IMA Journal of Numerical Analysis (2005) Page 1 of 15
doi: 10.1093/imanum/dri017

Non-Conforming Finite Element Discretisation of Convex Variational Problems

Christoph Ortner
Mathematical Institute,
24–29 St Giles', Oxford OX1 3LB, UK.

[Received on]

The Lavrentiev gap phenomenon is a well-known effect in the calculus of variations, related to singularities of minimizers. In its presence, conforming finite element methods are incapable of reaching the energy minimum. By contrast, it is shown in this work that, for convex variational problems, the non-conforming Crouzeix–Raviart finite element discretization always converges to the correct minimizer, and that the discrete energy converges to the correct limit.

Keywords: Calculus of Variations, Lavrentiev Phenomenon, non-conforming FEM.

1. Introduction

The *Lavrentiev gap phenomenon* is a well-known effect in the calculus of variations, related to singularities of minimizers. In its presence, conforming finite element methods are incapable of reaching the energy minimum, and consequently the numerical solutions converge to the wrong limit. The goal of this paper is to demonstrate that a standard non-conforming finite element method is succesful in approximating certain problems within this class.

Possibly the most well-known instance of the Lavrentiev phenomenon is the example discovered by Manià [22]. Suppose we want to minimize the functional

$$\mathcal{J}(u) = \int_0^1 u_x^6 (u^3 - x)^2 \,\mathrm{d}x \tag{1.1}$$

over the space $\mathrm{W}^{1,1}(0,1)$, subject to the constraints $u(0) = 0$ and $u(1) = 1$. The infimum is zero and it is attained for $u(x) = x^{1/3}$. The interesting feature of (1.1) is that the infimum of $\mathcal{J}$ over Lipschitz functions is strictly positive. This effect is commonly known as the *Lavrentiev gap phenomenon*, named after the first known example, in the work of Lavrentiev [19]. Manià's example is readily modified so that $\mathcal{J}$ becomes coercive in $\mathrm{W}^{1,p}(0,1)$ for some $p > 1$. Moreover, it was shown by Ball and Mizel [6] that the effect can even occur for uniformly elliptic integrands. Foss, Hrusa, and Mizel [17] gave a two-dimensional example where the integrand is autonomous and convex. This important class of problems is the focus of the present work.

The interest in the Lavrentiev phenomenon is due to its connection to the regularity of minimizers, as well as its relevance in mathematical models of solid mechanics [3, 4].

Let Ω be a connected bounded polyhedral Lipschitz domain in $\mathbb{R}^n$, $n \geqslant 2$; let $W : \mathbb{R}^{m\times n} \to [0, +\infty]$, $m \geqslant 1$, be a *convex* stored energy density, and let $f \in \mathrm{L}^\infty(\Omega)^m$. For $v \in \mathrm{W}^{1,1}(\Omega)^m$, we define

$$\mathcal{J}(v) = \int_\Omega \Big[W(\nabla v) - f \cdot v \Big] \,\mathrm{d}x. \tag{1.2}$$

Given $g \in \mathrm{W}^{1,\infty}(\Omega)^m$ and relatively open sets $\Gamma_i \subset \partial\Omega$ with $|\Gamma_i| > 0$, $i = 1, \dots, m$, we define the *admissible set*

$$\mathcal{A} = \big\{ u \in \mathrm{W}^{1,1}(\Omega)^m : u^{(i)} = g^{(i)} \text{ on } \Gamma_i, i = 1, \dots, m \big\}; \tag{1.3}$$

here and througout we use superscripts to denote components of a vector-valued function. For future reference we set $\mathcal{A}_p = \mathcal{A} \cap \mathrm{W}^{1,p}(\Omega)^m$, whenever $p \in [1, \infty]$.

The aim of this work is the numerical solution of the minimization problem

$$u \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_1). \tag{1.4}$$

To guarantee the well-posedness of (1.4), we need to add an additional assumption on W. We shall assume throughout that W has superlinear growth; more precisely, we assume that there exists a convex function $\phi : [0,\infty) \to [0,\infty)$ such that

$$W(\xi) \geqslant \phi(|\xi|) \quad \forall \xi \in \mathbb{R}^{m\times n} \quad \text{and} \quad \lim_{s\to\infty} \phi(s)/s = +\infty. \tag{1.5}$$

Assumption (1.5) allows us to extract weakly convergent subsequences from families of Sobolev functions with bounded energy.

Under the above conditions we have the following basic existence result. Its proof is a straightforward application of the direct method of the calculus of variations (see [13, Sec. 3.4.1.1] for a similar result which is readily generalized). Alternatively, it can be obtained as a consequence of Theorem 4.1.

PROPOSITION 1.1 There exists at least one solution to (1.4).

The example given by Foss, Hrusa and Mizel [17] shows that under the conditions stated above, it is possible that the minimization problem (1.4) may exhibit the *Lavrentiev gap phenomenon*

$$\inf \mathcal{J}(\mathcal{A}_\infty) > \inf \mathcal{J}(\mathcal{A}_1). \tag{1.6}$$

For the numerical solution of the minimization problem (1.4), (1.6) means that the energy of a solution to (1.4) *cannot* be approximated from a conforming finite element space. This effect will be explained in Section 3 where it is shown that the P_1-finite element method (P_1-FEM) converges to the global minimizer in $\mathcal{A}_1$ if and only if $\inf \mathcal{J}(\mathcal{A}_\infty) = \inf \mathcal{J}(\mathcal{A}_1)$. The main reason for this result is simple. If $(u_h)_{h\in(0,1]}$ is a family of discrete minimizers then $u_h \in \mathrm{W}^{1,\infty}(\Omega)^m$, and hence (1.6) makes it impossible that $\mathcal{J}(u_h) \to \inf \mathcal{J}(\mathcal{A}_1)$.

To overcome this difficulty, several methods have been proposed in the literature. While they differ in specific details, all methods use the same basic principle. First, one introduces a 'regularized' energy $\mathcal{J}_\varepsilon : \mathcal{A}_1 \to \mathbb{R}$ which is continuous in the strong $\mathrm{W}^{1,1}$-topology. It follows that $\mathcal{J}_\varepsilon$ can be approximated from a conforming finite element space, say $\mathrm{P}_1^{\mathrm{c}}(\mathcal{T}_h)^m$ (see Section 2.2 for its definition), where $\mathcal{T}_h$ denotes a finite element grid with global mesh size h. One then aims to prove that approximate minimizers of $\mathcal{J}_\varepsilon$, which can *in principle* be computed numerically, converge to a minimizer of $\mathcal{J}$ as $\varepsilon \searrow 0$. In this way one obtains convergence results of the following type: *for any sequence $\varepsilon_j \searrow 0$ there exists a sequence $h_j \searrow 0$ such that minimizers of $\mathcal{J}_{\varepsilon_j}$ in $\mathcal{A} \cap \mathrm{P}_1^{\mathrm{c}}(\mathcal{T}_{h_j})^m$ converge weakly to a minimizer u of $\mathcal{J}$ in $\mathcal{A}_1$, and $\mathcal{J}_{\varepsilon_j}(u_{h_j}) \to \mathcal{J}(u)$.* Methods of this type include the penalty method of Ball and Knowles [5, 18] and its extension to polyconvex integrands by Negron–Marrero [23], the element-removal method of Li [20, 21], the truncation method of Bai and Li [2], and the L^1-penalty method of Carstensen and Ortner [11].

The main advantage of these methods is their generality, as they are *in principle* appropriate for very general classes of minimization problems. However, they all share the same drawback, namely that the relationship between ε and h is entirely unkown a priori. All methods cited above are very sensitive to the choice of ε and h, and it is therefore difficult to devise robust algorithms which can compute the sequence (ε_j, h_j) and the corresponding minimizers [11]. For a more extensive discussion of numerical methods for computing singular minimizers see [4].

The novel contribution of the present work is the identification and analysis of a standard numerical method, the Crouzeix–Raviart finite element method [12] (CR-FEM) which does not require a regularization parameter. It should be noted from the outset that this method will not be succesful for general variational problems but is restricted to (minor extensions of) the class described above. This follows immediately from the fact that in one dimension the CR-FEM reduces to the conforming P_1-FEM, which is unable to approximate, for example, the Manià problem (1.1).

As explained earlier, the main difficulty for the P_1-FEM is that $\mathcal{J}(v_h) \to \mathcal{J}(v)$ may fail for some v, independent of the choice of the approximating sequence $v_h \in \mathrm{P}_1^{\mathrm{c}}(\mathcal{T}_h)$. However, for the non-conforming CR-finite element space $\mathrm{CR}(\mathcal{T}_h)$ (see Section 2.2 for its definition) one can define an interpolation operator $I_h : \mathrm{W}^{1,1}(\Omega)^m \to \mathrm{CR}(\mathcal{T}_h)^m$ which satisfies

$$\int_T \nabla v \,\mathrm{d}x = \int_T \nabla I_h v \,\mathrm{d}x \qquad \forall T \in \mathcal{T}_h. \tag{1.7}$$

Jensen's inequality immediately implies that

$$\int_T W(\nabla I_h v)\,\mathrm{d}x = \int_T W\Big(|T|^{-1}\int_T \nabla v(y)\,\mathrm{d}y\Big)\,\mathrm{d}x \leqslant \int_T W(\nabla v)\,\mathrm{d}x \qquad \forall T \in \mathcal{T}_h,$$

from which one easily obtains

$$\limsup_{h\to 0} \mathcal{J}(v_h) \leqslant \mathcal{J}(v) \qquad \forall v \in \mathrm{W}^{1,1}(\Omega)^m.$$

Property (1.7) is the crucial ingredient in the proof of convergence of the Crouzeix–Raviart finite element method for (1.4) (see Theorem 4.1).

Possibilities and challenges for extensions of this analysis will be discussed in the conclusion. Finally, it should be noted that the use of non-conforming finite element methods was first proposed by Ball [4].

2. Preliminaries

This section is intended to fix the notation and to state some auxiliary results.

2.1 *Function spaces*

Let A be an open subset of $\mathbb{R}^n$. We use $\mathrm{L}^p(A)$ and $\mathrm{W}^{1,p}(A)$ to denote the standard Lebesgue and Sobolev spaces and equip them with their usual norms. The space of distributions is denoted by $\mathcal{D}'(A)$ [1]. The distributional gradient operator is denoted D, while the weak gradient operator is denoted ∇. The spaces of continuously differentiable functions with compact support in A are denoted $\mathrm{C}_0^k(A)$.

In addition, we will also require the space of functions of bounded variation [15]. A function $u \in \mathrm{L}^1(\Omega)$ belongs to $\mathrm{BV}(\Omega)$ if its total variation,

$$|Du|(\Omega) = \sup_{\substack{\varphi \in \mathrm{C}_0^1(\Omega) \\ \|\varphi\|_{\mathrm{L}^\infty} = 1}} \int_\Omega u \,\mathrm{div}\varphi \,\mathrm{d}x,$$

is finite. $\mathrm{BV}(\Omega)$ equipped with the norm $\|u\|_{\mathrm{BV}} = \|u\|_{\mathrm{L}^1} + |Du|(\Omega)$ is a Banach space. We shall make use of two crucial properties of the space $\mathrm{BV}(\Omega)$. First, elements of $\mathrm{BV}(\Omega)$ may be discontinuous and therefore non-conforming finite element spaces are contained in it. Second, $\mathrm{BV}(\Omega)$ is compactly embedded in $\mathrm{L}^1(\Omega)$, i.e., if K is an index set and if $\sup_{k\in K}\|u_k\|_{\mathrm{BV}} < +\infty$ then there exists $u \in \mathrm{BV}(\Omega)$ and a sequence $(k_j) \subset K$ such that

$$u_{k_j} \to u \qquad \text{strongly in } \mathrm{L}^1(\Omega); \tag{2.1}$$

see [15, Section 5.2.3].

For example, we can use this compactness property to prove the following Poincaré–Friedrichs inequality.

LEMMA 2.1 There exists a constant C_p such that

$$\|u\|_{\mathrm{L}^1(\Omega)} \leqslant C_p\Big(|Du|(\Omega) + \sum_{i=1}^m \Big|\int_{\Gamma_i} u^{(i)}\,\mathrm{d}x\Big|\Big) \qquad \forall u \in \mathrm{BV}(\Omega)^m. \tag{2.2}$$

Proof. Suppose, for contradiction, that there exist $u_k \in \mathrm{BV}(\Omega)$ such that $\|u_k\|_{\mathrm{L}^1} = 1$ but $|Du_k|(\Omega)+\sum_{i=1}^m |\int_{\Gamma_i} u_k^{(i)}\,\mathrm{d}x| \leqslant 1/k$. Due to the aforementioned compactness result, we can assume, without loss of generality, that $u_k \to u$ strongly in $\mathrm{BV}(\Omega)^m$, and that $|\int_{\Gamma_i} u_k^{(i)}\,\mathrm{d}x| \to 0$ for all i. Since the trace operator is continuous from $\mathrm{BV}(\Omega)^m$ to $\mathrm{L}^1(\partial\Omega)^m$ [15, Sec. 5.3, Thm. 1], it follows that $\int_{\Gamma_i} u^{(i)}\,\mathrm{d}s = 0$, $i = 1, \dots, m$. Since $|Du|(\Omega) = 0$, u is constant in Ω, and Theorem 2 in [15, Sec. 5.3] shows that $u = 0$. Hence, we have arrived at a contradiction to our assumption that $\|u_k\|_{\mathrm{L}^1} = 1$. □

In order to guarantee that families with bounded energy only have accumulation points in $\mathrm{W}^{1,1}(\Omega)^m$, we have imposed superlinear growth of W in (1.5). This is related to the Dunford–Pettis criterion for compactness in the weak topology of $\mathrm{L}^1(\Omega)^m$ [14, Th. IV.8.9]. Namely, if $(v_j)_{j\in\mathbb{N}} \subset \mathrm{L}^1(\Omega)^k$ and if $\int_\Omega \phi(|v_j|)\,\mathrm{d}x$ is bounded then $(v_j)_{j\in\mathbb{N}}$ is precompact in the weak topology of $\mathrm{L}^1(\Omega)^k$. This result follows from the equi-integrability of the family $(v_j)_{j\in\mathbb{N}}$ which is an immediate consequence.

LEMMA 2.2 Suppose $(v_j)_{j\in\mathbb{N}} \subset \mathrm{L}^1(\Omega)^k$, and $\sup_{j\in\mathbb{N}} \int_\Omega \phi(|v_j|)\,\mathrm{d}x < +\infty$, then there exists a subsequence $j_r \to \infty$ and $v \in \mathrm{L}^1(\Omega)^k$ such that $v_{j_r} \rightharpoonup v$ weakly in $\mathrm{L}^1(\Omega)^k$.

In order to deduce strong convergence from weak convergence, we will use the following result.

LEMMA 2.3 Suppose that W is strictly convex. If $F_j \rightharpoonup F$ weakly in $\mathrm{L}^1(\Omega)^{m\times n}$ and if $\int_\Omega W(F_j)\,\mathrm{d}x \to \int_\Omega W(F)\,\mathrm{d}x$, then $F_j \to F$ strongly in $\mathrm{L}^1(\Omega)^{m\times n}$.

Proof. The result follows immediately from [26, Theorem 3(i)] upon noting that strict convexity of W implies that $(F(x), W(F(x)))$ is an extremal point of the epigraph of W for a.a. $x \in \Omega$. □

2.2 *Finite element spaces*

In this section, the finite element spaces used to discretize (1.4) are described briefly; see [7, 8] for further detail.

Let $(\mathcal{T}_h)_{h\in(0,1]}$ be a family of uniformly shape-regular partitions of $\bar\Omega$ into closed simplices T such that $h_T := \mathrm{diam}(T) \leqslant h$ for all $T \in \mathcal{T}_h$. As usual, we require that $\mathcal{T}_h$ has no hanging nodes in 2D, no hanging nodes or edges in 3D, and so forth. Let $\mathcal{E}_h$ denote the collection of $n-1$ dimensional faces of elements and let $\mathcal{N}_h^{\mathrm{c}}$ denote the set of all corners of elements. The collection of interior faces is denoted by $\mathcal{E}_h^{\mathrm{int}}$. For each face $E \in \mathcal{E}_h$ we set $h_E = \mathrm{diam}(E)$. Uniform shape regularity of the family $(\mathcal{T}_h)_{h\in(0,1]}$ implies the existence of a constant $c > 0$, independent of h, such that

$$c h_T^n \leqslant |T| \leqslant h_T^n \quad \forall T \in \mathcal{T}_h, \text{ and } \quad c h_E^{n-1} \leqslant |E| \leqslant h_E^{n-1} \quad \forall E \in \mathcal{E}_h. \tag{2.3}$$

We assume furthermore that for every h, the partition of the boundary induced by $\mathcal{T}_h$ respects the sets Γ_i, $i = 1, \dots, m$, i.e., up to a set of surface measure zero each of these sets can be written as a union of faces in $\mathcal{E}_h$.

The space of all piecewise affine functions relative to the partition $\mathcal{T}_h$ is denoted

$$\mathrm{P}_1(\mathcal{T}_h) = \big\{ v \in \mathrm{L}^1(\Omega) : v|_T \text{ is affine } \ \forall T \in \mathcal{T}_h \big\}.$$

The space of continuous P_1-finite element functions is denoted

$$\mathrm{P}_1^{\mathrm{c}}(\mathcal{T}_h) = \mathrm{P}_1(\mathcal{T}_h) \cap \mathrm{C}(\bar\Omega).$$

Let $\Pi_h : \mathrm{C}(\bar\Omega) \to \mathrm{P}_1^{\mathrm{c}}(\mathcal{T}_h)$ denote the nodal interpolation operator defined by

$$(\Pi_h v)(z) = v(z) \qquad \forall z \in \mathcal{N}_h^{\mathrm{c}}.$$

Let $\mathcal{N}_h^{\mathrm{nc}}$ denote the collection of all barycenters of faces,

$$\mathcal{N}_h^{\mathrm{nc}} = \big\{ |E|^{-1} \textstyle\int_E x\,\mathrm{d}s : E \in \mathcal{E}_h \big\},$$

and let $\mathrm{CR}(\mathcal{T}_h)$ denote the first-order Crouzeix–Raviart finite element space,

$$\mathrm{CR}(\mathcal{T}_h) = \{v_h \in \mathrm{P}_1(\mathcal{T}_h) : v_h \text{ is continuous in } \mathcal{N}_h^{\mathrm{nc}}\}.$$

Since elements of $\mathrm{CR}(\mathcal{T}_h)^m$ may be discontinuous we now use ∇v_h to denote the element-wise gradient of $v_h \in \mathrm{CR}(\mathcal{T}_h)^m$. We also require a notation for the jumps across interior faces. If $E = T^+ \cap T^- \in \mathcal{E}_h^{\mathrm{int}}$ and if $v_h^{\pm}$ denote the traces from $T^{\pm}$, and $\nu^{\pm}$ the outer unit normals to $T^{\pm}$, we set

$$[v_h] = v_h^+ \otimes \nu^+ + v_h^- \otimes \nu^-,$$

where $(a \otimes b)_{ij} = a_i b_j$. It follows that $|[v_h]| = |v_h^+ - v_h^-|$, where $|\cdot|$ denotes the Frobenius norm of a matrix or the ℓ^2-norm of a vector. With this notation, the distributional gradient of a function $v_h \in \mathrm{P}_1(\mathcal{T}_h)^m$ can be written as

$$\langle Dv_h, \varphi \rangle = -\int_\Omega v_h \cdot \mathrm{div}\varphi \,\mathrm{d}x = \int_\Omega \nabla v_h : \varphi \,\mathrm{d}x - \int_{\cup \mathcal{E}_h^{\mathrm{int}}} [v_h] : \varphi \,\mathrm{d}s \qquad \forall \varphi \in \mathrm{C}_0^1(\Omega)^{m \times n}. \tag{2.4}$$

For each $z \in \mathcal{N}_h^{\mathrm{nc}}$ let $E_z \in \mathcal{E}_h$ be the unique face which contains z. The interpolation operator $I_h : \mathrm{W}^{1,1}(\Omega)^m \to \mathrm{CR}(\mathcal{T}_h)^m$ is defined via

$$I_h v(z) = |E_z|^{-1} \int_{E_z} v \,\mathrm{d}s \qquad \forall z \in \mathcal{N}_h^{\mathrm{nc}}.$$

This operator was originally defined by Crouzeix and Raviart [12] and already used in a similar manner as we will use in the present work. We summarize its most important properties for our purpose in the following lemma.

Lemma 2.4 Let $v \in \mathrm{W}^{1,1}(\Omega)^m$, then

$$\|v - I_h v\|_{\mathrm{L}^1(T)} \leqslant C_a h_T \|\nabla v\|_{\mathrm{L}^1(T)} \qquad \forall T \in \mathcal{T}_h \quad \forall h \in (0,1], \tag{2.5}$$

where $C_a = 1/2 + 1/n$. Furthermore, it holds that

$$\int_T \nabla v \,\mathrm{d}x = \int_T \nabla I_h v \,\mathrm{d}x \qquad \forall T \in \mathcal{T}_h \quad \forall h \in (0,1]. \tag{2.6}$$

Proof. The first result is fairly standard and follows from the usual arguments for estimating interpolation errors. The constant $C_a = 1/2 + 1/n$ which is independent of the mesh quality can be found in [24].

To prove (2.6), we observe that

$$|E_z|^{-1} \int_{E_z} v \,\mathrm{d}s = I_h v(z) = |E_z|^{-1} \int_{E_z} I_h v \,\mathrm{d}s,$$

and hence,

$$\int_T \nabla v \,\mathrm{d}x = \int_{\partial T} v \otimes \nu \,\mathrm{d}s = \int_{\partial T} I_h v \otimes \nu \,\mathrm{d}s = \int_T \nabla I_h v \,\mathrm{d}x,$$

using the fact that ν is constant on each edge of T. □

3. Conforming Finite Element Methods

The purpose of this section is to show under which conditions conforming finite element methods converge to the global minimizer in $\mathcal{A}_1$, and to illustrate why they fail in the presence of the Lavrentiev phenomenon. The results contained here are well-understood by experts in the field (though not explicitly stated in the literature) and are included primarily for the purpose of motivating the subsequent analysis of non-conforming methods. Throughout this section, we make the simplifying assumption that W is continuous, i.e., that it cannot take the value $+\infty$.

Let $(\mathcal{T}_h)_{h\in(0,1]}$ be a uniformly shape regular family of finite element meshes as described in Section 2. With slight abuse of notation, we use $\Pi_h g$ to denote the piecewise affine boundary function with nodal values $g(z)$, $z \in \mathcal{N}_h^{\mathrm{c}} \cap \partial\Omega$. We define the set of admissible functions for the conforming P_1-FEM as

$$\mathcal{A}_h^{\mathrm{c}} = \big\{ v_h \in \mathrm{P}_1^{\mathrm{c}}(\mathcal{T}_h)^m : v_h^{(i)} = \Pi_h g^{(i)} \ \text{ on } \Gamma_i, i = 1, \dots, m \big\}.$$

The P_1-FEM is to find

$$u_h \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_h^{\mathrm{c}}). \tag{3.1}$$

It is easy to see that (3.1) has at least one solution, and it is fairly straightforward to obtain the following weak convergence result.

THEOREM 3.1 Assume that W is continuous in $\mathbb{R}^{m\times n}$, and that the Lavrentiev phenomenon (1.6) does *not* occur. For $h \in (0,1]$ let u_h be a solution of (3.1); then there exists a sequence $h_j \searrow 0$ and $u \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_1)$ such that $u_{h_j} \rightharpoonup u$ weakly in $\mathrm{W}^{1,1}(\Omega)^m$ and $\mathcal{J}(u_{h_j}) \to \mathcal{J}(u)$, as $j \to \infty$.

If $\#\operatorname{argmin} \mathcal{J}(\mathcal{A}_1) = 1$ then $u_h \rightharpoonup u$ weakly in $\mathrm{W}^{1,1}(\Omega)^m$. If W is strictly convex, then $u_h \to u$ strongly in $\mathrm{W}^{1,1}(\Omega)^m$.

Proof. We begin by proving an upper bound on $\mathcal{J}(u_h)$. Let $v_k \in \mathcal{A}_\infty$ such that $\mathcal{J}(v_k) \searrow \mathcal{J}(\mathcal{A}_1)$. Lemma 2.2 in [11] states that, for all $v \in \mathrm{W}^{1,\infty}(\Omega)^m$,

$$\begin{aligned} \|\Pi_h v\|_{\mathrm{W}^{1,\infty}} \leqslant C \|v\|_{\mathrm{W}^{1,\infty}}, &\quad \forall h \in (0,1], \\ \Pi_h v \to v \text{ strongly in } \mathrm{L}^\infty(\Omega)^m, &\quad \text{as } h \searrow 0, \quad \text{and} \\ \nabla \Pi_h v(x) \to \nabla v(x), &\quad \text{as } h \searrow 0, \quad \text{for a.e. } x \in \Omega. \end{aligned}$$

(The first and second result is established using standard interpolation error analysis. The third result is a consequence of Rademacher's theorem.) Since W is globally continuous, it follows that, for any fixed k, $W(\nabla v_k) \in \mathrm{L}^\infty(\Omega)$ and $W(\nabla \Pi_h v_k)$ is uniformly bounded in $\mathrm{L}^\infty(\Omega)$, for $h \in (0,1]$. The pointwise convergence of $\nabla \Pi_h v_k$ together with the dominated convergence theorem, and the strong convergence of $\Pi_h v_k$ in $\mathrm{L}^\infty(\Omega)^m$ then imply that

$$\mathcal{J}(\Pi_h v_k) \to \mathcal{J}(v_k) \quad \text{ as } h \searrow 0 \qquad \forall k \in \mathbb{N}.$$

Upon extracting a suitable diagonal sequence we find a family $w_h \in \mathcal{A}_h^{\mathrm{c}}$ such that

$$\mathcal{J}(w_h) \to \inf \mathcal{J}(\mathcal{A}_\infty) = \inf \mathcal{J}(\mathcal{A}_1) \quad \text{ as } h \searrow 0. \tag{3.2}$$

Suppose now that $u_h \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_h^{\mathrm{c}})$, then $\mathcal{J}(u_h) \leqslant \mathcal{J}(w_h) \leqslant C_1$ for some $C_1 \in \mathbb{R}$. Using also (2.2) and the fact that $g \in \mathrm{L}^\infty(\Omega)^m$, we have

$$\begin{aligned} \int_\Omega \phi(|\nabla u_h|)\,\mathrm{d}x \leqslant&\ C_1 + \|f\|_{\mathrm{L}^\infty} \|u_h\|_{\mathrm{L}^1} \\ \leqslant&\ C_1 + C_p \|f\|_{\mathrm{L}^\infty} \big(\|\nabla u_h\|_{\mathrm{L}^1} + |\partial\Omega| \|g\|_{\mathrm{L}^\infty} \big) \\ \leqslant&\ C_2 (1 + \|\nabla u_h\|_{\mathrm{L}^1}), \end{aligned}$$

for some $C_2 > 0$. Since ϕ is convex and $|\nabla u_h|$ is integrable, we can use Jensen's inequality to estimate

$$|\Omega| \phi\big(|\Omega|^{-1} \|\nabla u_h\|_{\mathrm{L}^1} \big) \leqslant C_2 \big(1 + \|\nabla u_h\|_{\mathrm{L}^1} \big).$$

Since ϕ is superlinear, it follows that $\|\nabla u_h\|_{\mathrm{L}^1}$ is uniformly bounded. Hence, we obtain

$$\|u_h\|_{\mathrm{L}^1} + \int_\Omega \phi(|\nabla u_h|)\,\mathrm{d}x \leqslant C_3 \qquad \forall h \in (0,1],$$

for some constant $C_3 \in \mathbb{R}$. In particular, Lemma 2.2 implies that $(u_h)_{h\in(0,1]}$ is precompact in the weak topology of $\mathrm{W}^{1,1}(\Omega)^m$, and that there exists $u \in \mathrm{W}^{1,1}(\Omega)^m$ and $h_j \searrow 0$ such that $u_{h_j} \rightharpoonup u$, weakly in $\mathrm{W}^{1,1}(\Omega)^m$.

To show that $u \in \mathcal{A}_1$ we use the fact that the trace operator is bounded from $\mathrm{W}^{1,1}(\Omega)^m$ to $\mathrm{L}^1(\Omega)^m$, and hence, for $i = 1, \dots, m$,

$$0 = u_h^{(i)} - \Pi_h g^{(i)} \rightharpoonup u^{(i)} - g^{(i)} \qquad \text{weakly in } \mathrm{L}^1(\Gamma_i).$$

Since $\mathcal{J}$ is lower-semicontinuous, and using (3.2), we can estimate

$$\mathcal{J}(u) \leqslant \liminf_{j\to\infty} \mathcal{J}(u_{h_j}) \leqslant \limsup_{j\to\infty} \mathcal{J}(u_{h_j}) \leqslant \limsup_{j\to\infty} \mathcal{J}(w_{h_j}) = \inf \mathcal{J}(\mathcal{A}_1). \tag{3.3}$$

It follows that $u \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_1)$ and that $\mathcal{J}(u_{h_j}) \to \mathcal{J}(u)$.

The remaining statements follow from standard arguments; they are contained, for example, in the proof of Theorem 4.1. □

Remark 3.1 If $\mathcal{A}_h^c \subset \mathcal{A}_\infty$, which essentially requires that $\Pi_h g = g$ for all $h \in (0, 1]$, then it follows from Theorem 3.1 that the P_1-FEM (3.1) converges to the original minimization problem (1.4) if, and only if, the Lavrentiev phenomenon (1.6) does *not* occur. See also [11, Sec. 2] for a more detailed discussion of this fact. □

Remark 3.2 The classical condition which is usually employed in order to avoid (1.6) is to impose the same growth on W from above and below, e.g.,

$$c_0(|F|^p - 1) \leqslant W(F) \leqslant c_1(|F|^p + 1) \tag{3.4}$$

for some $p \in (1, \infty)$. Namely, if (3.4) holds then $\mathcal{J}$ is continuous in the strong topology of $\mathrm{W}^{1,p}(\Omega)^m$ and density of smooth functions implies that $\inf \mathcal{J}(\mathcal{A}_1) = \inf \mathcal{J}(\mathcal{A}_p) = \inf \mathcal{J}(\mathcal{A}_\infty)$. □

Remark 3.3 The proof of Theorem 3.1 suggests that, to understand the conforming P_1-finite element method in the general case, we should study the lower semicontinuous envelope of the restriction of $\mathcal{J}$ to $\mathcal{A}_\infty$. This is defined, for $v \in \mathcal{A}_1$, by

$$\bar{\mathcal{J}}_\infty(v) = \inf \big\{ \liminf_{j\to\infty} \mathcal{J}(v_j) : v_j \in \mathcal{A}_\infty \text{ and } v_j \rightharpoonup v \text{ weakly in } \mathrm{W}^{1,1}(\Omega)^m \big\}.$$

By definition, $\bar{\mathcal{J}}_\infty$ is sequentially weakly lower-semicontinuous $\mathcal{A}_1$. Furthermore, since $\mathcal{J}$ was also lower semi-continuous, it follows that $\mathcal{J}(v) = \bar{\mathcal{J}}_\infty(v)$ for all $v \in \mathcal{A}_\infty$. Hence, there exists a minimizing sequence from $\mathcal{A}_\infty$, and therefore, the direct method of the calculus of variations guarantees the existence of at least one minimizer u of $\bar{\mathcal{J}}_\infty$ in $\mathcal{A}_1$.

Viewing the Lavrentiev phenomenon as a relaxation problem is not a new idea [10]. However, there seems to be no general representation for $\bar{\mathcal{J}}_\infty$ available and therefore this option was not exploited here. One interesting remark can be made, however. Assume again that $u_h \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_h^c)$. Provided that $\mathcal{A}_h^c \subset \mathcal{A}_\infty$ holds, it can be shown by repeating the proof of Theorem 3.1 almost verbatim that, for some seqence $h_j \searrow 0$, $u_{h_j} \rightharpoonup u$ weakly in $\mathrm{W}^{1,1}(\Omega)^m$ and $\mathcal{J}(u_{h_j}) \to \bar{\mathcal{J}}_\infty(u) = \inf \bar{\mathcal{J}}_\infty(\mathcal{A}_1)$. This shows that, in some sense, the conforming P_1-FEM approximates the wrong problem.

4. Non-Conforming Finite Element Methods

Let $(\mathcal{T}_h)_{h\in(0,1]}$ be a uniformly shape-regular family of partitions of Ω and let $\mathrm{CR}(\mathcal{T}_h)^m$ denote the Crouzeix–Raviart finite element spaces as described in Section 2.2. We can extend the definition of $\mathcal{J}$ to elements of $\mathrm{CR}(\mathcal{T}_h)^m$ by

$$\mathcal{J}(v_h) = \int_\Omega \Big[W(\nabla v_h) - f \cdot v_h \Big] \,\mathrm{d}x,$$

but note that ∇v_h now denotes the piecewise gradient.

Since $\mathrm{CR}(\mathcal{T}_h)$ is not a subspace of $\mathrm{W}^{1,1}(\Omega)$ we need to take care in defining the set of discrete admissible functions. For the sake of simplicity we shall use

$$\mathcal{A}_h^{\mathrm{nc}} = \{ I_h v : v \in \mathcal{A}_1 \}. \tag{4.1}$$

It can be easily seen that

$$\mathcal{A}_h^{\mathrm{nc}} = \Big\{ v_h \in \mathrm{CR}(\mathcal{T}_h)^m : \int_E (v_h^{(i)} - g^{(i)})\,\mathrm{d}s = 0 \quad \text{for } E \in \mathcal{E}_h, E \subset \Gamma_i, i = 1, \dots, m \Big\}.$$

The resulting CR-FEM is to compute

$$u_h \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_h^{\mathrm{nc}}). \tag{4.2}$$

Using the facts that W grows superlinearly, and that $\mathcal{A}_h^{\mathrm{nc}}$ is non-empty, it is easy to show that 4.2 has at least one solution.

THEOREM 4.1 For each $h \in (0, 1]$, let u_h be a solution to (4.2); then there exists a subsequence $h_j \searrow 0$ and $u \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_1)$ such that, as $j \to \infty$,

$$\begin{aligned} u_{h_j} &\to u && \text{strongly in } \mathrm{L}^1(\Omega)^m, \\ \nabla u_{h_j} &\rightharpoonup \nabla u && \text{weakly in } \mathrm{L}^1(\Omega)^{m\times n}, \text{ and} \\ \mathcal{J}(u_{h_j}) &\to \mathcal{J}(u). \end{aligned}$$

If $\#\operatorname{argmin} \mathcal{J}(\mathcal{A}_1) = 1$ then the entire family converges to the unique minimum. If W is strictly convex then $\nabla u_h \to \nabla u$ strongly in $\mathrm{L}^1(\Omega)^{m\times n}$.

The proof of Theorem 4.1 mimics the convergence proof for the P_1-FEM given in Section 3. However, we have used a number of tools such as Poincaré inequalities or the extraction of weakly convergent subsequences which are not readily available for elements of the space $\mathrm{CR}(\mathcal{T}_h)$. As a matter of fact, once these technical prerequisites are established, and bearing in mind the projection property of the CR-interpolant (2.6), the proof of Theorem 4.1 is a straightforward matter.

We begin with the elementary observation that the total variation of a CR-function can be bounded by the L^1-norm of its piecewise gradient.

LEMMA 4.1 There exists a constant C_g such that

$$|Dv_h|(\Omega) \leqslant C_g \|\nabla v_h\|_{\mathrm{L}^1} \qquad \forall v_h \in \mathrm{CR}(\mathcal{T}_h)^m \quad \forall h \in (0, 1]. \tag{4.3}$$

Proof. Recalling (2.4), we obtain

$$|Dv_h|(\Omega) \leqslant \|\nabla v_h\|_{\mathrm{L}^1(\Omega)} + \int_{\cup \mathcal{E}_h^{\mathrm{int}}} \big|[v_h]\big|\,\mathrm{d}s.$$

Thus, we need to bound the norm of the jumps in terms of $\|\nabla v_h\|_{\mathrm{L}^1}$ only. Let $E = T^+ \cap T^- \in \mathcal{E}_h^{\mathrm{int}}$ with midpoint z, and let $x \in E$, then

$$|[v_h(x)]| = |v_h^+(x) - v_h^-(x)| \leqslant |v_h^+(x) - v_h(z)| + |v_h(z) - v_h^-(x)| \leqslant h_E|\nabla v_h^+| + h_E|\nabla v_h^-|,$$

and consequently

$$\int_E \big|[v_h]\big|\,\mathrm{d}s \leqslant h_E^n\big(|\nabla v_h^+| + |\nabla v_h^-|\big) \leqslant h_{T^+}^n|\nabla v_h^+| + h_{T^-}^n|\nabla v_h^-| \leqslant c^{-1}\int_{T^+\cup T^-} |\nabla v_h|\,\mathrm{d}x,$$

where c is the shape-regularity constant from (2.3). Upon summing over $E \in \mathcal{E}_h^{\mathrm{int}}$, we obtain (4.3) with $C_g = (n + 1)/c + 1$. □

Combining Lemma 4.1 and Lemma 2.1, we obtain the broken Poincaré-Friedrichs inequality

$$\|v_h\|_{\mathrm{L}^1} \leqslant C_p' \Big(\|\nabla v_h\|_{\mathrm{L}^1} + \sum_{i=1}^{M} \Big| \int_{\Gamma_i} v_h^{(i)} \,\mathrm{d}s \Big| \Big) \qquad \forall v_h \in \mathrm{CR}(\mathcal{T}_h)^m \quad \forall h \in (0,1], \tag{4.4}$$

where $C_p' = C_p C_g$, which can be applied to $v_h = w_h - \Pi_h g$ and immediately implies

$$\|w_h\|_{\mathrm{L}^1} \leqslant C_p'' \big(1 + \|\nabla w_h\|_{\mathrm{L}^1}\big) \qquad \forall w_h \in \mathcal{A}_h^{\mathrm{nc}}. \tag{4.5}$$

Finally, before adressing the proof of Theorem 4.1, we demonstrate that 'weak' limits of Crouzeix–Raviart functions are weakly differentiable. This result and its proof are inspired by [9, Theorem 5.1].

THEOREM 4.2 Suppose that $u_h \in \mathrm{CR}(\mathcal{T}_h)^m$, $h \in (0,1]$, satisfy

$$\sup_{h \in (0,1]} \|u_h\|_{\mathrm{L}^1(\mathcal{T}_h)} + \int_\Omega \phi(|\nabla u_h|) \,\mathrm{d}x < +\infty; \tag{4.6}$$

then there exists a sequence $h_j \searrow 0$ and $u \in \mathrm{W}^{1,1}(\Omega)^m$ such that

$$u_{h_j} \to u \qquad \text{strongly in } \mathrm{L}^1(\Omega)^m, \quad \text{and} \tag{4.7}$$

$$\nabla u_{h_j} \rightharpoonup \nabla u \qquad \text{weakly in } \mathrm{L}^1(\Omega)^{m \times n}. \tag{4.8}$$

If $u_h \in \mathcal{A}_h^{\mathrm{nc}}$ for all $h \in (0,1]$, then $u \in \mathcal{A}_1$.

Proof. Using Lemma 4.1, it follows that $\|u_h\|_{\mathrm{BV}}$ is uniformly bounded. Hence, we can use the compactness theorem for the space $\mathrm{BV}(\Omega)^m$ (see Section 2.1 or [15, Sec. 5.2.3]), to deduce the existence of a subsequence $h_j \searrow 0$ and of $u \in \mathrm{L}^1(\Omega)^m$ such that

$$u_{h_j} \to u \qquad \text{strongly in } \mathrm{L}^1(\Omega)^m.$$

In particular, this implies (4.7).

Next, we show that u is weakly differentiable. We use (4.6) to extract a further subsequence (not relabelled) to obtain an $F \in \mathrm{L}^1(\Omega)^{m\times n}$ such that $\nabla u_{h_j} \rightharpoonup F$ weakly in $\mathrm{L}^1(\Omega)^{m\times n}$. We need to prove that $Du = F$ in the sense of distributions (or measures). To see this, fix $\varphi \in \mathrm{C}_0^1(\Omega)^{m\times n}$, and use (2.4) to obtain

$$\langle Du_{h_j}, \varphi \rangle = \int_\Omega \nabla u_{h_j} : \varphi \,\mathrm{d}x - \int_{\cup \mathcal{E}_{h_j}^{\mathrm{int}}} [u_{h_j}] : \varphi \,\mathrm{d}s. \tag{4.9}$$

For the first term on the right-hand side, we have

$$\int_\Omega \nabla u_{h_j} : \varphi \,\mathrm{d}x \to \int_\Omega F : \varphi \,\mathrm{d}x. \tag{4.10}$$

Furthermore, since $\int_E [u_{h_j}] \,\mathrm{d}s = 0$ for all $E \in \mathcal{E}_{h_j}^{\mathrm{int}}$, we can estimate the second term via

$$\begin{aligned} \Big| \sum_{E \in \mathcal{E}_{h_j}^{\mathrm{int}}} \int_E [u_{h_j}] : \varphi \,\mathrm{d}s \Big| &\leqslant \sum_{E \in \mathcal{E}_{h_j}^{\mathrm{int}}} \Big| \int_E [u_{h_j}] : \big(\varphi - (\varphi)_E\big) \,\mathrm{d}s \Big| \\ &\leqslant \sum_{E \in \mathcal{E}_{h_j}^{\mathrm{int}}} \int_E \big|[u_{h_j}]\big| \,\mathrm{d}s \, h_E \|\nabla\varphi\|_{\mathrm{L}^\infty} \\ &\leqslant h_j \|\nabla u_{h_j}\|_{\mathrm{L}^1} \, \|\nabla \varphi\|_{\mathrm{L}^\infty} \to 0. \end{aligned}$$

Combining this result with (4.9) and (4.10), we obtain

$$Du = \lim_{j\to\infty} Du_{h_j} = F$$

in the sense of distributions. This implies that u is weakly differentiable and that $F = \nabla u$.

To prove that $u \in \mathcal{A}_1$ we need to show that $u^{(i)} = g^{(i)}$ on Γ_i, for $i = 1, \dots, m$. Owing to the fact that $\mathrm{BV}(\Omega)$ is *not* compactly embedded in $\mathrm{L}^1(\partial\Omega)$, this turns out to be slightly tricky. First, we show that $u_h^{(i)}|_{\Gamma_i} \to g^{(i)}|_{\Gamma_i}$ in the sense of measures. If $\varphi \in \mathrm{C}_0^1(\mathbb{R}^n)$, then

$$\begin{aligned}\Big|\int_{\Gamma_i}(u_h^{(i)} - g^{(i)})\varphi\,\mathrm{d}s\Big| &\leqslant \sum_{E\subset\Gamma_i}\Big|\int_E (u_h^{(i)} - g^{(i)})(\varphi - (\varphi)_E)\,\mathrm{d}s\Big| \\ &\leqslant h\|u_h^{(i)} - g^{(i)}\|_{\mathrm{L}^1(\Gamma_i)}\|\nabla\varphi\|_{\mathrm{L}^\infty(\mathbb{R}^n)}.\end{aligned}$$

Since $\|u_h^{(i)} - g^{(i)}\|_{\mathrm{L}^1(\Gamma_i)} \leqslant C\|u_h - g\|_{\mathrm{BV}}$, which is uniformly bounded, we obtain said convergence.

In the second step, we show that $e_j := u_{h_j} - u \to 0$ in a similar sense. To this end, we extend e_j by zero to all of $\mathbb{R}^n$ to obtain a new function $\bar{e}_j \in \mathrm{BV}(\mathbb{R}^n)^m$. Since $\bar{e}_j \to 0$ strongly in $\mathrm{L}^1(\mathbb{R}^n)^m$ we have $\langle D\bar{e}_j, \varphi\rangle \to 0$ for all $\varphi \in \mathrm{C}_0^1(\mathbb{R}^n)^{m\times n}$. Moreover, as in (2.4), we have

$$\langle D\bar{e}_j, \varphi\rangle = \int_\Omega \nabla e_j : \varphi\,\mathrm{d}x - \int_{\cup\mathcal{E}_{h_j}^{\mathrm{int}}} [e_j] : \varphi\,\mathrm{d}s - \int_{\partial\Omega}(e_j\otimes\nu) : \varphi\,\mathrm{d}s. \tag{4.11}$$

We already know that the left-hand side, as well as first and second terms on the right-hand side of (4.11) converge to zero, and hence it follows that

$$\int_{\partial\Omega}(e_j\otimes\nu) : \varphi\,\mathrm{d}s \to 0 \qquad \forall\varphi \in \mathrm{C}_0^1(\mathbb{R}^n)^{m\times n}. \tag{4.12}$$

Combined with the previous step we obtain that, for $\varphi \in \mathrm{C}_0^1(\mathbb{R}^n)$,

$$\int_{\Gamma_i}(u^{(i)} - g^{(i)})\varphi\,\mathrm{d}s = \lim_{j\to\infty}\Big\{\int_{\Gamma_i}(u^{(i)} - u_{h_j}^{(i)})\varphi\,\mathrm{d}s + \int_{\Gamma_i}(u_{h_j}^{(i)} - g^{(i)})\varphi\,\mathrm{d}s\Big\} = 0.$$

Taking φ to be a mollified version of $(u^{(i)} - g^{(i)})/|u^{(i)} - g^{(i)}|$, and using the fact that Ω is a Lipschitz domain, we finally obtain that $u^{(i)} = g^{(i)}$ a.e. on Γ_i. □

Proof of Theorem 4.1. The proof of Theorem 4.1 is split into three parts. First, we complete the proof of the elementary but crucial approximation property which we have already outlined in the introduction. Second, we prove pre-compactness of numerical solutions. Finally, we use lower-semicontinuity of convex functionals to prove that any limit point is a minimizer.

Step 1: Upper Bound (Approximation). Let $v \in \mathcal{A}_1$ with $\mathcal{J}(v) < \infty$, and let $I_h v$ be the Crouzeix–Raviart interpolant of v. Then, by definition, $I_h v \in \mathcal{A}_h^{\mathrm{nc}}$. Since ∇v is integrable, we can use the mean value property (2.6) and Jensen's Inequality to deduce

$$\int_T W(\nabla I_h v)\,\mathrm{d}x = \int_T W\Big(|T|^{-1}\int_T \nabla v(y)\,\mathrm{d}y\Big)\,\mathrm{d}x \leqslant \int_T W(\nabla v)\,\mathrm{d}x.$$

Summing over $T \in \mathcal{T}_h$, we obtain

$$\int_\Omega \Big[W(\nabla I_h v) - f\cdot I_h v\Big]\,\mathrm{d}x \leqslant \int_\Omega \Big[W(\nabla v) - f\cdot v + f\cdot(v - I_h v)\Big]\,\mathrm{d}x \leqslant \mathcal{J}(v) + \|f\|_{\mathrm{L}^\infty}\|v - I_h v\|_{\mathrm{L}^1}.$$

From the interpolation error estimate (2.5) we can deduce the bound

$$\mathcal{J}(I_h v) \leqslant \mathcal{J}(v) + hC_a\|f\|_{\mathrm{L}^\infty}\|\nabla v\|_{\mathrm{L}^1}, \tag{4.13}$$

and in particular, that

$$\limsup_{h\to 0} \mathcal{J}(I_h v) \leqslant \mathcal{J}(v) \qquad \forall v \in \mathcal{A}_1. \tag{4.14}$$

Step 2: Compactness (Stability). Suppose now that, for $h \in (0,1]$, $u_h \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_h^{\mathrm{nc}})$. Due to the growth condition (1.5) and the upper bound (4.13), we have

$$\int_\Omega \phi(|\nabla u_h|)\,\mathrm{d}x \leqslant C_1(1 + \|u_h\|_{\mathrm{L}^1}), \tag{4.15}$$

for some constant $C_1 \in \mathbb{R}$. We apply Jensen's Inequality on the left-hand side of (4.15) and the broken Poincaré–Sobolev inequality (4.5) on its right-hand side to deduce

$$|\Omega|\phi(|\Omega|^{-1}\|\nabla u_h\|_{\mathrm{L}^1}) \leqslant C_1(1 + C_p''(1 + \|\nabla u_h\|_{\mathrm{L}^1})) \leqslant C_1'(1 + \|\nabla u_h\|_{\mathrm{L}^1}).$$

The superlinear growth of ϕ implies that $\|\nabla u_h\|_{\mathrm{L}^1} \leqslant C_2$ for some constant $C_2 < \infty$. Inserting this information back into (4.15), we find that there exists a constant $C_3 \in \mathbb{R}$ such that

$$\|u_h\|_{\mathrm{L}^1} + \int_\Omega \phi(|\nabla u_h|)\,\mathrm{d}x \leqslant C_3 \qquad \forall h \in (0,1].$$

We can now employ Theorem 4.2 to deduce the existence of a subsequence $h_j \searrow 0$ and of $u \in \mathcal{A}_1$ such that

$$u_{h_j} \to u \qquad \text{strongly in } \mathrm{L}^1(\Omega)^m, \text{ and} \tag{4.16}$$

$$\nabla u_{h_j} \rightharpoonup \nabla u \qquad \text{weakly in } \mathrm{L}^1(\Omega)^{m\times n}. \tag{4.17}$$

Step 3: Lower Bound (Convergence). As an immediate consequence of (4.16) and (4.17), and the convexity of W we have [13, Theorem 3.4]

$$\mathcal{J}(u) \leqslant \liminf_{j\to\infty} \mathcal{J}(u_{h_j}).$$

Recalling (4.13), we therefore obtain

$$\mathcal{J}(u) \leqslant \liminf_{j\to\infty} \mathcal{J}(u_{h_j}) \leqslant \limsup_{j\to\infty} \mathcal{J}(u_{h_j}) \leqslant \inf \mathcal{J}(\mathcal{A}_1),$$

which shows that $u \in \operatorname{argmin} \mathcal{J}(\mathcal{A}_1)$ and that $\mathcal{J}(u_{h_j}) \to \mathcal{J}(u)$.

Suppose now that the minimizer u is unique. To obtain the convergence of the entire family $(u_h)_{h\in(0,1]}$ as $h \searrow 0$ we note that we could have started the proof with an arbitrary subsequence. Thus, if there were any subsequence of $(u_h)_{h\in(0,1]}$ which is uniformly bounded away from u in the L^1-norm, we would immediately arrive at a contradiction.

Finally, if W is strictly convex then the minimizer is indeed unique and, using Lemma 2.3, we deduce strong convergence of the broken gradients. □

5. Computational Examples

We test the nonconforming finite element method on a modified version of the example given by Foss, Hrusa and Mizel [17]. Let $n = m = 2$ and define, for $\alpha > 0$ and $p \in (1,\infty)$,

$$W_{\alpha,p}(F) = \big(|F|^2 - 2\det F\big)^4 + \frac{\alpha}{p}|F|_p^p,$$

and $f \equiv 0$, where $|F|_p^p = \sum_{i,j=1}^2 |F_{ij}|^p$. We denote the resulting functional $\mathcal{J}_{\alpha,p}$. To see that $W_{\alpha,p}$ is convex, note that $F \mapsto (|F|^2 - 2\det F)$ is a non-negative quadratic form on $\mathbb{R}^{2\times 2}$ and that $s \mapsto s^4$ is monotone and convex.

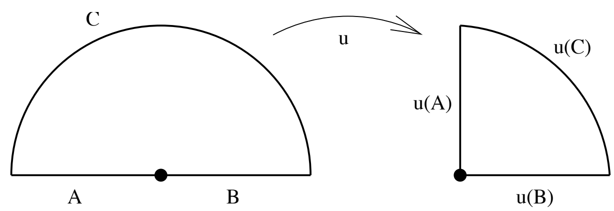

FIG. 1. Boundary conditions in the Foss–Hrusa–Mizel example: the portion A, B, C of the boundary of Ω are, respectively, deformed into $u(A), u(B), u(C)$.

Let Ω be the semi-circle

$$\Omega = \{|x| < 1, x_2 > 0\},$$

with boundary $\partial\Omega = A \cup B \cup C$ where

$$A = \{x_2 = 0, x_1 < 0\}, \quad B = \{x_2 = 0, x_1 > 0\}, \quad \text{and} \quad C = \partial\Omega \cap \{|x| = 1\},$$

and define

$$\mathcal{A} = \big\{u \in \mathrm{W}^{1,1}(\Omega)^2 : u^{(1)}|_A = 0, u^{(2)}|_B = 0, \text{ and } u(x) = \big(\cos(\theta/2), \sin(\theta/2)\big) \text{ for } x \in C\big\},$$

where $x = r(\cos(\theta), \sin(\theta))$. Admissible functions can be interpreted as deformations of the semi-circle Ω into a quarter circle (cf. Figure 1).

The idea of the example is that, for $\alpha = 0$, the minimizers of $\mathcal{J}$ in in $\mathcal{A}_1$ and $\mathcal{A}_\infty$ can be computed explicitly,

$$\begin{aligned} u &:= \operatorname{argmin} \mathcal{J}_{0,p}(\mathcal{A}_1) = r^{1/2}\big(\cos(\theta/2), \sin(\theta/2)\big), \\ \bar{u} &:= \operatorname{argmin} \mathcal{J}_{0,p}(\mathcal{A}_8) = r^{11/14}\big(\cos(\theta/2), \sin(\theta/2)\big). \end{aligned}$$

The corresponding energies are $\mathcal{J}_{0,p}(u) = 0$ and $\mathcal{J}_{0,p}(\bar{u}) = (2/7)^6\pi$, and hence the minimization problem $\min \mathcal{J}_{0,p}(\mathcal{A}_1)$, exhibits the Lavrentiev gap phenomenon. Note also that, since $\bar{u} \in \mathcal{A}_8$, and since $\mathcal{A}_\infty$ is dense in $\mathcal{A}_8$ and $\mathcal{J}_{0,p}$ continuous in $\mathrm{W}^{1,8}(\Omega)^2$, it follows that $\mathcal{J}_{0,p}(\bar{u}) = \inf \mathcal{J}_{0,p}(\mathcal{A}_\infty)$.

However, $W_{0,p}$ does not have superlinear growth, and in fact, the CR-FEM solution is unstable, that is, the CR-FEM minimizers are unbounded in the L^1-norm as $h \searrow 0$. However, upon observing that u is also the solution of Laplace's equation under the boundary conditions defined through $\mathcal{A}$, we see that

$$\begin{aligned} \mathcal{J}_{\alpha,2}(u) =& \; \mathcal{J}_{0,2}(u) + \tfrac{\alpha}{2}\|\nabla u\|_{\mathrm{L}^2}^2 \\ =& \; (\mathcal{J}_{0,2}(u) - \mathcal{J}_{0,2}(\bar{u})) + \mathcal{J}_{0,2}(\bar{u}) + \tfrac{\alpha}{2}\|\nabla u\|_{\mathrm{L}^2}^2 \\ \leqslant& \; (2/7)^7\pi + \mathcal{J}_{\alpha,2}(v) \qquad \forall v \in \mathcal{A}_\infty. \end{aligned}$$

Thus, the Lavrentiev gap phenomenon persists for $\alpha > 0$ and $p = 2$, with global minimizer u.

5.1 *Example 1*

For the first numerical experiment, we set $\alpha = 1$, $p = 2$, and solve the minimization problem using both the CR-FEM and the P_1-FEM. The radial components of the solutions are shown in Figure 2(a) where we see a significant gap. The convergence rate for $|\mathcal{J}(u_h) - \mathcal{J}(u)|$ is plotted in Figure 2(b). Since the exact solution has an $r^{1/2}$ singularity at the origin, we can at best expect an $O(h)$ convergence rate for the energy. This is precisely the rate observed in the experiment.

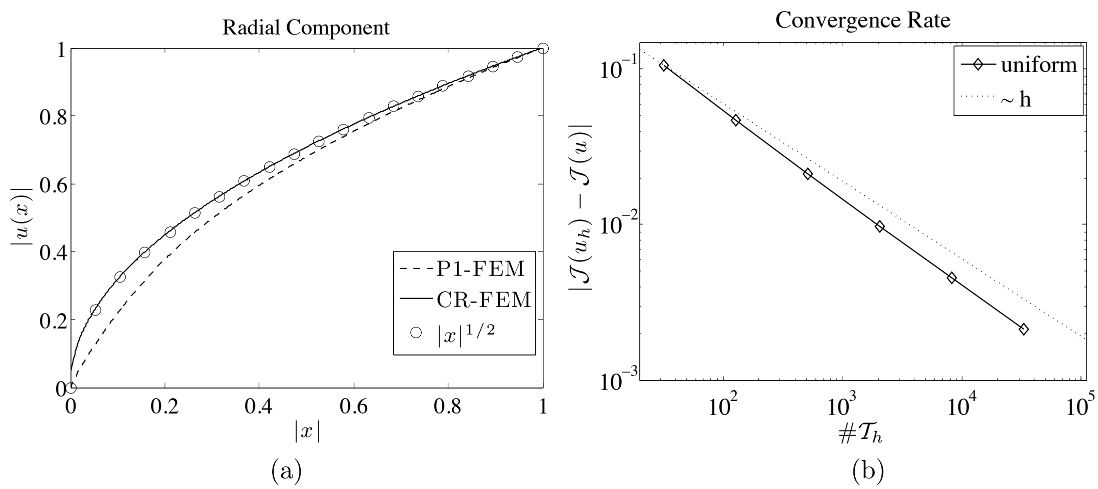

FIG. 2. (a) Comparison of exact, P_1-FEM and CR-FEM solutions for the modified Foss–Hrusa–Mizel example with $\alpha = 1$ and $p = 2$. (b) Convergence rate of the energy for the CR-FEM solution.

5.2 *Example 2*

In the second experiment, we choose $\alpha = 1/10$ and $p \in \{2, 3, 4, 6\}$, and try to predict in which cases the the resulting minimisation problem exhibits a Lavrentiev gap. To this end, we solve the problem both with the CR-FEM and the P_1-FEM and, in Figure 3, plot the difference in energy against the number of elements in the mesh. Since we expect convergence rates (though possibly fairly low rates) for both methods, we should be able to observe a flattening of the curves for those problems where a Lavrentiev gap occurs. For $p = 6$ we clearly observe a convergence rate, which suggests that no Lavrentiev phenomenon is present in this problem. This is consistent with (though it does not follow from) the results in [17]. For $p = 2$, we see the beginning of a flattening effect, which indicates that this problem possesses a Lavrentiev gap. This is again consistent with [17] and the discussion at the beginning of the section.

In the cases $p = 3, 4$, however, it is not clear Figure 3 whether the curves flatten or converge to zero. This difficulty is resolved in [24], where an adaptive mesh refinement algorithm for variational problems exhibiting the Lavrentiev phenomenon is formulated and analyzed.

Conclusion

For a small but important class of variational problems we have identified a numerical method, which is convergent even in the presence of the Lavrentiev phenomenon, and which does not require any regularization procedures. To conclude, we discuss various possibilities to extend the analysis in this paper.

A natural question to ask is, whether the analysis can be extended to stored energies of the type $W(x, u, F)$. It was already indicated in the introduction that this is not always possible. If the coupling between the variables u and ∇u is too strong (as in Manià's example), then the CR-FEM may fail. It should not pose great difficulties, however, to extend the analysis in the present paper to stored energies of the type $W(x, u, F)$ where

$$\widetilde{W}(F) + a(x) \leqslant W(x, u, F) \leqslant C(b(x) + |u|^q + \widetilde{W}(F)),$$

where $\widetilde{W}$ has superlinear growth, and where q is sufficiently small (in relation to $\widetilde{W}$), so that certain embedding results can be used to control the term $|u|^q$.

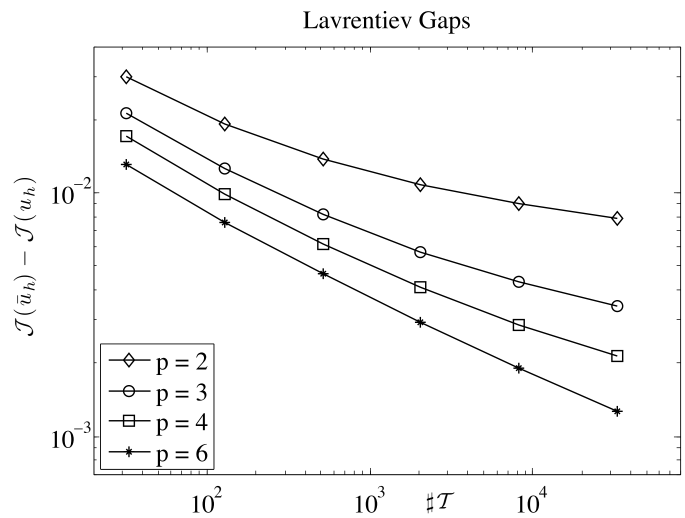

FIG. 3. Prediction of Lavrentiev gaps for the modified Foss–Hrusa–Mizel example with $\alpha = 1/10$ and $p \in \{2, 3, 4, 6\}$. Here, u_h denote the CR-FEM solutions and $\bar{u}_h$ the P_1-FEM solutions. The singularities in the exact solutions prevent a sufficiently rapid convergence using uniform mesh refinement so that a reliable prediction of the Lavrentiev gap cannot be made in all cases.

Another interesting question is whether the analysis can be extended to other non-conforming numerical methods. For example, it is straightforward to extend the convergence proof to the variational discontinuous Galerkin finite element method [9, 25]. To see this, simply note that the lower bound (lower-semicontinuity) is provided by the analysis in [9], while for the upper bound (approximating sequence) the Crouzeix–Raviart interpolant can be used, provided the discontinuous Galerkin finite element mesh is simplicial and contains no hanging nodes.

In fact, the interior penalty parameters of the discontinuous Galerkin finite element method provide an additional flexibility, so that that one may even attempt to explicitly control the function space in which to solve the minimization problem. For example, Foss [16] has given an example of a minimization problem where the the infimum of the energy in $\mathcal{A}_p$ depends continuously on the parameter p. It would be interesting to see whether, by adjusting the penalty term to have p-growth [9], the discontinuous Galerkin discretization converges to the solution in the correct function space. Establishing a rigorous theory for this case is, in all likelihood, a formidable challenge.

Finally, the most important extension, namely to quasi-convex or poly-convex integrands is completely open. It is fairly clear, however, that the Crouzeix–Raviart finite element method does not provide sufficient freedom to construct approximations which respect determinant constraints. One can easily construct examples (simply by trial and error) for which mean values of determinants are not preserved. This would, however, be necessary to extend the theory presented in this paper.

Acknowledgements

This work was initiated during a three-month research visit at the Hausdorff Centre for Mathematics in Bonn. The author was also supported by the EPSRC Critical Mass Programme 'New Frontiers in the Mathematics of Solids' (OxMOS).

The property that the Crouzeix–Raviart interpolant provides at the same time the L^2-projection for the gradient was pointed out to me by Carsten Carstensen.

REFERENCES

[1] R.A. Adams and J.F. Fournier. *Sobolev Spaces*. Elsevier Science, 2003.
[2] Y. Bai and Z.-P. Li. A truncation method for detecting singular minimizers involving the Lavrentiev

phenomenon. *Math. Models Methods Appl. Sci.*, 16(6):847–867, 2006.

[3] J. M. Ball. Discontinuous equilibrium solutions and cavitation in nonlinear elasticity. *Philos. Trans. Roy. Soc. London Ser. A*, 306(1496):557–611, 1982.

[4] J. M. Ball. Singularities and computation of minimizers for variational problems. In *Foundations of computational mathematics (Oxford, 1999)*, volume 284 of *London Math. Soc. Lecture Note Ser.*, pages 1–20. Cambridge Univ. Press, Cambridge, 2001.

[5] J. M. Ball and G. Knowles. A numerical method for detecting singular minimizers. *Numer. Math.*, 51(2):181–197, 1987.

[6] J. M. Ball and V. J. Mizel. One-dimensional variational problems whose minimizers do not satisfy the Euler-Lagrange equation. *Arch. Rational Mech. Anal.*, 90(4):325–388, 1985.

[7] D. Braess. *Finite Elements*. Cambridge University Press, Cambridge, third edition, 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.

[8] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.

[9] A. Buffa and C. Ortner. Compact embeddings of broken sobolev spaces and applications. to appear in IMA J. Numer. Anal.

[10] G. Buttazzo and V. J. Mizel. Interpretation of the Lavrentiev phenomenon by relaxation. *J. Funct. Anal.*, 110(2):434–460, 1992.

[11] C. Carstensen and C. Ortner. Gamma convergence in computational calculus of variations, 2009. OxMoS Preprint No. 17.

[12] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7(R-3):33–75, 1973.

[13] B. Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 1989.

[14] N. Dunford and J. T. Schwartz. *Linear operators. Part I*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[15] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[16] M. Foss. Examples of the Lavrentiev phenomenon with continuous Sobolev exponent dependence. *J. Convex Anal.*, 10(2):445–464, 2003.

[17] M. Foss, W. J. Hrusa, and V. J. Mizel. The Lavrentiev gap phenomenon in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 167(4):337–365, 2003.

[18] G. Knowles. Finite element approximation to singular minimizers, and applications to cavitation in non-linear elasticity. In *Differential equations and mathematical physics (Birmingham, Ala., 1986)*, volume 1285 of *Lecture Notes in Math.*, pages 236–247. Springer, Berlin, 1987.

[19] A. Lavrentiev. Sur quelques problémes du calcul des variations. *Ann. Mat. Pura Appl.*, 41:107–124, 1926.

[20] Z.-P. Li. Element removal method for singular minimizers in variational problems involving Lavrentiev phenomenon. *Proc. Roy. Soc. London Ser. A*, 439(1905):131–137, 1992.

[21] Z.-P. Li. Element removal method for singular minimizers in problems of hyperelasticity. *Math. Models Methods Appl. Sci.*, 5(3):387–399, 1995.

[22] B. Maniá. Sopra un esempio di Lavrentieff. *Boll. Un. Mat. Ital.*, 13:147–153, 1934.

[23] P. V. Negrón-Marrero. A numerical method for detecting singular minimizers of multidimensional problems in nonlinear elasticity. *Numer. Math.*, 58(2):135–144, 1990.

[24] C. Ortner and D. Praetorius. On the convergence of adaptive non-conforming finite element methods. Technical Report OxMOS Preprint no. 14, University of Oxford, 2008.

[25] A. Ten Eyck and A. Lew. Discontinuous Galerkin method for nonlinear elasticity. *Int. J. Numer. Methods Eng.*, 67:1204–1243, 2006.

[26] A. Visintin. Strong convergence results related to strict convexity. *Comm. Partial Differential Equations*, 9(5):439–466, 1984.