ESAIM: COCV 14 (2008) 802–824
DOI: 10.1051/cocv:2008011

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

EXTERNAL APPROXIMATION OF FIRST ORDER VARIATIONAL PROBLEMS *VIA* $W^{-1,p}$ ESTIMATES

Cesare Davini[1] and Roberto Paroni[2]

Abstract. Here we present an approximation method for a rather broad class of first order variational problems in spaces of piece-wise constant functions over triangulations of the base domain. The convergence of the method is based on an inequality involving $W^{-1,p}$ norms obtained by Nečas and on the general framework of Γ-convergence theory.

Mathematics Subject Classification. 65N12, 65N30, 46N10, 74K20, 74S05

Received December 28, 2006.
Published online January 30, 2008.

Introduction

A classical problem in the calculus of variations is: find the minimizers of the functional

$$\mathcal{F}(v) = \int_\Omega W(x, v, \nabla v)\,\mathrm{d}x - \langle f, v\rangle$$

among all functions $v \in W^{1,p}(\Omega)$ with trace equal to $w \in W^{1,p}(\Omega)$ over a subset (of positive length) $\partial_u\Omega$ of the boundary of Ω, where $1 < p < +\infty$, $\Omega \subset \mathbb{R}^2$ is an open bounded set with Lipschitz boundary, $f \in L^q(\Omega)$ and $W : \Omega \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ is a Carathéodory function convex in the last variable and satisfying a standard p-growth from below and above (see Sect. 5 for the precise requirements).

Different schemes have been developed in order to find an approximation of the minimizer(s) of the problem above. Probably, the most popular is the technique based on the use of continuous piece-wise affine finite elements. This simple (internal) approximation is particularly advantageous when W depends on ∇v only, because then the integrand is constant on each element of the triangulation of the base domain. Higher order approximants have also been used. These on one hand give a better rate of convergence but on the other hand make the numerical scheme more complex.

Our point of view here is to consider the space which makes the numerical scheme for general W as simple as possible, which is the space of piece-wise constant functions over triangulations of the base domain.

Our approach can be classified as a *discontinuous Galerkin* (DG) method, although the techniques and the functional framework we use are not common in that context. The DG methods were first proposed

Keywords and phrases. Numerical methods, non-conforming approximations, Γ-convergence.

[1] Dipartimento di Georisorse e Territorio, Via del Cotonificio 114, 33100 Udine, Italy; davini@uniud.it

[2] Dipartimento di Architettura e Pianificazione, Università di Sassari, Palazzo del Pou Salit, Piazza Duomo, 07041 Alghero, Italy; paroni@uniss.it

for hyperbolic problems by Reed and Hill [28]. Discontinuous approximations for elliptic or parabolic-elliptic problems were also introduced in about the same years by various authors, *e.g.*, Babuška [5], Babuška and Zlámal [6], Wheeler [29], Arnold [2], and evolved somewhat independently of those for the hyperbolic case. Unlike other standard finite element approaches, DG methods do not require continuity of the approximation functions across the interelement boundary, but use a penalty to enforce the matching of the values that the functions take on contiguous elements of the mesh. The idea was derived from Lions [25], who introduced a penalty term in order to impose the respect of rough Dirichlet data in elliptic problems, and it was imported in the context of the approximation theory by Aubin [4], Nitsche [27] and Babuška [5]. Allowing for discontinuities yields some flexibility under the point of view of both the functional setting and the computational efficiency. This is especially true for higher order problems. For the second order ones the advantages are less obvious, but the DG methods maintain interest because they provide "high order accuracy, high parallelizability, localizability and easy handling of complicated geometries", as noticed by Ye [30]. Thus, in spite of the fact that they were never proven to be more advantageous than the classical conforming finite element methods, these features have kept the interest for DG methods alive and have led to some revival of attention for their potential applications in recent years, see, *e.g.*, [7–10,14]. An account of the development of the DG methods is given by Cockburn, Karniadakis and Shu [15]. The various families of DG methods are characterized by different choices of the terms describing the non-conformity. Arnold *et al.* describe a common setting and discuss some relevant properties of a number of these methods in [3].

For the problem described at the beginning of the Introduction, in this paper we aim to construct a sequence of discrete functionals, defined in spaces of piece-wise constant functions, whose minima and relative minimizers converge to those of the original problem. We simply prove that this kind of approximation is possible without worrying, this will be done in a subsequent paper, to estimate the rate of convergence of the scheme. To achieve this goal we shall use the definition and the techniques of Γ-convergence's theory.

By using piece-wise constant functions the first problem at our hand is to define what we mean by gradient. At each nodal point x_i of the triangulation of the base domain we call *generalized gradient* of a piece-wise constant function a suitable mean of the distributional gradient on a dual element around the point x_i (see equation (3.1) and the remark after it). This notion is then extended to the full triangulation by taking the generalized gradient constant on each dual element (see (3.2)).

Despite the given name, the generalized gradient is not a gradient, even though we show that it has some of the properties which are peculiar to a gradient. In particular we prove that if a sequence of generalized gradients weakly converges in L^p then the weak limit is a gradient (see Th. 3.1). The generalized gradient instead does not have the "imbedding property" of a gradient, which is: a sequence of piece-wise constant functions which weakly converges together with the sequence of the generalized gradients does not necessarily strongly converges (see the example before Lem. 3.2). This property is recovered by requiring that a certain weighted L^p norm of the jumps of the piece-wise constant function across the edges of the mesh should tend to zero as the size of the triangulation goes to zero (see Th. 3.5). This is proved by strongly relying on an inequality due to Nečas. The lacking of this imbedding property strongly influences the definition of the discrete functionals which approximate the original one. Indeed in order to make the sequence of the approximating functionals coercive, in the appropriate norm, it is necessary to include a weighted norm of the jumps of the piece-wise constant functions in the definition.

On one side, our approach is close to the Finite Volume or Covolume methods. In particular, the introduction of a generalized gradient operator reminds the method by Andreianov, Gutnic and Wittbold [1], although a comparison is hardly made. More simply, here we carry over ideas discussed in [20,22,23], for second order variational problems, and in [18], for the Poisson problem in 2-dimensional domains.

The present paper generalizes the analysis of [18] under several respects. First of all, it takes into account mixed boundary conditions and extends the results from the Hilbertian case to convergence in L^p spaces, with $1 < p < +\infty$. We use an inequality obtained by Nečas and involving the norm in $W^{-1,p}$ as a basis for proving coerciveness of the approximating functionals. In particular, this extends the proof of convergence to the case of domains with Lipschitz boundary. Finally, we recourse to classical results from the direct methods of the Calculus of Variations to show how the method studied in [18] can be modified in order to cover functionals of the class described above.

As said above, we confine our attention to unqualified convergence and postpone the deduction of error estimates to a future work [21]. Some numerical applications of the method for the quadratic homogeneous case are found in [19].

1. Preliminaries and notation

Throughout the paper p will denote a real number such that $1 < p < +\infty$, $q := p/(p-1)$, and $\Omega \subset \mathbb{R}^2$ is an open bounded connected set with Lipschitz boundary. With $W^{-1,p}(\Omega)$ we denote the dual of $W_0^{1,q}(\Omega)$. The norm on $W^{-1,p}(\Omega)$ is defined by

$$\|f\|_{W^{-1,p}(\Omega)} = \sup\{|\langle f, v\rangle| : v \in W_0^{1,q}(\Omega), \text{ and } \|v\|_{W_0^{1,q}(\Omega)} \leq 1\}.$$

As $W_0^{1,q}(\Omega)$ is reflexive, a sequence $\{f_n\} \subset W^{-1,p}(\Omega)$ weakly converges to $f \in W^{-1,p}(\Omega)$ if

$$\langle f_n, v\rangle \to \langle f, v\rangle,$$

for every $v \in W_0^{1,q}(\Omega)$. We shall write $f_n \rightharpoonup f$ in $W^{-1,p}(\Omega)$.

If $f_n \rightharpoonup f$ in $W^{-1,p}(\Omega)$ then $\|f_n\|_{W^{-1,p}(\Omega)}$ is bounded and $\liminf \|f_n\|_{W^{-1,p}(\Omega)} \geq \|f\|_{W^{-1,p}(\Omega)}$. Moreover, if $v_n \to v$ in $W_0^{1,q}(\Omega)$, then $\langle f_n, v_n\rangle \to \langle f, v\rangle$.

We further recall that by the Banach-Alaoglu-Bourbaki theorem the set $\{f \in W^{-1,p}(\Omega) : \|f\|_{W^{-1,p}(\Omega)} \leq 1\}$ is weakly compact. Moreover for $1 < q < +\infty$ the spaces $W^{1,q}(\Omega)$ are separable and hence the unit ball in $W^{-1,p}(\Omega)$ is relatively sequentially weakly compact; thus a bounded sequence in $W^{-1,p}(\Omega)$ has a weakly converging subsequence.

Since the imbedding of $W_0^{1,q}(\Omega)$ in $L^q(\Omega)$ is compact, by Schauder's theorem, see Brezis [11], also the imbedding of $L^p(\Omega)$ in $W^{-1,p}(\Omega)$ is compact, and hence weakly converging sequences in $L^p(\Omega)$ are strongly converging in $W^{-1,p}(\Omega)$.

Let $v \in L^p(\Omega)$, from the definition of the norm in $W^{-1,p}(\Omega)$ it immediately follows that

$$\left(\|v\|_{W^{-1,p}(\Omega)} + \|Dv\|_{W^{-1,p}(\Omega)}\right) \leq c\,\|v\|_{L^p(\Omega)}$$

for some constant c independent of v. Nečas in [26] has proved that also the reverse inequality holds, more precisely he has shown that there exists a constant C, such that for every distribution v

$$\|v\|_{L^p(\Omega)} \leq C\left(\|v\|_{W^{-1,p}(\Omega)} + \|Dv\|_{W^{-1,p}(\Omega)}\right), \tag{1.1}$$

where the constant C does not depend on v.

In what follows it is convenient to regard functions v as defined in all $\mathbb{R}^2$ by extending them to zero outside of Ω, and Dv as elements of $W^{-1,p}(\mathbb{R}^2)$ with support in $\bar{\Omega}$. Let $\partial_t\Omega$ be a subset of $\partial\Omega$ and $W^{1,q}_{\partial_t\Omega}(\Omega) \subset W^{1,q}(\mathbb{R}^2)$ be the subspace of functions vanishing at $\partial_t\Omega$. By definition, for every $f \in (W^{1,q}_{\partial_t\Omega}(\Omega))'$

$$\|f\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} = \sup_{g\in W^{1,q}_{\partial_t\Omega}(\Omega)} \frac{\langle f, g\rangle}{\|g\|_{W^{1,q}(\Omega)}},$$

and, since $W_0^{1,q}(\Omega) \subset W^{1,q}_{\partial_t\Omega}(\Omega)$, we have

$$\|f\|_{W^{-1,p}(\Omega)} \leq \|f\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'},$$

where f denotes a continuous linear functional on $W^{1,q}_{\partial_t\Omega}(\Omega)$ in the right hand side, and its restriction to $W^{1,q}_0(\Omega)$ in the left hand side.

Rewriting equation (1.1) as

$$\|v\|_{L^p(\Omega)} \le C\left(\|v\|_{W^{-1,p}(\Omega)} + \|Dv\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'}\right), \tag{1.2}$$

we can prove the following lemma.

Hereafter, we denote by $\mathcal{H}^1$ the one-dimensional Hausdorff measure.

Lemma 1.1. *Let $\partial_t\Omega$ and $\partial_u\Omega$ be two disjoint subsets of $\partial\Omega$ such that*

$$\partial_t\Omega \cup \partial_u\Omega = \partial\Omega, \quad \textit{with } \mathcal{H}^1(\partial_u\Omega) > 0.$$

Then, there exists a constant C such that

$$\|v\|_{L^p(\Omega)} \le C\|Dv\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} \qquad \forall\, v \in L^p(\Omega). \tag{1.3}$$

Proof. Suppose not. Then there exists a sequence $\{v_n\} \subset L^p(\Omega)$ such that

$$\|v_n\|_{L^p(\Omega)} = 1 \quad \text{and} \quad \|Dv_n\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} \le \frac{1}{n}\cdot$$

Then, up to a subsequence, we have $v_n \rightharpoonup v$ in $L^p(\Omega)$ and therefore $v_n \to v$ in $W^{-1,p}(\Omega)$. Since $Dv_n \to 0$ in $(W^{1,q}_{\partial_t\Omega}(\Omega))'$, from inequality (1.2) we deduce that v_n is a Cauchy sequence in $L^p(\Omega)$ and therefore that $v_n \to v$ in $L^p(\Omega)$. But $0 = \liminf_n \|Dv_n\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} \ge \|Dv\|_{W^{-1,p}(\Omega)}$ and therefore v is constant almost everywhere, $v = \text{const.} =: k$. Then, from

$$\langle Dv, g\rangle = -\int_\Omega v\,\nabla g\,\mathrm{d}x,$$

which holds for every $g \in W^{1,q}_{\partial_t\Omega}(\Omega)$, it follows that

$$0 = \langle Dv, g\rangle = -\int_\Omega v\,\nabla g\,\mathrm{d}x = -k\int_{\partial_u\Omega} g\nu\,\mathrm{d}\mathcal{H}^1 \qquad \forall\, g \in W^{1,q}_{\partial_t\Omega}(\Omega)$$

which is false unless $k = 0$. But this contradicts the fact that $\|v\|_{L^p(\Omega)} = 1$. □

2. Discretization of the domain

We recall that $\Omega \subset \mathbb{R}^2$ denotes an open, bounded set with Lipschitz boundary. Let $\mathcal{T}_h := \{T_j\}_{j=1,\dots,P_h}$, with h taking values in some countable set $\mathcal{H}$ of real numbers, be a sequence of triangulations of Ω regular in the sense of Ciarlet [12], *i.e.*, such that the ratio between $\rho_h = \inf_j \sup\{\text{diam}\,(S) : S \text{ is a disk contained in } T_j\}$ and $h := \sup_j\{\text{diam}\,T_j\}$ is bounded away from zero by a constant independent of h. We shall call $\mathcal{T}_h$ *the primal mesh*. Let $\Omega_h := \overset{\circ}{\overline{\bigcup_{T_j\in\mathcal{T}_h} T_j}}$ approximate Ω from inside, *i.e.*, there exists a constant $C > 0$ such that $\text{dist}(x, \partial\Omega) \le Ch$ for every $x \in \partial\Omega_h$.

We denote by x_i the vertices of the triangles T_j and call them the nodes of the mesh. We indicate by $\mathcal{P}_h := \{1, 2, \dots, P_h\}$ and $\mathcal{N}_h := \{1, 2, \dots, N_h\}$ the sets of values taken by the indexes of the triangles and the mesh nodes, respectively. We will indicate by $\mathcal{I}_h$ and $\mathcal{B}_h$ the sets of the index values corresponding to the nodes in $\overset{\circ}{\Omega}_h$ and $\partial\Omega_h$, respectively.

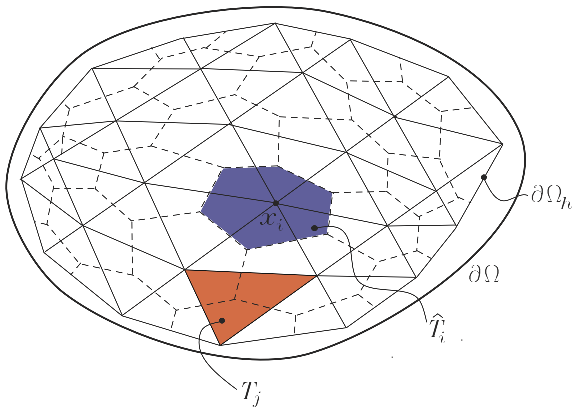

FIGURE 1. Ω, Ω_h, the primal and the dual mesh.

Following Davini and Pitacco [22,23], for each $h \in \mathcal{H}$ we also introduce a *dual mesh* $\hat{\mathcal{T}}_h := \left\{\hat{T}_i\right\}_{i=1,\ldots,N_h}$ consisting of disjoint open polygonal domains, each containing just one primal node, as shown in Figure 1, where the dual elements are drawn with dashed lines. We assume that the sequence of dual meshes is also regular and that $\Omega_h = \overset{\circ}{\overline{\bigcup_{\hat{T}_i \in \hat{\mathcal{T}}_h} \hat{T}_i}}$.

Let X_h be the space of functions which are affine on T_j and continuous on Ω_h (briefly, the *polyhedral functions* over $\mathcal{T}_h$), and let $X_{0h} \subset X_h$ denote the set of functions that vanish on $\partial\Omega_h$. We regard X_{0h} as a subspace of $H_0^1(\Omega)$ by extending the functions to zero in $\Omega \setminus \Omega_h$. $\hat{\varphi}_i$ will be the polyhedral splines in X_h defined by the condition that $\hat{\varphi}_i(x_j) = \delta_{ij}$ for $i, j = 1, \ldots, N_h$.

Although we never need to construct the dual mesh explicitly, as it will be clear later, we assume that the area of the dual elements satisfies the condition

$$|\hat{T}_i| = \int_{\Omega_h} \hat{\varphi}_i \,\mathrm{d}x \left(= \frac{1}{3}|\mathrm{supp}(\hat{\varphi}_i)|\right). \tag{2.1}$$

Finally, let us define

$$Y_h := \big\{v : v = \text{const. on } T_j \in \mathcal{T}_h, \text{ with } v = 0 \text{ in } \mathbb{R}^2 \setminus \Omega_h\big\}. \tag{2.2}$$

Throughout the paper we shall denote the functions of Y_h with an overline, to remind us that they are piecewise constant. For instance we will write $\bar{v}_h$.

3. Generalized gradient: definition and properties

From the definition it follows that $Y_h \subset L^p(\mathbb{R}^2)$. Let $\bar{v}_h \in Y_h$, then the distributional gradient $D\bar{v}_h$ belongs to $W^{-1,p}(\mathbb{R}^2)$ and has support in $\bar{\Omega}_h$, thus the following definition makes sense

$$\nabla_h \bar{v}_h(x_i) := \frac{\langle D\bar{v}_h, \hat{\varphi}_i\rangle}{|\hat{T}_i|} = -\frac{1}{|\hat{T}_i|}\int_{\Omega_h} \bar{v}_h \,\nabla\hat{\varphi}_i \,\mathrm{d}x \tag{3.1}$$

for all $i \in \mathcal{N}_h$. When $D\bar{v}_h$ is interpreted as a Radon measure on $\mathbb{R}^2$ and $\hat{T}_i$ cuts the sides of mesh $\mathcal{T}_h$ at the midpoints, a simple computation shows that

$$\nabla_h \bar{v}_h(x_i) = \frac{D\bar{v}_h(\hat{T}_i)}{|\hat{T}_i|}\cdot$$

Thus, $\nabla_h \bar{v}_h(x_i)$ is the mean value of the gradient on $\hat{T}_i$ and will be called then the *generalized gradient* of $\bar{v}_h$ in $\hat{T}_i$. Note that, while for the inner nodes these quantities have an intrinsic meaning, for the boundary elements they account for the extension of $\bar{v}_h$ to zero outside of Ω_h and are affected by the jumps across $\partial\Omega_h$. The implication of this in treating the boundary value problems will be considered in the following section.

Here we introduce the simple function

$$\mathring{\nabla}_h \bar{v}_h(x) := \sum_{i\in\mathcal{I}_h} \nabla_h \bar{v}_h(x_i)\chi_{\hat{T}_i}(x), \tag{3.2}$$

where χ_R denotes the characteristic function of the region R, *i.e.*, $\chi_R(x) = 1$ if $x \in R$, and otherwise equal to zero. We note that $\mathring{\nabla}_h \bar{v}_h \in L^p(\Omega)$.

For $i \in \mathcal{I}_h$, equation (3.1) can be written in an integral form with the use of the following definitions. Given a continuous function f we define

$$c_h f(x) := \sum_{i\in\mathcal{N}_h} f(x_i)\chi_{\hat{T}_i}(x), \tag{3.3}$$

$$\mathring{r}_h f(x) := \sum_{i\in\mathcal{I}_h} f(x_i)\hat{\varphi}_i(x), \quad \text{and} \quad r_h f(x) := \sum_{i\in\mathcal{N}_h} f(x_i)\hat{\varphi}_i(x). \tag{3.4}$$

We note that if f is continuous with compact support in Ω, $f \in C_0(\Omega)$, then $\mathring{r}_h f = r_h f$, for h small enough. Let $g \in C_0(\Omega)$. Then, multiplying (3.1) by $g(x_i)|\hat{T}_i|$ and summing over $i \in \mathcal{I}_h$ we obtain

$$\langle D\bar{v}_h, \mathring{r}_h g\rangle = \int_{\Omega_h} \mathring{\nabla}_h \bar{v}_h \, c_h g \, \mathrm{d}x. \tag{3.5}$$

Theorem 3.1. *Let $\bar{v}_h \in Y_h$ for every $h \in \mathcal{H}$. Assume that*

$$\sup_h \|\mathring{\nabla}_h \bar{v}_h\|_{L^p(\Omega)} < +\infty, \quad \textit{and} \quad \bar{v}_h \rightharpoonup v \textit{ in } L^p(\Omega).$$

Then

$$v \in W^{1,p}(\Omega) \quad \textit{and} \quad \mathring{\nabla}_h \bar{v}_h \rightharpoonup \nabla v \textit{ in } L^p(\Omega).$$

Proof. Since $\bar{v}_h \rightharpoonup v$ in $L^p(\Omega)$ we have that $D\bar{v}_h \rightharpoonup Dv$ in $W^{-1,p}(\Omega)$. Moreover, since $\sup_h \|\mathring{\nabla}_h \bar{v}_h\|_{L^p(\Omega)} < +\infty$, up to a subsequence, $\mathring{\nabla}_h \bar{v}_h \rightharpoonup A$ in $L^p(\Omega)$, for some $A \in L^p(\Omega)$ with values in $\mathbb{R}^2$. Let $g \in C_0^\infty(\Omega)$, then, since $\mathring{r}_h g \to g$ in $W^{1,q}(\Omega)$ and $c_h g \to g$ in $L^q(\Omega)$, by equation (3.5) we have

$$\langle Dv, g\rangle = \lim_{h\to 0}\langle D\bar{v}_h, \mathring{r}_h g\rangle = \lim_{h\to 0}\int_\Omega \mathring{\nabla}_h \bar{v}_h \, c_h g \, \mathrm{d}x = \int_\Omega A \, g \, \mathrm{d}x.$$

Hence $Dv = \nabla v = A$ and the proof is concluded. □

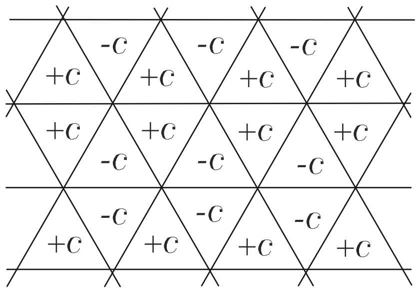

FIGURE 2. Representation of the sequence used in the example.

The generalized gradient satisfies also a Green type equality. In fact, for $g \in L^1(\Omega)$ we have

$$\begin{aligned}\int_\Omega \mathring{\nabla}_h \bar{v}_h \, g \,\mathrm{d}x &= \int_\Omega \sum_{i\in\mathcal{I}_h} \frac{\langle D\bar{v}_h, \hat{\varphi}_i\rangle}{|\hat{T}_i|} \chi_{\hat{T}_i} \, g \,\mathrm{d}x = \sum_{i\in\mathcal{I}_h} \langle D\bar{v}_h, \hat{\varphi}_i\rangle \frac{1}{|\hat{T}_i|} \int_{\hat{T}_i} g \,\mathrm{d}x \\ &= -\sum_{i\in\mathcal{I}_h} \int_{\Omega_h} \bar{v}_h \nabla\hat{\varphi}_i \,\mathrm{d}x \, \frac{1}{|\hat{T}_i|} \int_{\hat{T}_i} g \,\mathrm{d}x,\end{aligned}$$

and setting

$$\mathring{m}_h g(x) := \sum_{i\in\mathcal{I}_h} \frac{1}{|\hat{T}_i|} \int_{\hat{T}_i} g \,\mathrm{d}y \, \hat{\varphi}_i(x) =: \sum_{i\in\mathcal{I}_h} \fint_{\hat{T}_i,} g \,\mathrm{d}y \, \hat{\varphi}_i(x), \tag{3.6}$$

we obtain

$$\int_\Omega \mathring{\nabla}_h \bar{v}_h \, g \,\mathrm{d}x = \langle D\bar{v}_h, \mathring{m}_h g\rangle = -\int_\Omega \bar{v}_h \nabla \mathring{m}_h g \,\mathrm{d}x. \tag{3.7}$$

The theorem and the equality above show that some of the properties enjoyed by the gradient of a function hold also for the generalized gradient, of course with the appropriate modifications. The example below shows that it is not always so. Indeed, it is well known that if a sequence $v_h \rightharpoonup v$ in $L^p(\Omega)$, and $\nabla v_h \rightharpoonup \nabla v$ in $L^p(\Omega;\mathbb{R}^2)$ then by compactness $v_h \to v$ in $L^p(\Omega)$. The next example shows that this is not true if we replace the gradient with the generalized gradient.

Example. Let us consider a triangulation made of equilateral triangles of side length h. Let the dual mesh be the one obtained by joining the center of adjacent triangles. Furthermore, let $\bar{v}_h$ be the functions that take the values $+c$ and $-c$ as represented in Figure 2, with c a fixed constant. We then have that $\bar{v}_h \rightharpoonup 0$ in $L^p(\Omega)$, but not strongly. Also, an easy computation shows that $\mathring{\nabla}_h \bar{v}_h = 0$.

In the rest of the section we will look for a condition on the sequence $\bar{v}_h$ that guarantees the strong convergence from the weak convergence of $\bar{v}_h$ and the boundedness of the L^p norm of its generalized gradient. This will be achieved by using inequality (1.2) after having deduced an appropriate estimate of the $W^{-1,p}$ norm of the distributional gradient of $\bar{v}_h$. This last estimate will follow by appropriately manipulating equation (3.7) and after we have studied the convergence properties of $\mathring{m}_h g$.

Lemma 3.2. *Let $g \in W_0^{1,q}(\Omega)$. There exists a constant $c > 0$, independent of g, such that for all sufficiently small h we have*

$$\|g - \mathring{m}_h g\|_{L^q(\Omega)} \le ch\,\|\nabla g\|_{L^q(\Omega)}, \tag{3.8}$$

and

$$\|\nabla \mathring{m}_h g\|_{L^q(\Omega)} \le c\,\|\nabla g\|_{L^q(\Omega)}. \tag{3.9}$$

Proof. We start with the proof of inequality (3.9) and observe that if it holds for every $g \in C_0^1(\Omega)$ it holds also for every $g \in W_0^{1,q}(\Omega)$. This follows easily from the fact that if $g_\varepsilon \in C_0^1(\Omega)$ and $g_\varepsilon \to g$ in $W^{1,q}(\Omega)$, then $\mathring{m}_h g_\varepsilon \to \mathring{m}_h g$ in $W^{1,q}(\Omega)$, as ε goes to zero, since for fixed h the sum in (3.6) is finite.

So, let $g \in C_0^1(\Omega)$. Let $T \in \mathcal{T}_h$ be the generic triangle, and let i, j and k be the indexes of the nodes of the triangle T. The function $\mathring{m}_h g$ is affine on T and there exists a constant c such that

$$|\nabla \mathring{m}_h g(x)| \le \frac{c}{h}\,(|G_i - G_j| + |G_i - G_k| + |G_k - G_j|), \tag{3.10}$$

for every $x \in T$, where we have set

$$G_l := \fint_{\hat{T}_l} g(x)\,\mathrm{d}x, \quad \text{with } l = i, j, k.$$

Here we also used the regularity of the triangulation, which will be done without mention in the following. Assume, for the moment, that

$$|G_i - G_j|^q \le c\frac{h^{q+2}}{|\hat{T}_i|\,|\hat{T}_j|}\int_{S_{ij}} |\nabla g(x)|^q\,\mathrm{d}x \tag{3.11}$$

where S_{ij} denotes the convex hull of $\hat{T}_i \cup \hat{T}_j$, and that similar inequalities hold for the pairs of indexes $\{i, k\}$ and $\{k, j\}$.

Then,

$$\begin{aligned}\int_T \frac{|G_i - G_j|^q}{h^q}\,\mathrm{d}x &= \frac{|G_i - G_j|^q}{h^q}|T| \le c\frac{h^2}{|\hat{T}_i|\,|\hat{T}_j|}\,|T|\int_{S_{ij}} |\nabla g(x)|^q\,\mathrm{d}x\\ &\le c\int_{S_{ij}} |\nabla g(x)|^q\,\mathrm{d}x\end{aligned}$$

and hence, taking into account equation (3.10), we find

$$\int_T |\nabla \mathring{m}_h g(x)|^q\,\mathrm{d}x \le c\int_{S_T} |\nabla g(x)|^q\,\mathrm{d}x,$$

where $S_T := S_{ij} \cup S_{jk} \cup S_{ki}$ is the convex hull of $\hat{T}_i \cup \hat{T}_j \cup \hat{T}_k$. Summing the above equation over all triangles T in $\mathcal{T}_h$ we easily obtain equation (3.9). So it only remains to prove inequality (3.11), which is a Poincaré's kind of inequality and can be proved as Lemma 1 and Theorem 2 of Section 4.5.2 of Evans and Gariepy [24]. We sketch here the proof for completeness.

Let $z \in \hat{T}_j$. Then

$$\begin{aligned}
\int_{\hat{T}_i\cap\partial B(z,s)} |g(x)-g(z)|^q \,\mathrm{d}\mathcal{H}^1(x) &\le \int_{\hat{T}_l\cap\partial B(z,s)} \left|\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} g(z+t(x-z))\,\mathrm{d}t\right|^q \mathrm{d}\mathcal{H}^1(x),\\
&\le \int_0^1 \int_{\hat{T}_i\cap\partial B(z,s)} |\nabla g(z+t(x-z))|^q |x-z|^q \,\mathrm{d}\mathcal{H}^1(x)\,\mathrm{d}t,\\
&\le \int_0^1 \frac{1}{t}\int_{S_{ij}\cap\partial B(z,ts)} |\nabla g(y)|^q s^q \,\mathrm{d}\mathcal{H}^1(y)\,\mathrm{d}t,\\
&\le s^{q+1}\int_0^1 \int_{S_{ij}\cap\partial B(z,ts)} \frac{|\nabla g(y)|^q}{|y-z|}\,\mathrm{d}\mathcal{H}^1(y)\,\mathrm{d}t,\\
&\le s^q\int_0^s \int_{S_{ij}\cap\partial B(z,t)} \frac{|\nabla g(y)|^q}{|y-z|}\,\mathrm{d}\mathcal{H}^1(y)\,\mathrm{d}t,\\
&\le s^q\int_{S_{ij}\cap B(z,s)} \frac{|\nabla g(y)|^q}{|y-z|}\,\mathrm{d}y \le s^q\int_{S_{ij}} \frac{|\nabla g(y)|^q}{|y-z|}\,\mathrm{d}y,
\end{aligned}$$

and integrating the above inequality in ds between 0 and $\operatorname{diam}(\hat{T}_i \cup \hat{T}_j)$ we find

$$\int_{\hat{T}_i} |g(x)-g(z)|^q \,\mathrm{d}x \le ch^{q+1}\int_{S_{ij}} \frac{|\nabla g(y)|^q}{|y-z|}\,\mathrm{d}y.$$

Since the inequality above holds for every $z \in \hat{T}_j$, we have

$$\begin{aligned}
⨍_{\hat{T}_j} ⨍_{\hat{T}_i} |g(x)-g(z)|^q \,\mathrm{d}x\,\mathrm{d}z &\le c\frac{h^{q+1}}{|\hat{T}_i|\,|\hat{T}_j|}\int_{S_{ij}} |\nabla g(y)|^q \int_{\hat{T}_j}\frac{1}{|y-z|}\,\mathrm{d}z\,\mathrm{d}y,\\
&\le c\frac{h^{q+2}}{|\hat{T}_i|\,|\hat{T}_j|}\int_{S_{ij}} |\nabla g(y)|^q \,\mathrm{d}y,
\end{aligned}$$

where the last inequality follows by passing to polar coordinates in the integral $\int_{\hat{T}_j} \frac{1}{|y-z|}\,\mathrm{d}z$. Finally, using Jensen's inequality we deduce equation (3.11) and this concludes the proof of inequality (3.9).

We now prove inequality (3.8). Note that for $x \in \hat{T}_i$ we have

$$\mathring{m}_h g(x) = \mathring{m}_h g(x_i) + \nabla \mathring{m}_h g(x)\cdot(x-x_i) = ⨍_{\hat{T}_i} g(x)\,\mathrm{d}x + \nabla \mathring{m}_h g(x)\cdot(x-x_i).$$

Thus by Poincaré's inequality and equation (3.9) we have

$$\begin{aligned}
\|g-\mathring{m}_h g\|_{L^q(\hat{T}_i)} &\le \left\| g - ⨍_{\hat{T}_i} g(x)\,\mathrm{d}x\right\|_{L^q(\hat{T}_i)} + h\,\|\nabla \mathring{m}_h g\|_{L^q(\hat{T}_i)},\\
&\le (c+1)h\,\|\nabla g\|_{L^q(\hat{T}_i)},
\end{aligned}$$

and summing over i we deduce

$$\|g-\mathring{m}_h g\|_{L^q(\Omega_h)} \le ch\|\nabla g\|_{L^q(\Omega_h)}.$$

Taking into account that $\mathring{m}_h g = 0$ on $\Omega\setminus\Omega_h$ and, by Poincaré's inequality, that $\|g\|_{L^q(\Omega\setminus\Omega_h)} \le ch\|\nabla g\|_{L^q(\Omega\setminus\Omega_h)}$ we deduce equation (3.8). □

In the next theorem we deduce an estimate of the $W^{-1,p}$ norm of the distributional gradient of $\bar{v}_h$.

Theorem 3.3. *Let $\mathcal{T}_h = \{T_j\}_{j=1,\dots,P_h}$. There exists a constant c independent of h such that*

$$\|D\bar{v}_h\|_{W^{-1,p}(\Omega)} \leq \|\mathring{\nabla}_h \bar{v}_h\|_{W^{-1,p}(\Omega)} + c\left(h \int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1\right)^{1/p} \tag{3.12}$$

for all $\bar{v}_h \in Y_h$. Above, $[\![\bar{v}_h]\!]$ stands for the jump of $\bar{v}_h$ across the sides of the mesh.

Proof. Let $g \in W_0^{1,q}(\Omega;\mathbb{R}^2)$, with $\|g\|_{W^{1,q}(\Omega)} \leq 1$. Then, by(3.7),

$$\begin{aligned} \langle D\bar{v}_h, g\rangle &= \langle D\bar{v}_h, \mathring{m}_h g\rangle + \langle D\bar{v}_h, g - \mathring{m}_h g\rangle \\ &= \int_\Omega \mathring{\nabla}_h \bar{v}_h \cdot g \,\mathrm{d}x + \langle D\bar{v}_h, g - \mathring{m}_h g\rangle \end{aligned} \tag{3.13}$$

and hence

$$\|D\bar{v}_h\|_{W^{-1,p}(\Omega)} \leq \|\mathring{\nabla}_h \bar{v}_h\|_{W^{-1,p}(\Omega)} + \sup_{g \in W_0^{1,q}(\Omega)} \frac{|\langle D\bar{v}_h, g - \mathring{m}_h g\rangle|}{\|g\|_{W^{1,q}(\Omega)}}.$$

But

$$\begin{aligned} |\langle D\bar{v}_h, g - \mathring{m}_h g\rangle| &= \left| \int_{\cup_j \partial T_j} [\![\bar{v}_h]\!]\,(g - \mathring{m}_h g)\cdot \nu \,\mathrm{d}\mathcal{H}^1 \right| \\ &\leq \int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|\,|g - \mathring{m}_h g| \,\mathrm{d}\mathcal{H}^1 \\ &\leq \left(\int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1\right)^{1/p} \left(\int_{\cup_j \partial T_j} |g - \mathring{m}_h g|^q \,\mathrm{d}\mathcal{H}^1\right)^{1/q}, \end{aligned}$$

and since

$$\int_{\partial T_j} |g - \mathring{m}_h g|^q \,\mathrm{d}\mathcal{H}^1 \leq \frac{c}{h}\int_{T_j} |g - \mathring{m}_h g|^q \,\mathrm{d}x + ch^{q-1}\int_{T_j} |\nabla g - \nabla \mathring{m}_h g|^q \,\mathrm{d}x \tag{3.14}$$

where c does not depend on h, see Appendix, by Lemma 3.2 we deduce

$$\begin{aligned} \int_{\cup_j \partial T_j} |g - \mathring{m}_h g|^q \,\mathrm{d}\mathcal{H}^1 &\leq \frac{c}{h}\int_\Omega |g - \mathring{m}_h g|^q \,\mathrm{d}x + ch^{q-1}\int_\Omega |\nabla g - \nabla \mathring{m}_h g|^q \,\mathrm{d}x \\ &\leq \frac{c}{h} h^q \|g\|^q_{W^{1,q}(\Omega)} + ch^{q-1}\int_\Omega |\nabla g|^q \,\mathrm{d}x. \end{aligned}$$

Hence

$$\begin{aligned} |\langle D\bar{v}_h, g - \mathring{m}_h g\rangle| &\leq \left(\int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1\right)^{1/p} \left(ch^{q-1}\|g\|^q_{W^{1,q}(\Omega)}\right)^{1/q} \\ &= c\left(\int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1\right)^{1/p} h^{1/p}\|g\|_{W^{1,q}(\Omega)} \end{aligned}$$

and therefore we conclude that

$$\|D\bar{v}_h\|_{W^{-1,p}(\Omega)} \le \|\overset{\circ}{\nabla}_h \bar{v}_h\|_{W^{-1,p}(\Omega)} + c\left(h\int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1\right)^{1/p}. \qquad \square$$

The following lemma can be proved similarly, indeed it suffices to write an equation like (3.13) for k in place of h, subtract the equation in k from the one in h and proceed as above.

Lemma 3.4. *Let $\mathcal{T}_h = \{T_j\}_{j=1,\dots,P_h}$. There exists a constant c such that*

$$\begin{aligned}\|D\bar{v}_h - D\bar{v}_k\|_{W^{-1,p}(\Omega)} \le\ & \|\overset{\circ}{\nabla}_h \bar{v}_h - \overset{\circ}{\nabla}_k \bar{v}_k\|_{W^{-1,p}(\Omega)} \\ & + c\left(h\int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1\right)^{1/p} \\ & + c\left(k\int_{\cup_j \partial T_j} |[\![\bar{v}_k]\!]\,|^p \,\mathrm{d}\mathcal{H}^1\right)^{1/p}.\end{aligned}$$

Finally, the next theorem gives sufficient conditions to obtain strong convergence from weak convergence of the sequence $\{\bar{v}_h\}$. The theorem strengthens a similar result proved in [18], Theorem 3.

Theorem 3.5. *Let $\mathcal{T}_h = \{T_j\}_{j=1,\dots,P_h}$. If $\bar{v}_h \rightharpoonup v$ in $L^p(\Omega)$, $\overset{\circ}{\nabla}_h \bar{v}_h$ weakly converges in $L^p(\Omega)$ and*

$$\lim_{h\to 0} h\int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1 = 0,$$

then $\bar{v}_h \to v$ in $L^p(\Omega)$.

Proof. Since $\overset{\circ}{\nabla}_h \bar{v}_h$ weakly converges in $L^p(\Omega;\mathbb{R}^2)$ we have that it strongly converges in $W^{-1,p}(\Omega)$. Thus, from Lemma 3.4, we deduce that $\{D\bar{v}_h\}$ is a Cauchy sequence in $W^{-1,p}(\Omega)$. Hence, since $\bar{v}_h \rightharpoonup v$ in $L^p(\Omega)$ implies $D\bar{v}_h \rightharpoonup Dv$ in $W^{-1,p}(\Omega)$, we have $D\bar{v}_h \to Dv$ in $W^{-1,p}(\Omega)$. Also, since $\bar{v}_h \rightharpoonup v$ in $L^p(\Omega)$, we have $\bar{v}_h \to v$ in $W^{-1,p}(\Omega)$. The proof is concluded since

$$\|\bar{v}_h - v\|_{L^p(\Omega)} \le C\left(\|\bar{v}_h - v\|_{W^{-1,p}(\Omega)} + \|D\bar{v}_h - Dv\|_{W^{-1,p}(\Omega)}\right). \qquad \square$$

4. Generalized gradient at the boundary

In this section we want to study the approximation with piece-wise constant functions of a function $v \in W^{1,p}(\Omega)$ whose trace on $\partial_u\Omega$ equals that of a given function $w \in W^{1,p}(\Omega)$. To do so, we shall modify the definition of generalized gradient to the boundary of $\partial\Omega_h$ by taking into account that the inner value of $\bar{v}_h$ has to match w. We recall that $\partial\Omega$ is Lipschitz and divided into two complementary parts, $\partial_t\Omega$ and $\partial_u\Omega$, on which Neumann and Dirichlet boundary conditions are assigned, respectively. To avoid subtleties, we assume that $\partial_t\Omega$ and $\partial_u\Omega$ are finite unions of arcs, and, as previously done, that $\partial_u\Omega$ has strictly positive one-dimensional Hausdorff measure. Accordingly we define on $\partial\Omega_h$ two complementary parts as follows.

Let $\mathcal{B}_{h\partial_u}$ be a subset of $\mathcal{B}_h$ such that, if $\partial_u\Omega_h$ is the subset of $\partial\Omega_h$ generated by[1] the nodes x_i with $i \in \mathcal{B}_{h\partial_u}$, we have

$$\lim_{h\to 0}\left\{\sup_{x\in\partial_u\Omega} \operatorname{dist}(x,\partial_u\Omega_h) + \sup_{x\in\partial_u\Omega_h} \operatorname{dist}(x,\partial_u\Omega)\right\} = 0. \tag{4.1}$$

[1] With the sentence "$\partial_u\Omega_h$ is the subset of $\partial\Omega_h$ generated by the nodes x_i with $i \in \mathcal{B}_{h\partial_u}$" we mean that $\partial_u\Omega_h$ is the union of the sides of the triangles T_j lying on $\partial\Omega_h$ having a vertex in x_i with i in $\mathcal{B}_{h\partial_u}$.

For $i \in \mathcal{B}_{h\partial_u}$, we define

$$\nabla_h^{\partial_u\Omega}\bar{v}_h(x_i) := \nabla_h\bar{v}_h(x_i) + \frac{1}{|\hat{T}_i|}\int_{\partial_u\Omega_h} w\hat{\varphi}_i\nu\,\mathrm{d}\mathcal{H}^1, \tag{4.2}$$

where we recall, see equation (3.1), that

$$\nabla_h\bar{v}_h(x_i) = \frac{\langle D\bar{v}_h, \hat{\varphi}_i\rangle}{|\hat{T}_i|} = \frac{-1}{|\hat{T}_i|}\int_{\Omega_h}\bar{v}_h D\hat{\varphi}_i\,\mathrm{d}x. \tag{4.3}$$

The results of the previous section can be extended and in some case strengthened if we consider the functions

$$\nabla_h^{\partial_u\Omega}\bar{v}_h(x) := \sum_{i\in\mathcal{I}_h\bigcup\mathcal{B}_{h\partial_u}}\nabla_h^{\partial_u\Omega}\bar{v}_h(x_i)\chi_{\hat{T}_i}(x), \tag{4.4}$$

and, for $g \in L^1(\Omega)$,

$$m_h^{\partial_u\Omega}g(x) := \sum_{i\in\mathcal{I}_h\bigcup\mathcal{B}_{h\partial_u}} \fint_{\hat{T}_i,} g\,\mathrm{d}x\,\hat{\varphi}_i(x). \tag{4.5}$$

Then we obtain

$$\int_\Omega \nabla_h^{\partial_u\Omega}\bar{v}_h\, g\,\mathrm{d}x = \langle D\bar{v}_h, m_h^{\partial_u\Omega}g\rangle + \int_{\partial_u\Omega_h} w\, m_h^{\partial_u\Omega}g\,\nu\,\mathrm{d}\mathcal{H}^1. \tag{4.6}$$

Hence, by repeating the argument of Lemma 3.2 we get

$$\|g - m_h^{\partial_u\Omega}g\|_{L^q(\Omega)} \le ch\,\|\nabla g\|_{L^q(\Omega)}, \tag{4.7}$$

and

$$\|\nabla m_h^{\partial_u\Omega}g\|_{L^q(\Omega)} \le c\,\|\nabla g\|_{L^q(\Omega)}. \tag{4.8}$$

The following theorem is similar to Theorem 3.3.

Theorem 4.1. *Let $\mathcal{T}_h = \{T_j\}_{j=1,\dots,P_h}$. There exist positive constants c_1 and c_2 such that*

$$\begin{aligned}\|D\bar{v}_h\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} &\le \|\nabla_h^{\partial_u\Omega}\bar{v}_h\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} \\ &\quad + c_1\,\|w\|_{W^{1,p}(\Omega)} + c_2\left(h\int_{\cup_j\partial T_j} |\,[\![\bar{v}_h]\!]\,|^p\,\mathrm{d}\mathcal{H}^1\right)^{1/p}.\end{aligned} \tag{4.9}$$

Proof. We prove inequality (4.9) by taking tests function $g \in C^1_{\partial_t\Omega}(\Omega;\mathbb{R}^2)$, first, and then extending it to all $W^{1,q}_{\partial_t\Omega}(\Omega;\mathbb{R}^2)$ by continuity.

Let $g \in C^1_{\partial_t\Omega}(\Omega;\mathbb{R}^2)$, with $\|g\|_{W^{1,q}(\Omega)} \le 1$. Then using equation (4.6) we find

$$\begin{aligned}\langle D\bar{v}_h, g\rangle &= \langle D\bar{v}_h, m_h^{\partial_u\Omega}g\rangle + \langle D\bar{v}_h, g - m_h^{\partial_u\Omega}g\rangle \\ &= \int_\Omega \nabla_h^{\partial_u\Omega}\bar{v}_h\cdot g\,\mathrm{d}x - \int_{\partial_u\Omega_h} w\, m_h^{\partial_u\Omega}g\cdot\nu\,\mathrm{d}\mathcal{H}^1 + \langle D\bar{v}_h, g - m_h^{\partial_u\Omega}g\rangle\end{aligned}$$

and hence

$$\begin{aligned}\|D\bar{v}_h\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} \le \|\nabla_h^{\partial_u\Omega}\bar{v}_h\|_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} &+ \sup_{g\in W^{1,q}_{\partial_t\Omega}(\Omega)} \frac{|\int_{\partial_u\Omega_h} w\, m_h^{\partial_u\Omega}g\cdot\nu\,\mathrm{d}\mathcal{H}^1|}{\|g\|_{W^{1,q}(\Omega)}} \\ &+ \sup_{g\in W^{1,q}_{\partial_t\Omega}(\Omega)} \frac{|\langle D\bar{v}_h, g - m_h^{\partial_u\Omega}g\rangle|}{\|g\|_{W^{1,q}(\Omega)}}.\end{aligned}$$

The last term can be bounded from above exactly as is done in Theorem 3.3, but including now the sides on $\partial_u\Omega_h$. It remains to bound only the second term. Since $m_h^{\partial_u\Omega}g$ is equal to zero on $\partial\Omega_h \setminus \partial_u\Omega_h$ we have

$$\int_{\partial_u\Omega_h} w\, m_h^{\partial_u\Omega} g\cdot\nu\,\mathrm{d}\mathcal{H}^1 = \int_{\partial\Omega_h} w\, m_h^{\partial_u\Omega} g\cdot\nu\,\mathrm{d}\mathcal{H}^1 = \int_{\Omega_h} \operatorname{div}\,(w\, m_h^{\partial_u\Omega} g)\,\mathrm{d}x$$

and hence

$$\left|\int_{\partial_u\Omega_h} w\, m_h^{\partial_u\Omega} g\cdot\nu\,\mathrm{d}\mathcal{H}^1\right| \le c\,\|w\|_{W^{1,p}(\Omega)}\|m_h^{\partial_u\Omega}g\|_{W^{1,q}(\Omega)} \le c\,\|w\|_{W^{1,p}(\Omega)}\|g\|_{W^{1,q}(\Omega)}.$$

The proof of the lemma follows easily by combining the above inequalities. □

For any function $g\in C(\bar\Omega)$, from equations (4.2) and (4.3) we find, arguing as in the previous section

$$\int_{\Omega_h} \nabla_h^{\partial_u\Omega}\bar v_h\, c_h g\,\mathrm{d}x = -\int_{\Omega_h} \bar v_h\,\nabla r_h^{\partial_u\Omega} g\,\mathrm{d}x + \int_{\partial_u\Omega_h} w\, r_h^{\partial_u\Omega} g\,\nu\,\mathrm{d}\mathcal{H}^1, \tag{4.10}$$

where

$$r_h^{\partial_u\Omega}g(x) := \sum_{i\in\mathcal{I}_h\cup\mathcal{B}_{h\partial_u}} g(x_i)\hat\varphi_i(x).$$

In what follows the two function spaces will be useful

$$\begin{aligned} W^{1,p}_{\partial_u\Omega}(\Omega) &:= \{v\in W^{1,p}(\Omega) : v=0 \text{ on } \partial_u\Omega\},\\ C^\infty_{\partial_t,0}(\Omega) &:= \{v\in C^\infty(\bar\Omega) : v=0 \text{ in a neighborhood of } \partial_t\Omega\}, \end{aligned}$$

where $\bar\Omega$ denotes the closure of Ω. Note that if $g\in C^\infty_{\partial_t,0}(\Omega)$ then for all sufficiently small h we have $\|g - r_h^{\partial_u\Omega}g\|_{W^{1,\infty}(\Omega_h)} \le c(g)h$.

The following theorem, which is similar to Theorem 3.1, shows that by taking into account the generalized gradient up to the boundary it is possible to recover the desired boundary condition.

Theorem 4.2. *Let $\bar v_h\in Y_h$ for every $h\in\mathcal{H}$. Assume that*

$$\sup_h \|\nabla_h^{\partial_u\Omega}\bar v_h\|_{L^p(\Omega)} < +\infty, \quad \textit{and} \quad \bar v_h \rightharpoonup v \ \textit{in}\ L^p(\Omega).$$

Then

$$v\in (w + W^{1,p}_{\partial_u\Omega}(\Omega)) \quad \textit{and} \quad \nabla_h^{\partial_u\Omega}\bar v_h \rightharpoonup \nabla v \ \textit{in}\ L^p(\Omega).$$

Proof. Since $\sup_h \|\nabla_h^{\partial_u\Omega}\bar v_h\|_{L^p(\Omega)} < +\infty$ we have that, up to a subsequence, $\nabla_h^{\partial_u\Omega}\bar v_h$ weakly converges in $L^p(\Omega)$ and that $\sup_h \|\mathring{\nabla}_h\bar v_h\|_{L^p(\Omega)} < +\infty$. Let $g\in C_0^\infty(\Omega)$, then, taking into account Lemma 3.1, we have

$$\lim_{h\to 0}\int_\Omega \nabla_h^{\partial_u\Omega}\bar v_h\, g\,\mathrm{d}x = \lim_{h\to 0}\int_\Omega \mathring{\nabla}_h\bar v_h\, g\,\mathrm{d}x = \int_\Omega \nabla v\, g\,\mathrm{d}x;$$

hence $\nabla_h^{\partial_u\Omega}\bar v_h \rightharpoonup \nabla v$ in $L^p(\Omega)$.

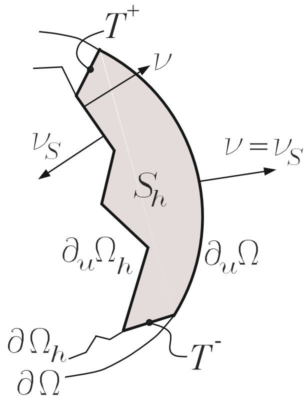

FIGURE 3. Schematic to illustrate the notation used.

Let $g \in C^\infty_{\partial_t,0}(\Omega)$. Taking the limit, as h goes to zero, in equation (4.10) we obtain

$$\int_\Omega \nabla v\, g\, \mathrm{d}x = -\int_\Omega v\, \nabla g\, \mathrm{d}x + \lim_{h\to 0} \int_{\partial_u\Omega_h} w\, r_h^{\partial_u\Omega} g\, \nu\, \mathrm{d}\mathcal{H}^1.$$

In Lemma 4.3, below, we show that

$$\lim_{h\to 0} \int_{\partial_u\Omega_h} w\, r_h^{\partial_u\Omega} g\, \nu\, \mathrm{d}\mathcal{H}^1 = \int_{\partial_u\Omega} w\, g\, \nu\, \mathrm{d}\mathcal{H}^1,$$

hence

$$\int_\Omega \nabla v\, g\, \mathrm{d}x = -\int_\Omega v\, \nabla g\, \mathrm{d}x + \int_{\partial_u\Omega} w\, g\, \nu\, \mathrm{d}\mathcal{H}^1.$$

Integrating by parts and simplifying we deduce

$$\int_{\partial_u\Omega} (v-w)\, g\, \nu\, \mathrm{d}\mathcal{H}^1 = 0,$$

which holds for every $g \in C^\infty_{\partial_t,0}(\Omega)$. Thus $v = w$ on $\partial_u\Omega$. □

Lemma 4.3. *Let $g \in C^\infty_{\partial_t,0}(\Omega)$ and $w \in W^{1,p}(\Omega)$. Let ν denote the outward unit normal to Ω_h and to Ω. Then,*

$$\lim_{h\to 0} \int_{\partial_u\Omega_h} w\, r_h^{\partial_u\Omega} g\, \nu\, \mathrm{d}\mathcal{H}^1 = \int_{\partial_u\Omega} w\, g\, \nu\, \mathrm{d}\mathcal{H}^1.$$

Proof. Let T^+ and T^- be the segments obtained by joining the end points of $\partial_u\Omega$ and $\partial_u\Omega_h$, see Figure 3. Let S_h be the region bounded by T^+, T^-, $\partial_u\Omega_h$ and $\partial_u\Omega$, and let ν_S be the outward normal to S_h. By the Gauss-Green theorem we have

$$\int_{\partial S_h} w\, g\, \nu_S\, \mathrm{d}\mathcal{H}^1 = \int_{S_h} \nabla(w\, g)\, \mathrm{d}x.$$

Equation (4.1) implies that $|S_h|$ approaches 0 as h goes to zero, thus

$$\lim_{h\to 0} \int_{\partial S_h} w\, g\, \nu_S\, \mathrm{d}\mathcal{H}^1 = \lim_{h\to 0} \int_{S_h} \nabla(w\, g)\, \mathrm{d}x = 0.$$

Equation (4.1) also implies that $\lim_{h\to 0}\mathcal{H}^1(T^+)=\lim_{h\to 0}\mathcal{H}^1(T^-)=0$, hence

$$0=\lim_{h\to 0}\int_{\partial S_h} w\, g\,\nu_S\,\mathrm{d}\mathcal{H}^1=\lim_{h\to 0}\int_{\partial_u\Omega_h\cup\partial_u\Omega} w\, g\,\nu_S\,\mathrm{d}\mathcal{H}^1.$$

Taking into account the relation between ν_S and ν, see Figure 3, the previous equation can be rewritten as

$$\lim_{h\to 0}\left(\int_{\partial_u\Omega} w\,g\,\nu\,\mathrm{d}\mathcal{H}^1-\int_{\partial_u\Omega_h} w\,g\,\nu\,\mathrm{d}\mathcal{H}^1\right)=0.$$

To conclude the proof it suffices to note that

$$\lim_{h\to 0}\int_{\partial_u\Omega_h} w\,g\,\nu\,\mathrm{d}\mathcal{H}^1=\lim_{h\to 0}\int_{\partial_u\Omega_h} w\,r_h^{\partial_u\Omega} g\,\nu\,\mathrm{d}\mathcal{H}^1. \qquad \square$$

We conclude the section by proving a density result.

Theorem 4.4. *For every $v\in w+W^{1,p}_{\partial_u\Omega}(\Omega)$ there is a sequence $\{\bar v_h\}$, with $v_h\in Y_h$, such that*

$$\begin{gathered}\bar v_h\to v \qquad \text{in } L^p(\Omega),\\ \nabla_h^{\partial_u\Omega}\bar v_h\to\nabla v \qquad \text{in } L^p(\Omega),\\ \int_{\cup_j\partial T_j\cap\Omega_h}|\,[\![\bar v_h]\!]\,|^p\,\mathrm{d}\mathcal{H}^1+\int_{\partial_u\Omega_h}|w-\bar v_h|^p\,\mathrm{d}\mathcal{H}^1+\frac{1}{|\ln h|}\int_{\partial_t\Omega_h}|\bar v_h|^p\,\mathrm{d}\mathcal{H}^1\to 0,\end{gathered}$$

where $\partial_t\Omega_h:=\partial\Omega_h\setminus\partial_u\Omega_h$.

Proof. Let us suppose, for the moment, that $v,w\in W^{1,p}(\Omega)\cap C^\infty(\bar\Omega)$ and that $v=w$ in a neighborhood of $\partial_u\Omega$. Let

$$\hat v_h(x):=r_h v(x).$$

Since $\hat v_h$ is affine on every triangle T_j it can be written as

$$\hat v_h(x)=\sum_{j=1}^{P_h}\big(\hat v_h(x_{G_j})+\nabla\hat v_h(x)\cdot(x-x_{G_j})\big)\,\chi_{T_j}(x),$$

where x_{G_j} is the center of mass of T_j, *i.e.*,

$$x_{G_j}:=\fint_{T_j} x\,\mathrm{d}x.$$

We define

$$\bar v_h(x)=\sum_{j=1}^{P_h}\hat v_h(x_{G_j})\chi_{T_j}(x).$$

We obviously have

$$\lim_{h\to 0}\|\bar v_h-v\|_{L^p(\Omega)}=\lim_{h\to 0}\|\bar v_h-\hat v_h\|_{L^p(\Omega)}=0.$$

Let us compute the generalized gradient of $\bar v_h$. Let $i\in\mathcal{I}_h\cup\mathcal{B}_{h\partial_u}$. Since $\nabla\hat\varphi_i$ is constant on T_j, we have

$$\int_{\Omega_h}\bar v_h\nabla\hat\varphi_i\,\mathrm{d}x=\sum_{j=1}^{P_h}\int_{T_j}\hat v_h(x_{G_j})\nabla\hat\varphi_i\,\mathrm{d}x=\sum_{j=1}^{P_h}\int_{T_j}\hat v_h\nabla\hat\varphi_i\,\mathrm{d}x=\int_{\Omega_h}\hat v_h\nabla\hat\varphi_i\,\mathrm{d}x.$$

Thus

$$\begin{aligned}\nabla_h^{\partial_u\Omega}\bar{v}_h(x_i) &= \frac{-1}{|\hat{T}_i|}\left(\int_{\Omega_h}\hat{v}_h\nabla\hat{\varphi}_i\,\mathrm{d}x - \int_{\partial_u\Omega_h} w\hat{\varphi}_i\nu\,\mathrm{d}\mathcal{H}^1\right)\\ &= \frac{1}{|\hat{T}_i|}\left(\int_{\Omega_h}\nabla\hat{v}_h\,\hat{\varphi}_i\,\mathrm{d}x + \int_{\partial_u\Omega_h}(w-\hat{v}_h)\hat{\varphi}_i\nu\,\mathrm{d}\mathcal{H}^1\right)\\ &= \frac{1}{|\hat{T}_i|}\left(\int_{\Omega_h}\nabla v\,\hat{\varphi}_i\,\mathrm{d}x + \int_{\Omega_h}\nabla(\hat{v}_h-v)\hat{\varphi}_i\,\mathrm{d}x + \int_{\partial_u\Omega_h}(w-\hat{v}_h)\hat{\varphi}_i\nu\,\mathrm{d}\mathcal{H}^1\right),\end{aligned}$$

and hence

$$\nabla_h^{\partial_u\Omega}\bar{v}_h = R_h(\nabla v) + A_h + B_h, \tag{4.11}$$

where we have set

$$\begin{aligned}R_h(\nabla v) &:= \sum_{i\in\mathcal{I}_h\cup\mathcal{B}_{h\partial_u}}\frac{1}{|\hat{T}_i|}\int_{\Omega_h}\nabla v\,\hat{\varphi}_i\,\mathrm{d}x\,\chi_{\hat{T}_i},\\ A_h &:= \sum_{i\in\mathcal{I}_h\cup\mathcal{B}_{h\partial_u}}\frac{1}{|\hat{T}_i|}\int_{\Omega_h}\nabla(\hat{v}_h-v)\hat{\varphi}_i\,\mathrm{d}x\,\chi_{\hat{T}_i},\\ B_h &:= \sum_{i\in\mathcal{B}_{h\partial_u}}\frac{1}{|\hat{T}_i|}\int_{\partial_u\Omega_h}(w-\hat{v}_h)\hat{\varphi}_i\nu\,\mathrm{d}\mathcal{H}^1\,\chi_{\hat{T}_i}.\end{aligned}$$

Note that $A_h \to 0$ in $L^p(\Omega)$, and since $v = w$ in a neighborhood of $\partial_u\Omega$ and $|v(x) - \hat{v}_h(x)| \le c(v)h^2$ we also have that $B_h \to 0$ in $L^p(\Omega)$. Under the assumption that $|\hat{T}_i| = \int_{\Omega_h}\hat{\varphi}_i\,\mathrm{d}x$, R_h is a mapping which preserves the constant functions, and, by a classical result in approximation theory, *cf.* Ciarlet [13], Theorem 15.3, we have that $R_h(\nabla v) \to \nabla v$ in $L^p(\Omega)$. Thus, from equation (4.11), we deduce

$$\lim_{h\to 0}\|\nabla_h^{\partial_u\Omega}\bar{v}_h - \nabla v\|_{L^p(\Omega)} = 0.$$

As the distance between the centers of mass is bounded by h, we have that $|\,[\![\bar{v}_h]\!]\,| \le c(v)h$ and hence

$$\int_{\cup_j\partial T_j\cap\Omega_h}|\,[\![\bar{v}_h]\!]\,|^p\,\mathrm{d}\mathcal{H}^1 \le c\frac{h^{p+1}}{h^2} \to 0,$$

while at $\partial_u\Omega_h$, where $|w - \bar{v}_h| \le c(v)\,h$, it is

$$\int_{\partial_u\Omega_h}|w-\bar{v}_h|^p\,\mathrm{d}\mathcal{H}^1 \le c\,h^p \to 0.$$

Finally, at $\partial_t\Omega_h$ we have

$$\frac{1}{|\ln h|}\int_{\partial_t\Omega_h}|\bar{v}_h|^p\,\mathrm{d}\mathcal{H}^1 \le c\,\frac{1}{|\ln h|} \to 0,$$

and this concludes the proof for the case of smooth functions v and w.

Let us now consider the general case. Let $z^k, w^k \in W^{1,p}(\Omega)\cap C^\infty(\Omega)$ such that $z^k \to v - w$ in $W^{1,p}(\Omega)$, $w^k \to w$ in $W^{1,p}(\Omega)$ and $z^k = 0$ in a neighborhood of $\partial_u\Omega$. Define $v^k := z^k + w^k$. Applying the argument above for each k and using a diagonalization procedure we conclude the proof. Indeed let

$$G_h^k(\bar{v}) := \int_{\cup_j\partial T_j\cap\Omega_h}|\,[\![\bar{v}]\!]\,|^p\,\mathrm{d}\mathcal{H}^1 + \int_{\partial_u\Omega_h}|w^k-\bar{v}|^p\,\mathrm{d}\mathcal{H}^1 + \frac{1}{|\ln h|}\int_{\partial_t\Omega_h}|\bar{v}|^p\,\mathrm{d}\mathcal{H}^1,$$

then, for each k there exists a H_k and $\{\bar{v}_h^k\} \in \cup_{h\in\mathcal{H}} Y_h$ such that for each $h \le H_k$ we have

$$\|\bar{v}_h^k - v^k\|_{L^p(\Omega)} + \|\nabla_h^{\partial_u\Omega}\bar{v}_h^k - \nabla v^k\|_{L^p(\Omega)} + G_h^k(\bar{v}_h^k) < 1/k.$$

Without loss of generality we may assume that $H_k > H_{k+1}$ for all values of k. To conclude the proof it suffices to set $\bar{v}_h = \bar{v}_h^k$ for each $H_k \ge h > H_{k+1}$. □

5. External approximations of convex functionals

We consider here functionals of the form

$$\mathcal{F}(v) := \int_\Omega W(x, v, \nabla v)\,\mathrm{d}x - \langle f, v\rangle \tag{5.1}$$

to be minimized in $w + W^{1,p}_{\partial_u\Omega}(\Omega)$, where $f \in L^q(\Omega)$. We assume that $W : \Omega \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ is a Carathéodory function (measurable in the first variable and continuous in the last two) satisfying the following requirements:

(H1) for almost every $x \in \Omega$ and every $s \in \mathbb{R}$ the function $W(x, s, \cdot)$ is convex on $\mathbb{R}^2$;

(H2) there exist four constants $c_1, c_2, b_1, b_2 > 0$ such that

$$-c_1 + b_1|\xi|^p \le W(x, s, \xi) \le c_2 + b_2(|s|^p + |\xi|^p)$$

for *a.e.* $x \in \Omega$ and for every $(s, \xi) \in \mathbb{R} \times \mathbb{R}^2$.

The goal of this and the next section is to study the approximation of this type of functionals in the spaces Y_h by using the generalized notion of gradient introduced above. Namely, we adopt the framework of Γ-convergence theory [17] in order to prove that a suitable sequence of functionals $\{\mathcal{F}_h\}$ defined in the spaces Y_h Γ-converges to the functional $\mathcal{F}$ in an appropriate topology. Then, according to the central property of Γ-convergence this implies, under suitable conditions, that:

$$\mathcal{F}(u^{\min}) = \min \mathcal{F}(v) = \lim_h \min \mathcal{F}_h(v) = \lim_h \mathcal{F}_h(\bar{u}_h^{\min})$$

and

$$\bar{u}_h^{\min} \to u^{\min},$$

with $\bar{u}_h^{\min}$ and $u^{\min}$ being the respective minimizers, see [17], Theorems 7.8 and 7.24. So, in particular, the $\bar{u}_h^{\min}$ provide an approximation of $u^{\min}$.

We extend $\mathcal{F}$ to $L^p(\Omega)$ by defining it equal to $+\infty$ in $L^p(\Omega) \setminus (w + W^{1,p}_{\partial_u\Omega,}(\Omega))$ and introduce the sequence of "discrete" functionals

$$\mathcal{F}_h(\bar{v}) := \begin{cases} \displaystyle\int_{\Omega_h} W(x, \bar{v}, \nabla_h^{\partial_u\Omega}\bar{v})\,\mathrm{d}x + \mathcal{J}_h(\bar{v}) - \langle f, \bar{v}\rangle & \text{if } \bar{v} \in Y_h, \\ +\infty & \text{if } \bar{v} \in L^p(\Omega) \setminus Y_h, \end{cases} \tag{5.2}$$

where

$$\mathcal{J}_h(\bar{v}) := \int_{\cup_j \partial T_j \cap \Omega_h} |[\![\,\bar{v}\,]\!]|^p\,\mathrm{d}\mathcal{H}^1 + \int_{\partial_u\Omega_h} |w - \bar{v}|^p\,\mathrm{d}\mathcal{H}^1 + \frac{1}{|\ln h|}\int_{\partial_t\Omega_h} |\bar{v}_h|^p\,\mathrm{d}\mathcal{H}^1. \tag{5.3}$$

In definition (5.3) the first integral on the right hand side is extended to the inner sides of the mesh $\mathcal{T}_h$.

Theorem 5.1. *The functional $\mathcal{F}_h$ sequentially Γ-converges to $\mathcal{F}$, with respect to the weak-$L^p(\Omega)$ topology, that is*

(1) [Lim inf inequality] *for every sequence of positive numbers h converging to 0 and for every sequence $\{\bar{v}_h\} \subset L^p(\Omega)$ and $v \in L^p(\Omega)$ such that $\bar{v}_h \rightharpoonup v$ in $L^p(\Omega)$,*

$$\liminf_{h\to 0} \mathcal{F}_h(\bar{v}_h) \ge \mathcal{F}(v);$$

(2) [Recovery sequence] *for every sequence of positive numbers h converging to 0 and for every $v \in L^p(\Omega)$ there exists a corresponding sequence of functions $\{\bar{v}_h\} \subset L^p(\Omega)$ such that $\bar{v}_h \rightharpoonup v$ in $L^p(\Omega)$, and*

$$\limsup_{h\to 0} \mathcal{F}_h(\bar{v}_h) \le \mathcal{F}(v).$$

Proof. We first prove the lim inf inequality. Let $\bar{v}_h$ be a sequence weakly converging in $L^p(\Omega)$ to v. If $\liminf_h \mathcal{F}_h(\bar{v}_h) = +\infty$ there is nothing to prove. Hence, by passing to a subsequence, if necessary, we may suppose $\liminf_h \mathcal{F}_h(\bar{v}_h) = \lim_h \mathcal{F}_h(\bar{v}_h) < +\infty$. Then, using (H2) and equation (5.3) we find

$$\mathcal{F}_h(\bar{v}_h) \ge -c_1|\Omega| + b_1\|\nabla_h^{\partial_u\Omega}\bar{v}_h\|^p_{L^p(\Omega)} + \mathcal{J}_h(\bar{v}_h) - \|f\|_{L^q(\Omega)}\|\bar{v}_h\|_{L^p(\Omega)}.$$

Since $\|\bar{v}_h\|_{L^p(\Omega)}$ is bounded, we have that

$$\sup_h \|\mathring{\nabla}_h\bar{v}_h\|_{L^p(\Omega)} \le \sup_h \|\nabla_h^{\partial_u\Omega}\bar{v}_h\|_{L^p(\Omega)} < +\infty$$

and, also,

$$\sup_h \mathcal{J}_h(\bar{v}_h) < \infty.$$

Observing that $|[\![\,\bar{v}_h\,]\!]| = |\bar{v}_h|$ at the boundary and using the trace theorem, for some positive constant c we have that

$$\mathcal{J}_h(\bar{v}_h) \ge c\left(\frac{1}{|\ln h|}\int_{\cup_j\partial T_j} |[\![\,\bar{v}_h\,]\!]|^p \,\mathrm{d}\mathcal{H}^1 - \|w\|^p_{W^{1,p}(\Omega)}\right). \tag{5.4}$$

From the last two inequalities it follows then

$$h\int_{\cup_j\partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1 \to 0.$$

By Theorem 4.2 we deduce that

$$v \in w + W^{1,p}_{\partial_u\Omega}(\Omega) \quad \text{and} \quad \nabla_h^{\partial_u\Omega}\bar{v}_h \rightharpoonup \nabla v \text{ in } L^p(\Omega),$$

while by Theorem 3.5 we have that

$$\bar{v}_h \to v, \text{ in } L^p(\Omega).$$

By Theorem 5.2 below and, without loss of generality, supposing that W is non-negative[2], we find

$$\begin{aligned}\liminf_h \mathcal{F}_h(\bar{v}_h) &\ge \liminf_h \int_\Omega W(x,\bar{v}_h,\nabla_h^{\partial_u\Omega}\bar{v}_h)\,\mathrm{d}x - \lim_h\langle f,\bar{v}_h\rangle \\ &\ge \int_\Omega W(x,v,\nabla v)\,\mathrm{d}x - \langle f, v\rangle.\end{aligned}$$

We now prove the recovery sequence condition. If v does not belong to $L^p(\Omega)\setminus(w + W^{1,p}_{\partial_u\Omega,}(\Omega))$ there is nothing to prove. So, let $v \in (w + W^{1,p}_{\partial_u\Omega,}(\Omega))$. By Theorem 4.4 there is a sequence $\{\bar{v}_h\} \in \cup_{h\in\mathcal{H}}Y_h$ such that

$$\begin{aligned}\bar{v}_h &\to v && \text{in } L^p(\Omega),\\ \nabla_h^{\partial_u\Omega}\bar{v}_h &\to \nabla v && \text{in } L^p(\Omega),\\ \mathcal{J}_h(\bar{v}_h) &\to 0.\end{aligned}$$

[2]If not, we can always apply the argument to the integrand $W + a$, with a a large enough constant.

Thus, thanks to the growth from above of W, see hypothesis (H2), and the dominated convergence theorem the proof is completed. □

We finally state the theorem on lower-semicontinuity used in the proof above.

Theorem 5.2 (see Dacorogna [16]). *Let Ω be a bounded open set of $\mathbb{R}^n$. Let $g : \Omega \times \mathbb{R}^m \times \mathbb{R}^N \to [0, +\infty]$ be a Carathéodory function. Let*

$$\mathcal{G}(\psi, \xi) := \int_\Omega g(x, \psi(x), \xi(x))\, \mathrm{d}x.$$

Assume that $g(x, \psi, \cdot)$ is convex that $\psi_k \to \psi$ in $L^p(\Omega)$ and that $\xi_k \rightharpoonup \xi$ in $L^p(\Omega)$ then

$$\liminf_k \mathcal{G}(\psi_k, \xi_k) \geq \mathcal{G}(\psi, \xi).$$

We conclude the section by observing that if W is continuous and not just measurable in x it is possible to localize the functional also in the x variable. Indeed one can define

$$\mathcal{F}_h(\bar{v}) := \begin{cases} \displaystyle\int_{\Omega_h} W(c_h x, \bar{v}, \nabla_h^{\partial_u \Omega} \bar{v})\, \mathrm{d}x + \mathcal{J}_h(\bar{v}) - \langle f, \bar{v} \rangle & \text{if } \bar{v} \in Y_h, \\ +\infty & \text{if } \bar{v} \in L^p(\Omega) \setminus Y_h, \end{cases} \tag{5.5}$$

where

$$c_h x(x) := \sum_{j=1}^{N_h} x_j \chi_{\hat{T}_j}(x).$$

The proof of Theorem 5.1 follows as before after observing that $c_h x \to id(x) := x$ in $L^\infty(\Omega)$, and that $\psi_h := (c_h x, \bar{v}_h) \to \psi := (id, v)$ in $L^p(\Omega; \mathbb{R}^3)$, so that Theorem 5.2 still applies.

6. Convergence of minima and minimizers

In this last section we prove the convergence of the minima and minimizers of the discretized functionals. We start by showing that the functionals $\mathcal{F}_h$ are equicoercive.

Theorem 6.1. *There exist two constants $k_1, k_2 > 0$, independent of h, such that*

$$\mathcal{F}_h(v) \geq -k_1 + k_2 \|v\|_{L^p(\Omega)},$$

for all $v \in L^p(\Omega)$.

Proof. From definition (5.2) of $\mathcal{F}_h$ we have that the theorem trivially holds for $v \in L^p(\Omega) \setminus Y_h$. So let $v = \bar{v} \in Y_h$. After noticing that for any $g \in L^p(\Omega)$ we have

$$\|g\|_{L^p(\Omega)} = \sup_{\psi \in L^q(\Omega)} \frac{\int_\Omega \psi g \,\mathrm{d}x}{\|\psi\|_{L^q(\Omega)}} \geq \sup_{\psi \in W^{1,q}_{\partial_t \Omega}(\Omega)} \frac{\int_\Omega \psi g \,\mathrm{d}x}{\|\psi\|_{W^{1,q}(\Omega)}} = \|g\|_{(W^{1,q}_{\partial_t \Omega}(\Omega))'}, \tag{6.1}$$

we deduce, from the definition of $\mathcal{F}_h$, property (H2), inequality (5.4) and Theorem 4.1 that

$$\begin{aligned}\mathcal{F}_h(\bar v) \quad \geq \quad & -c_1|\Omega| + b_1\|\nabla_h^{\partial_u\Omega}\bar v\|^p_{L^p(\Omega)} + c\frac{1}{|\ln h|}\int_{\cup_j \partial T_j} |[\![\,\bar v\,]\!]|^p \,\mathrm{d}\mathcal{H}^1 \\ & - c\,\|w\|^p_{W^{1,p}(\Omega)} - \|f\|_{L^q(\Omega)}\|\bar v\|_{L^p(\Omega)}, \\ \geq \quad & -c_1|\Omega| + b_1\|\nabla_h^{\partial_u\Omega}\bar v\|^p_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} + c\,h\int_{\cup_j \partial T_j} |\,[\![\bar v]\!]\,|^p \,\mathrm{d}\mathcal{H}^1 \\ & - c\,\|w\|^p_{W^{1,p}(\Omega)} - \|f\|_{L^q(\Omega)}\|\bar v\|_{L^p(\Omega)}, \\ \geq \quad & -c_1|\Omega| + c\,\|Dv\|^p_{(W^{1,q}_{\partial_t\Omega}(\Omega))'} - c\,\|w\|^p_{W^{1,p}(\Omega)} - \|f\|_{L^q(\Omega)}\|\bar v\|_{L^p(\Omega)}.\end{aligned} \tag{6.2}$$

Now applying Lemma 1.1 we find

$$\mathcal{F}_h(\bar v) \quad \geq \quad -c_1|\Omega| + \tilde c\,\|\bar v\|^p_{L^p(\Omega)} - c\,\|w\|^p_{W^{1,p}(\Omega)} - c\,\|f\|^q_{L^q(\Omega)} - \frac{\tilde c}{2}\,\|\bar v\|^p_{L^p(\Omega)},$$

and hence

$$\mathcal{F}_h(\bar v) \quad \geq \quad -k_1 + k_2\,\|\bar v\|^p_{L^p(\Omega)},$$

where the constant k_1 depends on f, w but not on h. □

Let us denote by $\bar u_h^{\min}$ a minimizer of $\mathcal{F}_h$, that is

$$\mathcal{F}_h(\bar u_h^{\min}) := \min_{\bar v\in Y_h} \mathcal{F}_h(\bar v).$$

By Theorem 6.1 and the inequality $\mathcal{F}_h(\bar u_h^{\min}) \leq \mathcal{F}_h(0)$ it follows that $\sup_h \|\bar u_h^{\min}\|_{L^p(\Omega)}$ is finite and therefore, up to subsequences, we have

$$\bar u_h^{\min} \rightharpoonup u^{\min} \quad \text{in } L^p(\Omega),$$

for some $u^{\min} \in L^p(\Omega)$. By the growth from below of W and the definition of $\mathcal{F}_h$ we deduce that

$$\sup_h \|\nabla_h^{\partial_u\Omega}\bar u_h^{\min}\|^p_{L^p(\Omega)} < +\infty.$$

Hence, from Theorem 4.2

$$u^{\min} \in w + W^{1,p}_{\partial_u\Omega}(\Omega) \quad \text{and} \quad \nabla_h^{\partial_u\Omega}\bar u_h^{\min} \rightharpoonup \nabla u^{\min} \text{ in } L^p(\Omega).$$

Let v be any function in $w + W^{1,p}_{\partial_u\Omega}(\Omega)$. Applying Theorem 5.1 we can find a recovery sequence $\{\bar v_h\} \subset L^p(\Omega)$ of v, *i.e.*, $\limsup_{h\to 0}\mathcal{F}_h(\bar v_h) \leq \mathcal{F}(v)$, and we have

$$\mathcal{F}(u^{\min}) \quad \leq \quad \liminf_{h\to 0}\mathcal{F}_h(\bar u_h^{\min}) \leq \limsup_{h\to 0}\mathcal{F}_h(\bar u_h^{\min}) \leq \limsup_{h\to 0}\mathcal{F}_h(\bar v_h) \leq \mathcal{F}(v).$$

This implies that

$$\mathcal{F}(u^{\min}) = \min_{v\in w+W^{1,p}_{\partial_u\Omega}(\Omega)} \mathcal{F}(v)$$

and also that

$$\lim_h \mathcal{F}_h(\bar u_h^{\min}) = \mathcal{F}(u^{\min})$$

when we take $v = u^{\min}$.

The last implication yields in particular that $h\int_{\cup_j \partial T_j \setminus \partial\Omega_h} |\,[\![\bar{u}_h^{\min}]\!]\,|^p \,\mathrm{d}\mathcal{H}^1 \to 0$. Thus, by Theorem 3.5 it follows that

$$\bar{u}_h^{\min} \to u^{\min} \text{ in } L^p(\Omega).$$

Theorem 6.2. *Let $\bar{u}_h^{min}$ be a minimizer of $\mathcal{F}_h$ over $L^p(\Omega)$. Then:*

(1) *$\{\bar{u}_h^{min}\}$ has at least a weakly convergent subsequence in $L^p(\Omega)$;*
(2) *if $u^{min} \in L^p(\Omega)$ is a weak limit of $\{\bar{u}_h^{min}\}$ (or a subsequence), then u^{min} is a minimizer of $\mathcal{F}$ over $L^p(\Omega)$ and*

$$\bar{u}_h^{min} \to u^{min} \quad \text{in } L^p(\Omega),$$

$$u^{min} \in (w + W^{1,p}_{\partial_u\Omega}(\Omega)) \quad \text{and} \quad \nabla_h^{\partial_u\Omega} \bar{u}_h^{min} \rightharpoonup \nabla u^{min} \text{ in } L^p(\Omega).$$

Furthermore,

$$\lim_h \mathcal{F}_h(\bar{u}_h^{min}) = \mathcal{F}(u^{min}).$$

The last part follows immediately from Theorem 5.1.

If the integrand W is strictly convex in the last variable, the convergence of the generalized gradient is indeed strong, *i.e.*,

$$\nabla_h^{\partial_u\Omega} \bar{u}_h^{\min} \to \nabla u^{\min} \text{ in } L^p(\Omega).$$

This can be proved, for instance, like Theorem 7 of [20]. We also note that if the minimizer of $\mathcal{F}$ is unique then the full sequence $\{\bar{u}_h^{\min}\}$ is convergent.

We conclude by observing that Theorems 3.3 and 4.1 might be written in a slightly different but equivalent form. For instance, we could write Theorem 3.3 as:

For every $\eta > 0$, there is a $\gamma > 0$ such that

$$\gamma\, \|D\bar{v}_h\|_{W^{-1,p}(\Omega)} \le \|\mathring{\nabla}_h \bar{v}_h\|_{W^{-1,p}(\Omega)} + \eta \left(h \int_{\cup_j \partial T_j} |\,[\![\bar{v}_h]\!]\,|^p \,\mathrm{d}\mathcal{H}^1 \right)^{1/p}, \tag{6.3}$$

arriving at the same conclusions seen above. Accordingly, we could take the functional $\eta\mathcal{I}_h$, instead of $\mathcal{I}_h$, in the definition of $\mathcal{F}_h$ and still get the results proved in Sections 5 and 6. While nothing changes from a theoretical point of view, for the efficiency of the method the free parameter η in the approximating functionals may be conveniently used, at least in some cases, to optimize the convergence of the numerical results, see for instance [19].

Appendix

In this appendix we prove inequality (3.14), that is: there exists a constant c independent of h and T_j such that

$$\int_{\partial T_j} |f|^q \,\mathrm{d}\mathcal{H}^1 \le \frac{c}{h} \int_{T_j} |f|^q \,\mathrm{d}x + c\, h^{q-1} \int_{T_j} |\nabla f|^q \,\mathrm{d}x, \tag{6.4}$$

for every $f \in W^{1,q}(T_j)$.

Consider any $T_j \in \mathcal{T}_h$ and let $p_h(y) := hy$ be the rescaling that maps the normalized triangle $\tilde{T}_j$ onto $T_j = h\tilde{T}_j$. For every j, it is always possible to define an affine transformation $B(j)$, continuous together with its inverse, that maps $\tilde{T}_j$ onto one and the same triangle T°. Moreover, from the regularity of the family of meshes, there are two constants α and β independent of j and h such that

$$\alpha < |B(j)|, \det B(j) < \beta. \tag{6.5}$$

Using the inverse of the map $p_h \circ B(j)$ to transform the domain T_j to T° and the continuity of the trace we easily get

$$\begin{aligned}\int_{\partial T_j} |f|^q \, \mathrm{d}\mathcal{H}^1 &\leq |B(j)|\, h \int_{\partial T^\circ} |f \circ p_h \circ B(j)|^q \, \mathrm{d}\mathcal{H}^1 \\ &\leq C_0\, h\, |B(j)|\, \|f \circ p_h \circ B(j)\|^q_{W^{1,q}(T^\circ)}.\end{aligned}$$

On the other hand, by [13], Theorem 15.1, we have:

$$|v \circ B(j)|_{W^{m,q}(T^\circ)} \leq C |B(j)|^m \, |\det B(j)|^{-1/q} \, |v|_{W^{m,q}(\tilde{T}_j)} \qquad \forall\, v \in W^{m,q}(\tilde{T}_j),$$

where C depends on m only. The symbol $|\cdot|_{W^{m,q}(\tilde{T}_j)}$ denotes the seminorm of $W^{m,q}(\tilde{T}_j)$. Then, taking (6.5) into account, we deduce

$$\begin{aligned}\int_{\partial T_j} |f|^q \, \mathrm{d}\mathcal{H}^1 &\leq C(\alpha,\beta)\, h \ \|f \circ p_h\|^q_{W^{1,q}(\tilde{T}_j)} \\ &= C(\alpha,\beta)\, h \left(\int_{\tilde{T}_j} |f \circ p_h|^q \, \mathrm{d}y + \int_{\tilde{T}_j} |h (\nabla f) \circ p_h|^q \, \mathrm{d}y \right) \\ &= C(\alpha,\beta)\, h \left(\frac{1}{h^2} \int_{T_j} |f|^q \, \mathrm{d}x + h^{q-2} \int_{T_j} |\nabla f|^q \, \mathrm{d}x \right),\end{aligned}$$

which is (6.4).

Acknowledgements. This work has been supported by the Italian Ministry of Education under the project PRIN 2005: "Modeling and approximation in advanced mechanical problems".

References

[1] B.A. Andreianov, M. Gutnic and P. Wittbold, Convergence of finite volume approximations for a nonlinear elliptic-parabolic problem: a "continuous" approach. *SIAM J. Numer. Anal.* **42** (2004) 228–251.

[2] D.N. Arnold, An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.* **19** (1982) 742–760.

[3] D.N. Arnold, F. Brezzi, B. Cockburn and L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* **39** (2001-2002) 1749–1779.

[4] J.P. Aubin, Approximation des problèmes aux limites non homogènes pour des opérateurs non linéaires. *J. Math. Anal. Appl.* **30** (1970) 510–521.

[5] I. Babuška, The finite element method with penalty. *Math. Comp.* **27** (1973) 221–228.

[6] I. Babuška and M. Zlámal, Nonconforming elements in the finite element method with penalty. *SIAM J. Numer. Anal.* **10** (1973) 863–875.

[7] I. Babuška, C.E. Baumann and J.T. Oden, A discontinuous *hp* finite element method for diffusion problems: 1-D analysis. *Comput. Math. Appl.* **37** (1999) 103–122.

[8] C.E. Baumann and J.T. Oden, *Advances and applications of discontinuous Galerkin methods in CFD*. Computational mechanics (Buenos Aires, 1998), Centro Internac. Métodos Numér. Ing., Barcelona (1998).

[9] C.E. Baumann and J.T. Oden, A discontinuous *hp* finite element method for convection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.* **175** (1999) 311–341.

[10] C.E. Baumann and J.T. Oden, An adaptive-order discontinuous Galerkin method for the solution of the Euler equations of gas dynamics. *Internat. J. Numer. Methods Engrg.* **47** (2000) 61–73.

[11] H. Brezis, *Analyse fonctionnelle: Théorie et applications*. Masson, Paris (1983).

[12] P.G. Ciarlet, *The finite element method for elliptic problems*. North Holland, Amsterdam (1978).

[13] P.G. Ciarlet, *Basic error estimates for elliptic problems*, in *Handbook of numerical analysis*, P.G. Ciarlet and J.-L. Lions Eds., North Holland, Amsterdam (1991).

[14] B. Cockburn and C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.* **35** (1998) 2440–2463.

[15] B. Cockburn, G.E. Karniadakis and C.-W. Shu, The development of discontinuous Galerkin methods, in *Discontinuous Galerkin methods* (Newport, RI, 1999), *Lect. Notes Comput. Sci. Eng.* **11** (2000) 3–50.
[16] B. Dacorogna, *Direct methods in the calculus of variations*. Springer-Verlag, New York (1989).
[17] G. Dal Maso, *An introduction to* Γ*-convergence*. Birkäuser, Boston (1993).
[18] C. Davini, Piece-wise constant approximations in the membrane problem. *Meccanica* **38** (2003) 555–569.
[19] C. Davini and F. Jourdan, Approximations of degree zero in the Poisson problem. *Comm. Pure Appl. Anal.* **4** (2005) 267–281.
[20] C. Davini and R. Paroni, Generalized Hessian and external approximations in variational problems of second order. *J. Elasticity* **70** (2003) 149–174.
[21] C. Davini and R. Paroni, *Error estimate of piece-wise constant approximations to the Poisson problem* (in preparation).
[22] C. Davini and I. Pitacco, Relaxed notions of curvature and a lumped strain method for elastic plates. *SIAM J. Numer. Anal.* **35** (1998) 677–691.
[23] C. Davini and I. Pitacco, An unconstrained mixed method for the biharmonic problem. *SIAM J. Numer. Anal.* **38** (2000) 820–836.
[24] L.C. Evans and R.F. Gariepy, *Measure theory and fine properties of functions, Studies in Advanced Mathematics*. CRC Press, Boca Raton (1992).
[25] J.-L. Lions, Problèmes aux limites non homogènes à données irrégulières : Une méthode d'approximation, in *Numerical Analysis of Partial Differential Equations* (C.I.M.E. 2 Ciclo, Ispra, 1967), Edizioni Cremonese, Rome (1968) 283–292.
[26] J. Nečas, *Équations aux dérivées partielles*. Presses de l'Université de Montréal (1965).
[27] J. Nitsche, Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg* **36** (1971) 9–15.
[28] W.H. Reed and T.R. Hill, *Triangular mesh method for neutron transport equation*. Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos (1973).
[29] M.F. Wheeler, An elliptic collocation-finite element method with interior penalties. *SIAM J. Numer. Anal.* **15** (1978) 152–161.
[30] X. Ye, A new discontinuous finite volume method for elliptic problems. *SIAM J. Numer. Anal.* **42** (2004) 1062–1072.