Journal of Convex Analysis
Volume 3 (1996), No. 2, 329–347

Numerical Approximation of Relaxed Variational Problems

T. Roubíček
*Mathematical Institute, Charles University, Sokolovská 83, CZ-186 00 Praha 8,
and Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic,
Pod vodárenskou věží 4, CZ-182 08 Praha 8, Czech Republic.
e-mail: roubicek@karlin.mff.cuni.cz*

Received January 8, 1995
Revised manuscript received July 3, 1995

Non-quasiconvex variational problems require, in general, a relaxation to ensure existence of their solutions. Here a relaxation by (generalized) Young functionals is used and then a finite-element method for a direct numerical approximation of the relaxed problems is developed. Beside a mere convergence, also some error estimates are analysed and a comparison with other methods is discussed.

Keywords : Vectorial variational problem, relaxation, generalized Young functionals, finite-element approximation.

1991 Mathematics Subject Classification: 65N30, 49M99, 73V25.

1. Introduction

In this paper we will treat the multidimensional vectorial variational problem

$$\text{minimize} \quad \Phi(u) \;=\; \int_\Omega \varphi(x,u(x),\nabla u(x))\,\mathrm{d}x \qquad \text{for} \;\; u\in W^{1,p}(\Omega;\mathbb{R}^m)\;, \tag{P}$$

where $\Omega\subset\mathbb{R}^n$ is a bounded Lipschitz domain, $\varphi:\Omega\times(\mathbb{R}^m\times\mathbb{R}^{mn})\to\mathbb{R}:(x,u,A)\mapsto\varphi(x,u,A)$ a Carathéodory function, $\min(m,n)<p<+\infty$, with $m,n\geq 1$, and $W^{1,p}(\Omega;\mathbb{R}^m)$ the Sobolev space of functions $u:\Omega\to\mathbb{R}^m$ with the norm $\|u\|_{W^{1,p}(\Omega;\mathbb{R}^m)}=\|u\|_{L^p(\Omega;\mathbb{R}^m)}+\|\nabla u\|_{L^p(\Omega;\mathbb{R}^{mn})}$. Throughout the paper, we will use the standard notation concerning function spaces C, C^0, $C^{0,\alpha}$, and L^p, standing respectively for continuous, continuous bounded, α-Hölder continuous, and p-integrable functions on the domains indicated. Moreover, $W^{\alpha,p}$ will denote the Sobolev (if α is an integer) or Sobolev-Slobodecki (if α is an non-integer) space.

We are especially interested in the case when $\varphi(x,u,\cdot):\mathbb{R}^{mn}\to\mathbb{R}$ is not quasiconvex (in the sense of Morrey [19]) so that the minimum in (P) is generally not achieved. In other words, (P) may have no solution in $W^{1,p}(\Omega;\mathbb{R}^m)$ and a need of its relaxation immediately follows. Here we will consider the relaxation not by weakly l.s.c. (=lower semicontinuous) envelope of Φ (which requires quasiconvexification of φ) but the relaxation by continuity,

ISSN 0944-6532 / $ 2.50 ⓒ Heldermann Verlag

using Young measures; cf. [2, 3, 9]. This relaxation has the advantages that it can avoid the generally difficult quasiconvexification and keeps a certain "limit information" about oscillations of the minimizing sequences for (P). Here we will use rather so-called generalized Young functionals, which are more appropriate especially in the case $\min(m, n) \geq 2$, and make the relaxation according to [25]. This is done in Sect. 2.

The aim of this paper is to propose a direct numerical approximation of the relaxed problem and to compare it with the usual approximation which uses standard finite-element discretization of the original problem (P) converging eventually to a solution of the relaxed problem, cf. [6, 7, 8, 11, 12, 13, 14, 17, 21]. The formulation of the approximate problem as well as the convergence analysis are performed in Sect. 3, while a more detailed analysis supporting the actual implementation is made in Sect. 4. Finally, some error estimates are obtained in Sect. 5 and a comparison is discussed in Sect. 6. Another approximation involving element-wise homogeneous Young measures has been recently proposed by Nicolaides and Walkington [20] but without any analysis. In view of some results by Dacorogna [15], the scheme from [20] is relevant if the quasiconvexification of $\varphi(x, u, \cdot)$ coincides with rank-1 convexification, while our scheme will basically apply if the quasiconvexification of $\varphi(x, u, \cdot)$ is polyconvex. In particular, for scalar ($m = 1$) or one-dimensional ($n = 1$) problems, the approximation theory presented here is fairly complete. The requirement of the polyconvex quasiconvexification is certainly very artificial and thus, in the case $\min(m, n) > 1$, our approximation theory is rather only a first step towards a very complicated problem; nevertheless, sometimes it is possible to verify experimentally that the polyconvex and the quasiconvex envelope do not differ much from each other, cf. [16].

2. Generalized Young functionals and relaxation of (P)

Let us introduce briefly some notation and some results from [25]. Following the original idea by L.C. Young [28], the generalized Young measures (or, more precisely, functionals) will be considered as certain linear continuous functionals on a suitable space H of Carathéodory integrands $\Omega \times \mathbb{R}^{mn} \to \mathbb{R}$, cf. also [26, 27]. Though there is quite a large freedom in the choice of H, we will take one particular and enough large space, namely

$$H = \left\{ h_0 + \sum_{s=0}^{\min(m,n)} g_s \otimes \mathrm{adj}_s;\ h_0 \in G \otimes V,\ g_s \in L^{p/(p-s)}(\Omega; \mathbb{R}^{\sigma(s)}) \right\}, \tag{2.1}$$

where $G \supset C^0(\bar{\Omega})$ is a separable subspace of $L^\infty(\Omega)$ closed under point-wise multiplication, and $\mathrm{adj}_s : \mathbb{R}^{mn} \to \mathbb{R}^{\sigma(s)}$ assigns each matrix $A \in \mathbb{R}^{mn}$ its cofactors of order s (i.e. the determinants of all $s \times s$-submatrices) with the convention $\sigma(0) = 1$ and $\mathrm{adj}_0 = 1$; clearly $\sigma(s) = \binom{m}{s}\binom{n}{s} = \frac{m!}{s!(m-s)!}\frac{n!}{s!(n-s)!}$. Of course, $g_s \otimes \mathrm{adj}_s$ abbreviates $\sum_{l=1}^{\sigma(s)} [g_s]_l \otimes [\mathrm{adj}_s]_l$ with "$\otimes$" denoting the usual tensorial product of functions, i.e. $[g \otimes v](x, A) = g(x)v(A)$. Obviously, $\mathrm{adj}_1 \cong \mathrm{id} : \mathbb{R}^{mn} \to \mathbb{R}^{mn}$ is the identity on $\mathbb{R}^{mn}$.

Furthermore, V in (2.1) is supposed to satisfy

$$V \text{ a linear subspace of } C(\mathbb{R}^{mn}), \tag{2.2a}$$

$$\forall v \in V: \sup_{A \in \mathbb{R}^{mn}} |v(A)|/(1+|A|^p) < +\infty, \tag{2.2b}$$

$$\forall\, 1 \le s \le \min(m,n): \ \mathrm{adj}_s \in V^{\sigma(s)}, \tag{2.2c}$$

$$\{A \mapsto v(A)/(1+|A|^p);\ v \in V\} \text{ is separable}, \tag{2.2d}$$

$$\forall v \in V: \ v^{\#} \ne -\infty \ \Rightarrow\ v^{\#} \text{ is polyconvex}, \tag{2.2e}$$

where $v^{\#}(A) = \inf_{u \in W_0^{1,\infty}(\Omega,\mathbb{R}^n)} \int_\Omega v(A+\nabla u(x))\, \mathrm{d}x$ is the quasiconvex envelope of v. Recall still that v is called polyconvex [1] if it can be expressed as a convex function of $(\mathrm{adj}_s A)_{s=1}^{\min(m,n)}$. For examples of V satisfying (2.2) we refer to [24]. Note that the only nontrivial condition (2.2e) is void if $\min(m,n) = 1$.

We can introduce a (possibly only semi-) norm on H from (2.1) by

$$\|h\|_H = \inf_{h_0 + \sum_s g_s \otimes \mathrm{adj}_s = h} \left(\operatorname*{ess\,sup}_{x \in \Omega} \sup_{A \in \mathbb{R}^{mn}} \frac{|h_0(x,A)|}{1+|A|^p} + \sum_{s=0}^{\min(m,n)} \|g_s\|_{L^{p/(p-s)}(\Omega;\mathbb{R}^{\sigma(s)})} \right). \tag{2.3}$$

This norm makes H separable because of (2.2d) and of the separability of G. Therefore the weak* topology of the dual space H^* is metrizable when restricted to bounded subsets. Furthermore, let us suppose that H contains a coercive integrand in the sense

$$\exists h_\mathrm{c} \in H\ \forall (x,A) \in \Omega \times \mathbb{R}^{mn}: \quad h_\mathrm{c}(x,A) \ge |A|^p, \tag{2.4}$$

and imbed $L^p(\Omega;\mathbb{R}^{mn})$ (strong,weak*)-continuously into H^* via the imbedding $i_H : L^p(\Omega;\mathbb{R}^{mn}) \to H^*$ defined by $\langle i_H(y), h\rangle = \int_\Omega h(x,y(x))\, \mathrm{d}x$ with $h \in H$. The elements of the set

$$Y_H^p(\Omega;\mathbb{R}^{mn}) \ = \ \text{w}^*\text{-cl}_{H^*} i_H(L^p(\Omega;\mathbb{R}^{mn})) \tag{2.5}$$

will be addressed as generalized Young functionals; it referes to the fact that for $H = L^1(0,T;C_0(\mathbb{R}^{mn}))$ the elements of $Y_H^p(\Omega;\mathbb{R}^{mn})$ attainable by bounded sequences can be identified with the classical Young measures. It was proved in [26, 27] that $Y_H^p(\Omega;\mathbb{R}^{mn})$ is a convex, closed, weakly* σ-compact, locally compact, and also locally sequentially compact subset of H^*.

Let us also denote the set of all so-called gradient generalized Young functionals by

$$G_H^p(\Omega;\mathbb{R}^{mn}) \ = \Big\{ \eta \in Y_H^p(\Omega;\mathbb{R}^{mn}); \tag{2.6}$$
$$\exists \text{ a sequence } \{u_k\}_{k\in\mathbb{N}} \subset W^{1,p}(\Omega;\mathbb{R}^m): \ \text{w}^* \text{ - } \lim_{k\to\infty} i_H(\nabla u_k) = \eta \Big\}\ .$$

Let us remark that $G_H^p(\Omega;\mathbb{R}^{mn})$ is *not* convex when $\min(m,n) \ge 2$. Indeed, this follows from [25, Corollary 2.1] when one realizes that, e.g., for $h = 1 \otimes v \in H$ with v quasiaffine but nonconvex, the functional $\Phi(u) = \int_\Omega h(x,\nabla u(x))\, \mathrm{d}x$ is weakly l.s.c.

on $W^{1,p}(\Omega;\mathbb{R}^m)$ but nonconvex, which can be seen if one makes a convex combination of $u_1, u_2 \in W^{1,p}(\Omega;\mathbb{R}^m)$ in the form $u_i(x) = A_i x$ with suitable A_i such that $\mathrm{Rank}(A_1 - A_2) \geq 2$.

For $k \in \mathbb{N}$ we define the bilinear mapping $(h,\eta) \mapsto h \bullet \eta : H^k \times H^* \to [G^*]^k$ by $\langle h \bullet \eta, g\rangle = \langle \eta, g \cdot h\rangle$ for any $g \in G^k$. We can understand $h \bullet \eta$ as a substitution of η into an $\mathbb{R}^k$-valued integrand h. It can be shown [26, 27] that $\eta \mapsto h \bullet \eta$ is an affine continuous extension of the Nemytskii mapping $y \mapsto h \circ y$.

The important property of a particular $\eta \in Y^p_H(\Omega;\mathbb{R}^{mn})$ is its possible "non-concentration": we say that $\eta \in Y^p_H(\Omega;\mathbb{R}^{mn})$ is p-nonconcentrating if there is a sequence $\{y_k\}_{k\in\mathbb{N}}$ such that η =w*-$\lim_{k\to\infty} i_H(y_k)$ and that the set $\{|y_k|^p;\ k \in \mathbb{N}\}$ is relatively weakly compact in $L^1(\Omega)$. Every such η admits a representation in terms of the classical Young measures, that means there is a (generally not uniquely determined) weakly* measurable mapping $\nu : x \mapsto \nu_x$, called a Young measure, from Ω to regular probability measures on $\mathbb{R}^{mn}$ such that

$$\forall h \in H : \quad h \bullet \eta \ = \ \langle \nu, h\rangle \qquad \text{in the sense of} \quad L^1(\Omega)\ , \tag{2.7}$$

where $\langle \nu, h\rangle$ abbreviates, as usual, the function $x \mapsto \langle \nu_x, h(x,\cdot)\rangle = \int_{\mathbb{R}^{mn}} h(x,A)\nu_x(\mathrm{d}A)$; cf. [25, Remark 4.2].

We will denote the extended potential by $\bar\Phi : W^{1,p}(\Omega;\mathbb{R}^m) \times G^p_H(\Omega;\mathbb{R}^{mn}) \to \mathbb{R}$ and always suppose that

$$\forall (u,\eta) \in \text{w*-cl}\, i(W^{1,p}(\Omega;\mathbb{R}^m)) : \ \bar\Phi(u,\eta) \ = \ \liminf_{i(u)\to(u,\eta)\ \text{weakly*}} \Phi(u)\ , \tag{2.8}$$

where $i : W^{1,p}(\Omega;\mathbb{R}^m) \to W^{1,p}(\Omega;\mathbb{R}^m) \times H^*$ is defined by $i(u) = (u, i_H(\nabla u))$. Then we can define the relaxed variational problem:

$$\left.\begin{array}{ll} \text{minimize} & \bar\Phi(u,\eta) \\ \text{subject to} & u \in W^{1,p}(\Omega;\mathbb{R}^m),\ \eta \in G^p_H(\Omega;\mathbb{R}^{mn})\ , \\ & \nabla u \ = \ (1 \otimes \mathrm{id}) \bullet \eta\ . \end{array}\right\} \tag{RP}$$

For all our considerations we will also need the coercivity of (P), this means

$$\exists a \in L^1(\Omega)\ \exists b,c > 0 : \quad \varphi(x,u,A) \geq a(x) + b|u|^p + c|A|^p\ . \tag{2.9}$$

Proposition 2.1. *Under the conditions* (2.1), (2.2a-d), (2.3), (2.4), (2.8) *and* (2.9), (RP) *is a proper relaxation of* (P) *in the sense that*

a) (RP) *has always a solution,*
b) $\min(\mathrm{RP}) = \inf(\mathrm{P})$,
c) *every cluster point of every minimizing sequence of* (P) *solves* (RP)*, and*
d) *every solution of* (RP) *can be attained by a sequence minimizing* (P).

Proof. See [25, Proposition 2.2 with Example 4.1] and realize that H is here supposed separable, which allows us to work in terms of sequences in $Y^p_H(\Omega;\mathbb{R}^{mn})$. □

3. Approximation of (RP)

We want to approximate (RP) directly. For this reason, we make a discretization of the admissible subset of $W^{1,p}(\Omega;\mathbb{R}^m)\times G^p_H(\Omega;\mathbb{R}^{mn})$ on which (RP) is defined and change the structure of the constraints to make them readily implementable.
We will use the finite-element technique, supposing, for simplicity, that Ω is polyhedral. For any $d>0$, let $\mathcal{T}_d$ be a triangulation of Ω consisting of elements (=simplexes) of the diameter not exceeding d. For $d\ge d'>0$, we suppose that $\mathcal{T}_d\subset\mathcal{T}_{d'}$, this means $\mathcal{T}_{d'}$ is a refinement of $\mathcal{T}_d$. Furthermore, we suppose that the family $\{\mathcal{T}_d\}_{d>0}$ is regular in the sense of [10], i.e. there is $\varepsilon>0$ such that each element of each $\mathcal{T}_d$ contains a ball of radius εd. Our aim is to approximate $W^{1,p}(\Omega;\mathbb{R}^m)$ by element-wise affine functions and $Y^p_H(\Omega;\mathbb{R}^{mn})$ by element-wise constant (=homogeneous) generalized Young functionals. As for the former discretization, we put simply

$$U_d \;=\; \left\{u\in W^{1,p}(\Omega;\mathbb{R}^m);\ \forall E\in\mathcal{T}_d:\ u|_E \text{ affine}\right\}\ . \tag{3.1}$$

The latter discretization is performed by the projection operator $P_d: H\to H$ defined by

$$[P_d h](x,y) \;=\; \frac{1}{\mathrm{meas}(E)}\int_E h(\tilde{x},y)\,\mathrm{d}\tilde{x} \quad \text{ if } \ x\in E\in\mathcal{T}_d\ . \tag{3.2}$$

Note that this projector is continuous with respect to the norm (2.3), and so is the adjoint projector $P^*_d: H^*\to H^*$. We put

$$Y_d \;=\; P^*_d Y^p_H(\Omega;\mathbb{R}^{mn})\ . \tag{3.3}$$

Proposition 3.1. *Y_d is convex, weakly* σ-compact, weakly* locally compact subset of $Y^p_H(\Omega;\mathbb{R}^{mn})$. If V is also finite-dimensional, then Y_d is strongly locally compact. Moreover, for all $d\ge d'>0$, we have $Y_d\subset Y_{d'}$.*

Proof. The (quite nontrivial) inclusion $Y_d\subset Y^p_H(\Omega;\mathbb{R}^{mn})$ was proved in [24, Lemma 3.2 with Sect. 4.1].
For the convexity and σ-compactness of Y_d we refer to [26, Section 3.5].
For the local compactness we refer to [27, proof of Theorem 2.1], if one realizes that $\int_\Omega h_{\mathrm{c}}(x,y(x))\,\mathrm{d}x\to+\infty$ for $\|y\|_{L^p(\Omega;\mathbb{R}^{mn})}\to+\infty$ for h_{c} from (2.4). Also note that P_dH with H from (2.1) and P_d from (3.2) is finite-dimensional provided V is so.
The monotonicity of the sequence $\{Y_d\}$ follows, as in [26, Section 3.5], from $P_d=P_d\circ P_{d'}$ which holds whenever $d\ge d'>0$ as a consequence of $\mathcal{T}_d\subset\mathcal{T}_{d'}$. $\square$

Supposing, for simplicity, that the extended potential $\bar{\Phi}(u,\eta)$ can be evaluated exactly for every u and η which are element-wise affine and constant, respectively, we can define our approximate problem:

$$\left.\begin{array}{ll} \text{minimize} & \bar{\Phi}(u,\eta) \\ \text{subject to} & u\in U_d,\ \eta\in Y_d, \\ & \mathrm{adj}_s\nabla u \;=\; (1\otimes\mathrm{adj}_s)\bullet\eta \quad \text{for all } 1\le s\le\min(m,n). \end{array}\right\} \qquad (\mathrm{RP}_d)$$

For the convergence analysis, we must inevitably strenghten (2.8). The realistic (cf. Remark 3.3 below) though a bit restrictive assumption seems to be the following:

$$\forall r \in \mathbb{R}^+ : \quad \bar{\Phi} : B_r \times G_H^p(\Omega; \mathbb{R}^{mn}) \to \mathbb{R} \ \text{ is weakly}\times\text{weakly* continuous} \ , \tag{3.4}$$

where $B_r = \{u \in W^{1,p}(\Omega; \mathbb{R}^m);\ \|u\|_{W^{1,p}(\Omega;\mathbb{R}^m)} \le r\}$. Note that (3.4) with (2.8) and (2.9) make $\bar{\Phi}$ the extension of Φ in the sense that $\bar{\Phi}(u, i_H(\nabla u)) = \Phi(u)$ for every $u \in W^{1,p}(\Omega; \mathbb{R}^m)$.

Proposition 3.2. *Under the conditions* (2.1)–(2.4), (2.8), *and* (2.9), (RP_d) *has a solution* (u_d, η_d). *Moreover, there is a subsequence of* $\{(u_d, \eta_d)\}_{d>0}$ *converging weakly* with* $d \to 0$ *and, if* (3.4) *is also valid, the limit of each such subsequence solves* (RP).

Proof. First, let us demonstrate that every $(u, \eta) \in U_d \times Y_d$ satisfying $\mathrm{adj}_s \nabla u = (1 \otimes \mathrm{adj}_s) \bullet \eta$ for all $s = 1, ..., \min(m, n)$ is admissible for (RP). Indeed, it suffices to show $\eta \in G_H^p(\Omega; \mathbb{R}^{mn})$, which, however, follows from [24, Corollary 4.], using the fact that $\eta \in Y_d$ is p-nonconcentrating thanks to the coercivity (2.9), cf. [25]. Therefore, we have proved $\inf(\mathrm{RP}_d) \ge \min(\mathrm{RP})$.

The coercivity (2.9) of φ implies that the set $M_d(c) = \{(u, \eta) \in U_d \times Y_d;\ \bar{\Phi}(u, \eta) \le c\ \&\ \mathrm{adj}_s \nabla u = (1 \otimes \mathrm{adj}_s) \bullet \eta,\ s = 1, ..., \min(m, n)\}$ is contained in w*-cl $i(B_r)$ with B_r a sufficiently large ball in $W^{1,p}(\Omega; \mathbb{R}^m)$. Then $M_d(c)$ is weakly* compact because w*-cl $i(B_r)$ is weakly* compact, $U_d \times Y_d$ is closed in w*-cl $i(B_r)$, $\bar{\Phi}$ is weakly* l.s.c., and the mapping $(u, \eta) \mapsto \mathrm{adj}_s \nabla u - (1 \otimes \mathrm{adj}_s) \bullet \eta : W^{1,p}(\Omega; \mathbb{R}^m) \times Y_H^p(\Omega; \mathbb{R}^{mn}) \to L^{p/s}(\Omega; \mathbb{R}^{\sigma(s)})$ is weakly* continuous; cf. [1, 22]. Also, $M_d(c)$ is surely non-empty for c sufficiently large, e.g. for $c \ge \Phi(0)$ because $(0, i_H(0))$ is certainly admissible for (RP_d). As $\bar{\Phi}$ is weakly* l.s.c., the existence of a solution to (RP_d) as well as the existence of a weakly* converging subsequence of these solutions follow by the standard compactness arguments.

Now we want to prove that every $(u, \eta) \in W^{1,p}(\Omega; \mathbb{R}^m) \times G_H^p(\Omega; \mathbb{R}^{mn})$ satisfying $\nabla u = (1 \otimes \mathrm{id}) \bullet \eta$ can be approximated by suitable admissible pairs for (RP_d) when $d \to 0$. First, there is a bounded sequence $\{u_k\} \subset W^{1,p}(\Omega; \mathbb{R}^m)$ such that $i_H(\nabla u_k) \to \eta$ weakly* in H^*. Also, $\{u_k\}$ converges weakly (possibly as a subsequence only) to some $\tilde{u} \in W^{1,p}(\Omega; \mathbb{R}^m)$. Then $i(u_k - \tilde{u} + u) \to (u, \eta)$. By mollifying (if necessary) each u_k, we can even assume $\{u_k\} \in C^\infty(\bar{\Omega})$. Let $\Pi_d u_k \in U_d$ be the linear interpolant of u_k on the triangulation $\mathcal{T}_d$. For k fixed and $d \to 0$, we have $\Pi_d u_k \to u_k$ strongly in $W^{1,p}(\Omega; \mathbb{R}^m)$ because of the regularity of u_k. Therefore $i_H(\nabla \Pi_d u_k) \to i_H(\nabla u_k)$ weakly* in H^* because $\int_\Omega [h(x, \nabla \Pi_d u_k(x)) - h(x, \nabla u_k(x))]\, \mathrm{d}x \to 0$ for any $h \in H$ as a consequence of the continuity of the Nemytskii mapping $y \mapsto h \circ y : L^p(\Omega; \mathbb{R}^{mn}) \to L^1(\Omega)$. At the same time, the pair $i(\Pi_d u_k) = (\Pi_d u_k, i_H(\nabla \Pi_d u_k))$ is admissible for (RP_d) because $(1 \otimes \mathrm{adj}_s) \bullet i_H(y) = \mathrm{adj}_s\, y$ for any $y \in L^p(\Omega; \mathbb{R}^{mn})$ and, in particular, also for $y = \nabla \Pi_d u_k$, and because $\Pi_d u_k \in U_d$ hence $i_H(\nabla \Pi_d u_k) \in Y_d$. Without any loss of generality, the sequence $\{u_k\}$ can be considered bounded in $W^{1,p}(\Omega; \mathbb{R}^m)$. Then also $\{\Pi_d u_k\} \subset B_r$ for r large enough and $\{i_H(\nabla \Pi_d u_k)\}$ is bounded in H^*. Realizing the metrizability of the weak* topology on $i(B_r)$ (recall that H is separable), we can select a subsequence such that $\Pi_d u_k \to u$ weakly in $W^{1,p}(\Omega; \mathbb{R}^m)$ and $i_H(\nabla \Pi_d u_k) \to \eta$ weakly* in H^*. Let us

suppose that (u,η) is a solution to (RP). We have $\bar{\Phi}(\Pi_d u_k, i_H(\nabla\Pi_d u_k)) \geq \min(\mathrm{RP}_d) \geq \min(\mathrm{RP})$. As $\bar{\Phi} : B_r \times G^p_H(\Omega;\mathbb{R}^{mn}) \to \mathbb{R}$ is weakly$\times$weakly* continuous, we obtain $\bar{\Phi}(\Pi_d u_k, i_H(\nabla\Pi_d u_k)) \to \bar{\Phi}(u,\eta) = \min(\mathrm{RP})$, which shows $\min(\mathrm{RP}_d) \to \min(\mathrm{RP})$ for $d \to 0$.
The rest of the assertion follows immediately by the standard compactness arguments, taking into account the coercivity of the problem. □

An example for $\bar{\Phi}$ satisfying (3.4) is

$$\bar{\Phi}(u,\eta) \;=\; \langle \eta, \varphi\circ u\rangle \;, \tag{3.5}$$

where $[\varphi\circ u](x,A) = \varphi(x,u(x),A)$, provided

$$\begin{aligned} &\forall u \in W^{1,p}(\Omega;\mathbb{R}^m): \;\; \varphi\circ u \in H \quad \& \\ &\forall r \in \mathbb{R}^+: \quad u \mapsto \varphi\circ u : (B_r,\mathrm{weak}) \to (H,\mathrm{norm}) \quad \text{is continuous}\;. \end{aligned} \tag{3.6}$$

Remark 3.3. Sometimes it may happen that for $\{\Pi_d u_k\}$, constructed in the proof of Proposition 3.2, $i_H(\nabla\Pi_d u_k)$ converges to η not only weakly* in H^* but even in the standard dual norm $\|\cdot\|_{H^*}$. Then Proposition 3.2 remains valid if (3.4) is weakened, so that Φ is to be continuous only from $(B_r,$weak$) \times (G^p_H(\Omega;\mathbb{R}^{mn}),$norm$) \to \mathbb{R}$. This appears in a trivial but important case $V = \{0\}$ and $\min(m,n) = 1$. Then $Y^p_H(\Omega;\mathbb{R}^{mn}) \cong L^p(\Omega;\mathbb{R}^{mn})$ and our theory basically coincides with the standard finite-element approximation of a convexified problem: minimize $\Phi^{**}(u) = \int_\Omega \varphi^{**}(x,u(x),\nabla u(x))\,\mathrm{d}x$ with $\varphi^{**}(x,u,\cdot)$ being a convex envelope of $\varphi(x,u,\cdot)$. Typically $\varphi^{**}\circ u \notin H$ and $\bar{\Phi}(u,y) = \int_\Omega \varphi^{**}(x,u(x),y(x))\,\mathrm{d}x$ is not weakly continuous but only norm continuous. More generally, the same effect is for $H = L^{p/(p-1)}(\Omega;\mathbb{R}^{mn}) \otimes \{\mathrm{id}\}$ with $\min(m,n)$ arbitrary. Then $\bar{\Phi}(u,y) = \int_\Omega \varphi^{\#}(x,u(x),y(x))\,\mathrm{d}x$. However, e.g. for the classical L^p-Young measures (i.e. for $H = L^1(\Omega;C_0(\mathbb{R}^{mn}))$) it is easy to see that the remainder $Y^p_H(\Omega;\mathbb{R}^{mn}) \setminus i_H(L^p(\Omega;\mathbb{R}^{mn}))$ cannot be reached from $L^p(\Omega;\mathbb{R}^{mn})$ with respect to the norm topology of H^* and therefore (3.4) cannot be weakened in general.

4. Implementation of a solution to (RP$_d$).

In principle, V can be finite-dimensional and then $P^*_d(H^*)$ is also finite-dimensional and therefore (RP$_d$), being defined on a subset of the finite-dimensional linear space $U_d \times P^*_d(H^*)$, can be implemented directly on digital computers. In the opposite case when V is infinite-dimensional, we must (and may) use special properties of the solutions of (RP$_d$) because our task is not to implement every admissible pair (u,η) for (RP$_d$) but only a subset of admissible pairs which contains (some of) the solutions to (RP$_d$). The basic tool here consists in enough selective and informative optimality conditions for (RP$_d$). Such conditions, of the Euler-Weierstrass type, were generalized for relaxed vectorial variational problems in [25]. Here we will adapt them for our (semi)discretized problem (RP$_d$). Throughout this section $d > 0$ will be fixed.

Let us introduce the "discrete" Hamiltonian $\mathcal{H}^d_{u,\lambda} \in H_d \equiv P_d(H)$ defined, for $\lambda = (\lambda_1, ..., \lambda_{\min(m,n)})$ with $\lambda_s \in L^\infty(\Omega;\mathbb{R}^{\sigma(s)})$ element-wise constant, by

$$\left[\mathcal{H}^d_{u,\lambda}\right](x,y) \;=\; -P_d(\varphi\circ u) + \sum_{s=1}^{\min(m,n)} \lambda_s \otimes \mathrm{adj}_s\ . \tag{4.1}$$

Moreover, we will assume that

$$\forall u \in W^{1,p}(\Omega;\mathbb{R}^m):\ \left\{\left[\frac{\partial\varphi}{\partial u}\circ u\right]\cdot\tilde u;\ \tilde u \in B_1\right\} \subset H \quad \text{bounded}\ , \tag{4.2a}$$

$$u \mapsto \left[\frac{\partial\varphi}{\partial u}\circ u\right]\cdot\tilde u : (W^{1,p}(\Omega;\mathbb{R}^m),\mathrm{weak}) \to (H,\mathrm{norm}) \quad \text{equi-continuous} \tag{4.2b}$$

with respect to $\tilde u$ from the unit ball B_1 of $W^{1,p}(\Omega;\mathbb{R}^m)$.

Proposition 4.1. *Let $p > \min(m,n)$, (2.9), (3.6) and (4.2) be valid, and (u,η) be a solution to* (RP$_d$). *Then there exists $\lambda = (\lambda_1, ..., \lambda_{\min(m,n)})$ with $\lambda_s \in L^\infty(\Omega;\mathbb{R}^{\sigma(s)})$ element-wise constant such that the following integral identity is valid:*

$$\int_\Omega \left(\sum_{s=1}^{\min(m,n)} \lambda_s(x)\cdot\frac{\partial\,\mathrm{adj}_s}{\partial A}(\nabla u(x))\cdot\nabla\tilde u(x) \;-\; \left[(\frac{\partial\varphi}{\partial u}\circ u)\bullet\eta\right](x)\ \tilde u(x)\right)\mathrm{d}x \;=\; 0 \tag{4.3}$$

for all $\tilde u \in U_d$ and the following "Weierstrass-type" maximum principle holds:

$$\left[\mathcal{H}^d_{u,\lambda}\bullet\eta\right](x) \;=\; \max_{A\in\mathbb{R}^{mn}} \mathcal{H}^d_{u,\lambda}(x,A) \quad \text{for a.a.}\ \ x \in \Omega\ . \tag{4.4}$$

Sketch of the proof. First, let us notice that (RP$_d$) has the form

$$\left.\begin{array}{ll} \text{minimize} & \bar\Phi(u,\eta) \\ \text{subject to} & u \in U_d,\ \eta \in Y_d\ \subset P_d^* H^*, \\ & N(u) \;=\; L\eta, \end{array}\right\} \tag{RP$'_d$}$$

with U_d and $P_d^*H^*$ Banach spaces, Y_d closed convex, $N\colon U_d \to \prod_{s=1}^{\min(m,n)} L_d(\Omega;\mathbb{R}^{\sigma(s)}) : u \mapsto (\mathrm{adj}_s\nabla u)_{s=1}^{\min(m,n)}$ smooth and (for $\min(m,n) \geq 2$) nonlinear with $L_d(\Omega;\mathbb{R}^{\sigma(s)}) \subset L^\infty(\Omega;\mathbb{R}^{\sigma(s)})$ denoting the space of element-wise constant functions $\Omega \to \mathbb{R}^{\sigma(s)}$, and $L : P_d^*(H^*) \to \prod_{s=1}^{\min(m,n)} L_d(\Omega;\mathbb{R}^{\sigma(s)}) : \eta \mapsto \Big((1\otimes\mathrm{adj}_s)\bullet\eta\Big)_{s=1}^{\min(m,n)}$ linear and continuous. Also, (4.2) make $\bar\Phi$ continuously Fréchet differentiable with $\bar\Phi'(u,\eta) = \Big((\frac{\partial\varphi}{\partial u}\circ u)\bullet\eta, \varphi\circ u\Big) \in U_d^* \times H$. Moreover, as in [25, Lemma 3.3], we can prove that L is surjective on Y_d. Identifying $L_d(\Omega;\mathbb{R}^{\sigma(s)})$ with its own dual, we can use the theorem by Zowe and

Kurczyusz [29], which gives the following necessary optimality conditions for (RP'_d): there is $\lambda = (\lambda_1, ..., \lambda_{\min(m,n)})$ with $\lambda_s \in L_d(\Omega; \mathbb{R}^{\sigma(s)})$ such that

$$[N'(u)]^*\lambda + \bar{\Phi}'_u(u,\eta) \;=\; 0 \;\in\; U_d^* \;, \tag{4.5}$$

$$\forall \tilde{\eta} \in Y_d : \quad \left\langle L^*\lambda - \bar{\Phi}'_\eta(u,\eta), \eta - \tilde{\eta} \right\rangle \geq 0 \;. \tag{4.6}$$

After routine calculations, (4.5) gives (4.3). As for (4.6), notice that $L^*\lambda - \bar{\Phi}'_\eta(u,\eta) = \sum_{s=1}^{\min(m,n)} \lambda_s \otimes \mathrm{adj}_s - \varphi \circ u$ and $\eta = P_d^*\eta$ and $\tilde{\eta} = P_d^*\tilde{\eta}$. Then $\left\langle L^*\lambda - \bar{\Phi}'_\eta(u,\eta), \eta - \tilde{\eta} \right\rangle = \left\langle \eta - \tilde{\eta}, P_d(\sum_{s=1}^{\min(m,n)} \lambda_s \otimes \mathrm{adj}_s - \varphi \circ u) \right\rangle = \left\langle \eta - \tilde{\eta}, \mathcal{H}^d_{u,\lambda} \right\rangle$. In other words, (4.6) means just $\left\langle \eta, \mathcal{H}^d_{u,\lambda} \right\rangle = \max_{\tilde{\eta} \in Y_d} \left\langle \tilde{\eta}, \mathcal{H}^d_{u,\lambda} \right\rangle$. The Weierstrass-type maximum principle (4.4) follows from this "integral" maximum principle by a usual localization technique, using also the assumption $p > \min(m,n)$; cf. [25, Theorem 3.2]. $\square$

So far we have the linear combination $\sum_{k=1}^K \theta_k \eta_k$ defined for $\eta_k \in H^*$ and $\theta_k \in \mathbb{R}$ only. However, we can generalize it naturally for $\theta_k \in G \subset L^\infty(\Omega)$ by means of the identity:

$$\forall h \in H : \quad \left\langle \sum_{k=1}^K \theta_k \eta_k \;,\; h \right\rangle \;=\; \sum_{k=1}^K \langle h \bullet \eta_k, \theta_k \rangle \;. \tag{4.7}$$

Of course, the left-hand-side duality is between H^* and H while the right-hand-side one is between G^* and G. Note that this extended definition has the previous meaning provided θ_k are constant on Ω.

Corollary 4.2. *Let the assumptions of* Proposition 4.1 *be satisfied and, for every* $x \in \Omega$, $\lambda = (\lambda_1, ...\lambda_{\min(m,n)})$ *and* $u \in U_d$, *let* $\left[\mathcal{H}^d_{u,\lambda}\right](x, \cdot) : \mathbb{R}^{mn} \to \mathbb{R}$ *attain the maximum at no more than* K *points in* $\mathbb{R}^{mn}$, *and let* (u,η) *solve* (RP_d). *Then* η *can be written in the form*

$$\eta \;=\; \sum_{k=1}^K \theta_k i_H(y_k) \;, \quad \theta_k \geq 0 \;, \quad \sum_{k=1}^K \theta_k = 1 \;, \tag{4.8}$$

with $\theta_k \in L^\infty(\Omega)$ *and* $y_k \in L^\infty(\Omega; \mathbb{R}^{mn})$ *element-wise constant. In other words,* η *admits a Young measure representation* (2.7) *with* ν *element-wise constant and* ν_x *being a convex combination of at most* K *Dirac measures.*

Proof. Let (u,η) be a solution to (RP_d). By the coercivity of the problem, η is p-nonconcentrating (cf. [25]), and thus it admits a Young-measure representation $\eta \cong \nu = \{\nu_x\}_{x \in \Omega}$, cf. (2.7), such that $x \mapsto \nu_x$ is element-wise constant. From the maximum principle (4.4) we can see that $\int_{\mathbb{R}^{mn}} \mathcal{H}^d_{u,\lambda}(x,A)\nu_x(\,\mathrm{d}A) = \max_{A \in \mathbb{R}^{mn}} \mathcal{H}^d_{u,\lambda}(x,A)$. In particular, the probability measure ν_x must be supported at the set on which $\mathcal{H}^d_{u,\lambda}(x,\cdot)$ attains it maximum over $\mathbb{R}^{mn}$ which is supposed to consist from no more than K points, let us denote them by $y_k(x) \in \mathbb{R}^{mn}$, $k = 1, ..., K$. Moreover, these points are independent of $x \in E$ for each particular element $E \in \mathcal{T}_d$ because $\mathcal{H}^d_{u,\lambda}(\cdot, A)$ is element-wise constant.

In other words, $\nu_x = \sum_{k=1}^{K} \theta_k(x)\delta_{y_k(x)}$ and both θ_k and y_k can be assumed element-wise constant, where δ_A denotes the Dirac measure supported at $A \in \mathbb{R}^{mn}$. As ν_x is a probability measure, $\theta_k \geq 0$ and $\sum_{k=1}^{K} \theta_k = 1$. Then (4.8) follows easily from the formulas (2.7) and (4.7) via the identity: $h \bullet \eta = \langle \nu, h\rangle = \sum_{k=1}^{K} \theta_k h(y_k) = \sum_{k=1}^{K} \theta_k(h \bullet i_H(y_k)) = h \bullet \left(\sum_{k=1}^{K} \theta_k i_H(y_k)\right)$, which holds for any $h \in H$. $\square$

Under the assumptions of Corollary 4.2 it is clear that every solution to (RP$_d$) can be implemented on digital computers because the number of elements in $\mathcal{T}_d$ as well as the maximal number of Dirac measures needed on each of them are finite. For an example of the implementation in the scalar case we refer to [23, Sect. 3].
However, it should be remarked that, especially in the case $\min(m,n) \geq 2$, the condition in Corollary 4.2 is not easy to be verified. On the other hand, mostly it is quite satisfactory to implement at least one of the solutions to (RP$_d$) because, even if we are actually able to implement all solutions, the nonlinear-programming algorithms we have eventually to use can normally find only one of them. From this standpoint, the following assertion makes a satisfactory basis for implementation of the problem (RP$_d$).

Corollary 4.3. *Let the assumptions of* Proposition 4.1 *be satisfied. Then there always exists at least one solution* (u,η) *to* (RP$_d$) *such that* η *is in the form* (4.8) *with* $\theta_k \in L^\infty(\Omega)$ *and* $y_k \in L^\infty(\Omega;\mathbb{R}^{mn})$ *element-wise constant and*

$$K \;=\; 1 + \sum_{s=1}^{\min(m,n)} \sigma(s)\ . \tag{4.9}$$

Sketch of the proof. Take (u_0,η_0) to be a solution to (RP$_d$) and $\lambda = (\lambda_1, ..., \lambda_{\min(m,n)})$ the corresponding Lagrange multipliers. Every pair $(u_0,\eta) \in U_d \times Y_d$ satisfying $(1 \otimes \mathrm{adj}_s) \bullet \eta = \mathrm{adj}_s \nabla u_0$ for $s = 1, ..., \min(m,n)$ and the maximum principle (4.4) with $u = u_0$ solves (RP$_d$). Indeed, such (u_0,η) is obviously an admissible pair for (RP$_d$) and also $\bar{\Phi}(u_0,\eta) = \min(\mathrm{RP}_d)$ because

$$\begin{aligned}
\langle \eta, \varphi \circ u_0\rangle \quad &= \quad \langle P_d^* \eta, \varphi \circ u_0\rangle \;=\; \langle \eta, P_d(\varphi \circ u_0)\rangle \\
&= \quad \langle \eta, \sum_{s=1}^{\min(m,n)} \lambda_s \otimes \mathrm{adj}_s - \mathcal{H}^d_{u_0,\lambda}\rangle \;=\; \langle \eta_0, \sum_{s=1}^{\min(m,n)} \lambda_s \otimes \mathrm{adj}_s - \mathcal{H}^d_{u_0,\lambda}\rangle \\
&= \quad \langle \eta_0, P_d(\varphi \circ u_0)\rangle \;=\; \langle P_d^* \eta_0, \varphi \circ u_0\rangle \;=\; \langle \eta_0, \varphi \circ u_0\rangle\ .
\end{aligned}$$

For this identity, we also used that

$$\langle \eta, \lambda_s \otimes \mathrm{adj}_s\rangle = \langle (1 \otimes \mathrm{adj}_s) \bullet \eta, \lambda_s\rangle = \langle \mathrm{adj}_s \nabla u_0, \lambda_s\rangle = \langle (1 \otimes \mathrm{adj}_s) \bullet \eta_0, \lambda_s\rangle = \langle \eta_0, \lambda_s \otimes \mathrm{adj}_s\rangle$$

and

$$\begin{aligned}
\langle \eta, \mathcal{H}^d_{u_0,\lambda}\rangle &= \int_\Omega \mathcal{H}^d_{u_0,\lambda} \bullet \eta \;\, \mathrm{d}x = \int_\Omega \max_{A \in \mathbb{R}^{mn}} \mathcal{H}^d_{u_0,\lambda}(x,A)\, \mathrm{d}x = \int_\Omega \mathcal{H}^d_{u_0,\lambda} \bullet \eta_0 \,\mathrm{d}x \\
&= \langle \eta_0, \mathcal{H}^d_{u_0,\lambda}\rangle.
\end{aligned}$$

As a result, η should only satisfy $\sum_{s=1}^{\min(m,n)} \sigma(s)$ conditions on the momenta $(1 \otimes \mathrm{adj}_s) \bullet \eta$, and also one additional momentum $\mathcal{H}^d_{u_0,\lambda} \bullet \eta$ which is located however at the maximum of $\mathcal{H}^d_{u_0,\lambda}$, so that actually the investigated momenta range a $(\sum_{s=1}^{\min(m,n)} \sigma(s))$-dimensional manifold. Then the assertion follows by the Carathéodory theorem applied to a Young-measure representation of η; see [26, Proposition 4.3.9] for details. $\square$

Let us remark that the estimate (4.9) cannot be improved in general; in other words, there exist relaxed problems that have no solutions of the form (4.8) with $K \le \sum_{s=1}^{\min(m,n)} \sigma(s)$. E.g., for $m = n = 1$, one has obviously $\sum_{s=1}^{\min(m,n)} \sigma(s) = 1$ but a one-dimensional scalar problem with a double well potential investigated in [11, 12, 18], having a unique two-atomic solution, has apparently no one-atomic solution.

5. Some error estimates.

The error estimates we are able to establish here differ from the usual results for elliptic equations mainly because the problem (RP) is not uniformly convex even if $\min(m, n) = 1$; for $\min(m, n) \ge 2$ it is in general nonconvex. This fact implies that we are able to establish only error estimates concerning $|\min(\mathrm{RP}) - \min(\mathrm{RP}_d)|$, which can be considered however as a significant indicator of the efficiency of the particular method.
In the literature one can occasionally find error estimates also for the solutions (in appropriate dual norms on H^* weaker than $\|\cdot\|_{H^*}$) but it is always for special scalar (i.e. $m = 1$) problems possessing a unique solution; cf. [11, 12, 17] for $n = 1$ or [7] for $n \ge 1$. Besides, except [20, 23], the discretization described in the literature concerns the original problem (P), hence the discrete problem looks as

$$\text{minimize} \quad \Phi(u) \; = \; \int_\Omega \varphi(x, u(x), \nabla u(x))\, \mathrm{d}x \quad \text{for} \;\; u \in U_d\,. \tag{P$_d$}$$

The first comparison of the standard discretization (P_d) with our discretization (RP_d) can be based on the simple observation that

$$\inf(\mathrm{P}) \; = \; \min(\mathrm{RP}) \; \le \; \min(\mathrm{RP}_d) \; \le \; \min(\mathrm{P}_d)\,, \tag{5.1}$$

which follows from the fact that every $i(u) = (u, i_H(\nabla u))$ with $u \in U_d$ is admissible for (RP_d). Therefore our method cannot be worse than the standard discretization as far as the rate of convergence of the minimum of the discrete problem towards the minimum of the original problem concerns. Both methods lead to nonconvex mathematical-programming problems, our method having slightly greater dimensionality but of the same order $\mathcal{O}(d^{-n})$ as the standard method provided Corollary 4.2 or 4.3 can be applied. The theoretical effort behind its implementation and slightly greater dimensionality are compensated by the fact that sometimes our method can give much faster convergence than the standard method; cf. Sect. 6 below.
Our aim is to get the convergence $\min(\mathrm{RP}_d) \to \min(\mathrm{RP})$ of the order $\mathcal{O}(d^\alpha)$ with some $\alpha > 0$. We will consider a subspace $\tilde{H} \subset H$ equiped with a norm stronger than the norm of H, so that the imbedding $\tilde{H} \to H$ is continuous. Thus $\tilde{H}^* \supset H^*$ and we can estimate

the generalized Young functionals from $Y^p_H(\Omega;\mathbb{R}^{mn})$ in the standard dual norm of $\tilde{H}^*$. Our error analysis will rely on the following general pattern:

Proposition 5.1. *Let (2.9) be valid, let the mapping $u \mapsto \varphi\circ u$ map every ball B_r in $W^{1,p}(\Omega;\mathbb{R}^m)$ into a bounded subset of $\tilde{H}$, i.e.*

$$\forall r\in\mathbb{R}^+\ \exists K_r\in\mathbb{R}^+\ \forall u\in B_r:\ \ \|\varphi\circ u\|_{\tilde{H}}\le K_r\ , \tag{5.2}$$

and let it be $(W^{1-\alpha,p}(\Omega;\mathbb{R}^m),H)$-Lipschitz continuous on each B_r, i.e.

$$\forall r\in\mathbb{R}^+\ \exists L_r\in\mathbb{R}^+\ \forall u,\tilde{u}\in B_r:\ \|\varphi\circ u-\varphi\circ\tilde{u}\|_H\le L_r\|u-\tilde{u}\|_{W^{1-\alpha,p}(\Omega;\mathbb{R}^m)}\ . \tag{5.3}$$

Also, let us assume that, for every $(u,\eta)\in W^{1,p}(\Omega;\mathbb{R}^m)\times Y^p_H(\Omega;\mathbb{R}^{mn})$ and every $d>0$, there exists an admissible pair (u_d,η_d) for (RP_d) such that the sequence $\{u_d\}_{d>0}$ is bounded in $W^{1,p}(\Omega;\mathbb{R}^m)$ and the sequence $\{(u_d,\eta_d)\}_{d>0}$ has the following approximation property

$$\|u-u_d\|_{W^{1-\alpha,p}(\Omega;\mathbb{R}^m)}\ \le\ C_1 d^\alpha\|u\|_{W^{1,p}(\Omega;\mathbb{R}^m)}\ , \tag{5.4a}$$

$$\|\eta-\eta_d\|_{\tilde{H}^*}\ \le\ C_2 d^\alpha\|\eta\|_{H^*}\ . \tag{5.4b}$$

Then $|\min(\mathrm{RP})-\min(\mathrm{RP}_d)|=\mathcal{O}(d^\alpha)$.

Proof. First, let us note that, thanks to (5.3), $\bar\Phi(u,\eta)=\langle\eta,\varphi\circ u\rangle$ actually defines the extension of Φ which is weakly* continuous on $B_r\times Y^p_H(\Omega;\mathbb{R}^{mn})$ for any $r\in\mathbb{R}^+$.
As always $\min(\mathrm{RP}_d)\ge\min(\mathrm{RP})$, we have only to estimate $\min(\mathrm{RP}_d)-\min(\mathrm{RP})$ from above.
First we will show that $\bar\Phi$ from (3.5) is $(W^{1-\alpha,p}(\Omega;\mathbb{R}^m)\times\tilde{H}^*,\mathbb{R})$-Lipschitz continuous on each $B_r\times H^*$ at any $(u,\eta)\in B_r\times H^*$ in the sense

$$|\bar\Phi(u,\eta)-\bar\Phi(\tilde{u},\tilde{\eta})|\ \le\ L_{r,\eta}\Big(\|u-\tilde{u}\|_{W^{1-\alpha,p}(\Omega;\mathbb{R}^m)}+\|\eta-\tilde{\eta}\|_{\tilde{H}^*}\Big) \tag{5.5}$$

for any $(\tilde{u},\tilde{\eta})\in B_r\times H^*$ and some $L_{r,\eta}\in\mathbb{R}^+$ depending possibly on r and η. Indeed, for $u,\tilde{u}\in B_r$ and $\eta,\tilde{\eta}\in H^*$ we can estimate

$$\begin{aligned}&|\bar\Phi(u,\eta)-\bar\Phi(\tilde{u},\tilde{\eta})|\ =\ |\langle\eta,\varphi\circ u\rangle-\langle\tilde{\eta},\varphi\circ\tilde{u}\rangle|\ \le\ |\langle\eta,\varphi\circ u-\varphi\circ\tilde{u}\rangle|\\&+|\langle\eta-\tilde{\eta},\varphi\circ\tilde{u}\rangle|\ \le\ \|\eta\|_{H^*}\|\varphi\circ u-\varphi\circ\tilde{u}\|_H\ +\ \|\eta-\tilde{\eta}\|_{\tilde{H}^*}\|\varphi\circ\tilde{u}\|_{\tilde{H}}\\&\le L_r\|\eta\|_{H^*}\|u-\tilde{u}\|_{W^{1-\alpha,p}(\Omega;\mathbb{R}^m)}\ +\ K_r\|\eta-\tilde{\eta}\|_{\tilde{H}^*}\ ,\end{aligned}$$

which gives (5.5) with $L_{r,\eta}=\max(L_r\|\eta\|_{H^*},K_r)$.
Eventually, let us take (u,η) a solution to (RP) and (u_d,η_d) as assumed. Putting $r=\sup_{d>0}\|u_d\|_{W^{1,p}(\Omega;\mathbb{R}^m)}$ and using (5.4) and (5.5), we can estimate:

$$\begin{aligned}&\min(\mathrm{RP}_d)-\min(\mathrm{RP})\ \le\ \bar\Phi(u_d,\eta_d)-\bar\Phi(u,\eta)\ \le\ L_{r,\eta}\|u-u_d\|_{W^{1-\alpha,p}(\Omega;\mathbb{R}^m)}\\&+\ L_{r,\eta}\|\eta-\eta_d\|_{\tilde{H}^*}\ \le\ L_{r,\eta}C_1d^\alpha\|u\|_{W^{1,p}(\Omega;\mathbb{R}^m)}\ +\ L_{r,\eta}C_2d^\alpha\|\eta\|_{H^*}\ =\ \mathcal{O}(d^\alpha)\ .\end{aligned}$$

□

Thus we have reduced the error-estimation problem to the verification of (5.2)–(5.4); in fact, it suffices to verify (5.4) only for some solution to (RP).

Let us first treat the one-dimensional case, which is especially simple. Recall that throughout this paper we suppose $p > 1$, hence we have now $W^{1,p}(\Omega;\mathbb{R}^m) \subset C^0(\Omega;\mathbb{R}^m)$ for $\Omega \subset \mathbb{R}^1$ and we can speak about an element-wise affine interpolation of a given $u \in W^{1,p}(\Omega;\mathbb{R}^m)$, denoted by $\Pi_d u$ as usual. In this one-dimensional case, it suffices to control the spatial smoothness only, using the parameter $0 < \alpha \le 1$. Thus we define naturaly the subspace $H_\alpha \subset H$ by

$$H_\alpha \;=\; \left\{ h_0 + \sum_{s=0}^{\min(m,n)} g_s \otimes \mathrm{adj}_s \in H;\; h_0 \in C^{0,\alpha}(\Omega)\otimes V,\; g_s \in W^{\alpha,p/(p-s)}(\Omega;\mathbb{R}^{\sigma(s)}) \right\}, \tag{5.6}$$

and endow it with the norm

$$\|h\|_{H_\alpha} \;=\; \inf_{h_0+\sum_s g_s\otimes\mathrm{adj}_s = h} \left(\|R_p h_0\|_{C^{\alpha,0}(\Omega;C^0(\mathbb{R}^{mn}))} \;+\; \sum_{s=0}^{\min(m,n)} \|g_s\|_{W^{\alpha,p/(p-s)}(\Omega;\mathbb{R}^{\sigma(s)})} \right), \tag{5.7}$$

where $R_p : H \to L^\infty(\Omega\times\mathbb{R}^{mn})$ is defined by $[R_p h](x,A) = h(x,A)/(1+|A|^p)$.

Lemma 5.2. *Let $n=1$ and let $(u,\eta) \in W^{1,p}(\Omega;\mathbb{R}^m)\times H^*$ be such that $(1\otimes\mathrm{id})\bullet\eta = \nabla u$. Putting $u_d = \Pi_d u$ and $\eta_d = P_d^*\eta$, we have $\|u_d\|_{W^{1,p}(\Omega;\mathbb{R}^m)} \le \|u\|_{W^{1,p}(\Omega;\mathbb{R}^m)}$ and (5.4) fulfilled for $\tilde{H} = H_\alpha$. Moreover, $(1\otimes\mathrm{id})\bullet\eta_d = \nabla u_d$, hence each pair (u_d,η_d) is admissible for (RP_d).*

Proof. The identity $(1\otimes\mathrm{id})\bullet\eta_d = \nabla u_d$ follows from $(1\otimes\mathrm{id})\bullet P_d^*\eta = A_d((1\otimes\mathrm{id})\bullet\eta)$ and $\nabla u_d = A_d\nabla u$, where A_d denotes the operator $L^p(\Omega;\mathbb{R}^m)\to L^\infty(\Omega;\mathbb{R}^m)$ which takes the mean over each element (here an interval, since $n=1$). The estimate $\|u_d\|_{W^{1,p}(\Omega;\mathbb{R}^m)} \le \|u\|_{W^{1,p}(\Omega;\mathbb{R}^m)}$ follows by easy calculations, while (5.4a) can be obtained by interpolation from standard results from the finite-element method; cf. [10]. The estimate (5.4b) follows from

$$\|h - P_d h\|_H \;\le\; C_2 d^\alpha \|h\|_{H_\alpha} \tag{5.8}$$

by transposition as in [26, Section 3.5]. In view of (2.3) and (5.7), the estimate (5.8) follows from $\|g - A_d g\|_{C^0(\Omega)} \le C d^\alpha \|g\|_{C^{0,\alpha}(\Omega)}$ and from the estimates $\|g - A_d g\|_{L^{p/(p-s)}(\Omega)} \le C d^\alpha \|g\|_{W^{\alpha,p/(p-s)}(\Omega)}$ for $s = 0,1$. The last estimate can be obtained by interpolation from the four estimates created by $\alpha = 0$ or 1 and $p/(p-s) = 1$ or ∞, which are obvious. $\square$

Lemma 5.3. *Let $n=1$, φ have the form $\varphi(x,u,A) = \varphi_0(x,u)+\varphi_1(x,A)$ with $\varphi_1 \in H_\alpha$ for some $0<\alpha\le 1$ and with $\varphi_0(x,\cdot):\mathbb{R}^m\to\mathbb{R}$ Lipschitz continuous on bounded subsets of $\mathbb{R}^m$ in the sense*

$$\begin{array}{c} \forall c\in\mathbb{R}^+\ \exists L_c \in L^{p/(p-1)}(\Omega)\ \forall x\in\Omega\ \forall u,\tilde{u}\in\mathbb{R}^m : \\ |u|\le c,\ |\tilde{u}|\le c \quad\Rightarrow\quad |\varphi_0(x,u)-\varphi_0(x,\tilde{u})| \le L_c(x)|u-\tilde{u}| \end{array} \tag{5.9}$$

and $\varphi_0(\cdot,0)\in L^1(\Omega)$. Then (5.2) with $\tilde{H} = H_\alpha$ and (5.3) are satisfied.

Proof. For a given $r \in \mathbb{R}^+$ there is $c = c(r) \in \mathbb{R}^+$ such that any $u \in B_r = \{u \in W^{1,p}(\Omega;\mathbb{R}^m);\ \|u\|_{W^{1,p}(\Omega;\mathbb{R}^m)} \le r\}$ take values bounded (in absolute values) by $c(r)$, hence φ_0 can be supposed Lipschitz with a Lipschitz constant $L_{c(r)}$, cf. (5.9).
As for (5.2), it follows from the estimates $\|\varphi \circ u\|_{H_\alpha} \le \|\varphi_0 \circ u\|_{W^{\alpha,1}(\Omega)} + \|\varphi_1\|_{H_\alpha}$ and

$$\begin{aligned}\|\varphi_0 \circ u\|_{W^{\alpha,1}(\Omega)} &\le \|\varphi_0 \circ u\|_{W^{1,1}(\Omega)} = \int_\Omega \left(|\varphi_0(x,u(x))| + \left|\frac{\mathrm{d}}{\mathrm{d}x}\varphi_0(x,u(x))\right|\right) \mathrm{d}x \\ &\le \int_\Omega \left(|\varphi_0(x,0)| + L_{c(r)}(x)|u(x)| + \left|\frac{\partial}{\partial u}\varphi_0(x,u(x))\frac{\mathrm{d}}{\mathrm{d}x}u(x)\right|\right) \mathrm{d}x \\ &\le \|\varphi_0(x,0)\|_{L^1(\Omega)} + \|L_{c(r)}\|_{L^{p/(p-1)}(\Omega)}\|u\|_{W^{1,p}(\Omega)} \ .\end{aligned}$$

Furthermore, we can estimate

$$\|\varphi \circ u - \varphi \circ \tilde{u}\|_H \ \le\ \|\varphi_0 \circ u - \varphi_0 \circ \tilde{u}\|_{L^1(\Omega)} \ \le\ \|L_{c(r)}\|_{L^{p/(p-1)}(\Omega)}\|u - \tilde{u}\|_{L^p(\Omega)} \ ,$$

which already implies (5.3) thanks to the continuous imbedding $W^{1-\alpha,p}(\Omega) \subset L^p(\Omega)$. □

Therefore, for $n = 1$, Proposition 5.1 together with Lemmas 5.2 and 5.3 yields the desired estimate $|\min(\mathrm{RP}) - \min(\mathrm{RP}_d)| = \mathcal{O}(d^\alpha)$. In the general case, this result is likely not to be much improvable because the data qualifications in Lemma 5.3 are actually very weak; e.g. we admit $\varphi(x,u,A) = g(x)v(A)$ with $v \in V$ only continuous and coercive. On the other hand, in special cases this guaranteed rate of error can be still pessimistic; sometimes even the convergence of the order $\mathcal{O}(d^\infty)$ can be achieved, cf. Section 6.
In the multidimensional case (i.e. $n \ge 2$) the construction $\eta_d = P_d^*\eta$ unfortunately does not work because $(1 \otimes \mathrm{id}) \bullet P_d^*\eta \in L^p(\Omega,\mathbb{R}^{mn})$ need not be the gradient of any function from $W^{1,p}(\Omega;\mathbb{R}^m)$ and, if also $m \ge 2$, $P_d^*\eta$ need not be any gradient Young measure even locally on each element $E \in \mathcal{T}_d$, this means $P_d\eta|_E \notin G_H^p(E;\mathbb{R}^{mn})$ in general. Nevertheless, in the scalar case (i.e. $m = 1$) a suitable η_d can be constructed by a more sophisticated way, using a shift operator $T_y : H \to H$ defined by $[T_y h](x,A) = h(x, A + y(x))$ with $y \in L^p(\Omega;\mathbb{R}^{mn})$. Obviously, T_y makes a "vertical shift" of the integrands from H and, for $y \in L^\infty(\Omega;\mathbb{R}^{mn})$, it is a linear continuous operator with the continuous inverse $T_y^{-1} = T_{-y}$; in particular, in view of (2.3) it is easy to verify

$$\begin{aligned}&\forall r \in \mathbb{R}^+\ \ \exists N_r \in \mathbb{R}^+\ \ \forall h \in H\ \ \forall y \in L^\infty(\Omega;\mathbb{R}^{mn}) : \\ &\|y\|_{L^\infty(\Omega;\mathbb{R}^{mn})} \le r \quad \Rightarrow \quad \|T_y h\|_H \ \le N_r\|h\|_H \ .\end{aligned} \tag{5.10}$$

The adjoint operator $T_y^* : H^* \to H^*$ maps $Y_H^p(\Omega;\mathbb{R}^{mn})$ onto $Y_H^p(\Omega;\mathbb{R}^{mn})$, making a "vertical shift" of the corresponding Young functionals. The philosophy of our construction of η_d is the following: first shift η in such a way that the first momentum of the shifted Young measure vanishes, then apply the operator P_d^* without falling out of the class $G_H^p(\Omega;\mathbb{R}^{mn})$, and eventually shift the resulted Young measure "almost" back; cf. (5.11) below. To realize this "almost", we will have to suppose a $W^{1+\alpha,\infty}$-regularity of a solution and a stronger data qualification, requiring a certain weight-Lipschitz continuity of $\varphi(x,u,\cdot)$ weaker than the usual Lipschitz continuity, which would not be a realistic requirement.

The corresponding subspace of H may be chosen as $H_{\rm Lip} = \{h \in H;\ \|h\|_{H_{\rm Lip}} < +\infty\}$ with the norm

$$\|h\|_{H_{\rm Lip}} \;=\; \operatorname*{ess\,sup}_{x\in\Omega} \sup_{A,\tilde A\in\mathbb{R}^{mn}} \frac{|h(x,A)-h(x,\tilde A)|}{|A-\tilde A|(1+|A|^p+|\tilde A|^p)}\ .$$

We will use Proposition 5.1 with $\tilde H = H_\alpha \cap H_{\rm Lip}$ equiped with the norm $\|h\|_{H_\alpha\cap H_{\rm Lip}} = \|h\|_{H_\alpha} + \|h\|_{H_{\rm Lip}}$.

Lemma 5.4. *Let $m=1$, $0<\alpha\le 1$, and let $(u,\eta)\in W^{1+\alpha,\infty}(\Omega)\times Y^p_H(\Omega;\mathbb{R}^n)$ be such that $(1\otimes{\rm id})\bullet\eta = \nabla u$. Putting*

$$u_d \;=\; \Pi_d u \quad\text{and}\quad \eta_d \;=\; T^*_{\nabla u_d} P^*_d T^*_{-\nabla u}\eta\ , \tag{5.11}$$

we have got $\|u_d\|_{W^{1,p}(\Omega)} \le \|u\|_{W^{1,p}(\Omega)}$ and (5.4) with $\tilde H = H_\alpha\cap H_{\rm Lip}$ fulfilled. Moreover, $\eta_d\in Y_d$ and $(1\otimes{\rm id})\bullet\eta_d = \nabla u_d$, hence each pair (u_d,η_d) is admissible for (RP$_d$).

Proof. First, $(1\otimes{\rm id})\bullet\eta_d = (1\otimes{\rm id})\bullet T^*_{\nabla u_d}\big(P^*_d T^*_{-\nabla u}\eta\big) = \nabla u_d + (1\otimes{\rm id})\bullet P^*_d T^*_{-\nabla u}\eta = \nabla u_d$ because $(1\otimes{\rm id})\bullet T^*_{-\nabla u}\eta = -\nabla u + (1\otimes{\rm id})\bullet\eta = 0$, and therefore also $(1\otimes{\rm id})\bullet P^*_d T^*_{-\nabla u}\eta = 0$. Secondly, let us notice that T_y and P_d commute with each other provided y is element-wise constant on $\mathcal{T}_d$. In particular, $T_{\nabla u_d}P_d = P_d T_{\nabla u_d}$ and we can write alternatively $\eta_d = P^*_d T^*_{\nabla u_d} T^*_{-\nabla u}\eta = P^*_d T^*_{\nabla(u_d-u)}\eta$. It shows that $\eta_d\in Y_d$.
The estimate $\|u_d\|_{W^{1,p}(\Omega)} \le \|u\|_{W^{1,p}(\Omega)}$ is obvious. The estimate (5.4a) follows by interpolation from the standard estimates $\|u-u_d\|_{W^{1,p}(\Omega)} \le 2\|u\|_{W^{1,p}(\Omega)}$ and $\|u-u_d\|_{L^p(\Omega)} \le Cd\|u\|_{W^{1,p}(\Omega)}$, cf. [10, Theorem 16.1]. As for (5.4b), we will estimate it as

$$\begin{aligned}\|\eta-\eta_d\|_{[H_\alpha\cap H_{\rm Lip}]^*} \le{}& \|\eta - T^*_{\nabla(u_d-u)}\eta\|_{[H_\alpha\cap H_{\rm Lip}]^*}\\ &+ \|T^*_{\nabla(u_d-u)}\eta - P^*_d T^*_{\nabla(u_d-u)}\eta\|_{[H_\alpha\cap H_{\rm Lip}]^*} \equiv I_1 + I_2.\end{aligned}$$

Let us estimate I_1. For any $r\in\mathbb{R}^+$ there is C_r such that

$$|h(x,A)-h(x,A+\tilde A)| \;\le\; |\tilde A|(1+|A|^p+|A+\tilde A|^p)\|h\|_{H_{\rm Lip}} \;\le\; C_r|\tilde A|(1+|A|^p)\|h\|_{H_{\rm Lip}}$$

for any $h\in H_{\rm Lip}$ and $A,\tilde A\in\mathbb{R}^n$ provided $|\tilde A|\le r$. This gives the estimate

$$\|h-T_y h\|_H \;\le\; \sup_{x\in\Omega}\sup_{A\in\mathbb{R}^n} \frac{|h(x,A)-h(x,A+y(x))|}{1+|A|^p} \;\le\; C_r\|y\|_{L^\infty(\Omega;\mathbb{R}^n)}\|h\|_{H_{\rm Lip}}$$

for any $h\in H_{\rm Lip}$ and $y\in L^\infty(\Omega;\mathbb{R}^n)$ with $\|y\|_{L^\infty(\Omega;\mathbb{R}^n)}\le r$. From this, for $y=\nabla(u_d-u)$, we get by transposition $I_1 \le \|\eta - T^*_{\nabla(u_d-u)}\eta\|_{H^*_{\rm Lip}} \le C_r\|\nabla(u_d-u)\|_{L^\infty(\Omega;\mathbb{R}^n)}\|h\|_{H^*}$ provided $\nabla(u_d-u)$ is small enough, namely $\|\nabla(u_d-u)\|_{L^\infty(\Omega;\mathbb{R}^n)}\le r$. Afterwards we can use the estimate $\|\nabla(u_d-u)\|_{L^\infty(\Omega;\mathbb{R}^n)} \le Cd^\alpha\|u\|_{W^{1+\alpha,\infty}(\Omega)}$, which follows by interpolation from [10, Theorem 16.1]. This shows $I_1 = \mathcal{O}(d^\alpha)$.

By transposition of (5.10), we get $\|T_y^*\eta\|_{H^*} \leq N_r\|\eta\|_{H^*}$. Using (5.8), we can eventually estimate the term I_2 as follows:

$$I_2 \;\leq\; \|T^*_{\nabla(u_d-u)}\eta - P_d^* T^*_{\nabla(u_d-u)}\eta\|_{H^*_\alpha} \;\leq\; C_2 d^\alpha \|T^*_{\nabla(u_d-u)}\eta\|_{H^*} \;\leq\; C_2 N_r d^\alpha \|\eta\|_{H^*}$$

provided $\|\nabla(u_d-u)\|_{L^\infty(\Omega)} \leq r$. It gives $I_2 = \mathcal{O}(d^\alpha)$ whenever $u \in W^{1,\infty}(\Omega)$; recall that even $u \in W^{1+\alpha,\infty}(\Omega)$ had to be supposed. □

As for (5.2) and (5.3), we can use readily Lemma 5.3 also for $n > 1$ with H_α replaced by $H_\alpha \cap H_{\mathrm{Lip}}$ if $p > n$, which ensures $W^{1,p}(\Omega) \subset L^\infty(\Omega)$. Then, for $m = 1$, Proposition 5.1 together with Lemma 5.3 (modified as outlined) and Lemma 5.4 enables us to get $|\min(\mathrm{RP}) - \min(\mathrm{RP}_d)| = \mathcal{O}(d^\alpha)$ if (RP) possesses at least one solution (u,η) with $u \in W^{1+\alpha,\infty}(\Omega)$ and if φ fulfils the data qualification from Lemma 5.3 with H_α replaced by $H_\alpha \cap H_{\mathrm{Lip}}$. If $p \leq n$, then (5.2) with $\tilde{H} = H_\alpha \cap H_{\mathrm{Lip}}$ and (5.3) must be verified for each particular case. E.g., if $\varphi(x,u,A) = \varphi_1(x,A)$, then they reduce to the only requirement $\varphi_1 \in H_\alpha \cap H_{\mathrm{Lip}}$. Let us also remark that, supposing additionally (3.6), requirement (5.3) can be weakened to $\|\varphi\circ u - \varphi\circ\tilde{u}\|_H \leq L_r\|u-\tilde{u}\|_{W^{1,p}(\Omega;\mathbb{R}^m)}$ because, due to the regularity assumption $u \in W^{1+\alpha,\infty}(\Omega)$, we can use $\|u - u_d\|_{W^{1,p}(\Omega)} \leq C_1 d^\alpha \|u\|_{W^{1+\alpha,p}(\Omega)}$ instead of (5.4a).

The error estimates for the case $\min(m,n) \geq 2$ remains open because no effective formula to construct an admissible pair (u_d,η_d) sufficiently near to (u,η) is known; however, in very special cases some error estimates for (P_d) have been obtained by Chipot, Collins and Kinderlehrer [8], which yields via (5.1) certain (probably rather pessimistic) error estimates for (RP_d), too.

6. A comparison and concluding remarks.

Let us compare our approximation scheme (RP_d) with the standard scheme (P_d). As already mentioned, we have always $|\min(\mathrm{RP}) - \min(\mathrm{RP}_d)| \leq |\min(\mathrm{RP}) - \min(\mathrm{P}_d)|$, hence our method cannot be worse than (P_d) even in case $\min(m,n) \geq 2$. Sometimes, it might be even considerably better. The only analysis of (P_d) in multidimensional scalar case under data qualification similarly weak as in Section 5 was performed by Chipot and Collins, see [6, 7]. For $m = 1$, Dirichlet boundary conditions, and $\varphi(x,u,A) = \varphi_1(A) + \varphi_0(u)$ with either $\varphi_0 \equiv 0$ in [6] or some special φ_0 having the growth at most $\mathcal{O}(|u|^q)$ in [7], the error $|\min(\mathrm{RP}) - \min(\mathrm{P}_d)|$ is of the order either $\mathcal{O}(d^{1/2})$ if $q \geq 1$ or $\mathcal{O}(d^{q/(q+1)})$ if $0 < q < 1$. A similar result, namely the rate $\mathcal{O}(d^{1/2})$ for $\varphi_0 \equiv 0$, was obtained also by Brighi and Chipot [5] in a bit different context. Let us remind that, in contrast to [5, 6, 7], we admitted φ to be spatially dependent and our guaranteed rate was $\mathcal{O}(d^\alpha)$ with $0 < \alpha \leq 1$ depending on the spatial smoothness of φ and, if $n \geq 2$, also on $W^{1+\alpha,\infty}$-regularity of at least one solution to (RP).

Moreover, there are several results concerning the one-dimensional scalar case $n = m = 1$, see [4, 11, 12, 13, 17, 21]. For $\varphi(x,u,A) = \varphi_1(A) + \varphi_0(x,u)$ with a special φ_0 and with φ_1 having two wells (=local minima) of the same value and a quadratic growth in neighbourhoods of these wells, the rate of $|\min(\mathrm{RP}) - \min(\mathrm{P}_d)|$ was shown to be $\mathcal{O}(d^2)$. It is better than our guaranteed general rate, but (5.1) ensures in this special case the order also at least $\mathcal{O}(d^2)$. Moreover, the problems in [11, 12, 13, 17] are so special that even $\min(\mathrm{RP}_d) = \min(\mathrm{RP})$ because (RP) possesses in these special cases the solution (u,η) with u affine and η homogeneous, which is apparently admissible also for our approximate problems (RP_d) for any $d > 0$; therefore, in fact, even the rate $\mathcal{O}(d^\infty)$ was achieved.

A similar effect can happen even for multidimensional vectorial problems, where no rate-of-error analysis is disposable for (RP_d), neither for (P_d). We can demonstrate it on the potential proposed by Ericksen and James and used by Collins, Luskin and Riordan [13, 14] for numerical experiments in the case $n = m = 2$:

$$\varphi(x,u,A) \;=\; \varphi_1(A) \;=\; k_1(b_{11}+b_{22}-2)^2 + k_2 b_{12}^2 + k_3\left(\left(\frac{b_{11}-b_{22}}{2}\right)^2 - \varepsilon^2\right)^2 , \quad (6.1)$$

where $B = A^{\mathrm{T}}A = [b_{ij}]$ and $k_1, k_2, k_3, \varepsilon$ are positive, $\varepsilon \le 1$. It is known, cf. [13, 14], that the relaxed problem completed by suitable Dirichlet boundary conditions possesses the solution $u(x) = (\frac{1}{2}A_1 + \frac{1}{2}A_2)x$ and η homogeneous having a classical-Young-measure representation as a two-atomic Young measure $\frac{1}{2}\delta_{A_1} + \frac{1}{2}\delta_{A_2}$ with δ_A denoting the Dirac distribution at $A \in \mathbb{R}^{2\times 2}$ and

$$A_1 \;=\; \begin{pmatrix} \sqrt{1-\varepsilon} & 0 \\ 0 & \sqrt{1+\varepsilon} \end{pmatrix} , \qquad A_2 \;=\; \begin{pmatrix} (1+\varepsilon)\sqrt{1-\varepsilon} & \varepsilon\sqrt{1-\varepsilon} \\ -\varepsilon\sqrt{1+\varepsilon} & (1-\varepsilon)\sqrt{1+\varepsilon} \end{pmatrix}$$

Obviously, $\mathrm{Rank}(A_1 - A_2) = 1$ and $(1\otimes\mathrm{id})\bullet\eta = \nabla u$. It implies that (u,η) is actually admissible for (RP) and obviously also for (RP_d) with any $d > 0$, as well. Therefore we have got again $\min(\mathrm{RP}) = \min(\mathrm{RP}_d) = 0$, hence the rate $\mathcal{O}(d^\infty)$ was achieved.

One important remark should be however made to (RP_d) with φ from (6.1) because it is not known whether φ_1 from (6.1) has a polyconvex quasiconvexification, thus $\varphi\circ u = 1\otimes\varphi_1$ need not be contained in any H from (2.1) with V satisfying (2.2). Taking V satisfying only (2.2a-d) and containing φ_1, the extension by (3.5) is made possible but (RP_d) in general does not approximate (RP) but only the following auxiliary problem

$$\left.\begin{array}{lll} \text{minimize} & \bar{\Phi}(u,\eta) & \\ \text{subject to} & u \in W^{1,p}(\Omega;\mathbb{R}^m),\ \eta \in Y^p_H(\Omega;\mathbb{R}^{mn}), & \\ & \mathrm{adj}_s\nabla u \;=\; (1\otimes\mathrm{adj}_s)\bullet\eta \ \text{ for all } 1 \le s \le \min(m,n). & \end{array}\right\} \quad \text{(AP)}$$

It may happen that $\min(\mathrm{AP}) < \min(\mathrm{RP})$ because not every η admissible for (AP) belongs to $G^p_H(\Omega;\mathbb{R}^{mn})$ if (2.2e) is not satisfied. However, by the results from [15] it can be shown that, if (AP) possesses a solution (u,η) such that the poly- and the quasi-convex envelopes of φ_1 from (6.1) coincides with each other at every point $A = \nabla u(x)$ with $x \in \Omega$, then $\min(\mathrm{AP}) = \min(\mathrm{RP})$. This actually applies here because both the poly- and the quasi-convex envelopes of φ_1 from (6.1) are zero at $A = \frac{1}{2}A_1 + \frac{1}{2}A_2 = \nabla u$.

Additionally, the exponents in (6.1) should be better modified a bit because the present form of φ_1 ensures the coercivity in $L^4(\Omega;\mathbb{R}^{2\times 2})$ via the first term while the third term can take values $+\infty$ on this space.

Acknowledgements. The author is thankful to Professors Karl-Heinz Hoffmann and Stefan Müller for many discussions. This work has been carried out during author's stay at the Institute for Applied Mathematics and Statistics, TU München.

References

[1] J.M. Ball: On the calculus of variations and sequentially weakly continuous maps. In: Proc. Conf. Ordinary and Partial Differential Equations. (W.N.Everitt, B.D.Sleeman, eds.), Lecture Notes in Math. 564, Springer, Berlin (1976) 13–25.

[2] J.M. Ball, R.D. James: Fine phase mixtures as minimizers of energy, Archive Rat. Mech. Anal. 100 (1988) 13–52.

[3] J.M. Ball, R.D. James: Proposed experimental tests of a theory of fine microstructure and the two-well problem, Phil. Trans. Royal Soc. London A 338 (1992) 389–450.

[4] B. Brighi, M. Chipot: Approximation in nonconvex problems. In: "Progress in PDEs: Calculus of Variations, Applications" (Eds. C.Bandle et al.), Pitman Res. Notes in Math. Sci. 267 (1992) Longmann, 150–157.

[5] B. Brighi, M. Chipot: Approximated convex envelope of a function. A preprint.

[6] M. Chipot: Numerical analysis of oscillations in conconvex problems, Numer. Math. 59 (1991) 747–767.

[7] M. Chipot, C. Collins: Numerical approximations in variational problems with potential wells, SIAM J. Numer. Anal. 29 (1992) 1002–1019.

[8] M. Chipot, C. Collins and D. Kinderlehrer: Numerical analysis of oscillations in multiple well problems, Numer. Math. 70 (1995) 259–282.

[9] M. Chipot, D. Kinderlehrer: Equilibrium configurations of crystals, Arch. Rational Mech. Anal. 103 (1988) 237–277.

[10] P.G. Ciarlet: Basic error estimates for elliptic problems, In: Handbook of Numerical Analysis, Vol. II, Finite element methods (Part 1), P.G. Ciarlet, J.L. Lions, eds., North-Holland, Amsterdam, 1991.

[11] C. Collins, D. Kinderlehrer and M. Luskin: Numerical approximation of the solution of a variational problem with a double well potential, SIAM J. Numer. Anal. 28 (1991) 321–332.

[12] C. Collins, M. Luskin: Optimal order error estimates for finite element approximation of the solution of a nonconvex variational problem, Math. Comp. 57 (1991) 621–637.

[13] C. Collins, M. Luskin: Numerical modelling of the microstructure of crystals with symmetry-related variants. In: US-Japan Workshop on Smart/Inteligent Materials and Systems (Eds. I.Ahmad et al.), Technomic Publ. Comp., Lancaster (1990) 309–318.

[14] C. Collins, M. Luskin and J. Riordan: Computational results for a two-dimensional model of crystalline microstructure. In: IMA Vol. in Math. and Applications "Microstructure and Phase Transitions" (Eds. R.James, D.Kinderlehrer, M.Luskin), Springer, New York (1993) 51–56.

[15] B. Dacorogna: Direct Methods in the Calculus of Variations, Springer, Berlin, 1989.

[16] M. Kružík: Numerical approach to double-well problems (submitted).

[17] M. Luskin: Numerical analysis of microstructure for crystals with a nonconvex energy density. In: The Metz Days Surveys 1989-90 (M.Chipot, J. Saint Jean Paulin, eds.), Pitman Res. Notes in Math., Longman, Harlow (1991) 156–165.

[18] S. Müller: Minimizing sequences for nonconvex functionals, phase transitions and singular perturbations. In: Lecture Notes in Physics 359 (K. Kirchgässner, ed.), Springer, Berlin (1988) 31–44.

[19] C.B. Morrey, Jr.: Quasi-convexity and the lower semicontinuity of multiple integrals. Pacific J. Math. 2 (1952) 25–53.

[20] R.A. Nicolaides, N.J. Walkington: Computation of microstructure utilizing Young measure representations. In: Recent Advances in Adaptive and Sensory Materials and their Applications (C.A.Rogers, R.A.Roegers, eds.), Technomic Publ. (1992) 131–141.

[21] R.A. Nicolaides, N.J. Walkington: Strong convergence of numerical solutions to degenerate variational problems, Res. Report 92-NA-027, Center for Nonlinear Analysis, Carnegie Mellon University, 1992.

[22] Y.G. Reshetnyak: On the stability of conformal mappings in multidimensional spaces. Siberian Math. J. 8 (1967) 69–85.

[23] T. Roubíček: Finite element approximation of a microstructure evolution, Math. Methods in the Appl. Sci. 17 (1994) 377–393.

[24] T. Roubíček: Effective characterization of generalized Young measures generated by gradients. Bollettino Uni. Mat. Italiana 9-B (1995) 755–779.

[25] T. Roubíček: Relaxation of vectorial variational problems, Matematica Bohemica 120 (1995) 411–430.

[26] T. Roubíček: Relaxation in Optimization Theory and Variational Calculus. W. de Gruyter, Berlin, 1997.

[27] T. Roubíček, K.-H. Hoffmann: Theory of convex local compactifications with applications to Lebesgue spaces, Nonlinear Analysis, Th., Methods, Appl. 20 (1995) 607–628.

[28] L.C. Young: Generalized curves and the existence of an attained absolute minimum in the calculus of variations, Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, Classe III 30 (1937) 212–234.

[29] J. Zowe, S. Kurcyusz: Regularity and stability for the mathematical programming problem in Banach spaces, Appl. Math. Optim. 5 (1979) 49–62.

HIER :

Leere Seite
348