1,p}(\mathscr{T}_h)$) Let $p \in (1, \infty)$. For each $h \in (0,1]$, let $u_h \in W^{1,p}(\mathscr{T}_h)^m$ such that

$$\sup_{h\in(0,1]} \left[\|u_h\|_{L^1(\Omega)} + |u_h|_{W^{1,p}(\mathscr{T}_h)}\right] < +\infty. \tag{5.3}$$

Then, there exists a sequence $h_j \downarrow 0$ and a function $u \in W^{1,p}(\Omega)^m$ such that

$$u_{h_j} \overset{*}{\rightharpoonup} u \quad \text{in } \mathrm{BV}(\Omega)^m \quad \text{and}$$

$$\nabla u_{h_j} + R\left(u_{h_j}\right) \rightharpoonup \nabla u \quad \text{in } L^p(\Omega)^{m\times n}.$$

Proof. From Lemma 2, it follows that $\|u_h\|_{\mathrm{BV}}$ is bounded. Hence, there exists a subsequence (which is not relabelled for notational convenience) and a function $u \in \mathrm{BV}(\Omega)^m$ such that $u_h \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega)^m$. Using the boundedness of the penalty term and applying Lemma 7, we also see that ∇u_h and $R(u_h)$ are bounded in $L^p(\Omega)^{m\times n}$ which implies their weak compactness. Upon extracting a further subsequence (again not relabelled), we obtain

$$\nabla u_h \rightharpoonup F_a \quad \text{and} \quad R(u_h) \rightharpoonup F_j$$

as $h \to 0$, where $F_a, F_j \in L^p(\Omega)^{m\times n}$. We show now that Du_h converges to $F_a + F_j$ in the sense of distributions. Since $\nabla u_h \rightharpoonup F_a$, we only need to show that the jumps generate F_j in the limit, i.e. that

$$-\int_{\Gamma_{\mathrm{int}}} [\![u_h]\!] \cdot \varphi \, \mathrm{d}s \to \int_\Omega F_j \cdot \varphi \, \mathrm{d}x \quad \forall \varphi \in C_{\mathrm{c}}^\infty(\Omega)^{m\times n}. \tag{5.4}$$

To this end, we add and subtract a function $\varphi_h \in S^l(\mathscr{T}_h)^{m\times n}$, then use the definition of $R(u_h)$ and subtract φ again. This procedure gives

$$\begin{aligned}
-\int_{\Gamma_{\mathrm{int}}} [\![u_h]\!] \cdot \varphi \, \mathrm{d}s &= -\int_{\Gamma_{\mathrm{int}}} [\![u_h]\!] \cdot \{\varphi - \varphi_h\} \mathrm{d}s - \int_{\Gamma_{\mathrm{int}}} [\![u_h]\!] \cdot \{\varphi_h\} \mathrm{d}s \\
&= -\int_{\Gamma_{\mathrm{int}}} [\![u_h]\!] \cdot \{\varphi - \varphi_h\} \mathrm{d}s + \int_\Omega R(u_h) \cdot \varphi_h \, \mathrm{d}x \\
&= -\int_{\Gamma_{\mathrm{int}}} [\![u_h]\!] \cdot \{\varphi - \varphi_h\} \mathrm{d}s + \int_\Omega R(u_h) \cdot (\varphi_h - \varphi) \mathrm{d}x + \int_\Omega R(u_h) \cdot \varphi \, \mathrm{d}x.
\end{aligned}$$

Using Lemma 7, it follows immediately that if we choose φ_h in such a way that $\|\varphi - \varphi_h\|_{L^\infty} \to 0$, for example $\varphi_h = (\varphi)_\kappa$ in κ, then the first and second term tend to zero as $h \to 0$. Since $R(u_h)$ converges weakly to F_j, it follows that Du_h converges to $F_a + F_j$ in the sense of distributions. Since Du_h converges also to Du in the sense of distributions, it follows that $Du = (F_a + F_j)\mathrm{d}x$. Therefore, the singular part of Du is zero, and hence u has a weak derivative $\nabla u = F_a + F_j \in L^p(\Omega)^{m\times n}$. Poincaré's inequality implies that $u \in L^p(\Omega)^m$ and hence $u \in W^{1,p}(\Omega)$. □

LEMMA 8 (Compact embeddings) Under the conditions of Theorem 5.2, it also holds that

$$u_{h_j} \to u \ \text{ in } L^q(\Omega)^m \quad \forall q : 1 \leqslant q < p^* \quad \text{and} \tag{5.5}$$

$$u_{h_j} \to u \ \text{ in } L^q(\partial\Omega)^m \quad \forall q : 1 \leqslant q < q^*, \tag{5.6}$$

where $q^* = (n-1)p/(n-p)$ if $p < n$ and $q^* = \infty$ if $p \geqslant n$.

Proof. For the proof of strong L^q-convergence (5.5), it is sufficient to use the compactness of the embedding $\mathrm{BV}(\Omega)^m \subset L^1(\Omega)^m$ and Riesz' interpolation theorem to lift the strong convergence to the L^q spaces indicated. To make this precise, suppose that $u_{h_j} \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega)^m$, then $u_{h_j} \to u$ strongly in $L^1(\Omega)^m$. Furthermore, if $\|u_{h_j}\|_{L^1} + |u_{h_j}|_{W^{1,p}(\mathscr{T}_h)}$ is bounded, then, by (4.2), $\|u_{h_j}\|_{L^{p*}}$ is bounded and, by Theorem 5.2, $u \in W^{1,p}(\Omega)^m \subset L^{p*}(\Omega)$. Hence, using Riesz' interpolation theorem, we can estimate

$$\|u - u_{h_j}\|_{L^q(\Omega)} \leqslant \|u - u_{h_j}\|_{L^{p*}(\Omega)}^{(1-\theta)} \|u - u_{h_j}\|_{L^1(\Omega)}^{\theta} \leqslant C \|u - u_{h_j}\|_{L^1(\Omega)}^{\theta}$$

for some $\theta \in (0, 1)$. The right-hand side in this inequality tends to zero.

Unfortunately, the trace operator presented in Theorem 4.2 is not compact and thus, we must revert to using the continuous reconstruction operator Q_h to prove the second result. From (3.8), it follows that for each face $e \subset \partial\Omega \cap \mathscr{E}_h$,

$$\|u_h - Q_h u_h\|^q_{L^q(e)} \lesssim h_e^{n-1-\frac{nq}{p}+q} |u_h|^q_{W^{1,p}(\mathscr{T}_h \cap T_e)}. \tag{5.7}$$

We prove (5.6) only for $q \in [p, q^*)$, where q^* is defined as above, the other cases being an immediate consequence of the statement for, e.g. $q = p$. Set $\alpha = n - 1 - nq/p + q > 0$. Summing (5.7) over the faces on the boundary, we obtain the following:

$$\|u_h - Q_h u_h\|^q_{L^q(\partial\Omega)} \lesssim h^\alpha \sum_{e \subset \partial\Omega} |u_h|^q_{W^{1,p}(\mathscr{T}_h \cap T_e)}.$$

Since $q \geqslant p$, we can use $\|\cdot\|_{\ell^q} \leqslant \|\cdot\|_{\ell^p}$ and Assumption 1(b) to deduce that

$$\begin{aligned}\|u_h - Q_h u_h\|^q_{L^q(\partial\Omega)} &\lesssim h^\alpha \sum_{e \subset \partial\Omega} |u_h|^q_{W^{1,p}(\mathscr{T}_h \cap T_e)} \\ &\lesssim h^\alpha \left(\sum_{e \subset \partial\Omega} |u_h|^p_{W^{1,p}(\mathscr{T}_h \cap T_e)} \right)^{q/p} \lesssim h^\alpha |u_h|^q_{W^{1,p}(\mathscr{T}_h)}.\end{aligned}$$

This implies that

$$\|u_h - Q_h u_h\|_{L^q(\partial\Omega)} \to 0 \quad \text{as } h \to 0. \tag{5.8}$$

Since the trace operator from $W^{1,p}(\Omega)^m$ to $L^q(\partial\Omega)^m$ is compact (Adams & Fournier, 2003, Theorem 6.3) and $Q_h u_h$ is bounded in $W^{1,p}(\Omega)^m$, it follows that $Q_h u_h \to u$ in $L^q(\partial\Omega)^m$ and therefore, by virtue of (5.8), $u_h \to u$ in $L^q(\partial\Omega)^m$. □

6. Variational DG approximation of minimization problems

Let Ω be a domain in $\mathbb{R}^n$ with boundary $\partial\Omega = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}}$, $\Gamma_{\mathrm{D}} \cap \Gamma_{\mathrm{N}} = \emptyset$, where Γ_{D} has positive surface measure. Let $f\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Carathéodory function, i.e. measurable in its first and continuous in its second and third argument. Suppose, further, that f satisfies the p-growth condition

$$c_0(|F|^p - |u|^r + a_0(x)) \leqslant f(x, u, F) \leqslant c_1(|F|^p + |u|^q + a_1(x)), \tag{6.1}$$

where $a_i \in L^1(\Omega)$. We furthermore require that $p \in (1, \infty)$, $r < p$ and $r \leqslant q < p^*$. Let $g\colon \Gamma_{\mathrm{N}} \times \mathbb{R}^m \to \mathbb{R}$ be a Carathéodory function which satisfies the growth condition

$$|g(x, u)| \leqslant c_2(|u|^r + a_2(x)), \tag{6.2}$$

where $a_2 \in L^1(\Gamma_{\mathrm{N}})$ and r is the same index as in (6.1).

We define the functional $\mathscr{I} : W^{1,p}(\Omega)^m \to \mathbb{R}$ by

$$\mathscr{I}(u) = \int_\Omega f(x, u, \nabla u)\mathrm{d}x + \int_{\Gamma_{\mathrm{N}}} g(x, u)\mathrm{d}s, \quad u \in W^{1,p}(\Omega)^m. \tag{6.3}$$

Fix $u_D \in W^{1,p}(\Omega)^m$ and define the set of admissable trial functions $\mathscr{A}$ to be the closed, affine subspace of $W^{1,p}(\Omega)^m$ given by

$$\mathscr{A} = \left\{ u \in W^{1,p}(\Omega)^m : u|_{\Gamma_D} = u_D \right\}.$$

We consider the problem of finding a minimizer of $\mathscr{I}$ in $\mathscr{A}$. If f is convex in its third component, then the existence of minimizers follows from the direct method of the calculus of variations, see for example Dacorogna (1989, Theorems 3.1, 3.4 and 4.1). Note in particular that, if either $m = 1$ or $n = 1$, then convexity of f in its third argument is a necessary and sufficient condition for $\mathscr{I}$ to be sequentially weakly lower semicontinuous (Dacorogna, 1989, Theorem 3.1), which is a necessary condition for the direct method to apply to our problem. However, if $\min(m, n) \geqslant 2$, then a more general notion of convexity should be allowed (Dacorogna, 1989).

Before proposing a discretization strategy, we summarize the most important technical facts about (6.3) which we use in the convergence proof.

LEMMA 9 Let f and g be Carathéodory functions which respectively satisfy the growth conditions (6.1) and (6.2).

(i) If $u_j \to u$ strongly in $L^q(\Omega)^m$ and $F_j \to F$ strongly in $L^p(\Omega)^{m\times n}$, then

$$\int_\Omega f(x, u_j, F_j)\mathrm{d}x \to \int_\Omega f(x, u, F)\mathrm{d}x \quad \text{as } j \to \infty.$$

(ii) If $u_j \to u$ strongly in $L^r(\Gamma_N)^m$, then

$$\int_{\Gamma_N} g(x, u_j)\mathrm{d}s \to \int_{\Gamma_N} g(x, u)\mathrm{d}s \quad \text{as } j \to \infty.$$

(iii) If $u_j \to u$ strongly in $L^q(\Omega)^m$, $F_j \rightharpoonup F$ weakly in $L^p(\Omega)^{m\times n}$ and f is convex in the third argument, then

$$\int_\Omega f(x, u, F)\mathrm{d}x \leqslant \liminf_{j\to\infty} \int_\Omega f(x, u_j, F_j)\mathrm{d}x.$$

Items (i) and (ii) follow from Fatou's lemma while item (iii) is an application of Dacorogna (1989, Theorem 3.4).

We now turn to the discretization of the functional (6.3). To this end, we chose a polynomial degree $l \geqslant 0$ and then define the lifting operator $R\colon W^{1,p}(\mathscr{T}_h)^m \to S^l(\mathscr{T}_h)^{m\times n}$ as in (5.1). The lifting $R(u)$ is a bulk representation of the jump contribution to the distributional gradient of u. The polynomial degree l is a method parameter and can be chosen arbitrarily.

We propose the following discrete functional:

$$\begin{aligned}\mathscr{I}_h(u_h) = &\int_\Omega f(x, u_h, \nabla u_h + R(u_h))\mathrm{d}x + \int_{\Gamma_N} g(x, u_h)\mathrm{d}s \\ &+ \int_{\Gamma_D} \mathsf{h}^{1-p}|u_h - u_D|^p\,\mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \mathsf{h}^{1-p}|[\![u_h]\!]|^p\,\mathrm{d}s,\end{aligned} \tag{6.4}$$

and our discrete problem is to find a minimizer of (6.4) among all possible vector fields in $S^k(\mathscr{T}_h)^m$. In the tradition of the literature on DGFEMs, we chose to label this variational method as variational interior penalty DGFEM. We note that the fourth term in (6.4) weakly imposes the Dirichlet boundary condition and it is therefore not necessary to impose this condition on the approximation space.

A closely related DGFE discretization (with $p = 2$ but allowing a more general definition of the flux) was defined by Ten Eyck & Lew (2006) for applications in finite elasticity. We refer to their paper for a linearized stability analysis and very promising numerical results. An error analysis for smooth solutions of the Euler–Lagrange equations was given by Ortner & Süli (2007).

Note that, despite its appearance, (6.4) is in fact fairly straightforward to implement. The definition of the lifting operator (5.1) allows the construction of $R(u_h)$ locally in each element, taking into account only the degrees of freedom on the edges of the element. For example, if $R(u_h)$ is chosen to be piecewise constant (which is sufficient to obtain convergence), then

$$R(u_h)|_\kappa = |\kappa|^{-1} \int_{\partial\kappa\setminus\partial\Omega} [\![u_h]\!] \mathrm{d}s \quad \forall \kappa \in \mathscr{T}_h. \tag{6.5}$$

Our first step in the analysis of (6.4) is to prove that families with bounded energies are bounded in the broken $W^{1,p}$-norm.

LEMMA 10 (Coercivity) Suppose that the energy densities f and g satisfy, respectively, (6.1) and (6.2). Then, there exists a constant C, independent of the mesh size, such that

$$\|u\|^p_{W^{1,p}(\mathscr{T}_h)} \leqslant C(\mathscr{I}_h(u) + 1) \quad \forall u \in S^k(\mathscr{T}_h)^m, \ \ \forall h \in (0, 1].$$

Proof. Let $u \in S^k(\mathscr{T}_h)^m$. By the growth hypotheses (6.1) and (6.2) and the Trace Theorem 4.4, we have

$$\begin{aligned}\mathscr{I}_h(u) \geqslant c_0 &\left(\|\nabla u + R(u)\|^p_{L^p(\Omega)} - \|u\|^r_{L^r(\Omega)} - \|a_0\|_{L^1(\Omega)}\right) \\ &-c_2\left(\|u\|^r_{L^r(\Omega)} + |u|^r_{W^{1,r}(\mathscr{T}_h)} + \|a_2\|_{L^1(\Gamma_\mathrm{N})}\right) \\ &+ \int_{\Gamma_\mathrm{int}} \mathsf{h}^{1-p} |[\![u]\!]|^p \,\mathrm{d}s + \int_{\Gamma_\mathrm{D}} \mathsf{h}^{1-p}|u - u_\mathrm{D}|^p \,\mathrm{d}s.\end{aligned}$$

Since $r < p$, for any $\varepsilon > 0$, we can estimate

$$\|u\|^r_{L^r(\Omega)} \lesssim \|u\|^r_{L^p(\Omega)} \leqslant \frac{\varepsilon}{p/r}\|u\|^p_{L^p(\Omega)} + \frac{1}{\varepsilon(p/r)'} \lesssim \varepsilon^{-1} + \varepsilon\|u\|^p_{L^p(\Omega)}.$$

Treating the term $|u|^r_{W^{1,r}(\mathscr{T}_h)}$ in a similar fashion, we obtain

$$\begin{aligned}\mathscr{I}_h(u) + C(\varepsilon) \geqslant c_0 &\left(\|\nabla u + R(u)\|^p_{L^p(\Omega)} - \varepsilon\|u\|^p_{L^p(\Omega)} - \varepsilon|u|^p_{W^{1,p}(\mathscr{T}_h)}\right) \\ &+ \int_{\Gamma_\mathrm{int}} \mathsf{h}^{1-p}|[\![u]\!]|^p \,\mathrm{d}s + \int_{\Gamma_\mathrm{D}} \mathsf{h}^{1-p}|u - u_\mathrm{D}|^p \,\mathrm{d}s.\end{aligned}$$

An application of the broken Friedrichs inequality, Corollary 4.3, gives

$$\begin{aligned}\mathscr{I}_h(u) + C(\varepsilon) \geqslant \ c_0 &\left(\|\nabla u + R(u)\|^p_{L^p(\Omega)} - \varepsilon(1 + 2^{p-1}C^p_{\mathrm{BF}})\left(\|u\|^p_{L^p(\Gamma_\mathrm{D})} + |u|^p_{W^{1,p}(\mathscr{T}_h)}\right)\right) \\ &+ \int_{\Gamma_\mathrm{int}} \mathsf{h}^{1-p}|[\![u]\!]|^p \,\mathrm{d}s + \int_{\Gamma_\mathrm{D}} \mathsf{h}^{1-p}|u - u_\mathrm{D}|^p \,\mathrm{d}s.\end{aligned}$$

To shorten the notation, in what follows, we rename $\varepsilon = \varepsilon(1 + 2^{p-1}C_{\mathrm{BF}}^p)$. For a given $\delta \in (0, 1]$, we estimate the first and last terms on the right-hand side, respectively, by

$$\|\nabla u + R(u)\|_{L^p(\Omega)}^p \geqslant \delta\|\nabla u + R(u)\|_{L^p(\Omega)}^p \geqslant 2^{1-p}\delta\|\nabla u\|_{L^p(\Omega)}^p - \delta\|R(u)\|_{L^p(\Omega)}^p \quad \text{and}$$
$$\int_{\Gamma_{\mathrm{D}}} \mathsf{h}^{1-p}|u - u_{\mathrm{D}}|^p \,\mathrm{d}s \geqslant \int_{\Gamma_{\mathrm{D}}} |u - u_{\mathrm{D}}|^p \,\mathrm{d}s \geqslant 2^{1-p}\int_{\Gamma_{\mathrm{D}}} |u|^p \,\mathrm{d}s - \int_{\Gamma_{\mathrm{D}}} |u_{\mathrm{D}}|^p \,\mathrm{d}s,$$

and hence deduce

$$\mathscr{I}_h(u) + C(\varepsilon) \geqslant c_0\left((2^{1-p}\delta - \varepsilon)\|\nabla u\|_{L^p(\Omega)}^p - \delta\|R(u)\|_{L^p(\Omega)}^p - \varepsilon\int_{\Gamma_{\mathrm{D}}} |u|^p \,\mathrm{d}s\right)$$
$$+ \int_{\Gamma_{\mathrm{int}}} \mathsf{h}^{1-p}|[\![u]\!]|^p \,\mathrm{d}s + \int_{\Gamma_{\mathrm{D}}} |u|^p \,\mathrm{d}s$$

We now fix $\delta = \frac{1}{2c_0C_R^p}$, where C_R is the constant appearing in Lemma 7, so that penalty integral dominates $\delta\|R(u)\|_{L^p(\Omega)}^p$. Finally, we obtain

$$\mathscr{I}_h(u) + C(\varepsilon) \geqslant c_0(2^{1-p}\delta - \varepsilon)\|\nabla u\|_{L^p(\Omega)}^p + (1/2 - c_0\varepsilon)\left(\int_{\Gamma_{\mathrm{int}}} \mathsf{h}^{1-p}|[\![u]\!]|^p \,\mathrm{d}s + \int_{\Gamma_{\mathrm{D}}} |u|^p \mathrm{d}s\right),$$

which provides the required bound after choosing, e.g. $\varepsilon = \min\{1/4c_0, 2^{-p}\delta\}$ and then applying Corollary 4.3. □

Together, Lemma 10 and Theorem 5.2 establish the compactness of any family of DGFEM functions u_h for which $\mathscr{I}_h(u_h)$ is bounded. This allows us to use a direct method-related technique (namely Γ-convergence, see De Giorgi & Franzoni, 1975; Dal Maso, 1993) to prove the convergence of discrete minimizers to a minimizer of $\mathscr{I}$ in $\mathscr{A}$.

THEOREM 6.1 (Convergence) Suppose that f and g are Carathéodory functions which, respectively, satisfy (6.1) and (6.2) and f is convex in its third argument.

For each $h \in (0, 1]$, let $u_h \in \operatorname{argmin}_{S^k(\mathscr{T}_h)^m} \mathscr{I}_h$. Then, there exists a subsequence $h_j \downarrow 0$ and $u \in \mathrm{BV}(\Omega)^m$ such that $u_{h_j} \overset{*}{\rightharpoonup} u$. Any such accumulation point u is a minimizer of $\mathscr{I}$ in $\mathscr{A}$ (in particular, $u \in W^{1,p}(\Omega)^m$) and satisfies

$$u_{h_j} \to u \quad \text{in } L^q(\Omega)^m \;\; \forall q < p^*, \tag{6.6}$$

$$\nabla u_{h_j} \rightharpoonup \nabla u \quad \text{in } L^p(\Omega)^{m\times n}, \tag{6.7}$$

$$\mathscr{I}_{h_j}\left(u_{h_j}\right) \to \mathscr{I}(u) \quad \text{and} \tag{6.8}$$

$$\int_{\Gamma_{\mathrm{D}}} \mathsf{h}_j^{1-p}\left|u_{h_j} - u_{\mathrm{D}}\right|^p \,\mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \mathsf{h}_j^{1-p}\left|[[u_{h_j}]]\right|^p \,\mathrm{d}s \to 0 \tag{6.9}$$

as $j \to \infty$. If f is strictly convex in its third argument, then, in addition,

$$\left|u - u_{h_j}\right|_{W_{\mathrm{D}}^{1,p}(\mathscr{T}_{h_j})} \to 0 \quad \text{as } j \to \infty.$$

If the minimizer is unique, then the entire family u_h converges.

Proof. By the growth condition (6.1), any family (u_h) which is bounded in $W^{1,p}(\mathscr{T}_h)^m$ has bounded energy $\mathscr{I}_h(u_h)$ and conversely, by Lemma 10, if $\mathscr{I}_h(u_h)$ is bounded, then $\|u_h\|_{W^{1,p}(\mathscr{T}_h)}$ is bounded as well.

From the compactness result, Theorem 5.2, we therefore deduce the existence of a subsequence $h_j \downarrow 0$ and a limit point $u \in W^{1,p}(\Omega)^m$ such that $u_{h_j} \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega)^m$.

Assume now that $\left(u_{h_j}\right)$ is any minimizing sequence for $\mathscr{I}_{h_j}$ converging weakly-$*$ to some $u \in \mathrm{BV}(\Omega)^m$. From the boundedness of the energy and the broken Friedrichs inequality, we can again deduce the boundedness of $\left|u_{h_j}\right|_{W^{1,p}(\mathscr{T}_{h_j})}$ and therefore can employ Theorem 5.2 to deduce that $u \in W^{1,p}(\Omega)^m$ as well as

$$\nabla u_{h_j} + R\left(u_{h_j}\right) \rightharpoonup \nabla u \quad \text{weakly in } L^p(\Omega)^{m\times n}. \tag{6.10}$$

Lemma 8 implies (6.6).

Since the boundary penalty terms

$$\int_{\Gamma_{\mathrm{D}}} \mathsf{h}_j^{1-p} \left|u_{h_j} - u_{\mathrm{D}}\right|^p \mathrm{d}s$$

are bounded, using also Lemma 8, it follows that

$$\|u - u_{\mathrm{D}}\|_{L^p(\Gamma_{\mathrm{D}})} \leqslant \left\|u - u_{h_j}\right\|_{L^p(\Gamma_{\mathrm{D}})} + \left\|u_{h_j} - u_{\mathrm{D}}\right\|_{L^p(\Gamma_{\mathrm{D}})} \to 0$$

as $j \to \infty$ and hence $u \in \mathscr{A}$.

Lemma 8 also implies the strong convergence of u_{h_j} to u in $L^r(\partial\Omega)^m$, and therefore, it follows from Lemma 9(ii) that the surface integral converges, i.e.

$$\int_{\Gamma_{\mathrm{N}}} g\left(x, u_{h_j}\right) \mathrm{d}s \to \int_{\Gamma_{\mathrm{N}}} g(x,u)\mathrm{d}s \quad \text{as } j \to \infty.$$

As a consequence, using (6.10) and Lemma 9(iii), we deduce that

$$\mathscr{I}(u) \leqslant \liminf_{j\to\infty} \left[\int_\Omega f\left(x, u_{h_j}, \nabla u_{h_j} + R\left(u_{h_j}\right)\right) \mathrm{d}x + \int_{\Gamma_{\mathrm{N}}} g\left(x, u_{h_j}\right) \mathrm{d}s\right].$$

To see that $u \in \operatorname{argmin}_{\mathscr{A}} \mathscr{I}$, fix $v \in \mathscr{A}$ and let $v_h \in S^k(\mathscr{T}_h)^m$ converge strongly to v in the $\|\cdot\|_{L^{p^*}(\Omega)}$ as well as the $|\cdot|_{W^{1,p}(\mathscr{T}_h)}$-norm (see Lemma 3). From Lemma 9 (using also the Trace Theorem 4.4), we therefore obtain $\mathscr{I}_h(v_h) \to \mathscr{I}(v)$, which allows us to estimate

$$\begin{aligned}\mathscr{I}(u) &\leqslant \liminf_{j\to\infty} \left[\int_\Omega f\left(x, u_{h_j}, \nabla u_{h_j} + R\left(u_{h_j}\right)\right) \mathrm{d}x + \int_{\Gamma_{\mathrm{N}}} g\left(x, u_{h_j}\right) \mathrm{d}s\right] \\ &\leqslant \limsup_{j\to\infty} \mathscr{I}_{h_j}\left(u_{h_j}\right) \leqslant \limsup_{j\to\infty} \mathscr{I}_{h_j}(v_{h_j}) \leqslant \mathscr{I}(v).\end{aligned}$$

Since v was arbitrary, it follows that $\mathscr{I}(u) \in \operatorname{argmin}_{\mathscr{A}} \mathscr{I}$. By choosing $v = u$, we find that all inequalities are equalities from which we can infer that $\mathscr{I}_{h_j}\left(u_{h_j}\right) \to \mathscr{I}(u)$ and that the penalty terms converge to zero as $h_j \to 0$, i.e. that (6.9) holds. As a consequence, we also have $R\left(u_{h_j}\right) \to 0$ strongly which implies (6.7).

If f is strictly convex in its third argument, then the theory of Young measures shows that weak convergence together with convergence of the energy implies strong convergence. For example, the proof of Theorem 3.16 in the monograph of Pedregal (1997) can be immediately adapted to give our result. See also Lemma A3 in the appendix.

The last point follows from a straightforward uniqueness argument. □

7. Optimal embedding constants

In this final section, we present a second application of the compactness results of Section 5. Under suitable conditions, we shall deduce that in the limit as $h \to 0$, the optimal embedding constant in the broken Sobolev–Poincaré inequality (4.1) is the same as the embedding constant for the classical Sobolev space. We demonstrate the technique only on the example of the Sobolev–Poincaré inequality, but we believe that it should apply to any compact embedding of a Sobolev space. Throughout this section, we take $m = 1$.

Unfortunately, our results are incomplete for the particular broken seminorm which we have chosen. Instead, we analyse the equivalent norm

$$|u|_{W_1^{1,p}(\mathscr{T}_h)} = \|\nabla u\|_{L^p(\Omega)} + \alpha \left(\int_{\Gamma_{\rm int}} \mathrm{h}^{1-p} |[\![u]\!]|^p \, \mathrm{d}s \right)^{1/p}, \tag{7.1}$$

where α is some fixed positive constant.

From norm equivalence in $\mathbb{R}^2$, it follows immediately that $|\cdot|_{W^{1,p}(\mathscr{T}_h)}$ and $|\cdot|_{W_1^{1,p}(\mathscr{T}_h)}$ are equivalent; more precisely, there exists a constant $c_\alpha > 0$ such that

$$c_\alpha |u|_{W^{1,p}(\mathscr{T}_h)} \leqslant |u|_{W_1^{1,p}(\mathscr{T}_h)} \leqslant \frac{1}{c_\alpha} |u|_{W^{1,p}(\mathscr{T}_h)} \quad \forall\, u \in W^{1,p}(\mathscr{T}_h), \quad \forall\, h \in (0, 1]. \tag{7.2}$$

We can now study the Poincaré constants of the newly defined broken seminorm. Fix $p \in (1, \infty)$, $q \in [1, p^*)$ and let $V = \{v \in L^1(\Omega) : (v)_\Omega = 0\}$. From (7.2), it follows that we can replace $|\cdot|_{W^{1,p}(\mathscr{T}_h)}$ by $|\cdot|_{W_1^{1,p}(\mathscr{T}_h)}$ in (4.1) to obtain

$$\|u_h - (u_h)_\Omega\|_{L^q(\Omega)} \leqslant C_h(p, q) |u_h|_{W_1^{1,p}(\mathscr{T}_h)} \quad \forall\, u_h \in S^k(\mathscr{T}_h), \tag{7.3}$$

which is the discrete counterpart of the Sobolev–Poincaré inequality

$$\|u - (u)_\Omega\|_{L^q(\Omega)} \leqslant C(p, q) \|\nabla u\|_{L^p(\Omega)} \quad \forall\, u \in W^{1,p}(\Omega). \tag{7.4}$$

We begin by noting that the optimal constants $C_h(p, q)$ and $C(p, q)$ in (7.3) and (7.4) are, respectively, given by

$$\frac{1}{C(p, q)} = \inf_{\substack{u \in W^{1,p}(\Omega) \cap V, \\ \|u\|_{L^q(\Omega)} = 1}} \|\nabla u\|_{L^p(\Omega)} \tag{7.5}$$

and

$$\frac{1}{C_h(p, q)} = \inf_{\substack{u_h \in S^k(\mathscr{T}_h) \cap V, \\ \|u_h\|_{L^q(\Omega)} = 1}} |u_h|_{W_1^{1,p}(\mathscr{T}_h)}.$$

In particular, the latter can be viewed as a discretization to the minimization problem defining $C(p,q)$ and we can therefore employ a similar type of analysis as in Section 6 to obtain the following result.

We note for future reference that both infima $1/C(p,q)$ and $1/C_h(p,q)$ are attained. This statement is trivial for the latter and for the former, it follows from the fact that the set over which we minimize in (7.5) is weakly closed in $W^{1,p}(\Omega)$.

PROPOSITION 7.1 There exists a constant $\hat{\alpha} > 0$ such that

$$\lim_{h\downarrow 0} C_h(p,q) = C(p,q) \ \text{ if } \alpha \geqslant \hat{\alpha} \quad \text{and}$$

$$\liminf_{h\downarrow 0} C_h(p,q) > C(p,q) \ \text{ if } 0 < \alpha < \hat{\alpha}.$$

Proof. We begin by investigating the case where α is large. Suppose that $u_h \in S^k(\mathscr{T}_h) \cap V$, $h \in (0,1]$; $\|u_h\|_{L^q(\Omega)} = 1$ and $|u_h|_{W_1^{1,p}(\mathscr{T}_h)} = C_h(p,q)^{-1}$. From Lemma 3 and norm equivalence, it follows that $|u_h|_{W^{1,p}(\mathscr{T}_h)}$ is bounded and hence we can extract a subsequence u_{h_j} converging weakly-$*$ in $\mathrm{BV}(\Omega)$ and strongly in $L^q(\Omega)$ to a function $u \in W^{1,p}(\Omega)$. In particular, $\|u\|_{L^q(\Omega)} = 1$ and we have

$$\|\nabla u\|_{L^p(\Omega)} \leqslant \liminf_{j\to\infty} \left\|\nabla u_{h_j} + R\left(u_{h_j}\right)\right\|_{L^p(\Omega)}$$

$$\leqslant \liminf_{j\to\infty} \left(\left\|\nabla u_{h_j}\right\|_{L^p(\Omega)} + \left\|R\left(u_{h_j}\right)\right\|_{L^p(\Omega)}\right).$$

If α is sufficiently large (e.g. if $\alpha > C_R$), it follows from Lemma 7 that

$$\|\nabla u\|_{L^p(\Omega)} \leqslant \liminf_{j\to\infty} \left|u_{h_j}\right|_{W_1^{1,p}(\mathscr{T}_{h_j})}$$

and therefore $\liminf_{h\downarrow 0} C_h(p,q)^{-1} \geqslant C(p,q)^{-1}$. From Lemma 3, we obtain $\lim_{h\downarrow 0} C_h(p,q) = C(p,q)$.

Now assume that α is small. Let $u \in W^{1,p}(\Omega) \cap V$ such that $\|u\|_{L^q(\Omega)} = 1$ and $\|\nabla u\|_{L^p(\Omega)} = C(p,q)^{-1}$. For each $h \in (0,1]$, let u_h be defined by

$$u_h(x) = (u)_\kappa \quad \forall x \in \kappa, \ \ \forall \kappa \in \mathscr{T}_h.$$

Clearly, $u_h \in S^k(\mathscr{T}_h) \cap V$ and $\|u_h - u\|_{L^q(\Omega)} \to 0$ as $h \downarrow 0$. Furthermore, we can bound the seminorm $|u_h|_{W_1^{1,p}(\mathscr{T}_h)}$ in terms of $\|\nabla u\|_{L^p(\Omega)}$ as follows:

$$\alpha^{-p}|u_h|^p_{W_1^{1,p}(\mathscr{T}_h)} = \sum_{e\subset\Gamma_{\text{int}}} h_e^{1-p}\mathscr{H}^{n-1}(e)|(u)_{\kappa^+} - (u)_{\kappa^-}|^p$$

$$\lesssim \sum_{e\subset\Gamma_{\text{int}}} h_e^{n-p}[|(u)_{\kappa^+} - \pi| + |(u)_{\kappa^-} - \pi|]^p \tag{7.6}$$

for any $\pi \in \mathbb{R}$.

We construct π in a similar fashion as the local projection operators in Section 3.1. Fix $e = \kappa^+ \cap \kappa^- \in \mathscr{E}_h$. Assumption 2.1 implies the existence of $z \in e$ and $\rho \approx h_e$ such that $B(z, \rho) \subset K := \kappa^+ \cup \kappa^-$. In particular, K is star shaped with respect to z. Hence, we can set $\pi = (u)_B$ and use Lemma 4 to deduce that

$$|(u)_{\kappa^+} - \pi| + |(u)_{\kappa^-} - \pi| \lesssim h_{\kappa^+}^{-n} \|u - \pi\|_{L^1(\kappa^+)} + h_{\kappa^-}^{-n} \|u - \pi\|_{L^1(\kappa^-)} \lesssim h_e^{-n+1} \|\nabla u\|_{L^1(K)}.$$

Upon taking pth powers and applying Jensen's inequality, we obtain

$$[|(u)_{\kappa^+} - \pi| + |(u)_{\kappa^-} - \pi|]^p \lesssim h_e^{p-np} \|\nabla u\|_{L^1(K)}^p \lesssim h_e^{p-n} \|\nabla u\|_{L^p(K)}^p.$$

Combined with (7.6) and the contact regularity assumptions, this gives

$$\alpha^{-p} |u_h|_{W_1^{1,p}(\mathscr{T}_h)}^p \lesssim \|\nabla u\|_{L^p(\Omega)}^p = C(p, q)^{-1}.$$

In summary, we have obtained that there exists a constant $\tilde{\alpha}$ which is independent of h such that

$$\alpha^{-1} |u_h|_{W_1^{1,p}(\mathscr{T}_h)} \leqslant \tilde{\alpha} C(p, q)^{-1}.$$

Hence, for $\alpha < 1/\tilde{\alpha}$, it follows that

$$C_h(p, q)^{-1} \leqslant |u_h|_{W_1^{1,p}(\mathscr{T}_h)} \leqslant \alpha \tilde{\alpha} C(p, q)^{-1} < C(p, q)^{-1},$$

and, as a consequence, we obtain that $\liminf_{h \downarrow 0} C_h(p, q) > C(p, q)$.

Finally, we note that if the latter property holds for a specific $\alpha = \alpha'$, then it also holds for all $\alpha < \alpha'$ and hence the proposition follows. □

REMARK 7.2 We conclude our analysis of optimal Sobolev–Poincaré imbedding constants with a remark on a modification of the seminorm $|\cdot|_{W^{1,p}(\mathscr{T}_h)}$. If we redefine it as

$$|u|_{W^{1,p}(\mathscr{T}_h)} = \left(\|\nabla u\|_{L^p(\Omega)}^p + \alpha \int_{\Gamma_{\text{int}}} \mathsf{h}^{1-p} |[\![u]\!]|^p \, \mathrm{d}s \right)^{1/p},$$

with Sobolev–Poincaré constant $\tilde{C}_h(p, q)$, then we can obviously use the construction of a *recovery sequence* for the $|\cdot|_{W_1^{1,p}(\mathscr{T}_h)}$-seminorm in the proof of Proposition 7.1 to deduce that, if α is sufficiently small, then $\liminf_{h \downarrow 0} \tilde{C}_h(p, q) > C(p, q)$. However, we have a gap for large α.

For sufficiently large α, we can deduce from Proposition 7.1 that

$$\limsup_{h \downarrow 0} \tilde{C}_h(p, q) \leqslant 2^{1/p} C(p, q),$$

which is a good bound but not optimal. Setting $a = \|\nabla u_h\|_{L^p}$ and $b = \left(\int_{\Gamma_{\text{int}}} \mathsf{h}^{1-p} |[\![u_h]\!]|^p \, \mathrm{d}s \right)^{1/p}$ in the following inequality:

$$(|a| + |b|)^p \leqslant (1 + \varepsilon)|a|^p + B_\varepsilon |b|^p,$$

where B_ε depends only on ε and on p, we can strengthen this result to

$$\lim_{\alpha\to\infty} \limsup_{h\downarrow 0} \tilde{C}_h(p,q) = C(p,q).$$

However, we are unable to prove that $\limsup_{h\downarrow 0} C_h(p,q) = C(p,q)$ for any sufficiently large (but fixed) α. In fact, our numerical experiments suggest that this is not the case.

Funding

A. Buffa was supported by Italian FIRB (RBNE01C79T0-04, 2002-2006 to A.B.); C. Ortner was supported by European research (HPRN-CT-2002-00284: New Materials, Adaptive Systems and their Nonlinearities. Modelling, Control and Numerical Simulation to C.O.); and by EPSRC Critical Mas Project "New frontiers in the mathematics of solids".

REFERENCES

ACOSTA, G. & DURÁN, R. G. (2004) An optimal Poincaré inequality in L^1 for convex domains. *Proc. Am. Math. Soc.*, **132**, 195–202.

ADAMS, R. A. & FOURNIER, J. F. (2003) *Sobolev Spaces*. Amsterdam: Elsevier Science.

BALL, J. M. (2001) Singularities and computation of minimizers for variational problems. *Foundations of Computational Mathematics (Oxford, 1999)* (R. A. DeVore, A. Iserles & E. Süli eds.) London Mathematical Society Lecture Note Series, vol. 284. Cambridge: Cambridge University Press, pp. 1–20.

BRAIDES, A. (2002) *Γ-Convergence for Beginners*. Oxford Lecture Series in Mathematics and its Applications, vol. 22. Oxford: Oxford University Press.

BRENNER, S. C. (2003) Poincaré-Friedrichs inequalities for piecewise H^1 functions. *SIAM J. Numer. Anal.*, **41**, 306–324 (electronic).

BRENNER, S. C. (2004) Korn's inequalities for piecewise H^1 vector fields. *Math. Comput.*, **73**, 1067–1087 (electronic).

CIARLET, PH. (1978) *The Finite Element Method for Elliptic Problems*. Amsterdam: North-Holland.

DACOROGNA, B. (1989) *Direct Methods in the Calculus of Variations*. Applied Mathematical Sciences, vol. 78. Berlin: Springer.

DAL MASO, G. (1993) *An Introduction to Γ-Convergence*. Progress in Nonlinear Differential Equations and their Applications, vol. 8. Boston, MA: Birkhäuser Boston Inc.

DE GIORGI, E. & FRANZONI, T. (1975) Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat.*, **58**, 842–850.

EVANS, L. C. & GARIEPY, R. F. (1992) *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics. Boca Raton, FL: CRC Press.

KARAKASHIAN, O. A. & PASCAL, F. (2003) A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.*, **41**, 2374–2399 (electronic).

KUFNER, A., JOHN, O. & FUČÍK, S. (1977) *Function Spaces*, Monographs and Textbooks on Mechanics of Solids and Fluids; Mechanics: Analysis. Leyden: Noordhoff International Publishing.

LASIS, A. & SÜLI, E. (2003) Poincaré-type inequalities for broken sobolev spaces. *Technical Report 03/10*. Oxford: Oxford University Computing Laboratory.

LASSER, C. & TOSELLI, A. (2003) An overlapping domain decomposition preconditioner for a class of discontinuous Galerkin approximations of advection-diffusion problems. *Math. Comput.*, **72**, 1215–1238 (electronic).

LE TALLEC, P. (1994) Numerical methods for nonlinear three-dimensional elasticity. *Handbook of Numerical Analysis* (P. G. Ciarlet & J.-L. Lions eds.) vol. III. Amsterdam: North-Holland, pp. 465–622.

LEW, A., NEFF, P., SULSKY, D. & ORTIZ, M. (2004) Optimal BV estimates for a discontinuous Galerkin method for linear elasticity. *AMRX Appl. Math. Res. Express*, **3**, 73–106.

ORTNER, C. & SÜLI, E. (2007) Discontinuous Galerkin finite element approximation of nonlinear second-order elliptic and hyperbolic systems. *SIAM J. Numer. Anal.*, **45**, 1370–1397.

PEDREGAL, P. (1997) *Parametrized Measures and Variational Principles*. Progress in Nonlinear Differential Equations and their Applications, vol. 30. Basel: Birkhäuser.

PEDREGAL, P. (2000) *Variational Methods in Nonlinear Elasticity*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

TEN EYCK, A. & LEW, A. (2006) Discontinuous Galerkin method for nonlinear elasticity. *Int. J. Numer. Methods Eng.*, **67**, 1204–1243.

VERFÜRTH, R. (1999) Error estimates for some quasi-interpolation operators. *M2AN Math. Model. Numer. Anal.*, **33**, 695–713.

Appendix A

A.1 *Proof of Lemma 4*

This proof is a modification of the proof of Verfürth (1999, Lemma 4.1). Throughout, we set $\gamma = \rho_2/\rho_1$.

Using the local approximation of BV functions by smooth functions (Evans & Gariepy, 1992, Section 5.2.2), there exists a sequence $u_j \in \mathrm{BV}(K) \cap C^\infty(K)$ such that $u_j \to u$ strictly in BV, i.e. $u_j \to u$ strongly in L^1 and $|Du_j|(K) \to |Du|(K)$ as $j \to \infty$. Hence, we can assume without loss of generality that $u \in C^\infty(\Omega) \cap W^{1,1}(\Omega)$.

We write

$$\|u\|_{L^1(K)} = \|u\|_{L^1(B)} + \|u\|_{L^1(K\setminus B)}.$$

Let Σ be the unit sphere in $\mathbb{R}^n$ and, for each $\sigma \in \Sigma$, let $x_0 + r(\sigma)\sigma \in \partial K$. For the second term, we compute

$$\begin{aligned}\|u\|_{L^1(K\setminus B)} &= \int_\Sigma \int_{\rho_1}^{r(\sigma)} t^{n-1}|u(t\sigma)|\mathrm{d}t\,\mathrm{d}s(\sigma)\\ &\leqslant \int_\Sigma \int_{\rho_1}^{r(\sigma)} t^{n-1}|u(t\sigma) - u(\rho_1\sigma)|\mathrm{d}t + \int_\Sigma \int_{\rho_1}^{r(\sigma)} t^{n-1}|u(\rho_1\sigma)|\mathrm{d}t\,\mathrm{d}s(\sigma)\\ &=: S_1 + S_2.\end{aligned}$$

To obtain a bound on S_1, consider

$$\begin{aligned}S_1 &= \int_\Sigma \int_{\rho_1}^{r(\sigma)} t^{n-1}\left|\int_{\rho_1}^{t} \partial_r u(r\sigma)\mathrm{d}r\right| \mathrm{d}t\,\mathrm{d}s(\sigma)\\ &\leqslant {\rho_1}^{1-n}\int_\Sigma \int_{\rho_1}^{r(\sigma)} t^{n-1}\int_{\rho_1}^{t} r^{n-1}|\partial_r u(r\sigma)|\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}s(\sigma)\\ &\leqslant \frac{1}{n}{\rho_1}^{1-n}(\rho_2^n - {\rho_1}^n)\int_\Sigma \int_{\rho_1}^{r(\sigma)} r^{n-1}|\partial_r u(r\sigma)|\mathrm{d}r\,\mathrm{d}s(\sigma)\\ &\leqslant \frac{\rho_1}{n}(\gamma^n - 1)\|\nabla u\|_{L^1(K\setminus B)}.\end{aligned}$$

For S_2, we estimate

$$\begin{aligned} S_2 &= \frac{1}{n}\int_{\Sigma}(r(\sigma)^n - {\rho_1}^n)|u(\rho_1\sigma)|\mathrm{d}s(\sigma) \\ &\leqslant \frac{\rho_1}{n}\int_{\Sigma}\left[\frac{\rho_2^n}{{\rho_1}^n} - 1\right]{\rho_1}^{n-1}|u(\rho_1\sigma)|\mathrm{d}s(\sigma) \\ &= \frac{\rho_1}{n}(\gamma^n - 1)\int_{\Sigma}{\rho_1}^{n-1}|u(\rho_1\sigma)|\mathrm{d}s(\sigma) \\ &= \frac{\rho_1}{n}(\gamma^n - 1)\|u\|_{L^1(\partial B)}. \end{aligned}$$

We bound $\|u\|_{L^1(\partial B)}$ as follows:

$$\begin{aligned} \|u\|_{L^1(\partial B)} &= \int_{\Sigma}{\rho_1}^{n-1}|u(\rho_1\sigma)|\mathrm{d}s(\sigma) \\ &= \int_{\Sigma}{\rho_1}^{n-1}\left|\int_0^{\rho_1}\partial_r\left[\left(\frac{r}{\rho_1}\right)^n u(r\sigma)\right]\mathrm{d}r\right|\mathrm{d}s(\sigma) \\ &= \int_{\Sigma}{\rho_1}^{n-1}\left|\int_0^{\rho_1}\left[\left(\frac{r}{\rho_1}\right)^n\partial_r u(r\sigma) + \frac{nr^{n-1}}{{\rho_1}^n}u(r\sigma)\right]\mathrm{d}r\right|\mathrm{d}s(\sigma) \\ &\leqslant \int_{\Sigma}{\rho_1}^{-1}\int_0^{\rho_1}r^n|\partial_r u(r\sigma)|\mathrm{d}r\,\mathrm{d}s(\sigma) + n\int_{\Sigma}{\rho_1}^{-1}\int_0^{\rho_1}r^{n-1}|u(r\rho_1)|\mathrm{d}r\,\mathrm{d}s(\sigma) \\ &\leqslant \|\nabla u\|_{L^1(B)} + \frac{n}{\rho_1}\|u\|_{L^1(B)}. \end{aligned}$$

Combining all our estimates, we obtain

$$\begin{aligned} \|u\|_{L^1(K)} \leqslant \|u\|_{L^1(B)} &+ \frac{\rho_1}{n}(\gamma^n - 1)\|\nabla u\|_{L^1(K\setminus B)} \\ &+ \frac{\rho_1}{n}(\gamma^n - 1)\|\nabla u\|_{L^1(B)} + (\gamma^n - 1)\|u\|_{L^1(B)} \\ &= \gamma^n\|u\|_{L^1(B)} + \frac{\rho_1}{n}(\gamma^n - 1)\|\nabla u\|_{L^1(K)} \end{aligned}$$

which gives (3.3).

To obtain the second result, we note that the Poincaré inequality on balls takes the form (see Acosta & Durán (2004), where this is proved for arbitrary convex sets)

$$\|u\|_{L^1(B)} \leqslant \rho_1\|\nabla u\|_{L^1(B)} \quad \forall\, u \in W^{1,1}(B), \;\; (u)_B = 0. \tag{A.1}$$

Thus, (3.4) follows immediately from (3.3).

A.2 *Auxiliary results*

LEMMA A1 Let $(\mathscr{T}_h)_{h\in(0,1]}$ be a family of partitions of Ω satisfying Assumption 2.1. Then, for each $p, q \in [1, \infty]$, there exists a constant $C > 0$, independent of h, such that for any $\kappa \in \mathscr{T}_h$,

$$h_\kappa^{-\frac{n}{p}} \|v\|_{L^p(\kappa)} \leqslant C h_\kappa^{-\frac{n}{q}} \|v\|_{L^q(\kappa)} \quad \forall\, v \in S^k(\mathscr{T}_h), \ \ \forall\, h \in (0, 1].$$

Moreover, for any $\widetilde{\kappa} \in \widetilde{\mathscr{T}}_h$,

$$h_{\widetilde{\kappa}}^{-\frac{n}{p}} \|v\|_{L^p(\widetilde{\kappa})} \leqslant C h_{\widetilde{\kappa}}^{-\frac{n}{q}} \|v\|_{L^q(\widetilde{\kappa})} \quad \forall\, v \in S^1(\widetilde{\mathscr{T}}_h) + S^k(\mathscr{T}_h), \ \ \forall\, h \in (0, 1].$$

Proof. Let $\kappa \in \mathscr{T}_h$, $\hat{\kappa}$ its corresponding reference element and $F_\kappa : \hat{\kappa} \to \kappa$ the associated mapping. We set $J = |\det \nabla F_\kappa|$. Since F_κ is bi-Lipschitz, we have $C^{-1} h_\kappa^n \leqslant J \leqslant C h_\kappa^n$ for some constant C which is independent of κ. From the area formula (cf. Evans & Gariepy, 1992), we have

$$\int_\kappa |u|^p \,\mathrm{d}x = \int_{\hat{\kappa}} J |u \circ F_\kappa|^p \,\mathrm{d}x \approx h_\kappa^n \int_{\hat{\kappa}} |u \circ F_\kappa|^p \,\mathrm{d}x.$$

Using norm equivalence in finite-dimensional spaces, we obtain

$$\int_\kappa |u|^p \,\mathrm{d}x \approx h_\kappa^n \left(\int_{\hat{\kappa}} |u \circ F_\kappa|^q \,\mathrm{d}x \right)^{p/q} \approx h_\kappa^{n - np/q} \left(\int_\kappa |u|^q \,\mathrm{d}x \right)^{p/q}.$$

The first equivalence follows by taking the p root.

The second equivalence is proved with the same technique, after noting that given $v \in S^1(\widetilde{\mathscr{T}}_h) + S^k(\mathscr{T}_h)$, then $v|_{\widetilde{\kappa}}$ is a polynomial of degree k. Thus, the previous reasoning applies. □

LEMMA A2 Let $S^k(\mathscr{T}_h)$ be defined as in Section 2 and let the mesh family satisfy Assumption 2.1. Then, for each $p \in [1, \infty)$, there exists a constant C, independent of h, such that

$$\inf_{u \in S^k(\mathscr{T}_h)} \sup_{v \in S^k(\mathscr{T}_h)} \frac{\int_\Omega uv \,\mathrm{d}x}{\|u\|_{L^p(\Omega)} \|v\|_{L^{p'}(\Omega)}} \geqslant C > 0.$$

Proof. For a given $u \in L^p(\Omega)$, set $v = |u|^{p-2} u$ so that $\int_\Omega uv = \|u\|_{L^p(\Omega)} \|v\|_{L^{p'}(\Omega)}$. At the discrete level, if $u \in S^k(\mathscr{T}_h)$, the choice $v = |u|^{p-2} u$ is not allowed, in general. Instead, we set $v = \Pi_k(|u|^{p-2} u)$, where Π_k denotes the L^2-projection onto $S^k(\mathscr{T}_h)$ (note that this is a projection element by element), and therefore

$$\|\Pi_k u\|_{L^2(\kappa)}^2 = \int_\kappa u \Pi_k u \,\mathrm{d}x \leqslant \|u\|_{L^{p'}(\kappa)} \|\Pi_k u\|_{L^p(\kappa)} \quad \forall\, \kappa \in \mathscr{T}_h.$$

Using Lemma A1, we obtain

$$\|\Pi_k u\|_{L^{p'}(\kappa)} \leqslant C_\Pi \|u\|_{L^{p'}(\kappa)} \quad \forall\, \kappa \in \mathscr{T}_h,$$

where C_Π is independent of h and κ. Moreover, by the definition of Π_k, it holds that $\int_\Omega u \Pi_k v \,\mathrm{d}x = \int_\Omega uv \,\mathrm{d}x$ for all $u \in S^k(\mathscr{T}_h)$. A possible value for the constant C in the statement is therefore given by $1/C_\Pi$. □

The last result which we prove in this appendix allows us deduce strong convergence of a sequence from its weak convergence together with convergence of a strictly convex energy. This result is well known and the proof is a straightforward adaption of Pedregal (1997, Theorem 3.16). However, we did not find a precise statement suited for our specific needs and therefore prefer to give a sketch of the proof.

LEMMA A3 Let $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^k$ be a Carathéodory function satisfying the growth condition

$$|f(x,u,v)| \leqslant c(1+|u|^q+|v|^p)$$

and such that $f(x,u,\cdot)$ is strictly convex for a.a. $x \in \Omega$ and for all $u \in \mathbb{R}^m$.

If $u_j \to u$ strongly in $L^q(\Omega)^m$ and $v_j \rightharpoonup v$ weakly in $L^p(\Omega)^k$ and if

$$\lim_{j\to\infty}\int_\Omega f(x,u_j,v_j)\mathrm{d}x = \int_\Omega f(x,u,v)\mathrm{d}x,$$

then $v_j \to v$ strongly in $L^p(\Omega)^k$.

Proof. The proof requires the machinery of Young measures which we cannot introduce at this point. A nice introduction is given in Pedregal (2000). Suffice to say that Young measures give a more precise description of weak limits and, when the functional (6.3) is extended in a suitable way, it becomes continuous under weak convergence.

Let $(\mu_x)_{x\in\Omega}$ be the Young measure generated by (a subsequence of) $(v_j)_{j\in\mathbb{N}}$. Then, $(\delta_{u(x)}\otimes\mu_x)_{x\in\Omega}$ is the Young measure generated by the pairs $(u_j,v_j)_{j\in\mathbb{N}}$. Using the assumptions of the lemma and Pedregal (2000, Corollary 5.7), we can estimate

$$\begin{aligned}
\lim_{j\to\infty}\int_\Omega f(x,u_j,v_j)\mathrm{d}x &= \int_\Omega f(x,u(x),v(x))\mathrm{d}x\\
&= \int_\Omega f\left(x,u(x),\int_{\mathbb{R}^k} z\,\mathrm{d}\mu_x(z)\right)\mathrm{d}x\\
&\leqslant \int_\Omega\int_{\mathbb{R}^k} f(x,u(x),z)\,\mathrm{d}\mu_x(z)\mathrm{d}x\\
&= \int_\Omega\int_{\mathbb{R}^m\otimes\mathbb{R}^k} f(x,z',z)\mathrm{d}(\delta_{u(x)}\otimes\mu_x)(z',z)\mathrm{d}x\\
&= \lim_{j\to\infty}\int_\Omega f(x,u_j,v_j)\mathrm{d}x.
\end{aligned}$$

Thus, equality must hold in the inequality of line three, which means that

$$f\left(x,u(x),\int_{\mathbb{R}^k} z\,\mathrm{d}\mu_x(z)\right) = \int_{\mathbb{R}^k} f(x,u(x),z)\mathrm{d}\mu_x(z) \quad \text{for a.a. } x\in\Omega.$$

By assumption, $f(x,u(x),\cdot)$ is strictly convex for a.e. x and hence $\mu_x = \delta_{\overline{\mu_x}} = \delta_{v(x)}$.

Now, we can use Pedregal (2000, Corollary 5.7) again to deduce that

$$\lim_{j\to\infty}\int_\Omega |v_j|^p\,\mathrm{d}x = \int_\Omega\int_{\mathbb{R}^k}|z|^p\mu_x(\mathrm{d}z)\mathrm{d}x = \int_\Omega |v|^p\,\mathrm{d}x,$$

and therefore $v_j \to v$ strongly in L^p (see also Pedregal, 2000, Lemma 5.8). □