CONSTRAINED NONSMOOTH PROBLEMS OF THE CALCULUS OF VARIATIONS*

MAKSIM DOLGOPOLIK**

Abstract. The paper is devoted to an analysis of optimality conditions for nonsmooth multidimensional problems of the calculus of variations with various types of constraints, such as additional constraints at the boundary and isoperimetric constraints. To derive optimality conditions, we study generalised concepts of differentiability of nonsmooth functions called codifferentiability and quasidifferentiability. Under some natural and easily verifiable assumptions we prove that a nonsmooth integral functional defined on the Sobolev space is continuously codifferentiable and compute its codifferential and quasidifferential. Then we apply general optimality conditions for nonsmooth optimisation problems in Banach spaces to obtain optimality conditions for nonsmooth problems of the calculus of variations. Through a series of simple examples we demonstrate that our optimality conditions are sometimes better than existing ones in terms of various subdifferentials, in the sense that our optimality conditions can detect the non-optimality of a given point, when subdifferential-based optimality conditions fail to disqualify this point as non-optimal.

1. INTRODUCTION

Nonsmooth problems of the calculus of variations arise in various applications, such as optimisation of hydrothermal systems [3–5] and nonsmooth modelling in mechanics and engineering (see monograph [29]). Their theoretical study was started by Rockafellar in the convex case in [74–76], where some existence and duality results, as well as optimality conditions in terms of subdifferentials, were obtained. In these optimality conditions the classical Euler-Lagrange equation and transversality condition for the problem of Bolza

$$\min\, l(x(0), x(T)) + \int_0^T L(t, x(t), \dot{x}(t))\, \mathrm{d}t \qquad (1.1)$$

were replaced by the following inclusions:

$$(\dot{p}(t), p(t)) \in \partial L(t, x(t), \dot{x}(t)) \quad \text{for a.e. } t \in (0, T), \quad (p(0), -p(T)) \in \partial l(x(0), x(T),$$

Institute for Problems in Mechanical Engineering, Russian Academy of Sciences, Saint Petersburg, 199178 Russia.

** Corresponding author: dmv@ipme.ru

© The authors. Published by EDP Sciences, SMAI 2021

where "∂" stands for subdifferential in the sense of convex analysis [37, 50]. Note that if the function L is differentiable, then $p(t) = L'_{\dot{x}}(t, x(t), \dot{x}(t))$ and the first inclusion is reduced to the Euler-Lagrange equation.

Further research was devoted to relaxing the convexity assumptions made by Rockafellar and replacing the subdifferential in the sense of convex analysis by some other subdifferential defined for nonconvex functions. Important steps in this direction were made by Clarke [13, 14, 16, 17], who studied problems with locally Lipschitz continuous functions l and L and replaced the subdifferentials in the sense of convex analysis with what now is known as the Clarke subdifferential. Apart from optimality conditions in the form of the Euler-Lagrange inclusion, Clarke also obtained optimality conditions in the Hamiltonian form. Clarke's results were sharpened and extended to more general variational problems by Loewen and Rockafellar [56, 57], while equivalence between Euler-Lagrange and Hamiltonian forms of optimality conditions for nonsmooth variational problems was studied in [7, 18, 46, 77]. Nonlocal optimality conditions for nonconvex problems of the calculus of variations in terms of subdifferentials in the sense of convex analysis and their connections to the existence of minimisers were studied by Marcelli *et al.* [21, 60–62].

First optimality conditions for nonsmooth variational problems involving *nonconvex* subdifferentials were obtained by Mordukhovich [63, 64] (see also [65, 66]). Later, optimality conditions for a nonsmooth problem of Bolza in terms of limiting proximal and limiting Fréchet subdifferentials were studied by Loewen and Rockafellar [55, 58, 59], Ioffe and Rockafellar [49], Vinter and Zheng [80, 81], Bellaassali [6], and Jourani [51].

A different approach to an analysis of optimality conditions for nonsmooth problems of the calculus of variations based on the use of *codifferentials* was developed by the author in [31]. Codifferentials of nonsmooth functions were introduced by Demyanov [23–25] in the late 1980s. A general theory of codifferentiable functions, closely related to the theory of Demyanov-Rubinov-Polyakova quasidifferentials [26, 28, 29], was developed in the finite dimensional case in [27]. Its infinite dimensional generalisations were studied in [30, 32, 33, 82, 83]. In [31] it was shown that optimality conditions for problem (1.1) in terms of codifferentials are sometimes better than subdifferential-based optimality conditions. Let us also mention a completely different approach to the derivation of optimality conditions and numerical solution of nonsmooth problems of the calculus of variations based on the Chebyshev pseudospectral method [79].

It should be noted that in most of the aforementioned papers nonsmooth problems of the calculus of variations were not studied by themselves, but in the context of variational problems for differential inclusions. It seems that since the mid-90s nonsmooth problems of the calculus of variations became just an auxiliary tool for the derivation of optimality conditions for nonsmooth optimal control problems and nonsmooth variational problems involving differential inclusions (*cf.* [47, 48, 51]). As a result, relatively little attention has been paid to nonsmooth multidimensional problems of the calculus of variations, as well as problems with additional constraints, such as nonsmooth isoperimetric problems and problems with additional constraints at the boundary. Nonsmooth multidimensional problems of the calculus of variations were first studied by Clarke [15] for locally Lipschitz continuous integrands. Improved versions of the Clarke's first result were later published in monographs [12, 17], while Bousquet [10] showed that one can significantly relax the growth conditions on the integrand imposed in the Clarke's work. Bonfanti and Cellina [9] obtained optimality conditions for the problem

$$\min \int_\Omega L(x, v(x), \nabla v(x)) \, dx \quad \text{subject to} \quad v \in v_0 + W_0^{1,1}(\Omega)$$

in the case when the integrand $L(x, v, \xi)$ is differentiable in v and convex in ξ. Optimality conditions for nonsmooth multidimensional problems of the calculus of variations in terms of codifferentials were obtained in the author's paper [31], while optimality conditions for such problems in terms of the so-called K-subdifferential were obtained in [67]. Finally, nonsmooth variational problems with additional constraints have been explicitly studied only by Clarke [12, 17] and Bellaassali [6].

The main goal of this paper is to present a general theory of necessary optimality conditions for nonsmooth multidimensional problems of the calculus of variations with various types of additional constraints, such as problems with constraints at the boundary and problems with isoperimetric constraints. To this end, we significantly improve our earlier results from [31] and prove the codifferentiability of a nonsmooth integral functional

defined on the Sobolev space under natural and easily verifiable assumptions on the integrand. In comparison with our previous paper [31], we get rid of the obscure and hard to verify assumption on the uniform codifferentiability of the integrand with respect to the Sobolev space and do not impose any assumptions on the domain of integration, thus extending the results of [31] to the case of unbounded domains and domains with irregular boundary (see Sect. 3 for more details). Furthermore, under natural assumptions we prove the continuity of a codifferential of the integral functional in the general case (in [31] the continuity was proved only in the case $p = +\infty$). Continuity is an important property for an analysis of discretisation of variational problems, approximation methods, and convergence of numerical methods. In particular, in the general case the continuity of codifferential is necessary for the global convergence of optimisation methods based on codifferentials [33].

With the use of the general result on the codifferentiability of an integral functional obtained in this paper and necessary optimality conditions for nonsmooth mathematical programming problems in Banach spaces in terms of quasidfferentials from [34, 35] we derive optimality conditions for unconstrained nonsmooth problems of the calculus of variations, as well as problems with additional constraints at the boundary and isoperimetric constraints. Each of these optimality conditions is illustrated by a simple example, in which existing optimality conditions in terms of various subdifferentials are satisfied at a non-optimal point, while our optimality conditions are able to detect the non-optimality of this point. Thus, the optimality conditions obtained in this paper are in some cases better than existing subdifferential-based optimality conditions.

The paper is organised as follows. The codifferentiability of an integral functional defined on the Sobolev space is studied in Section 3. Section 4 is devoted to derivation of necessary optimality conditions for constrained nonsmooth problems of the calculus of variations in terms of codifferentials. This section also contains several examples illustrating advantages of optimality conditions in terms of codifferentials in comparison with subdifferential-based optimality conditions. Finally, Section 2 contains some auxiliary definitions from nonsmooth analysis that are necessary for understanding the paper (apart from Examples 4.2, 4.5, and 4.7, whose understanding requires some familiarity with the Clarke subdifferential [17], the limiting proximal subdifferential [20, 81], and the limiting Fréchet subdifferential [65, 66]).

2. Codifferentiable and quasidifferentiable functions

In what follows, let X be a real Banach space. Its topological dual space is denoted by X^*, while the canonical duality pairing between X and X^* is denoted by $\langle \cdot, \cdot \rangle$, *i.e.* $\langle x^*, x \rangle = x^*(x)$ for all $x^* \in X^*$ and $x \in X$. The standard topology on $\mathbb{R}$ is denoted by $\tau_\mathbb{R}$ and the weak* topology on X^* is denoted by w^* or $\sigma(X^*, X)$.

We equip the Cartesian product $\mathbb{R} \times X$ with the norm $\|(a, x)\| = \sqrt{|a|^2 + \|x\|^2}$ for all $(a, x) \in \mathbb{R} \times X$. It is easily seen that the topological dual space $(\mathbb{R} \times X)^*$ endowed with the weak* topology is isomorphic (in the category of topological vector spaces) to the space $\mathbb{R} \times X^*$ endowed with the product topology $\tau_\mathbb{R} \times w^*$. Utilising this fact (or arguing directly) one can check that a subset of the topological vector space $(\mathbb{R} \times X^*, \tau_\mathbb{R} \times w^*)$ is compact if and only if it is closed in the topology $\tau_\mathbb{R} \times w^*$ and bounded with respect to the norm $\|(a, x^*)\| = \sqrt{|a|^2 + \|x^*\|^2}$, $(a, x^*) \in \mathbb{R} \times X^*$ (see, *e.g.* [30], Thm. 2.1).

Definition 2.1. Let $U \subseteq X$ be an open set. A function $f: U \to \mathbb{R}$ is called *codifferentiable* at a point $x \in U$, if there exists a pair $Df(x) = [\underline{d}f(x), \overline{d}f(x)]$ of convex sets $\underline{d}f(x), \overline{d}f(x) \subset \mathbb{R} \times X^*$ that are compact in the product topology $\tau_\mathbb{R} \times w^*$ and satisfy the equalities $\Phi_f(x, 0) = \Psi_f(x, 0) = 0$ and

$$\lim_{\alpha \to +0} \frac{1}{\alpha} \Big| f(x + \alpha \Delta x) - f(x) - \Phi_f(x, \alpha \Delta x) - \Psi_f(x, \alpha \Delta x) \Big| = 0 \quad \forall \Delta x \in X, \tag{2.1}$$

where

$$\Phi_f(x, \Delta x) = \max_{(a, x^*) \in \underline{d}f(x)} (a + \langle x^*, \Delta x \rangle), \quad \Psi_f(x, \Delta x) = \min_{(b, y^*) \in \overline{d}f(x)} (b + \langle y^*, \Delta x \rangle) \quad \forall \Delta x \in X. \tag{2.2}$$

The pair $Df(x) = [\underline{d}f(x), \overline{d}f(x)]$ is called a *codifferential* of f at x, the set $\underline{d}f(x)$ is called a *hypodifferential* of f at x, while the set $\overline{d}f(x)$ is referred to as a *hyperdifferential* of f at x.

Remark 2.2. Note that the equalities $\Phi_f(x,0) = \Psi_f(x,0) = 0$ simply mean that $a \leq 0$ for all $(a,x^*) \in \underline{d}f(x)$ and $\max_{(a,x^*)\in\underline{d}f(x)} a = 0$, while $b \geq 0$ for all $(b,y^*) \in \overline{d}f(x)$ and $\min_{(b,y^*)\in\overline{d}f(x)} b = 0$. Note also that the maximum in the definition of Φ_f and the minimum in the definition of Ψ_f are attained due to the fact that the sets $\underline{d}f(x)$ and $\overline{d}f(x)$ are compact in the product topology $\tau_{\mathbb{R}} \times w^*$.

Let us comment on the definition of codifferentiability. Observe that the function $\Phi_f(x,\cdot)$ from this definition is convex, while the function $\Psi_f(x,\cdot)$ is concave. Thus, in the definition of codifferentiable function one approximates the increment of a nonsmooth function f with the use of the DC (difference-of-convex) function $\Phi_f(x,\cdot) + \Psi_f(x,\cdot)$ (see (2.1)). One can check that f is codifferentiable at a point x if and only if its increment $f(x + \alpha\Delta x) - f(x)$ can be approximated in this way by *some* continuous DC function (see [32], Exam. 3.10). Hence, in particular, any continuous DC function $f\colon X \to \mathbb{R}$ is codifferentiable at every point $x \in X$. Let us note that the benefit of using codifferentiable functions in comparison with DC functions consists in the existence of a well-developed codifferential calculus, which allows one to easily compute codifferentials of many nonsmooth functions appearing in applications (see, *e.g.* monograph [27]). In contrast, while it is usually fairly easy to prove theoretically that a given function is DC, in some cases it might be very problematic to find an explicit DC representation of a DC function.

Observe that codifferential is not uniquely defined. For instance, it is easily seen that if $Df(x)$ is a codifferential of f at x, then for any convex compact subset C of the space $(\mathbb{R} \times X^*, \tau_{\mathbb{R}} \times w^*)$ the pair $[\underline{d}f(x) + C, \overline{d}f(x) - C]$ is a codifferential of f at x as well.

Recall that for any two nonempty subsets A and B of a metric space (M, d) *the Hausdorff metric* $d_H(A, B)$ is defined by

$$d_H(A, B) = \max\{\sup_{x\in A} \inf_{y\in B} d(x,y), \sup_{y\in B} \inf_{x\in A} d(x,y)\}. \tag{2.3}$$

A multifunction F between metric spaces (M_1, d_1) and (M_2, d_2) is called *Hausdorff continuous* (or continuous with respect to the Hausdorff metric) at a point $x \in M_1$, if for any $\varepsilon > 0$ one can find $\delta > 0$ such that for all $y \in M_1$ with $d_1(y, x) < \delta$ one has $d_H(F(y), F(x)) < \varepsilon$. Let, as above, $U \subseteq X$ be an open set.

Definition 2.3. A function $f\colon U \to \mathbb{R}$ is said to be *continuously codifferentiable* at a point $x \in U$, if f is codifferentiable at every point in a neighbourhood $V \subseteq U$ of x and there exists a codifferential mapping $Df(\cdot)$ defined in V and such that the corresponding set-valued mappings $\underline{d}f(\cdot)$ and $\overline{d}f(\cdot)$ are Hausdorff continuous at x (in this case we say that $Df(\cdot)$ is Hausdorff continuous at x). Finally, f is called continuously codifferentiable on a set $A \subset U$, if it is codifferentiable at every point of this set and there exists a codifferential mapping $Df(\cdot)$ defined on A and such that the corresponding mappings $\underline{d}f(\cdot)$ and $\overline{d}f(\cdot)$ are Hausdorff continuous on A.

The class of continuously codifferentiable functions is closed under addition, multiplication, pointwise maximum and minimum of finite families of functions, and composition with smooth functions (see [27, 30, 32, 33] for more details).

Remark 2.4. Let us note that in the nonsmooth case the standard necessary optimality condition $\nabla f(x) = 0$ takes the form of set-theoretic inclusion $0 \in \partial f(x)$, involving some subdifferential of the objective function f. For codifferentiable functions, necessary optimality conditions are formulated as set-theoretic inclusions involving hypo- and hyperdifferentials:

$$0 \in \underline{d}f(x) + (0, y^*) \quad \forall(0, y^*) \in \overline{d}f(x).$$

As a result, in the nonsmooth case optimality conditions no longer play the role of an equation for finding minimisers and are often used only for verifying whether a given point is optimal and constructing optimisation

methods. As we will see in the following sections, the situation is precisely the same in the case of nonsmooth problems of the calculus of variations. For nonsmooth variational problems, the Euler-Lagrange equation, which in the smooth case can be used to find potential extremals, is replaced by a certain inclusion. This inclusion no longer allows one to directly find extremals and can usually be used only to check whether a given point is potentially optimal and to study numerical procedures for finding potential extremals.

The class of codifferentiable nonsmooth functions is closely related to the class of quasidifferentiable functions. Recall that a function $f \colon U \to \mathbb{R}$ is called *quasidifferentiable* at a point $x \in U$, if f is directionally differentiable at x, *i.e.* for any $v \in X$ there exists the finite limit

$$f'(x, v) = \lim_{\alpha \to +0} \frac{f(x + \alpha v) - f(x)}{\alpha},$$

and the function $f'(x, \cdot)$ can be represented as the difference of continuous sublinear functions or, equivalently, if there exists a pair $\mathscr{D}f(x) = [\underline{\partial} f(x), \overline{\partial} f(x)]$ of convex weak* compact sets $\underline{\partial} f(x), \overline{\partial} f(x) \subset X^*$ such that

$$f'(x, v) = \max_{x^* \in \underline{\partial} f(x)} \langle x^*, v \rangle + \min_{y^* \in \overline{\partial} f(x)} \langle y^*, v \rangle \quad \forall v \in X.$$

The set $\underline{\partial} f(x)$ is called a *subdifferential* of f at x, while the set $\overline{\partial} f(x)$ is referred to as a *superdifferential* of f at x. Note that, just like codifferential, quasidifferential is not uniquely defined. The interesting problem of finding a minimal (in some sense) quasidifferential of a given nonsmooth function was studied in [39, 41–45, 68, 69, 78].

Remark 2.5. In what follows, we denote a codifferential of a nonsmooth function f at a point x by $Df(x)$, while a quasidifferential of this function is denoted by $\mathscr{D}f(x)$.

Finally, one says that a function $f \colon U \to \mathbb{R}$ is *Hadamard quasidifferentiable at x*, if f is quasidifferentiable at x and Hadamard directionally differentiable at this point, that is,

$$f'(x, v) = \lim_{[\alpha, v'] \to [+0, v]} \frac{f(x + \alpha v') - f(x)}{\alpha} \tag{2.4}$$

(see [40] for a discussion of the notation under this limit). In other words, for any $\varepsilon > 0$ and $v \in X$ one can find $\delta > 0$ such that $|(f(x + \alpha v') - f(x))/\alpha - f'(x, v)| < \varepsilon$ for all $0 < \alpha < \delta$ and $v' \in B(v, \delta) = \{v' \in X \mid \|v' - v\| \le \delta\}$. We will need the following result stating that every continuously codifferentiable function is, in fact, Hadamard quasidifferentiable and indicating how one can compute a quasidifferential of this function. A relationship between continuously codifferentiable functions and quasidifferentiable functions having outer semicontinuous subdifferential and superdifferential mappings $\underline{\partial} f(\cdot)$ and $\overline{\partial} f(\cdot)$ was analysed in the finite dimensional case by Kuntz [53].

Lemma 2.6. *Let a function $f \colon U \to \mathbb{R}$ be continuously codifferentiable at a point $x \in U$. Then f is Hadamard quasidifferentiable at this point and for any codifferential $Df(x)$ of f at x the pair $\mathscr{D}f(x) = [\underline{\partial} f(x), \overline{\partial} f(x)]$ defined by*

$$\underline{\partial} f(x) = \left\{ x^* \in X^* \mid (0, x^*) \in \underline{\mathrm{d}} f(x) \right\}, \quad \overline{\partial} f(x) = \left\{ y^* \in X^* \mid (0, y^*) \in \overline{\mathrm{d}} f(x) \right\} \tag{2.5}$$

is a quasidifferential of f at x.

Proof. Let $Df(x)$ be any codifferential of f at x and the functions Φ_f and Ψ_f be defined as in (2.2). Denote $g(\cdot) = \Phi_f(x, \cdot) + \Psi_f(x, \cdot)$. Recall that by the definition of codifferential one has $g(0) = 0$ (see Def. 2.1).

Suppose at first that g is directionally differentiable at zero and fix any $v \in X$ and $\varepsilon > 0$. Then there exists $\delta_1 > 0$ such that $|g(\alpha v)/\alpha - g'(0, v)| < \varepsilon/2$ for all $\alpha \in (0, \delta_1)$. By the definition of codifferential there exists $\delta_2 > 0$

such that for all $\alpha \in (0, \delta_2)$ one has $|f(x + \alpha v) - f(x) - g(\alpha v)| < \varepsilon \alpha / 2$. Therefore, for any $0 < \alpha < \min\{\delta_1, \delta_2\}$ one has

$$\left| \frac{f(x + \alpha v) - f(x)}{\alpha} - g'(0, v) \right| \leq \frac{1}{\alpha} \left| f(x + \alpha v) - f(x) - g(\alpha v) \right| + \left| \frac{g(\alpha v)}{\alpha} - g'(0, v) \right| < \varepsilon,$$

i.e. f is directionally differentiable at x and $f'(x, \cdot) = g'(0, \cdot)$.

Let us now show that both functions $\Phi_f(x, \cdot)$ and $\Psi_f(x, \cdot)$ are directionally differentiable at zero and compute their directional derivatives. Then one obtains that f is directionally differentiable at x and $f'(x, v) = \Phi'_f(x, \cdot)(0, v) + \Psi'_f(x, \cdot)(0, v)$ for all $v \in X$.

As was pointed out above, the sets $\underline{d}f(x)$ and $\overline{d}f(x)$ are norm-bounded due to the fact that they are compact in the product topology $\tau_{\mathbb{R}} \times w^*$. Thus, there exists $K > 0$ such that $|a| + \|x^*\| \leq K$ for any $(a, x^*) \in \underline{d}f(x) \cup \overline{d}f(x)$. Hence

$$|\Phi_f(x, \Delta x)| \leq K + K\|\Delta x\|, \quad |\Psi_f(x, \Delta x)| \leq K + K\|\Delta x\| \quad \forall \Delta x \in X,$$

i.e. the functions $\Phi_f(x, \cdot)$ and $\Psi_f(x, \cdot)$ are bounded on bounded sets. Therefore, the convex function $\Phi_f(x, \cdot)$ is continuous ([37], Prop. I.2.5), subdifferentiable on X ([37], Prop. I.5.2), everywhere directionally differentiable ([50], Prop. 4.1.4), its subdifferential at zero has the form

$$\partial \Phi_f(x, \cdot)(0) = \{x^* \in X^* \mid (0, x^*) \in \underline{d}f(x)\}$$

by Thereom 4.2.3 of [50] (recall that by definition $\Phi_f(x, 0) = 0$; see Def. 2.1), and for all $v \in X$ one has $\Phi'(x, \cdot)(0, v) = \max_{x^* \in \partial \Phi_f(x, \cdot)(0)} \langle x^*, v \rangle$ by Proposition 4.1.1 of [50]. Similarly, the concave function $\Psi_f(x, \cdot)$ is everywhere directionally differentiable and

$$\Psi'_f(x, \cdot)(0, v) = \min_{y^* \in \partial \Psi_f(x, \cdot)(0)} \langle y^*, v \rangle \quad \forall v \in X,$$

where $\partial \Psi_f(x, \cdot)(0) = \{y^* \in X^* \mid (0, y^*) \in \overline{d}f(x)\}$. Thus, one can conclude that f is quasidifferentiable at x and the pair (2.5) is a quasidifferential of f at x.

Let us finally show that f is Hadamard directionally differentiable at x. Indeed, by Corollary 2 [33] the function f is Lipschitz continuous near x, *i.e.* there exist $r > 0$ and $L > 0$ such that $|f(x_1) - f(x_2)| \leq L\|x_1 - x_2\|$ for all $x_1, x_2 \in B(x, r)$.

Fix any $\varepsilon > 0$ and $v \in X$. Note that $f'(x, 0) = 0$ and $x + \alpha v' \in B(x, r)$ for any $\alpha \in (0, \sqrt{r})$ and $v' \in X$ with $\|v'\| \leq \sqrt{r}$. Therefore for $\delta = \min\{\sqrt{r}, \varepsilon / 2L\}$ one has

$$\left| \frac{f(x + \alpha v') - f(x)}{\alpha} - f'(x, 0) \right| \leq L\|v'\| < \varepsilon \quad \forall \alpha \in (0, \delta), \ v' \in B(0, \delta),$$

i.e. (2.4) with $v = 0$ holds true. Thus, one can suppose that $v \neq 0$.

By the definition of directional derivative there exists $\delta_0 > 0$ such that

$$\left| \frac{f(x + \alpha v) - f(x)}{\alpha} - f'(x, v) \right| < \frac{\varepsilon}{2} \quad \forall \alpha \in (0, \delta_0).$$

Observe that if $0 < \alpha < r/(2\|v\|)$ and for some $v' \in X$ one has $\|v' - v\| < \|v\|$, then $x + \alpha v \in B(x, r/2)$ and $\|\alpha v' - \alpha v\| = \alpha \|v' - v\| < r/2$, *i.e.* $x + \alpha v' \in B(x, r)$. Therefore, put $\delta = \min\{\delta_0, \varepsilon/2L, r/(2\|v\|), \|v\|\}$. Then for

any $\alpha \in (0, \delta)$ and $v' \in B(v, \delta)$ one has $x + \alpha v' \in B(x, r)$ and

$$\left| \frac{f(x + \alpha v') - f(x)}{\alpha} - f'(x, v) \right| \leq \left| \frac{f(x + \alpha v) - f(x)}{\alpha} - f'(x, v) \right| + \frac{1}{\alpha} \left| f(x + \alpha v') - f(x + \alpha v) \right| \leq \frac{\varepsilon}{2} + L \|v' - v\| < \varepsilon.$$

Thus, f is Hadamard quasidifferentiable at x. $\square$

Remark 2.7. From the proof of the lemma above it follows that if the function f is codifferentiable, but *not* continuously codifferentiable at x, then f is still quasidifferentiable at x and for any codifferential $Df(x)$ of f at x the pair (2.5) is a quasidifferential of f at x.

3. CODIFFERENTIABILITY OF INTEGRAL FUNCTIONALS

In this section we present simple sufficient conditions for the codifferentiability of integral functional

$$\mathcal{I}(u) = \int_\Omega f(x, u(x), \nabla u(x)) \, dx, \quad u = (u_1, \ldots, u_m) \in W^{1,p}(\Omega; \mathbb{R}^m)$$

and compute its codifferential and quasidifferential. Here $\Omega \subseteq \mathbb{R}^d$ is an open set (not necessarily bounded), $f: \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to \mathbb{R}$, $f = f(x, u, \xi)$, is a given function, $W^{1,p}(\Omega; \mathbb{R}^m)$ is the Cartesian product of m copies of the Sobolev space $W^{1,p}(\Omega)$ with $1 \leq p \leq +\infty$. The space $W^{1,p}(\Omega; \mathbb{R}^m)$ is endowed with the norm $\|u\|_{1,p} = \left(\|u\|_p^p + \|\nabla u\|_p^p \right)^{1/p}$ in the case $1 \leq p < +\infty$ and $\|u\|_{1,\infty} = \max\{\|u\|_\infty, \|\nabla u\|_\infty\}$, where $\|\cdot\|_p$ is the standard norm in $L^p(\Omega; \mathbb{R}^k)$ for any $k \in \mathbb{N}$, *i.e.* $\|u\|_p = \left(\int_\Omega |u(x)|^p \, dx \right)^{\frac{1}{p}}$ in the case $1 \leq p < +\infty$, and $\|u\|_\infty = \operatorname{ess\,sup}_{x \in \Omega} |u(x)|$ (here $|\cdot|$ is the Euclidean norm). Denote by p' the conjugate exponent of p, *i.e.* $1/p + 1/p' = 1$.

Below we assume that for a.e. $x \in \Omega$ the function $(u, \xi) \mapsto f(x, u, \xi)$ is codifferentiable, *i.e.* for a.e. $x \in \Omega$ and for all $(u, \xi) \in \mathbb{R}^m \times \mathbb{R}^{m \times d}$ there exist compact convex sets $\underline{d}_{u,\xi} f(x, u, \xi), \overline{d}_{u,\xi} f(x, u, \xi) \subset \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^{m \times d}$ such that for any $(\Delta u, \Delta \xi) \in \mathbb{R}^m \times \mathbb{R}^{m \times d}$ one has

$$\lim_{\alpha \to +0} \frac{1}{\alpha} \left| f(x, u + \alpha \Delta u, \xi + \alpha \Delta \xi) - f(x, u, \xi) - \Phi_f(x, u, \xi; \alpha \Delta u, \alpha \Delta \xi) - \Psi_f(x, u, \xi; \alpha \Delta u, \alpha \Delta \xi) \right| = 0$$

and $\Phi_f(x, u, \xi; 0, 0) = \Psi_f(x, u, \xi; 0, 0) = 0$, where

$$\Phi_f(x, u, \xi; \Delta u, \Delta \xi) = \max_{(a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u, \xi)} \left(a + \langle v_1, \Delta u \rangle + \langle v_2, \Delta \xi \rangle \right), \tag{3.1}$$

$$\Psi_f(x, u, \xi; \alpha \Delta u, \alpha \Delta \xi) = \min_{(b, w_1, w_2) \in \overline{d}_{u,\xi} f(x, u, \xi)} \left(b + \langle w_1, \Delta u \rangle + \langle w_2, \Delta \xi \rangle \right), \tag{3.2}$$

and $\langle \cdot, \cdot \rangle$ is the inner product in $\mathbb{R}^k$. We denote a codifferential of this function at a point (x, u, ξ) by $D_{u,\xi} f(x, u, \xi) = [\underline{d}_{u,\xi} f(x, u, \xi), \overline{d}_{u,\xi} f(x, u, \xi)]$. Finally, recall that a multifunction $F: \Omega \times Y \rightrightarrows Z$, where Y and Z are metric spaces, is called *a Carathéodory map*, if for every $y \in Y$ the map $F(\cdot, y)$ is measurable and for every $x \in \Omega$ the map $F(x, \cdot)$ is continuous (see [2], Def. 8.2.7). As is well-known, in the case when $Z = \mathbb{R}^k$ and F is compact-valued, the map $F(x, \cdot)$ is continuous iff it is Hausdorff continuous.

The following definition describes natural assumptions on the integrand $f = f(x, u, \xi)$ ensuring the codifferentiability of the functional $\mathcal{I}$.

Definition 3.1. We say that f satisfies *the codifferentiability conditions* of order p, if

(1) f is a Carathéodory function satisfying *the growth condition* of order p, *i.e.* there exist an a.e. nonnegative function $\beta \in L^1(\Omega)$ and $C > 0$ such that

$$|f(x, u, \xi)| \leq \beta(x) + C \left(|u|^p + |\xi|^p \right)$$

for a.e. $x \in \Omega$ and for all $(u, \xi) \in \mathbb{R}^m \times \mathbb{R}^{m \times d}$ in the case $1 \le p < +\infty$, and for any $N > 0$ there exists an a.e. nonnegative function $\beta_N \in L^1(\Omega)$ such that $|f(x, u, \xi)| \le \beta_N(x)$ for a.e. $x \in \Omega$ and all $(u, \xi) \in \mathbb{R}^m \times \mathbb{R}^{m \times d}$ with $\max\{|u|, |\xi|\} \le N$ in the case $p = +\infty$;

(2) for a.e. $x \in \Omega$ the function $(u, \xi) \mapsto f(x, u, \xi)$ is codifferentiable on $\mathbb{R}^m \times \mathbb{R}^{m \times d}$ and its codifferential mapping $D_{u,\xi} f(\cdot)$ is a Carathéodory map (*i.e.* both $\underline{d}_{u,\xi} f(\cdot)$ and $\overline{d}_{u,\xi} f(\cdot)$ are Carathéodory maps) satisfying *the growth condition* of order p, *i.e.* there exist $C > 0$ and a.e. nonnegative functions $\beta \in L^1(\Omega)$ and $\gamma \in L^{p'}(\Omega)$ such that for a.e. $x \in \Omega$ and for all $(u, \xi) \in \mathbb{R}^m \times \mathbb{R}^{m \times d}$, $(a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u, \xi)$, and $(b, w_1, w_2) \in \overline{d}_{u,\xi} f(x, u, \xi)$ one has

$$\max\left\{|a|, |b|\right\} \le \beta(x) + C(|u|^p + |\xi|^p), \quad \max\left\{|v_1|, |v_2|, |w_1|, |w_2|\right\} \le \gamma(x) + C(|u|^{p-1} + |\xi|^{p-1})$$

in the case $1 \le p < +\infty$, and for any $N > 0$ there exists an a.e. nonnegative function $\beta_N \in L^1(\Omega)$ such that $\max\left\{|a|, |v_1|, |v_2|, |b|, |w_1|, |w_2|\right\} \le \beta_N(x)$ for a.e. $x \in \Omega$ and for all $(a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u, \xi)$, $(b, w_1, w_2) \in \overline{d}_{u,\xi} f(x, u, \xi)$, and $(u, \xi) \in \mathbb{R}^m \times \mathbb{R}^{m \times d}$ with $\max\{|u|, |\xi|\} \le N$ in the case $p = +\infty$.

Remark 3.2. In the case $1 \le p < +\infty$ the growth conditions from the previous definition can be weakened with the use of the Sobolev imbedding theorem (*cf.* [22], Sect. 3.4.2). In particular, if $d = 1$, then it is sufficient to suppose that for any $N > 0$ there exists $C_N > 0$ and a.e. nonnegative functions $\beta_N \in L^1(\Omega)$ and $\gamma_N \in L^{p'}(\Omega)$ such that for a.e. $x \in \Omega$, for all $(u, \xi) \in \mathbb{R}^m \times \mathbb{R}^m$ with $|u| \le N$ and for all $(a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u, \xi) \cup \overline{d}_{u,\xi} f(x, u, \xi)$ one has $|a| \le \beta_N(x) + C_N |\xi|^p$ and $\max\{|v_1|, |v_2|\} \le \gamma_N(x) + C_N |\xi|^{p-1}$.

Our aim is to prove that the codifferentiability conditions from the definition above guarantee that the functional $\mathcal{I}$ is codifferentiable on $W^{1,p}(\Omega; \mathbb{R}^m)$. Due to some technical difficulties, in the case $p = 1$ we need to assume that the set Ω has *the segment property* [1], p. 53–54, *i.e.* that for every x from the boundary of Ω there exist a neighbourhood $U_x \subset \mathbb{R}^d$ of x and a nonzero vector $y_x \in \mathbb{R}^d$ such that for any $z \in \text{cl}\,\Omega \cap U_x$ one has $z + t y_x \in \Omega$ for all $t \in (0, 1)$. The segment property ensures that the set Ω has a $(d-1)$-dimensional boundary and cannot simultaneously lie on both sides of any given part of its boundary (*i.e.* there are no cuts). Furthermore, it ensures that the space of continuously differentiable functions is dense in $W^{1,p}(\Omega)$ (see, *e.g.* [1], Thm. 3.18).

Theorem 3.3. *Let f satisfy the codifferentiability conditions of order $p \in [1, +\infty]$ and let either $1 < p \le +\infty$ or the set Ω be bounded and have the segment property. Then the functional $\mathcal{I}$ is correctly defined on $W^{1,p}(\Omega; \mathbb{R}^m)$, codifferentiable at every $u \in W^{1,p}(\Omega; \mathbb{R}^m)$, and the pair $D\mathcal{I}(u) = [\underline{d}\mathcal{I}(u), \overline{d}\mathcal{I}(u)]$ with*

$$\underline{d}\mathcal{I}(u) = \left\{ (A, x^*) \in \mathbb{R} \times (W^{1,p}(\Omega; \mathbb{R}^m))^* \,\Big|\, A = \int_\Omega a(x)\,dx, \quad \langle x^*, h \rangle = \int_\Omega \left(\langle v_1(x), h(x) \rangle + \langle v_2(x), \nabla h(x) \rangle \right) dx \right.$$

$$\left. \forall h \in W^{1,p}(\Omega; \mathbb{R}^m), \quad (a(\cdot), v_1(\cdot), v_2(\cdot)) \text{ is a measurable selection of the map } \underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot)) \right\} \quad (3.3)$$

and

$$\overline{d}\mathcal{I}(u) = \left\{ (B, y^*) \in \mathbb{R} \times (W^{1,p}(\Omega; \mathbb{R}^m))^* \,\Big|\, B = \int_\Omega b(x)\,dx, \quad \langle y^*, h \rangle = \int_\Omega \left(\langle w_1(x), h(x) \rangle + \langle w_2(x), \nabla h(x) \rangle \right) dx \right.$$

$$\left. \forall h \in W^{1,p}(\Omega; \mathbb{R}^m), \quad (b(\cdot), w_1(\cdot), w_2(\cdot)) \text{ is a measurable selection of the map } \overline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot)) \right\}.$$

is a codifferential of $\mathcal{I}$ at u. Furthermore, the multifunctions $\underline{d}\mathcal{I}(\cdot)$ and $\overline{d}\mathcal{I}(\cdot)$ are Hausdorff continuous, i.e. the functional $\mathcal{I}$ is continuously codifferentiable on $W^{1,p}(\Omega; \mathbb{R}^m)$, provided either $1 < p \le +\infty$ or the set-valued maps $\underline{d}_{u,\xi} f(\cdot)$ and $\overline{d}_{u,\xi} f(\cdot)$ have the form

$$\underline{d}_{u,\xi} f(x, u, \xi) = \text{co}\left\{ (f_i(x, u, \xi), v_{1i}, v_{2i}) \mid i \in I \right\}, \quad \overline{d}_{u,\xi} f(x, u, \xi) = \text{co}\left\{ (g_j(x, u, \xi), w_{1j}, w_{2j}) \mid j \in J \right\} \quad (3.4)$$

for some vectors $v_{1i}, w_{1j} \in \mathbb{R}^m$, $v_{2i}, w_{2j} \in \mathbb{R}^{m \times d}$ and Carathéodory functions $f_i, g_j \colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to \mathbb{R}$, where $i \in I = \{1, \ldots, \ell\}$ and $j \in J = \{1, \ldots, r\}$.

Remark 3.4. The assumption that in the case $p = 1$ the set-valued maps $\underline{d}_{u,\xi} f(\cdot)$ and $\overline{d}_{u,\xi} f(\cdot)$ have the form (3.4) might seem unnatural at first glance. However, it should be noted that this assumption is satisfied in many particular examples. Furthermore, with the use of the codifferential calculus [27] one can easily show that this assumption is preserved under addition, multiplication by scalar, pointwise maximum, and pointwise minimum. For example, if $d = m = 1$, $\Omega = (\alpha, \beta)$, and

$$\mathcal{I}(u) = \int_\alpha^\beta \max\{|u'(x)| - |u(x)|, 0\} \, \mathrm{d}x,$$

then $f(x, u, \xi) = \max\{|\xi| - |u|, 0\} = \max\{|\xi|, |u|\} + \min\{u, -u\}$ and by applying the codifferential calculus [27] one gets

$$\underline{d}_{u,\xi} f(x, u, \xi) = \mathrm{co}\left\{\begin{pmatrix} \pm\xi - \max\{|\xi|, |u|\} \\ 0 \\ \pm 1 \end{pmatrix}, \begin{pmatrix} \pm u - \max\{|\xi|, |u|\} \\ \pm 1 \\ 0 \end{pmatrix}\right\}, \quad \overline{d}_{u,\xi} f(x, u, \xi) = \mathrm{co}\left\{\begin{pmatrix} \pm u + |u| \\ \pm 1 \\ 0 \end{pmatrix}\right\},$$

i.e. assumption (3.4) is satisfied.

We split the proof of Theorem 3.3 into four parts, each of which is formulated as a separate lemma. Before we proceed to these lemmas, it should be remarked that Theorem 3.3 significantly improves ([31], Thm. 5.1), since it states that the functional $\mathcal{I}$ is *continuously* codifferentiable for any $1 \le p \le +\infty$ and demonstrates that the rather restrictive and obscure assumption on the uniform codifferentiability of the integrand f with respect to the space $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ (see [31], Def. 4.18) is redundant. Furthermore, in the case $1 < p \le +\infty$ Theorem 3.3 extends ([31], Thm. 5.1) to the case of unbounded domains Ω and domains not having a segment property. Finally, Theorem 3.3 gives a positive answer to the second question raised by the author in Remark 4.22 of [31].

We start with a simple technical lemma on the function Φ_f defined in (3.1).

Lemma 3.5. *Suppose that for a.e. $x \in \Omega$ and for all $(u, \xi) \in \mathbb{R}^m \times \mathbb{R}^{m \times d}$ the function $(u, \xi) \mapsto f(x, u, \xi)$ is codifferentiable and let Φ_f be defined as in (3.1). Then for any $u, h \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$, for a.e. $x \in \Omega$, and for all $t \in \mathbb{R}$ the function $g(t) = \Phi_f(x, u(x), \nabla u(x); th(x), t\nabla h(x))$ is codifferentiable and for any $t \in \mathbb{R}$ the pair $Dg(t) = [\underline{d}g(t), \{0\}]$ is a codifferential of g at t, where*

$$\underline{d}g(t) = \Big\{ (a_g, v_g) \in \mathbb{R} \times \mathbb{R} \ \Big| \ a_g = a + \langle v_1, th(x) \rangle + \langle v_2, t\nabla h(x) \rangle - g(t),$$
$$v_g = \langle v_1, h(x) \rangle + \langle v_2, \nabla h(x) \rangle, \ (a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u(x), \nabla u(x)) \Big\}.$$

Proof. By the definition of Φ_f for any $t, \Delta t \in \mathbb{R}$ one has

$$g(t + \Delta t) - g(t) = \max\big(a + \langle v_1, (t + \Delta t)h(x) \rangle + \langle v_2, (t + \Delta t)\nabla h(x) \rangle\big) - g(t)$$
$$= \max\Big(\big[a + \langle v_1, th(x) \rangle + \langle v_2, t\nabla h(x) \rangle - g(t)\big] + \Delta t\langle v_1, h(x) \rangle + \Delta t\langle v_2, \nabla h(x) \rangle\Big) = \max_{(a_g, v_g) \in \underline{d}g(t)}(a_g + v_g\Delta t),$$

where the first two maximums are taken over all $(a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u(x), \nabla u(x))$ and the set $\underline{d}g(t)$ is defined in the formulation of the lemma. The set $\underline{d}g(t)$ is obviously convex and compact as the image of the set $\underline{d}_{u,\xi} f(x, u(x), \nabla u(x))$ under the affine map

$$(a, v_1, v_2) \mapsto \Big(a + \langle v_1, th(x) \rangle + \langle v_2, t\nabla h(x) \rangle - g(t), \langle v_1, h(x) \rangle + \langle v_2, \nabla h(x) \rangle\Big) \in \mathbb{R} \times \mathbb{R}.$$

Moreover, one has

$$\max_{(a_g,v_g)\in\underline{d}g(t)} a_g = \max\big(a + \langle v_1, th(x)\rangle + \langle v_2, t\nabla h(x)\rangle\big) - g(t) = g(t) - g(t) = 0,$$

where the second maximum is taken over all $(a, v_1, v_2) \in \underline{d}_{u,\xi}f(x, u(x), \nabla u(x))$. Thus, the function g is codifferentiable at every $t \in \mathbb{R}$ and the pair $Dg(\cdot) = [\underline{d}g(\cdot), \{0\}]$ is its codifferential (see Def. 2.1). $\square$

Next we show that the increment $\mathcal{I}(u + \alpha h) - \mathcal{I}(u)$ of the functional $\mathcal{I}$ can be approximated by a DC function defined via the sets $\underline{d}\mathcal{I}(u)$ and $\overline{d}\mathcal{I}(u)$ from Theorem 3.3. Note that these sets are nonempty, since measurable selections of the multifunctions $\underline{d}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$ and $\overline{d}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$ exist ([2], Thms. 8.1.3 and 8.2.8).

The statement of the following lemma coincides with that of Lemma 5.2 in [31] with the only difference being the fact that here were remove the obscure assumption on uniform codifferentiability of the integrand with respect to the Sobolev space. In [31] this assumption ensured that one can pass to the limit under the integral sign. Below we prove that one can pass to the limit without this assumption by applying Lebesgue's dominated convergence theorem and obtaining necessary estimates with the use of the mean value theorem for codifferentiable functions ([33], Prop. 2).

Lemma 3.6. *Let f satisfy the codifferentiability conditions of order $p \in [1, +\infty]$ and the sets $\underline{d}\mathcal{I}(u)$ and $\overline{d}\mathcal{I}(u)$ be defined as in Theorem 3.3. Then the functional $\mathcal{I}$ is correctly defined on $W^{1,p}(\Omega; \mathbb{R}^m)$, $\underline{d}\mathcal{I}(u), \overline{d}\mathcal{I}(u) \subset \mathbb{R} \times (W^{1,p}(\Omega; \mathbb{R}^m))^*$, and*

$$\lim_{\alpha \to +0} \frac{1}{\alpha}\bigg|\mathcal{I}(u + \alpha h) - \mathcal{I}(u) - \max_{(A,x^*)\in\underline{d}\mathcal{I}(u)}(A + \langle x^*, \alpha h\rangle) - \min_{(B,y^*)\in\overline{d}\mathcal{I}(u)}(B + \langle y^*, \alpha h\rangle)\bigg| = 0. \tag{3.5}$$

for all $u, h \in W^{1,p}(\Omega; \mathbb{R}^m)$.

Proof. Fix any $u \in W^{1,p}(\Omega; \mathbb{R}^m)$. By our assumption f is a Carathéodory function satisfying the growth condition (see Def. 3.1). Therefore, as is well-known, the function $f(\cdot, u(\cdot), \nabla u(\cdot))$ is measurable and belongs to $L^1(\Omega)$, which implies that $\mathcal{I}(u)$ is correctly defined and finite.

Let us verify that the sets $\underline{d}\mathcal{I}(u)$ and $\overline{d}\mathcal{I}(u)$ are correctly defined. Indeed, fix any measurable selection $(a(\cdot), v_1(\cdot), v_2(\cdot))$ of the multifunction $\underline{d}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$. By the growth condition on $D_{u,\xi}f(\cdot)$ (see Def. 3.1) there exist $C > 0$ and a.e. nonnegative functions $\beta \in L^1(\Omega)$ and $\gamma \in L^{p'}(\Omega)$ such that

$$|a(x)| \le \beta(x) + C\big(|u(x)|^p + |\nabla u(x)|^p\big), \quad \max\big\{|v_1(x)|, |v_2(x)|\big\} \le \gamma(x) + C\big(|u(x)|^{p-1} + |\nabla u(x)|^{p-1}\big)$$

for a.e. $x \in \Omega$ in the case $1 \le p < +\infty$, and there exists $\beta_N \in L^1(\Omega)$ such that for a.e. $x \in \Omega$ one has $\max\{|a(x)|, |v_1(x)|, |v_2(x)|\} \le \beta_N(x)$ in the case $p = +\infty$ (here $N = \|u\|_{1,\infty}$). Hence with the use of Hölder's inequality one obtains that $a(\cdot) \in L^1(\Omega)$, $v_1(\cdot) \in L^{p'}(\Omega; \mathbb{R}^m)$, and $v_2(\cdot) \in L^{p'}(\Omega; \mathbb{R}^{m \times d})$. Therefore the integral $A = \int_\Omega a(x)\mathrm{d}x$ is correctly defined and finite, while the functional x^* defined as

$$\langle x^*, h\rangle = \int_\Omega \big(\langle v_1(x), h(x)\rangle + \langle v_2(x), \nabla h(x)\rangle\big)\mathrm{d}x \quad \forall h \in W^{1,p}(\Omega; \mathbb{R}^m)$$

is a continuous linear functional on $W^{1,p}(\Omega; \mathbb{R}^m)$, *i.e.* the set $\underline{d}\mathcal{I}(u)$ is correctly defined and $\underline{d}\mathcal{I}(u) \subset \mathbb{R} \times (W^{1,p}(\Omega; \mathbb{R}^m))^*$. The fact that $\overline{d}\mathcal{I}(u) \subset \mathbb{R} \times (W^{1,p}(\Omega; \mathbb{R}^m))^*$ is proved in the same way.

Choose any $h \in W^{1,p}(\Omega; \mathbb{R}^m)$ and a sequence $\{\alpha_n\} \subset (0, +\infty)$ converging to zero. Let us prove that

$$\lim_{n \to \infty} \frac{1}{\alpha_n}\bigg|\mathcal{I}(u + \alpha_n h) - \mathcal{I}(u) - \int_\Omega \Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big)\mathrm{d}x$$

$$- \int_\Omega \Psi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big) \, dx \bigg| = 0, \quad (3.6)$$

where the functions Φ_f and Ψ_f are defined in (3.1), (3.2). Indeed, for any $n \in \mathbb{N}$ and $x \in \Omega$ denote

$$g_n(x) = \frac{1}{\alpha_n}\Big(f(x, u(x) + \alpha_n h(x), \nabla u(x) + \alpha_n \nabla h(x)) - f(x, u(x), \nabla u(x))$$
$$- \Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big) - \Psi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big)\Big). \quad (3.7)$$

Our aim is to prove (3.6) by applying Lebesgue's dominated convergence theorem to the sequence of functions $\{g_n\}$. Firstly, note that by the definition of codifferential $g_n(x) \to 0$ as $n \to \infty$ for a.e. $x \in \Omega$. Next, we show that $g_n \in L^1(\Omega)$ for all $n \in \mathbb{N}$.

From the fact that the integrand f satisfied the growth condition it obviously follows that the first two terms in the definition of g_n belong to $L^1(\Omega)$. Let us check that the function $\eta_n(x) = \Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big)$, $x \in \Omega$, belongs to $L^1(\Omega)$ as well. The proof of this fact for the function Ψ_f is exactly the same.

By the codifferentiability conditions $\underline{d}_{u,\xi} f(\cdot)$ is a Carathéodory map, which by Thereom 8.2.8 of [2] implies that the multifunction $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ is measurable. The map $(x, (a, v_1, v_2)) \mapsto \langle a + \alpha_n \langle v_1, h(x) \rangle + \alpha_n \langle v_2, \nabla h(x) \rangle$ is obviously a Carathéodory function. Hence by the definitions of Φ_f (see (3.1)) and η_n and the theorem on the measurability of marginal functions ([2], Thm. 8.2.11) one obtains that the function η_n is measurable. Moreover, by the growth condition on the codifferential mapping $D_{u,\xi} f(\cdot)$ (see Def. 3.1) there exist $C > 0$ and a.e. nonnegative functions $\beta \in L^1(\Omega)$ and $\gamma \in L^{p'}(\Omega)$ such that

$$|\eta_n(x)| = \big|\Phi_f(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x))\big| \leq \beta(x) + C\big(|u(x)|^p + |\nabla u(x)|^p\big)$$
$$+ \alpha_n\big(\gamma(x) + C(|u(x)|^{p-1} + |\nabla u(x)|^{p-1})\big)\big(|h(x)| + |\nabla h(x)|\big)$$

for a.e. $x \in \Omega$ in the case $1 \leq p < +\infty$, and there exists an a.e. nonnegative function $\beta_N \in L^1(\Omega)$ such that

$$|\eta_n(x)| \leq \beta_N(x)\big(1 + \alpha_n |h(x)| + \alpha_n |\nabla h(x)|\big)$$

for a.e. $x \in \Omega$ in the case $p = +\infty$ (here $N = \|u\|_{1,\infty}$). Hence taking into account the fact that $u, h \in W^{1,p}(\Omega; \mathbb{R}^m)$ and applying Hölder's inequality in the case $1 < p < +\infty$ one obtains that $\eta_n \in L^1(\Omega)$ for all $n \in \mathbb{N}$, which implies that $g_n \in L^1(\Omega)$ for all $n \in \mathbb{N}$ as well.

Now we prove that the sequence $\{g_n\}$ is dominated by some integrable function. Indeed, by the mean value theorem for codifferentiable functions ([33], Prop. 2) for any $n \in \mathbb{N}$ and for a.e. $x \in \Omega$ one can find $\theta_n(x) \in (0, \alpha_n)$, and triplets $(0, v_{1n}(x), v_{2n}(x)) \in \underline{d}_{u,\xi} f(x, u(x) + \theta_n(x)h(x), \nabla u(x) + \theta_n(x)\nabla h(x))$, and $(0, w_{1n}(x), w_{2n}(x)) \in \overline{d}_{u,\xi} f(x, u(x) + \theta_n(x)h(x), \nabla u(x) + \theta_n(x)\nabla h(x))$ such that

$$\frac{1}{\alpha_n}\big(f(x, u(x) + \alpha_n h(x), \nabla u(x) + \alpha_n \nabla h(x)) - f(x, u(x), \nabla u(x))\big)$$
$$= \langle v_{1n}(x) + w_{1n}(x), h(x) \rangle + \langle v_{2n}(x) + w_{2n}(x), \nabla h(x) \rangle.$$

Hence by the growth condition on $D_{u,\xi} f(\cdot)$ (see Def. 3.1) there exist $C > 0$ and a.e. nonnegative function $\gamma \in L^{p'}(\Omega)$ such that

$$\frac{1}{\alpha_n}\big|f(x, u(x) + \alpha_n h(x), \nabla u(x) + \alpha_n \nabla h(x)) - f(x, u(x), \nabla u(x))\big|$$

$$\leq 2\big(\gamma(x) + C(|u(x) + \theta_n(x)h(x)|^{p-1} + |\nabla u(x) + \theta_n(x)\nabla h(x)|^{p-1})\big)\big(|h(x)| + |\nabla h(x)|\big)$$

for a.e. $x \in \Omega$ in the case $1 \leq p < +\infty$, and there exists an a.e. nonnegative function $\beta_N \in \mathrm{L}^1(\Omega)$ such that

$$\frac{1}{\alpha_n}\big|f(x, u(x) + \alpha_n h(x), \nabla u(x) + \alpha_n \nabla h(x)) - f(x, u(x), \nabla u(x))\big| \leq 2\beta_N(x)\big(|h(x)| + |\nabla h(x)|\big)$$

for a.e. $x \in \Omega$ in the case $p = +\infty$, where $N = \|u\|_{1,\infty} + \alpha_*\|h\|_{1,\infty}$ and $\alpha_* = \max_{n \in \mathbb{N}} \alpha_n$. Now, taking into account the fact that $u, h \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ and applying Hölder's inequality in the case $1 < p < +\infty$ one gets that the first two terms in (3.7) are dominated by an integrable function independent of n.

Let us now turn to the third term in (3.7). The fact that the last term is dominated by an integrable function can be proved in exactly the same way. By applying the mean value theorem for codifferentiable functions and Lemma 3.5 one obtains that for any $n \in \mathbb{N}$ and for a.e. $x \in \Omega$ there exist $\theta_n(x) \in (0, \alpha_n)$ and $(a_n(x), v_{1n}(x), v_{2n}(x)) \in \underline{\mathrm{d}}_{u,\xi} f(x, u(x), \nabla u(x))$ such that

$$\Phi_f(x, u(x), \nabla u(x); \theta_n(x)h(x), \theta_n(x)\nabla h(x)) = a_n(x) + \langle v_{1n}(x), \theta_n(x)h(x)\rangle + \langle v_{2n}(x), \theta_n(x)\nabla h(x)\rangle$$

$$\frac{1}{\alpha_n}\Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big) = \langle v_{1n}(x), h(x)\rangle + \langle v_{2n}(x), \nabla h(x)\rangle$$

(here we used the fact that $\Phi_f(x, u(x), \nabla u(x); 0, 0) = 0$ by the definition of codifferential). Hence utilising the growth condition on $D_{u,\xi}f(\cdot)$ in the same way as above one can easily verify that the third term in (3.7) is dominated by an integrable function independent of n as well. Consequently, applying Lebesgue's dominated convergence theorem one obtains that $\int_\Omega g_n(x)\,\mathrm{d}x \to 0$ as $n \to \infty$ or, equivalently, (3.6) holds true (see the definition of g_n, formula (3.7)).

Let us check that

$$\int_\Omega \Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big)\,\mathrm{d}x = \max_{(A, x^*) \in \underline{\mathrm{d}}\mathcal{I}(u)}\big(A + \langle x^*, \alpha_n h\rangle\big) \tag{3.8}$$

for all $n \in \mathbb{N}$, where $\underline{\mathrm{d}}\mathcal{I}(u)$ is defined in Theorem 3.3. The validity of a similar equality involving Ψ_f and $\overline{\mathrm{d}}\mathcal{I}(u)$ can be proved in the same way. Then applying (3.6) one obtains that equality (3.5) holds true and the proof is complete.

By the definition of Φ_f for any measurable selection $(a(\cdot), v_1(\cdot), v_2(\cdot))$ of the multifunction $\underline{\mathrm{d}}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ one has

$$a(x) + \alpha_n \langle v_1(x), h(x)\rangle + \alpha_n \langle v_2(x), \nabla h(x)\rangle \leq \Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big)$$

for a.e. $x \in \Omega$ and for all $n \in \mathbb{N}$, which obviously implies that the inequality

$$\sup_{(A, x^*) \in \underline{\mathrm{d}}\mathcal{I}(u)}\big(A + \langle x^*, \alpha_n h\rangle\big) \leq \int_\Omega \Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big)\,\mathrm{d}x$$

holds true for all $n \in \mathbb{N}$ (see (3.3)). On the other hand, observe that by definition

$$\Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big) \in \Big\{a + \alpha_n\langle v_1, h(x)\rangle + \alpha_n\langle v_2, \nabla h(x)\rangle \in \mathbb{R} \ \Big|$$

$$(a, v_1, v_2) \in \underline{\mathrm{d}}_{u,\xi} f(x, u(x), \nabla u(x))\Big\}$$

for a.e. $x \in \Omega$ and for all $n \in \mathbb{N}$. As was noted above, from the codifferentiability conditions in follows that the multifunction $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ is measurable. Therefore, by Filippov's theorem ([2], Thm. 8.2.10) for any $n \in \mathbb{N}$ there exists a measurable selection $(a(\cdot), v_1(\cdot), v_2(\cdot))$ of the set-valued map $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ such that

$$\Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big) = a(x) + \alpha_n \langle v_1(x), h(x) \rangle + \alpha_n \langle v_2(x), \nabla h(x) \rangle$$

for a.e. $x \in \Omega$, which implies that for the corresponding element $(A, x^*) \in \underline{d}\mathcal{I}(u)$ (see (3.3)) one has

$$\int_\Omega \Phi_f\big(x, u(x), \nabla u(x); \alpha_n h(x), \alpha_n \nabla h(x)\big)\, \mathrm{d}x = A + \langle x^*, \alpha_n h \rangle.$$

Thus, equality (3.8) holds true and the proof is complete. $\qquad\square$

Next we prove that the pair $D\mathcal{I}(u) = [\underline{d}\mathcal{I}(u), \overline{d}\mathcal{I}(u)]$ defined in Theorem 3.3 is indeed a codifferential of $\mathcal{I}$ at u. According to the definition of codifferential (see Def. 2.1), we need to prove that both sets $\underline{d}\mathcal{I}(u)$ and $\overline{d}\mathcal{I}(u)$ are convex and compact in the corresponding product topology. A proof of this result in the case when Ω is bounded and has the segment property was given in Lemmas 5.4 and 5.6 of [31]. Therefore, below we give a proof of the case $1 < p \le +\infty$ only.

Let us note that that we managed to remove the assumptions on the set Ω in the case $1 < p \le +\infty$ by using a completely different proof technique. Instead of reducing the proof of the compactness of $\underline{d}\mathcal{I}(u)$ and $\overline{d}\mathcal{I}(u)$ to the proof of the closedness of an auxiliary Aumann integral as it is done in [31], here we prove that the sets of measurable selections of the multifunctions $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ and $\overline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ are compact in a suitable topology and then conclude that the sets $\underline{d}\mathcal{I}(u)$ and $\overline{d}\mathcal{I}(u)$ are compact in the product topology as continuous images of the corresponding sets of measurable selections.

Lemma 3.7. *Let f satisfy the codifferentiability conditions of order $p \in [1, +\infty]$ and let either $1 < p \le +\infty$ or the set Ω be bounded and have the segment property. Then for any $u \in W^{1,p}(\Omega; \mathbb{R}^m)$ the sets $\underline{d}\mathcal{I}(u)$ and $\overline{d}\mathcal{I}(u)$ defined in Theorem 3.3 are convex and compact in the topology $\tau_\mathbb{R} \times w^*$. Furthermore, the equalities $\max\{A\colon (A, x^*) \in \underline{d}\mathcal{I}(u)\} = \min\{B\colon (B, y^*) \in \overline{d}\mathcal{I}(u)\} = 0$ hold true.*

Proof. Fix any $u \in W^{1,p}(\Omega; \mathbb{R}^m)$. We prove this lemma only for the hypodifferential $\underline{d}\mathcal{I}(u)$, since the proof for the hyperdifferential $\overline{d}\mathcal{I}(u)$ is exactly the same.

Choose any $(A_1, x_1^*), (A_2, x_2^*) \in \underline{d}\mathcal{I}(u)$, and let $z_i(\cdot) = (a_i(\cdot), v_{1i}(\cdot), v_{2i}(\cdot))$ be a measurable selection of the set-valued mapping $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ corresponding to (A_i, x_i^*), $i \in \{1, 2\}$. For a.e. $x \in \Omega$ the set $\underline{d}_{u,\xi} f(x, u(x), \nabla u(x))$ is convex by definition. Consequently, for any $\alpha \in [0, 1]$ the map $\alpha z_1(\cdot) + (1 - \alpha) z_2(\cdot)$ is a measurable selection of the set-valued map $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$, which obviously corresponds to the pair $Z_\alpha = (\alpha A_1 + (1 - \alpha) A_2, \alpha x_1^* + (1 - \alpha) x_2^*)$. Therefore, $Z_\alpha \in \underline{d}\mathcal{I}(u)$ for any $\alpha \in [0, 1]$ and one can conclude that the set $\underline{d}\mathcal{I}(u)$ is convex.

By the definition of codifferential for a.e. $x \in \Omega$ and for all $(a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u(x), \nabla u(x))$ one has $a \le 0$, which obviously implies that for any $(A, x^*) \in \underline{d}\mathcal{I}(u)$ one has $A \le 0$. Furthermore, by definition $\max\{a \mid (a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u(x), \nabla u(x))\} = 0$ for a.e. $x \in \Omega$, that is, for a.e. $x \in \Omega$ one has $0 \in \{a \mid (a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u(x), \nabla u(x))\}$. As was noted in the proof of Lemma 3.6, the multifunction $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ is measurable. Therefore, by Filippov's theorem ([2], Thm. 8.2.10) there exists a measurable selection $(a_0(\cdot), v_{10}(\cdot), v_{20}(\cdot))$ of $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ such that $a_0(x) = 0$ for a.e. $x \in \Omega$. Consequently, one has $(0, x^*) \in \underline{d}\mathcal{I}(u)$ for x^* defined as

$$\langle x^*, h \rangle = \int_\Omega \big(\langle v_{10}(x), h(x) \rangle + \langle v_{20}(x), \nabla h(x) \rangle\big)\, \mathrm{d}x \quad \forall h \in W^{1,p}(\Omega; \mathbb{R}^m),$$

which yields $\max\{A\colon (A, x^*) \in \underline{d}\mathcal{I}(u)\} = 0$.

Now we turn to the proof of the compactness of $\underline{d}\mathcal{I}(u)$. We consider two cases.

Case $1 < p < +\infty$. Denote by $\mathcal{F}$ the set of all measurable selections of the set-valued map $\underline{d}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$. By the codifferentiability conditions (see Def. 3.1) there exist $C > 0$ and a.e. nonnegative functions $\beta \in L^1(\Omega)$ and $\gamma \in L^{p'}(\Omega)$ such that for any $(a(\cdot), v_1(\cdot), v_2(\cdot)) \in \mathcal{F}$ and for a.e. $x \in \Omega$ one has

$$|a(x)| \le \beta(x) + C\big(|u(x)|^p + |\nabla u(x)|^p\big), \quad \max\{|v_1(x)|, |v_2(x)|\} \le \gamma(x) + C\big(|u(x)|^{p-1} + |\nabla u(x)|^{p-1}\big). \tag{3.9}$$

Observe that the right-hand side of the first inequality belongs to $L^1(\Omega)$, while the right-hand side of the second inequality belongs to $L^{p'}(\Omega)$ by virtue of the facts that $u \in W^{1,p}(\Omega; \mathbb{R}^m)$ and $p'(p-1) = p$. Thus, $\mathcal{F}$ is a bounded subset of the space $X = L^1(\Omega) \times L^{p'}(\Omega; \mathbb{R}^m) \times L^{p'}(\Omega; \mathbb{R}^{m \times d})$.

For any $(a(\cdot), v_1(\cdot), v_2(\cdot)) \in \mathcal{F}$ denote by $\mathcal{T}(a(\cdot), v_1(\cdot), v_2(\cdot))$ the pair $(A, x^*) \in \mathbb{R} \times (W^{1,p}(\Omega; \mathbb{R}^m))^*)$ such that $A = \int_\Omega a(x)\, \mathrm{d}x$ and

$$\langle x^*, h \rangle = \int_\Omega \big(\langle v_1(x), h(x) \rangle + \langle v_2(x), \nabla h(x) \rangle\big)\, \mathrm{d}x \quad \forall h \in W^{1,p}(\Omega; \mathbb{R}^m).$$

Clearly, $\mathcal{T}(\mathcal{F}) = \underline{d}\mathcal{I}(u)$ (see (3.3)). Furthermore, one can easily verify that $\mathcal{T}$ is a continuous linear operator from the vector space $X = L^1(\Omega) \times L^{p'}(\Omega; \mathbb{R}^m) \times L^{p'}(\Omega; \mathbb{R}^{m \times d})$ endowed with the weak topology to the space $(\mathbb{R} \times Y^*, \tau_\mathbb{R} \times \sigma(Y^*, Y))$ with $Y = W^{1,p}(\Omega; \mathbb{R}^m)$. Therefore, it is sufficient to check that the set $\mathcal{F}$ is weakly compact. Then one can conclude that the set $\underline{d}\mathcal{I}(u)$ is compact in the topology $\tau_\mathbb{R} \times \sigma(Y^*, Y)$ as the image of the compact set $\mathcal{F}$ under the continuous map $\mathcal{T}$.

By the by the Eberlein-Šmulian theorem it suffice to verify that $\mathcal{F}$ is weakly sequentially compact. Choose any sequence $z_n(\cdot) = (a_n(\cdot), v_{1n}(\cdot), v_{2n}(\cdot)) \in \mathcal{F}$, $n \in \mathbb{N}$. From the second inequality in (3.9) it follows that the sequence $\{(v_{1n}(\cdot), v_{2n}(\cdot))\}$ is bounded in $L^{p'}(\Omega; \mathbb{R}^m) \times L^{p'}(\Omega; \mathbb{R}^{m \times d})$. Hence taking into account the fact that the space $L^{p'}(\Omega)$ is reflexive (recall that $1 < p < +\infty$, which yields $1 < p' < +\infty$) one obtains that there exists a subsequence $\{(v_{1n_k}(\cdot), v_{2n_k}(\cdot))\}$ weakly converging to some $(v_1(\cdot), v_2(\cdot))$ in $L^{p'}(\Omega; \mathbb{R}^m) \times L^{p'}(\Omega; \mathbb{R}^{m \times d})$.

Let us now turn to the sequence $\{a_n(\cdot)\}$. Denote $g(\cdot) = \beta(\cdot) + C(|u(\cdot)|^p + |\nabla u(\cdot)|^p)$ (see (3.9)). Clearly, $g \in L^1(\Omega)$ and $\int_{|a_{n_k}| > g} |a_{n_k}(x)| \mathrm{d}x = 0 < \varepsilon$ for any $\varepsilon > 0$ and all $k \in \mathbb{N}$. Therefore, by Theorem 4.7.20 of [8] the closure of the set $\{a_{n_k}(\cdot)\}_{n \in \mathbb{N}}$ in the weak topology is weakly compact in $L^1(\Omega)$ or, equivalently, weakly sequentially compact in $L^1(\Omega)$ by the Eberlein-Šmulian theorem. Consequently, one can extract a subsequence of the sequence $\{a_{n_k}(\cdot)\}$, which we denote again by $\{a_{n_k}(\cdot)\}$, weakly converging to some $a \in L^1(\Omega)$.

Observe that the subsequence $z_{n_k}(\cdot) = (a_{n_k}(\cdot), v_{1n_k}(\cdot), v_{2n_k}(\cdot))$, $k \in \mathbb{N}$, weakly converges to the function $z(\cdot) = (a(\cdot), v_1(\cdot), v_2(\cdot))$ in $X = L^1(\Omega) \times L^{p'}(\Omega; \mathbb{R}^m) \times L^{p'}(\Omega; \mathbb{R}^{m \times d})$. By Mazur's lemma there exists a sequence $\widehat{z}_k(\cdot)$ of convex combinations of elements of the sequence $z_{n_k}(\cdot)$ strongly converging to $z(\cdot)$. As is well-known (see, *e.g.* [38], Exer. 6.9), one can extract a subsequence $\{\widehat{z}_{k_l}(\cdot)\}$ that converges to $z(\cdot)$ almost everywhere. From the convexity of the hypodifferential $\underline{d}_{u,\xi}f(\cdot)$ it follows that $\widehat{z}_k(\cdot)$ is a measurable selection of the multifunction $\underline{d}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$ for any $k \in \mathbb{N}$. Hence bearing in mind the fact that the hypodifferential $\underline{d}_{u,\xi}f(x, u(x), \nabla u(x))$ is closed for a.e. $x \in \Omega$ one obtains that $z(\cdot)$ is a measurable selection of $\underline{d}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$, *i.e.* $z(\cdot) \in \mathcal{F}$. Thus, we found a subsequence of the original sequence $\{z_n(\cdot)\} \subset \mathcal{F}$ weakly converging to an element of $\mathcal{F}$. In other words, $\mathcal{F}$ is weakly sequentially compact.

Case $p = +\infty$. Let, as above, $\mathcal{F}$ be the set of all measurable selections of the map $\underline{d}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$. By the codifferentiability conditions (Def. 3.1) there there exists an a.e. nonnegative function $\beta_N \in L^1(\Omega)$ such that for any $(a(\cdot), v_1(\cdot), v_2(\cdot)) \in \mathcal{F}$ one has

$$\max\big\{|a(x)|, |v_1(x)|, |v_2(x)|\big\} \le \beta_N(x) \tag{3.10}$$

for a.e. $x \in \Omega$ (here $N = \|u\|_{1,\infty}$). Thus, $\mathcal{F}$ is a bounded subset of the space $X = L^1(\Omega; \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^{m \times d})$.

Let the operator $\mathcal{T}$ be defined as in the case $1 < p < +\infty$. Then $\mathcal{T}(\mathcal{F}) = \underline{d}\mathcal{I}(u)$ and, as is easily seen, $\mathcal{T}$ is a continuous linear operator from the space X equipped with the weak topology to the space $(\mathbb{R} \times Y^*, \tau_\mathbb{R} \times$

$\sigma(Y^*, Y))$ with $Y = \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$. Therefore, it suffice to check that the set $\mathcal{F}$ is weakly compact in X. Then one can conclude that the set $\underline{\mathrm{d}}\mathcal{I}(u)$ is compact as the continuous image of a compact set.

From (3.10) it follows that for any $\varepsilon > 0$ and for all $(a(\cdot), v_1(\cdot), v_2(\cdot)) \in \mathcal{F}$ one has

$$\int_{|a|>\beta_N} |a|\,\mathrm{d}\mu = 0 < \varepsilon, \quad \int_{|v_1|>\beta_N} |v_1|\,\mathrm{d}\mu = 0 < \varepsilon, \quad \int_{|v_2|>\beta_N} |v_2|\,\mathrm{d}\mu = 0 < \varepsilon,$$

where μ is the Lebesgue measure. Consequently, by Theorem 4.7.20 of [8] the closure of the set $\mathcal{F}$ in the weak topology is weakly compact in X, which by the Eberlein-Šmulian theorem implies that it is weakly sequentially compact. Let us verify that the set $\mathcal{F}$ itself is weakly sequentially compact. Then by applying the Eberlein-Šmulian theorem once again we arrive at the desired result.

Indeed, let $\{z_n\} \subset \mathcal{F}$ be an arbitrary sequence. By the weak sequential compactness of the weak closure of $\mathcal{F}$ there exists a subsequence $\{z_{n_k}\}$ weakly converging to some $z \in X$. By Mazur's lemma there exists a sequence of convex combinations $\{\hat{z}_k\}$ of elements of the sequence $\{z_{n_k}\}$ strongly converging to z, which implies that there exists a subsequence $\{\hat{z}_{k_l}\}$ converging to z almost everywhere. Observe that each triplet $\hat{z}_{k_l}(\cdot)$ is a measurable selection of $\underline{\mathrm{d}}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ due to the definition of $\mathcal{F}$ and the fact that this multifunction is convex-valued. Therefore, bearing in mind the fact that the set $\underline{\mathrm{d}}_{u,\xi} f(x, u(x), \nabla u(x))$ is closed for a.e. $x \in \Omega$ one obtains that $z(\cdot)$ is a measurable selection of the multifunction $\underline{\mathrm{d}}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$. Thus, $z \in \mathcal{F}$, i.e. the subsequence $\{z_{n_k}\}$ weakly converges to an element of the set $\mathcal{F}$, which means that this set is weakly sequentially compact. $\qquad\square$

Let us finally prove that the functional $\mathcal{I}$ is, in fact, *continuously* codifferentiable. For any subset C of a metric space (M, d) and a point $x \in M$ denote $\mathrm{dist}(x, C) = \inf_{y \in C} d(x, y)$.

Let us underline that in our earlier paper [31] the continuous codifferentiability of the functional $\mathcal{I}$ was proved only in the case when the set Ω is bounded and $p = +\infty$ (see [31], Thm. 5.7 and Rem. 5.8). Here we extend this result to the case of unbounded domains and arbitrary $p \in [1, +\infty]$ by utilising Vitali's theorem characterising convergence in L^p-spaces (see, *e.g.* [36], Thm. III.6.15), instead of relying on certain compactness arguments as it is done in [31].

Lemma 3.8. *Let f satisfy the codifferentiability conditions of order $p \in [1, +\infty]$ and suppose that either $1 < p \leq +\infty$ or the set-valued maps $\underline{\mathrm{d}}_{u,\xi} f(\cdot)$ and $\overline{\mathrm{d}}_{u,\xi} f(\cdot)$ have the form (3.4) for some vectors $v_{1i}, w_{1j} \in \mathbb{R}^m$, $v_{2i}, w_{2j} \in \mathbb{R}^{m \times d}$, and Carathéodory functions f_i and g_j, $i \in I = \{1, \ldots, \ell\}$ and $j \in J = \{1, \ldots, r\}$. Then the set-valued mappings $\underline{\mathrm{d}}\mathcal{I}(\cdot)$ and $\overline{\mathrm{d}}\mathcal{I}(\cdot)$ defined in Theorem 3.3 are Hausdorff continuous.*

Proof. We prove the statement of the lemma only for the hypodifferential mapping $\underline{\mathrm{d}}\mathcal{I}(\cdot)$, since the proof of the lemma for $\overline{\mathrm{d}}\mathcal{I}(\cdot)$ is exactly the same.

Arguing by reductio ad absurdum, suppose that the multifunction $\underline{\mathrm{d}}\mathcal{I}(\cdot)$ is not Hausdorff continuous at a point $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$. Then there exist $\theta > 0$ and a sequence $\{u_n\} \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ converging to u such that $d_H(\underline{\mathrm{d}}\mathcal{I}(u_n), \underline{\mathrm{d}}\mathcal{I}(u)) > \theta$ for all $n \in \mathbb{N}$. Replacing, if necessary, the sequence $\{u_n\}$ with its subsequence, one can suppose that u_n converges to u almost everywhere and ∇u_n converges to ∇u almost everywhere.

By the definition Hausdorff distance (see (2.3)), two cases are possible. Namely, there exists a subsequence, which we denote again by $\{u_n\}$, such that one of the following inequalities hold true:

$$\sup_{(B,y^*)\in\underline{\mathrm{d}}\mathcal{I}(u_n)} \inf_{(A,x^*)\in\underline{\mathrm{d}}\mathcal{I}(u)} \sqrt{|B-A|^2 + \|y^*-x^*\|^2} > \theta \tag{3.11}$$

$$\sup_{(A,x^*)\in\underline{\mathrm{d}}\mathcal{I}(u)} \inf_{(B,y^*)\in\underline{\mathrm{d}}\mathcal{I}(u_n)} \sqrt{|B-A|^2 + \|y^*-x^*\|^2} > \theta. \tag{3.12}$$

We start with the first case.

Case I. From (3.11) it follows that for any $n \in \mathbb{N}$ there exists $(A_n, x_n^*) \in \underline{\mathrm{d}}\mathcal{I}(u_n)$ satisfying the inequality $\mathrm{dist}((A_n, x_n^*), \underline{\mathrm{d}}\mathcal{I}(u)) \geq \theta$. Denote by $z_n(\cdot) = (a_n(\cdot), v_{1n}(\cdot), v_{2n}(\cdot))$ a measurable selection of the multifunction

$\underline{\mathrm{d}}_{u,\xi}f(\cdot, u_n(\cdot), \nabla u_n(\cdot))$ corresponding to the pair (A_n, x_n^*) (see (3.3)). We consider the cases $p > 1$ and $p = 1$ separately

Case I, $1 < p = +\infty$. Recall that $\underline{\mathrm{d}}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$ is a convex and compact-valued multifunction. Furthermore, as was shown in the proof of Lemma 3.6, the codifferentiability conditions guarantee that this multifunction is measurable. Therefore, for any $n \in \mathbb{N}$ and for a.e. $x \in \Omega$ the set

$$
R_n(x) = \Big\{(a, v_1, v_2) \in \underline{\mathrm{d}}_{u,\xi}f(x, u(x), \nabla u(x)) \;\Big|
$$
$$
\mathrm{dist}\,\big(z_n(x), \underline{\mathrm{d}}_{u,\xi}f(x, u(x), \nabla u(x))\big)^2 = |a_n(x) - a|^2 + |v_{1n}(x) - v_1|^2 + |v_{2n}(x) - v_2|^2\Big\}
$$

(*i.e.* $R_n(x)$ is the set of points at which the infimum in the definition of the distance between $z_n(x)$ and the set $\underline{\mathrm{d}}_{u,\xi}f(x, u(x), \nabla u(x))$ is attained) is nonempty and the set-valued mapping $R_n(\cdot)$ is measurable by Theorem 8.2.11 [2].

Let $z_n^0(\cdot)$ be any measurable selection of the multifunction $R_n(\cdot)$, which exists by Theorem 8.1.3 of [2]. Define function $\widehat{z}_n(\cdot) = (\widehat{a}_n(\cdot), \widehat{v}_{1n}(\cdot), \widehat{v}_{2n}(\cdot))$ as follows:

$$
\widehat{z}_n(x) = \begin{cases} z_n^0(x), & \text{if } z_n(x) \notin \underline{\mathrm{d}}_{u,\xi}f(x, u(x), \nabla u(x)), \\ z_n(x), & \text{othwerwise.} \end{cases}
$$

Clearly, $\widehat{z}_n(\cdot)$ is a selection of the multifunction $\underline{\mathrm{d}}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$. Furthermore, it is measurable due to the fact that the set of all those $x \in \Omega$ for which $z_n(x) \notin \underline{\mathrm{d}}_{u,\xi}f(x, u(x), \nabla u(x))$ is measurable ([2], Cor. 8.2.13, part 2).

By the codifferentiability conditions (see Def. 3.1) the multifunction $\underline{\mathrm{d}}_{u,\xi}f(\cdot)$ is a Carathéodory map. Thus, for a.e. $x \in \Omega$ the set-valued map $(u, \xi) \mapsto \underline{\mathrm{d}}_{u,\xi}f(x, u, \xi)$ is continuous. Therefore, for a.e. $x \in \Omega$ one has

$$
\lim_{n \to \infty} d_H(\underline{\mathrm{d}}_{u,\xi}f(x, u_n(x), \nabla u_n(x)), \underline{\mathrm{d}}_{u,\xi}f(x, u(x), \nabla u(x))) = 0.
$$

Hence, in particular, $\mathrm{dist}(z_n(x), \underline{\mathrm{d}}_{u,\xi}f(x, u(x), \nabla u(x))) \to 0$ as $n \to \infty$, which implies that the sequence $\{z_n - \widehat{z}_n\}$ converges to zero almost everywhere. Let us prove that this sequence converges to zero in $\mathrm{L}^1(\Omega) \times \mathrm{L}^{p'}(\Omega; \mathbb{R}^m) \times \mathrm{L}^{p'}(\Omega; \mathbb{R}^{m \times d})$. To this end, we shall utilise Vitali's theorem characterising convergence in L^p-spaces with $1 \le p < +\infty$ (see, *e.g.* [36], Thm. III.6.15). Note that $1 \le p' < +\infty$, since we consider the case $1 < p \le +\infty$.

Fix any $\varepsilon > 0$. By the growth condition on the codifferential mapping $D_{u,\xi}f$ (see Def. 3.1) there exist $C > 0$ and an a.e. nonnegative function $\beta \in \mathrm{L}^1(\Omega)$ such that

$$
|a_n(x) - \widehat{a}_n(x)| \le 2\beta(x) + C\big(|u(x)|^p + |u_n(x)|^p + |\nabla u(x)|^p + |\nabla u_n(x)|^p\big) \tag{3.13}
$$

for a.e. $x \in \Omega$ in the case $1 < p < +\infty$, and there exists an a.e. nonnegative function $\beta_N \in \mathrm{L}^1(\Omega)$ such that $|a_n(x) - \widehat{a}_n(x)| \le 2\beta_N(x)$ for a.e. $x \in \Omega$ in the case $p = +\infty$ (here $N = \max_{n \in \mathbb{N}}\{\|u\|_{1,\infty}, \|u_n\|_{1,\infty}\}$). If $p = +\infty$, then the sequence $\{a_n - \widehat{a}_n\}$ converges to zero in $\mathrm{L}^1(\Omega)$ by Lebesgue's dominated convergence theorem. Therefore, let us consider the case $1 < p < +\infty$.

By the absolute continuity of the Lebesgue integral there exists $\delta_1 > 0$ such that for any measurable set $D \subseteq \Omega$ with $\mu(D) < \delta_1$ (here μ is the Lebesgue measure) one has

$$
\int_D \beta \mathrm{d}\mu < \frac{\varepsilon}{10}, \quad \int_D |u|^p \mathrm{d}\mu < \frac{\varepsilon}{5C}, \quad \int_D |\nabla u|^p \mathrm{d}\mu < \frac{\varepsilon}{5C}.
$$

Moreover, by the "only if" part of the Vitali convergence theorem, the convergence of u_n to u in $W^{1,p}(\Omega; \mathbb{R}^m)$ implies that there exists $\delta_2 > 0$ such that for any measurable set $D \subseteq \Omega$ with $\mu(D) < \delta_2$ one has

$$\int_D |u_n|^p \mathrm{d}\mu < \frac{\varepsilon}{5C}, \quad \int_D |\nabla u_n|^p \mathrm{d}\mu < \frac{\varepsilon}{5C} \quad \forall n \in \mathbb{N}.$$

Hence with the use of (3.13) one obtains that for any measurable set $D \subseteq \Omega$ with $\mu(D) < \min\{\delta_1, \delta_2\}$ one has $\int_D |a_n - \widehat{a}_n| \mathrm{d}\mu < \varepsilon$ for all $n \in \mathbb{N}$.

Denote $\Omega_N = \{x \in \Omega \mid |x| \leq N\}$. From the fact that $\beta \in L^1(\Omega)$ and $u \in W^{1,p}(\Omega; \mathbb{R}^m)$ it follows that there exists $N \in \mathbb{N}$ such that

$$\int_{\Omega \setminus \Omega_N} \beta \mathrm{d}\mu < \frac{\varepsilon}{10}, \quad \int_{\Omega \setminus \Omega_N} |u|^p \mathrm{d}\mu < \frac{\varepsilon}{5C}, \quad \int_{\Omega \setminus \Omega_N} |\nabla u|^p \mathrm{d}\mu < \frac{\varepsilon}{5C}$$

(see, *e.g.* [8], Prop. 2.6.2). Furthermore, by the "only if" part of the Vitali convergence theorem there exists a measurable set $E_\varepsilon \subseteq \Omega$ such that $\mu(E_\varepsilon) < +\infty$ and

$$\int_{\Omega \setminus E_\varepsilon} |u_n|^p \mathrm{d}\mu < \frac{\varepsilon}{5C}, \quad \int_{\Omega \setminus E_\varepsilon} |\nabla u_n|^p \mathrm{d}\mu < \frac{\varepsilon}{5C} \quad \forall n \in \mathbb{N}.$$

Therefore, by applying (3.13) one obtains that $\int_{\Omega \setminus \Omega_\varepsilon} |a_n - \widehat{a}_n| \mathrm{d}\mu < \varepsilon$ for all $n \in \mathbb{N}$, where $\Omega_\varepsilon = \Omega_N \cup E_\varepsilon$. Hence with the use of the "if" part of the Vitali convergence theorem one concludes that the sequence $\{a_n - \widehat{a}_n\}$ converges to zero in $L^1(\Omega)$.

Let us now consider the sequence $\{v_{1n} - \widehat{v}_{1n}\}$. By the growth condition on the codifferential mapping $D_{u,\xi}f$ (see Def. 3.1) there exist $C > 0$ and a.e. nonnegative function $\gamma \in L^{p'}(\Omega)$ such that

$$|v_{1n}(x) - \widehat{v}_{1n}(x)|^{p'} \leq 2^{p'} \left(|v_{1n}(x)|^{p'} + |\widehat{v}_{2n}(x)|^{p'} \right)$$
$$\leq 2^{p'} 3^{p'} \left(2|\gamma(x)|^{p'} + C^{p'} \left(|u(x)|^p + |\nabla u(x)|^p + |u_n(x)|^p + |\nabla u_n(x)|^p \right) \right)$$

for a.e. $x \in \Omega$ in the case $1 < p < +\infty$, and there exists an a.e. nonnegative function $\beta_N \in L^1(\Omega)$ such that $|v_{1n}(x) - \widehat{v}_{1n}(x)| \leq 2\beta_N(x)$ for a.e. $x \in \Omega$ in the case $p = +\infty$. Now, arguing in the same way as above and applying Vitali's convergence theorem in the case $1 < p < +\infty$ and Lebesgue's dominated convergence theorem in the case $p = +\infty$ one can readily verify that $\{v_{1n} - \widehat{v}_{1n}\}$ converges to zero in $L^{p'}(\Omega; \mathbb{R}^m)$. The convergence of $\{v_{2n} - \widehat{v}_{2n}\}$ to zero in $L^{p'}(\Omega; \mathbb{R}^{m \times d})$ is proved in exactly the same way.

Denote by $(\widehat{A}_n, \widehat{x}_n^*)$ the element of $\underline{\mathrm{d}}\mathcal{I}(u)$ corresponding to the selection $\widehat{z}_n(\cdot) = (\widehat{a}_n(\cdot), \widehat{v}_{1n}(\cdot), \widehat{v}_{2n}(\cdot))$ of the multifunction $\underline{\mathrm{d}}_{u,\xi}f(\cdot, u(\cdot), \nabla u(\cdot))$ (see (3.3)). Let us check that $|A_n - \widehat{A}_n| + \|x_n^* - \widehat{x}_n^*\| \to 0$ as $n \to \infty$. Indeed, for any $n \in \mathbb{N}$ one has

$$|A_n - \widehat{A}_n| \leq \int_\Omega |a_n(x) - \widehat{a}_n(x)| \, \mathrm{d}x = \|a_n - \widehat{a}_n\|_1,$$

which implies that $|A_n - \widehat{A}_n| \to 0$ as $n \to \infty$. Similarly, with the use of Hölder's inequality for any $h \in W^{1,p}(\Omega; \mathbb{R}^m)$ one has

$$\left| \langle x_n^* - \widehat{x}_n^*, h \rangle \right| \leq \int_\Omega \left| \langle v_{1n}(x) - \widehat{v}_{1n}(x), h(x) \rangle \right| \mathrm{d}x + \int_\Omega \left| \langle v_{2n}(x) - \widehat{v}_{2n}(x), \nabla h(x) \rangle \right| \mathrm{d}x$$
$$\leq \left(\|v_{1n} - \widehat{v}_{1n}\|_{p'} + \|v_{2n} - \widehat{v}_{2n}\|_{p'} \right) \|h\|_{1,p},$$

which implies that $\|x_n^* - \widehat{x}_n^*\| \le \|v_{1n} - \widehat{v}_{1n}\|_{p'} + \|v_{2n} - \widehat{v}_{2n}\|_{p'}$ for all $n \in \mathbb{N}$, and $\|x_n^* - \widehat{x}_n^*\| \to 0$ as $n \to \infty$. Consequently, bearing in mind the fact that $(\widehat{A}_n, \widehat{x}_n^*) \in \underline{d}\mathcal{I}(u)$ for all $n \in \mathbb{N}$ one obtains that $\mathrm{dist}((A_n, x_n^*), \underline{d}\mathcal{I}(u)) \to 0$ as $n \to \infty$, which contradicts the inequality $\mathrm{dist}((A_n, x_n^*), \underline{d}\mathcal{I}(u)) \ge \theta$. Thus, the proof of the first case for $1 < p \le +\infty$ is complete.

Case I, $p = 1$. Let S^ℓ be the standard (probability) simplex in $\mathbb{R}^\ell$, *i.e.*

$$ S^\ell = \left\{ \alpha = (\alpha^{(1)}, \dots, \alpha^{(\ell)}) \in \mathbb{R}^\ell \;\middle|\; \alpha^{(1)} + \dots + \alpha^{(\ell)} = 1,\ \alpha^{(i)} \ge 0 \ \forall i \in \{1, \dots, \ell\} \right\}. $$

For any $\alpha \in \mathbb{R}^\ell$, $x \in \Omega$, $u \in \mathbb{R}^m$, and $\xi \in \mathbb{R}^{m \times d}$ define

$$ g(x, u, \xi, \alpha) = \sum_{i=1}^{l} \alpha^{(i)} (f_i(x, u, \xi), v_{1i}, v_{2i}). \tag{3.14} $$

It is easily seen that g is a Carathéodory map and $g(x, u, \xi, S^\ell) = \underline{d}_{u,\xi} f(x, u, \xi)$ for all (x, u, ξ) by the definition of convex hull (see (3.4)).

Recall that $z_n(\cdot) = (a_n(\cdot), v_{1n}(\cdot), v_{2n}(\cdot))$ is a measurable selection of the set-valued map $\underline{d}_{u,\xi} f(\cdot, u_n(\cdot), \nabla u_n(\cdot))$ such that for the corresponding pair $(A_n, x_n^*) \in \underline{d}\mathcal{I}(u_n)$ one has $\mathrm{dist}((A_n, x_n^*), \underline{d}\mathcal{I}(u)) \ge \theta$ for all $n \in \mathbb{N}$. By definition for any $n \in \mathbb{N}$ and a.e. $x \in \Omega$ one has $z_n(x) \in g(x, u_n(x), \nabla u_n(x), S^\ell)$, which by Filippov's theorem ([2], Thm. 8.2.10) implies that for any $n \in \mathbb{N}$ there exists a measurable function $\alpha_n \colon \Omega \to S^\ell$ such that $z_n(x) = g(x, u_n(x), \nabla u_n(x), \alpha_n(x))$ for a.e. $x \in \Omega$. Define

$$ \widehat{z}_n(\cdot) = (\widehat{a}_n(\cdot), \widehat{v}_{1n}(\cdot), \widehat{v}_{2n}(\cdot)) = g(\cdot, u(\cdot), \nabla u(\cdot), \alpha_n(\cdot)). $$

Clearly, $\widehat{z}_n$ is a measurable selection of the multifunction $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$. Denote by $(\widehat{A}_n, \widehat{x}_n^*)$ the element of $\underline{d}\mathcal{I}(u)$ corresponding to this selection (see (3.3)).

From the definition of g (see (3.14)) and the definition of $\widehat{z}_n$ it follows that $x_n^* = \widehat{x}_n^*$ for all $n \in \mathbb{N}$. Furthermore, for all $n \in \mathbb{N}$ and a.e. $x \in \Omega$ one has

$$ |a_n(x) - \widehat{a}_n(x)| \le \sum_{i=1}^{\ell} \alpha_n^{(i)}(x) \big| f_i(x, u_n(x), \nabla u_n(x)) - f_i(x, u(x), \nabla u(x)) \big|. $$

Hence $|a_n(x) - \widehat{a}_n(x)| \to 0$ as $n \to \infty$ for a.e. $x \in \Omega$, since by our assumptions $u_n \to u$ and $\nabla u_n \to \nabla u$ almost everywhere, and f_i are Carathéodory functions.

By the growth condition on $D_{u,\xi} f(\cdot)$ (see Def. 3.1) there exist $C > 0$ and an a.e. nonnegative function $\beta \in \mathrm{L}^1(\Omega)$ such that

$$ |a_n(x) - \widehat{a}_n(x)| \le \beta(x) + C \big(|u(x)| + |\nabla u(x)| + |u_n(x)| + |\nabla u_n(x)| \big) $$

for a.e. $x \in \Omega$. With the use of this inequality and Vitali's convergence theorem one can check that $|a_n - \widehat{a}_n|$ converges to zero in $\mathrm{L}^1(\Omega)$ as in the case $1 < p < +\infty$. Hence $|A_n - \widehat{A}_n| \le \int_\Omega |a_n - \widehat{a}_n| \mathrm{d}\mu$ converges to zero as $n \to \infty$, *i.e.* $|A_n - \widehat{A}_n| + \|x_n^* - \widehat{x}_n^*\| \to 0$ as $n \to \infty$. Therefore, $\mathrm{dist}((A_n, x_n^*), \underline{d}\mathcal{I}(u)) \to 0$ as $n \to \infty$, which once again contradicts the inequality $\mathrm{dist}((A_n, x_n^*), \underline{d}\mathcal{I}(u)) \ge \theta$.

Case II. Suppose now that (3.12) holds true. Then for any $n \in \mathbb{N}$ there exists $(A_n, x_n^*) \in \underline{d}\mathcal{I}(u)$ such that $\mathrm{dist}((A_n, x_n^*), \underline{d}\mathcal{I}(u_n)) \ge \theta$. For any $n \in \mathbb{N}$ denote by $z_n(\cdot) = (a_n(\cdot), v_{1n}(\cdot), v_{2n}(\cdot))$ a measurable selection of the set-valued mapping $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ corresponding to (A_n, x_n^*) (see (3.3)).

Case II, $1 < p \leq +\infty$**.** Denote by $z_n^0(\cdot) = (a_n^0(\cdot), v_{1n}^0(\cdot), v_{2n}^0(\cdot))$ any measurable selection of the multifunction $\underline{d}_{u,\xi} f(\cdot, u_n(\cdot), \nabla u_n(\cdot))$ such that

$$\text{dist}\left(z_n(x), \underline{d}_{u,\xi} f(x, u_n(x), \nabla u_n(x))\right)^2 = |a_n(x) - a_n^0(x)|^2 + |v_{1n}(x) - v_{1n}^0(x)|^2 + |v_{2n}(x) - v_{2n}^0(x)|^2$$

for a.e. $x \in \Omega$. The existence of such selection can be proved in the same way it is done in **Case I**. Finally, define mapping $\widehat{z}_n(\cdot) = (\widehat{a}_n(\cdot), \widehat{v}_{1n}(\cdot), \widehat{v}_{2n}(\cdot))$ as follows:

$$\widehat{z}_n(x) = \begin{cases} z_n^0(x), & \text{if } z_n(x) \notin \underline{d}_{u,\xi} f(x, u_n(x), \nabla u_n(x)), \\ z_n(x), & \text{othwerwise.} \end{cases}$$

Then $\widehat{z}_n$ is a measurable selection of the set-valued mapping $\underline{d}_{u,\xi} f(\cdot, u_n(\cdot), \nabla u_n(\cdot))$.

By the codifferentiability conditions (see Def. 3.1) the multifunction $\underline{d}_{u,\xi} f(\cdot)$ is a Carathéodory map, *i.e.* for a.e. $x \in \Omega$ the set-valued map $(u, \xi) \mapsto \underline{d}_{u,\xi} f(x, u, \xi)$ is continuous. Therefore, for a.e. $x \in \Omega$ one has

$$\lim_{n \to \infty} d_H(\underline{d}_{u,\xi} f(x, u_n(x), \nabla u_n(x)), \underline{d}_{u,\xi} f(x, u(x), \nabla u(x))) = 0.$$

Hence, in particular, $\text{dist}(z_n(x), \underline{d}_{u,\xi} f(x, u_n(x), \nabla u_n(x)) \to 0$ as $n \to \infty$, which implies that the sequence $\{z_n - \widehat{z}_n\}$ converges to zero almost everywhere. Applying the growth condition on the codifferential mapping $D_{u,\xi} f$ and arguing in the same way as in **Case I** one can check that this sequence converges to zero in $L^1(\Omega) \times L^{p'}(\Omega; \mathbb{R}^m) \times L^{p'}(\Omega; \mathbb{R}^{m \times d})$. With the use of this fact it is easy to show that $|A_n - \widehat{A}_n| + \|x_n^* - \widehat{x}_n^*\| \to 0$ as $n \to \infty$, where $(\widehat{A}_n, \widehat{x}_n^*)$ is the element of $\underline{d}\mathcal{I}(u_n)$ corresponding to the selection $\widehat{z}_n$. Therefore, $\text{dist}((A_n, x_n), \underline{d}\mathcal{I}(u_n)) \to 0$ as $n \to \infty$, which contradicts the inequality $\text{dist}((A_n, x_n^*), \underline{d}\mathcal{I}(u_n)) \geq \theta$.

Case II, $p = 1$**.** Arguing in the same way as in **Case I** and applying Filippov's theorem, for any $n \in \mathbb{N}$ one can find a measurable function $\alpha_n \colon \Omega \to S^\ell$ such that $z_n(x) = g(x, u(x), \nabla u(x), \alpha_n(x))$ for a.e. $x \in \Omega$. Define

$$\widehat{z}_n(x) = (\widehat{a}_n(x), \widehat{v}_{1n}(x), \widehat{v}_{2n}(x)) = g(x, u_n(x), \nabla u_n(x), \alpha_n(x))$$

for a.e. $x \in \Omega$. Then $\widehat{z}_n(\cdot)$ is a measurable selection of the set-valued mapping $\underline{d}_{u,\xi} f(\cdot, u_n(\cdot), \nabla u_n(\cdot))$. Denote by $(\widehat{A}_n, \widehat{x}_n^*)$ the element of $\underline{d}\mathcal{I}(u_n)$ corresponding to this selection. Then $x_n^* = \widehat{x}_n^*$ for all $n \in \mathbb{N}$, and arguing in the same way as in **Case I** one can check that $|A_n - \widehat{A}_n| \to 0$ as $n \to \infty$. Therefore $\text{dist}((A_n, x_n), \underline{d}\mathcal{I}(u_n)) \to 0$ as $n \to \infty$, which once again contradicts the inequality $\text{dist}((A_n, x_n^*), \underline{d}\mathcal{I}(u_n)) \geq \theta$. $\qquad \square$

Applying Theorem 3.3, ([33], Cor. 2), and Lemma 2.6 one obtains that in the case when the integrand f satisfies the codifferentiability conditions, the functional $\mathcal{I}$ is locally Lipschitz continuous and Hadamard quasidifferentiable.

Corollary 3.9. *Let f satisfy the codifferentiability conditions of order $p \in [1, +\infty]$, and let either $1 < p \leq +\infty$ or the set Ω be bounded and have the segment property, and the set-valued maps $\underline{d}_{u,\xi} f(\cdot)$ and $\overline{d}_{u,\xi} f(\cdot)$ have the form (3.4). Then the functional $\mathcal{I}$ is locally Lipschitz continuous, Hadamard quasidifferentiable at every $u \in W^{1,p}(\Omega; \mathbb{R}^m)$, and the pair $\mathscr{D}\mathcal{I}(u) = [\underline{\partial}\mathcal{I}(u), \overline{\partial}\mathcal{I}(u)]$ with*

$$\underline{\partial}\mathcal{I}(u) = \{x^* \in (W^{1,p}(\Omega; \mathbb{R}^m))^* \mid (0, x^*) \in \underline{d}\mathcal{I}(u)\}, \quad \overline{\partial}\mathcal{I}(u) = \{y^* \in (W^{1,p}(\Omega; \mathbb{R}^m))^* \mid (0, y^*) \in \overline{d}\mathcal{I}(u)\},$$

is a quasidifferential of $\mathcal{I}$ at u, where the sets $\underline{d}\mathcal{I}(u)$ and $\overline{d}\mathcal{I}(u)$ are defined in Theorem 3.3.

Remark 3.10. Recall that by the definition of codifferential one has $a \leq 0$ for any $(a, v_1, v_2) \in \underline{d}_{u,\xi} f(x, u, \xi)$. Hence with the use of Theorem 3.3 and the corollary above one obtains that $x^* \in \underline{\partial}\mathcal{I}(u)$ if and only if there

exists a measurable selection $(0, v_1(\cdot), v_2(\cdot))$ of the multifunction $\underline{d}_{u,\xi} f(\cdot, u(\cdot), \nabla u(\cdot))$ such that

$$\langle x^*, h \rangle = \int_\Omega \big(\langle v_1(x), h(x) \rangle + \langle v_2(x), \nabla h(x) \rangle \big)\, dx \quad \forall h \in W^{1,p}(\Omega; \mathbb{R}^m).$$

A similar statement holds true for $\overline{\partial} \mathcal{I}(u)$ as well.

As usual, denote by $W_0^{1,p}(\Omega; \mathbb{R}^m)$ the closure of the space $C_c^\infty(\Omega; \mathbb{R}^m)$ of infinitely differentiable functions $\varphi \colon \Omega \to \mathbb{R}^m$ with compact support in the Sobolev space $W^{1,p}(\Omega; \mathbb{R}^m)$. To derive optimality conditions for problems with prescribed boundary conditions we will utilise the following corollary on the quasidifferentiability of the restriction of $\mathcal{I}$ to the space $W_0^{1,p}(\Omega; \mathbb{R}^m)$. This result is almost trivial. Nevertheless, we briefly outline its proof for the sake of completeness and mathematical rigour.

Corollary 3.11. *Let f satisfy the codifferentiability conditions of order $p \in [1, +\infty]$, $u_0 \in W^{1,p}(\Omega; \mathbb{R}^m)$ be fixed, and let either $1 < p \le +\infty$ or the set Ω be bounded and have the segment property, and the set-valued maps $\underline{d}_{u,\xi} f(\cdot)$ and $\overline{d}_{u,\xi} f(\cdot)$ have the form (3.4). Then the functional $\mathcal{J} \colon W_0^{1,p}(\Omega; \mathbb{R}^m) \to \mathbb{R}$, $\mathcal{J}(u) = \mathcal{I}(u_0 + u)$ is correctly defined, locally Lipschitz continuous, and Hadamard quasidifferentiable at every $u \in W_0^{1,p}(\Omega; \mathbb{R}^m)$. Furthermore, the pair $\mathscr{D}\mathcal{J}(u) = [\underline{\partial}\mathcal{J}(u), \overline{\partial}\mathcal{J}(u)]$ with*

$$\underline{\partial}\mathcal{J}(u) = \Big\{ x^* \in (W_0^{1,p}(\Omega; \mathbb{R}^m))^* \,\Big|\, \langle x^*, h \rangle = \int_\Omega \big(\langle v_1(x), h(x) \rangle + \langle v_2(x), \nabla h(x) \rangle \big)\, dx \ \ \forall h \in W_0^{1,p}(\Omega; \mathbb{R}^m),$$

$$(0, v_1(\cdot), v_2(\cdot)) \text{ is a measurable selection of the set-valued map } \underline{d}_{u,\xi} f(\cdot, u_0(\cdot) + u(\cdot), \nabla u_0(\cdot) + \nabla u(\cdot)) \Big\} \quad (3.15)$$

and

$$\overline{\partial}\mathcal{J}(u) = \Big\{ y^* \in (W_0^{1,p}(\Omega; \mathbb{R}^m))^* \,\Big|\, \langle y^*, h \rangle = \int_\Omega \big(\langle w_1(x), h(x) \rangle + \langle w_2(x), \nabla h(x) \rangle \big)\, dx \ \ \forall h \in W_0^{1,p}(\Omega; \mathbb{R}^m),$$

$$(0, w_1(\cdot), w_2(\cdot)) \text{ is a measurable selection of the set-valued map } \overline{d}_{u,\xi} f(\cdot, u_0(\cdot) + u(\cdot), \nabla u_0(\cdot) + \nabla u(\cdot)) \Big\} \quad (3.16)$$

is a quasidifferential of $\mathcal{J}$ at u.

Proof. The fact that the functional $\mathcal{J}$ is correctly defined and locally Lipschitz continuous follows directly from its definition and Corollary 3.9. Let us prove that it is Hadamard quasidifferentiable.

Denote $X = W^{1,p}(\Omega; \mathbb{R}^m)$ and $X_0 = W_0^{1,p}(\Omega; \mathbb{R}^m)$. Introduce the linear operator $\mathcal{T} \colon X^* \to X_0^*$ that maps $x^* \in X^*$ to its restriction to X_0, *i.e.* $\mathcal{T}(x^*) = x^*|_{X_0}$. It is easily seen that $\mathcal{T}$ is a continuous operator from X^* endowed with the weak* topology to X_0^* endowed with the weak* topology, since $X_0 \subset X$.

Observe that by definitions (see Cor 3.9 and Rem. 3.10) one has $\underline{\partial}\mathcal{J}(u) = \mathcal{T}(\underline{\partial}\mathcal{I}(u_0 + u))$ and $\overline{\partial}\mathcal{J}(u) = \mathcal{T}(\overline{\partial}\mathcal{I}(u_0 + u))$. Therefore, $\underline{\partial}\mathcal{J}(u)$ and $\overline{\partial}\mathcal{J}(u)$ are convex and weak* compact convex subsets of X_0^* due to the fact that $\underline{\partial}\mathcal{I}(u_0 + u)$ and $\overline{\partial}\mathcal{I}(u_0 + u)$ are convex and weak* compact convex subsets of X^*.

Fix any $u \in W_0^{1,p}(\Omega; \mathbb{R}^m)$. With the use of Corollary 3.9 and the fact that for any $\alpha > 0$ and $h \in W_0^{1,p}(\Omega; \mathbb{R}^m)$ one has $\mathcal{J}(u + \alpha h) = \mathcal{I}(u_0 + u + \alpha h)$ one obtains that

$$\lim_{[\alpha, h'] \to [+0, h]} \frac{\mathcal{J}(u + \alpha h') - \mathcal{J}(u)}{\alpha} = \lim_{[\alpha, h'] \to [+0, h]} \frac{\mathcal{I}(u_0 + u + \alpha h') - \mathcal{J}(u_0 + u)}{\alpha}$$

$$= \mathcal{I}'(u_0 + u; h) = \max_{x^* \in \underline{\partial}\mathcal{I}(u_0 + u)} \langle x^*, h \rangle + \min_{y^* \in \overline{\partial}\mathcal{I}(u_0 + u)} \langle y^*, h \rangle = \max_{x^* \in \underline{\partial}\mathcal{J}(u)} \langle x^*, h \rangle + \min_{y^* \in \overline{\partial}\mathcal{J}(u)} \langle y^*, h \rangle$$

for all $h \in W_0^{1,p}(\Omega; \mathbb{R}^m)$ (here $h' \in W_0^{1,p}(\Omega; \mathbb{R}^m)$ as well). Thus, the functional $\mathcal{J}$ is Hadamard quasidifferentiable, and the pair (3.15), (3.16) is its quasidifferential. $\qquad\square$

Remark 3.12. By Theorem 3.3, the assumption that in the case $p = 1$ the set-valued mappings $\underline{d}_{u,\xi}f(\cdot)$ and $\overline{d}_{u,\xi}f(\cdot)$ have the form (3.4) is needed only to ensure the continuity of the multifunctions $\underline{d}\mathcal{I}(\cdot)$ and $\overline{d}\mathcal{I}(\cdot)$, *i.e.* to ensure that the functional $\mathcal{I}$ is *continuously* codifferentiable. Therefore, in the case when the function f satisfies the codifferentiability conditions of order $p = 1$ and Ω is bounded and has the segment property, but the set-valued mappings $\underline{d}_{u,\xi}f(\cdot)$ and $\overline{d}_{u,\xi}f(\cdot)$ do *not* have the form (3.4), the functional $\mathcal{J}$ from Corollary 3.11 is still quasidifferentiable and the pair (3.15), (3.16) is a quasidifferential of $\mathcal{J}$ at u by Remark 2.7 and Theorem 3.3.

4. Constrained nonsmooth problems of the calculus of variations

In this section we derive optimality conditions in terms of codifferentials for nonsmooth problems of the calculus of variations with nonsmooth isoperimetric constraints and nonsmooth constraints at the boundary of the domain. By means of several simple examples we also demonstrate that in some cases optimality conditions in terms of codifferentials are better than optimality conditions in terms of various subdifferentials.

4.1. Unconstrained problems

We start with an unconstrained problem of the form

$$\min \, \mathcal{I}(u) = \int_{\Omega} f(x, u(x), \nabla u(x)) \, dx, \quad u \in u_0 + \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m). \tag{4.1}$$

Here, as in the previous section, $\Omega \subseteq \mathbb{R}^d$ is an open set, $f \colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to \mathbb{R}$, $f = f(x, u, \xi)$, is a nonsmooth function, while $u_0 \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ is a fixed function.

In essence, problem (4.1) can be viewed as the classical problem of minimising $\mathcal{I}(u)$ over the set of all those $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ for which $u|_{\partial\Omega} = \psi$ for some prespecified function ψ, where $\partial\Omega$ is the boundary of Ω (simply put $\psi = u_0|_{\partial\Omega}$). However, to avoid the usage of trace operators and corresponding assumptions on the domain Ω, we pose this classical "boundary value problem" in the abstract form (4.1).

In the case when the domain Ω is bounded and has the segment property, optimality conditions for this problem in terms of codifferentials were first obtained by the author in [31]. Here we rederive this conditions in the general case to help the reader more readily understand the derivation of optimality conditions for constrained problems, as well as due to the fact the optimality conditions for problem (4.1) are closely related to a natural constraint qualification for isoperimetric constraints.

Recall that a function $v \in \mathrm{L}_{loc}^1(\Omega)$ is called a weak divergence of a vector field $u \in \mathrm{L}^1(\Omega; \mathbb{R}^d)$, if

$$\int_{\Omega} v\varphi \, dx = -\int_{\Omega} \langle u, \nabla\varphi \rangle \, dx \quad \forall \varphi \in \mathrm{C}_c^{\infty}(\Omega).$$

In this case we write $v = \operatorname{div} u$. Denote by $\mathrm{L}^p(\Omega; \mathbb{R}^{m \times d}; \operatorname{div})$ the space of all those functions $u \in \mathrm{L}^p(\Omega; \mathbb{R}^{m \times d})$ for which there exists the weak divergence $\operatorname{div} u = (\operatorname{div}(u_{11}, \ldots, u_{1d}), \ldots, \operatorname{div}(u_{m1}, \ldots, u_{md}))$ and $\operatorname{div} u \in \mathrm{L}^p(\Omega; \mathbb{R}^m)$. Note that in the one-dimensional case (*i.e.* when $d = 1$) the weak divergence $\operatorname{div} u$ coincides with the weak derivative u', which implies that the space $\mathrm{L}^p(\Omega; \mathbb{R}^{m \times 1}; \operatorname{div})$ coincides with the Sobolev space $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$.

Theorem 4.1. *Let f satisfy the codifferentiability conditions of order $p \in [1, +\infty]$, and let either $1 < p \le +\infty$ or the set Ω be bounded and have the segment property. Let also u_* be a locally optimal solution of problem (4.1). Then for any measurable selection $(0, w_1(\cdot), w_2(\cdot))$ of the set-valued map $\overline{d}_{u,\xi}f(\cdot, u_*(\cdot), \nabla u_*(\cdot))$ there exists $\zeta \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^{m \times d}; \operatorname{div})$ satisfying the Euler-Lagrange inclusion*

$$(0, \operatorname{div}(\zeta)(x), \zeta(x)) \in \underline{d}_{u,\xi}f(x, u_*(x), \nabla u_*(x)) + (0, w_1(x), w_2(x)) \quad \textit{for a.e. } x \in \Omega. \tag{4.2}$$

Proof. Define $\mathcal{J}(h) = \mathcal{I}(u_* + h)$ for any $h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m)$. By Corollary 3.11 and Remark 3.12 the functional $\mathcal{J}$ is quasidifferentiable at $u = 0$, *i.e.* its directional derivative at this point has the form

$$\mathcal{J}'(0,h) = \max_{x^* \in \underline{\partial}\mathcal{J}(0)} \langle x^*, h \rangle + \min_{y^* \in \overline{\partial}\mathcal{J}(0)} \langle y^*, h \rangle, \tag{4.3}$$

where the pair $[\underline{\partial}\mathcal{J}(0), \overline{\partial}\mathcal{J}(0)]$ is from Corollary 3.11.

Fix any measurable selection $(0, w_1(\cdot), w_2(\cdot))$ of the set-valued mapping $x \mapsto \overline{\mathrm{d}}_{u,\xi} f(x, u_*(x), \nabla u_*(x))$ and define a linear functional y_0^* as follows:

$$\langle y_0^*, h \rangle = \int_\Omega \left(\langle w_1(x), h(x) \rangle + \langle w_2(x), \nabla h(x) \rangle \right) \mathrm{d}x \quad \forall h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m).$$

Observe that $y_0^* \in \overline{\partial}\mathcal{J}(0)$ by Corollary 3.11.

Recall that u_* is a locally optimal solution of problem (4.1). Therefore, $h = 0$ is a point of local minimum of the functional $\mathcal{J}$, which obviously implies that $\mathcal{J}'(0, h) \geq 0$ for all $h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m)$. Hence by applying (4.3) one obtains that

$$\max_{x^* \in \underline{\partial}\mathcal{J}(0)} \langle x^*, h \rangle + \langle y_0^*, h \rangle \geq 0 \quad \forall h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m).$$

Consequently, $0 \in \underline{\partial}\mathcal{J}(0) + y_0^*$, since otherwise utilising the separation theorem in the space $(\mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m))^*$ equipped with the weak* topology one can find $h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m)$ such that $\max\{\langle x^*, h \rangle \mid x^* \in \underline{\partial}\mathcal{J}(0) + y_0^*\} < 0$, which is impossible. Thus, there exists $x_0^* \in \underline{\partial}\mathcal{J}(0)$ such that $x_0^* + y_0^* = 0$. Hence by Corollary 3.11 there exists a measurable selection $(0, v_1(\cdot), v_2(\cdot))$ of the multifunction $\underline{\mathrm{d}}_{u,\xi} f(\cdot, u_*(\cdot), \nabla u_*(\cdot))$ such that

$$\int_\Omega \left(\langle v_1(x) + w_1(x), h(x) \rangle + \langle v_2(x) + w_2(x), \nabla h(x) \rangle \right) \mathrm{d}x = 0 \quad \forall h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m).$$

Define $\zeta = v_2 + w_2$. Then the equality above implies that there exists the weak divergence of ζ and $\operatorname{div} \zeta = v_1 + w_1$. From the growth condition on the codifferential mapping $D_{u,\xi} f(\cdot)$ (see the definition of codifferentiability conditions, Def. 3.1) it obviously follows that $v_2 + w_2 \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^{m \times d})$ and $v_1 + w_1 \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^m)$. Thus, $\zeta \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^{m \times d}; \operatorname{div})$, and (4.2) holds true by the definition of ζ. $\square$

Let us give an example illustrating optimality conditions from the theorem above.

Example 4.2. Let $d = 2$, $m = 1$, $p = 1$, and $\Omega = (-1, 1) \times (-1, 1)$. Consider the following problem:

$$\min_{u \in \mathrm{W}^{1,1}(\Omega)} \mathcal{I}(u) = \int_\Omega \left(|u'_{x^{(1)}}(x)| - |u'_{x^{(2)}}(x)| \right) \mathrm{d}x, \quad u|_{\partial\Omega} = 0. \tag{4.4}$$

In this case $f(x, u, \xi) = |\xi^{(1)}| - |\xi^{(2)}|$, and we define $u_0 = 0$ (see problem (4.1)). We want to know whether the function $u_* = 0$ is an optimal solution of problem (4.4).

Let us apply optimality conditions in term of the Clarke subdifferential first ([17], Thm. 4.6.1) (see also [12], Sect. 20). Denote $L(u, \xi) = |\xi^{(1)}| - |\xi^{(2)}|$. As is easily seen, the Clarke subdifferential of this function at the origin has the form:

$$\partial_{Cl} L(0, 0) = \operatorname{co}\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} \right\}.$$

Therefore, for the function $\zeta = 0$ one has $(\operatorname{div}\zeta(x), \zeta(x)) \in \partial_{Cl}L(u_*(x), \nabla u_*(x))$ for all $x \in \Omega$, *i.e.* optimality conditions in terms of the Clarke subdifferential ([17], Thm. 4.6.1) are satisfied at $u_* = 0$. One can verify that optimality conditions in terms of K-subdifferential from [67] are satisfied for $u_* = 0$ as well.

Let us now check optimality conditions from Theorem 4.1. With the use of the codifferential calculus [27] one gets

$$\underline{\mathrm{d}}_{u,\xi}f(x,u,\xi) = \operatorname{co}\left\{(\pm\xi^{(1)} - |\xi^{(1)}|, 0, \pm 1, 0)\right\}, \quad \overline{\mathrm{d}}_{u,\xi}f(x,u,\xi) = \operatorname{co}\left\{(\pm\xi^{(2)} + |\xi^{(2)}|, 0, 0, \pm 1)\right\}$$

(here the first coordinate is a, the second is v_1, while the third and fourth ones are v_2 in the notation of the previous section). Therefore, as is readily seen, the integrand f satisfies the codifferentiability conditions of order $p = 1$.

For any $n \in \mathbb{N}$ define

$$\rho_n(t) = \begin{cases} 1, & \text{if } t \in \left[-1 + \frac{2k-2}{2n}, -1 + \frac{2k-1}{2n}\right), k \in \{1, \ldots, 2n\} \\ -1, & \text{if } t \in \left[-1 + \frac{2k-1}{2n}, -1 + \frac{2k}{2n}\right), k \in \{1, \ldots, 2n\}. \end{cases} \tag{4.5}$$

Clearly, the mapping $(0, w_1(\cdot), w_2(\cdot))$ with $w_1(x) = 0$ and $w_2 = (0, \rho_n(x^{(2)}))$ for all $x = (x^{(1)}, x^{(2)}) \in \Omega$ is a measurable selection of the multifunction $\overline{\mathrm{d}}_{u,\xi}f(\cdot, u_*(\cdot), \nabla u_*(\cdot))$ for all $n \in \mathbb{N}$. To verify whether the optimality conditions from Theorem 4.1 hold true, suppose that there exists $\zeta \in \mathrm{L}^\infty(\Omega; \mathbb{R}^{1\times 2}; \operatorname{div})$ such that

$$(0, \operatorname{div}(\zeta)(x), \zeta(x)) \in \underline{\mathrm{d}}_{u,\xi}f(x, u_*(x), \nabla u_*(x)) + (0, w_1(x), w_2(x))\mathrm{d}x = \operatorname{co}\left\{(0, 0, \pm 1, \rho_n(x^{(2)}))\right\}$$

for a.e. $x \in \Omega$. Hence $\operatorname{div}(\zeta)(x) = 0$, $|\zeta_1(x)| \leq 1$, and $\zeta_2(x) = \rho_n(x^{(2)})$ for a.e. $x \in \Omega$. Consequently, by the definition of weak divergence one has

$$\int_\Omega \left(\zeta_1(x)\varphi'_{x^{(1)}}(x) + \rho_n(x^{(2)})\varphi'_{x^{(2)}}(x)\right)\mathrm{d}x = 0 \quad \forall \varphi \in \mathrm{C}_c^\infty(\Omega). \tag{4.6}$$

Since both ρ_n and ζ_1 belong to $\mathrm{L}^\infty(\Omega)$, the equality above holds true for all $\varphi \in \mathrm{W}_0^{1,1}(\Omega)$. Define $\psi_n(t) = 2n\int_{-1}^t \rho_n(\tau)\,d\tau$ for all $t \in (-1, 1)$, and for any $n \in \mathbb{N}$ put $\varphi_n(x) = (-(x^{(1)})^2 + 1)\psi_n(x^{(2)})$. Observe that $\varphi_n \in \mathrm{W}_0^{1,1}(\Omega)$ due to the fact that $\varphi_n(x) = 0$ for all $x \in \partial\Omega$. Hence with the use of (4.6) one gets that

$$0 = \int_\Omega \langle \zeta(x), \nabla\varphi_n(x)\rangle \,\mathrm{d}x = \int_\Omega \left(-2\zeta_1(x)x^{(1)}\psi_n(x^{(2)}) + 2n(-(x^{(1)})^2 + 1)\right)\mathrm{d}x$$
$$= -\int_\Omega 2\zeta_1(x)x^{(1)}\psi_n(x^{(2)})\,\mathrm{d}x + \frac{16n}{3} \geq -4 + \frac{16n}{3} > 0 \quad \forall n \geq 1,$$

which is impossible (the penultimate inequality follows from the fact that $|\zeta_1(x)| \leq 1$ and $\psi_n(x^{(2)}) \in [0, 1]$ for any $x \in \Omega$; see (4.5)). Thus, the optimality conditions from Theorem 4.1 are not satisfied at $u_* = 0$, unlike optimality conditions in terms of the Clarke subdifferential. For the sake of completeness, let us finally note that, in actuality, the functional $\mathcal{I}$ is unbounded below on $\mathrm{W}_0^{1,1}(\Omega)$, which can be easily verified directly or by noting that $u_* = 0$ is not an optimal solution of problem (4.4) and $\mathcal{I}$ is positively homogeneous of degree one.

4.2. Problems with constraints at the boundary

Next we turn to problems with additional constraints at the boundary. For the sake of simplicity we study only the one dimensional case (*i.e.* $d = 1$). Our aim is to obtain optimality conditions for the problem

$$\min \mathcal{I}(u) = \int_\alpha^\beta f(x, u(x), u'(x)) \, \mathrm{d}x + g_0(u(\alpha), u(\beta)), \quad u \in \mathrm{W}^{1,p}((\alpha, \beta); \mathbb{R}^m) \tag{4.7}$$

$$\text{subject to} \quad g_i(u(\alpha), u(\beta)) \le 0, \, i \in I, \quad g_j(u(\alpha), u(\beta)) = 0, \, j \in J.$$

Here $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$ (*i.e.* $\Omega = (\alpha, \beta)$), $f : (\alpha, \beta) \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, and $g_i : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, $i \in I \cup J \cup \{0\}$ are given nonsmooth functions, $I = \{1, \dots, \ell_1\}$ and $J = \{\ell_1 + 1, \dots, \ell_2\}$ for some $\ell_1, \ell_2 \in \mathbb{N} \cup \{0\}$. Observe also that the set $\Omega = (\alpha, \beta)$ is obviously bounded and has the segment property.

For any $u \in \mathrm{W}^{1,p}((\alpha, \beta); \mathbb{R}^m)$ denote $I(u) = \{i \in I \mid g_i(u(\alpha), u(\beta)) = 0\}$. For any subset C of a real vector space E denote by

$$\operatorname{cone} C = \Big\{ \sum_{i=1}^n \alpha_i x_i \; \Big| \; x_i \in C, \; \alpha_i \ge 0, \; i \in \{1, \dots, n\}, \; n \in \mathbb{N} \Big\}$$

the *conic hull* of C (*i.e.* the smallest convex cone containing the set C).

To derive optimality conditions for problem (4.7) we will use general optimality conditions for nonsmooth mathematical programming problems in infinite dimensional spaces in terms of quasidifferentials [34, 35]. To this end, we will suppose that the equality constraints are *polyhedrally codifferentiable*, that is, they are codifferentiable and the sets $\underline{\mathrm{d}}g_j(u(\alpha), u(\beta))$ and $\overline{\mathrm{d}}g_j(u(\alpha), u(\beta))$ are *polytopes* (*i.e.* convex hulls of a finite number of points). This assumption is needed to ensure that certain cones generated by these sets are closed. It should be noted that this assumption can be replaced by a more restrictive constraint qualification (see [34, 35] for more details). For the sake of shortness, we do not consider this alternative assumption and leave it to the interested reader.

Let us also point out that most of codifferentiable functions appearing in applications are, in fact, polyhedrally codifferentiable (see numerous examples in [27]). Thus, the assumption that the equality constraints are polyhedrally codifferentiable is not very restrictive.

Theorem 4.3. *Let f satisfy the codifferentiability conditions of order $p \in [1, +\infty]$, the set-valued maps $\underline{\mathrm{d}}_{u,\xi} f(\cdot)$ and $\overline{\mathrm{d}}_{u,\xi} f(\cdot)$ have the form (3.4) in the case $p = 1$, and u_* be a locally optimal solution of problem (4.7). Suppose also that the functions g_i, $i \in I \cup J \cup \{0\}$, are continuously codifferentiable at the point $(u_*(\alpha), u_*(\beta))$, and the sets $\underline{\mathrm{d}}g_j(u_*(\alpha), u_*(\beta))$ and $\overline{\mathrm{d}}g_j(u_*(\alpha), u_*(\beta))$, $j \in J$, are polytopes. Let finally vectors $(0, s_{1i}, s_{2i}) \in \overline{\mathrm{d}}g_i(u_*(\alpha), u_*(\beta))$, $i \in I$, $(0, s_{1j}, s_{2j}) \in \overline{\mathrm{d}}g_j(u_*(\alpha), u_*(\beta))$, and $(0, r_{1j}, r_{2j}) \in \underline{\mathrm{d}}g_j(u_*(\alpha), u_*(\beta))$, $j \in J$, be such that the following constraint qualification holds true:*

$$C_j \cap \operatorname{cone} \big\{ -C_k \mid k \in J \setminus \{j\} \big\} = \emptyset \quad \forall j \in J \tag{4.8}$$

$$\operatorname{co} \big\{ \underline{\partial} g_i(u_*(\alpha), u_*(\beta)) + (s_{1i}, s_{2i}) \mid i \in I(u_*) \big\} \cap \operatorname{cone} \big\{ -C_j \mid j \in J \big\} = \emptyset, \tag{4.9}$$

where $C_j = \{\underline{\partial} g_j(u_*(\alpha), u_*(\beta)) + (s_{1j}, s_{2j})\} \cup \{-(r_{1j}, r_{2j}) - \overline{\partial} g_j(u_*(\alpha), u_*(\beta))\}$, and the sets $\underline{\partial} g_j(u_*(\alpha), u_*(\beta))$ and $\overline{\partial} g_j(u_*(\alpha), u_*(\beta))$ are defined as in Lemma 2.6.

Then for all $(0, s_{10}, s_{20}) \in \overline{\mathrm{d}}g_0(u_*(\alpha), u_*(\beta))$ and for any measurable selection $(0, w_1(\cdot), w_2(\cdot))$ of the multifunction $\overline{\mathrm{d}}_{u,\xi} f(\cdot, u_*(\cdot), u_*'(\cdot))$ there exist an absolutely continuous function $\zeta \in \mathrm{W}^{1,p'}((\alpha, \beta); \mathbb{R}^m)$, $\lambda_i \ge 0$, $i \in I$, and $\underline{\mu}_j, \overline{\mu}_j \ge 0$, $j \in J$, such that $\lambda_i g_i(u_*(\alpha), u_*(\beta)) = 0$ for all $i \in I$, the Euler-Lagrange inclusion

$$(0, \zeta'(x), \zeta(x)) \in \underline{\mathrm{d}}_{u,\xi} f(x, u_*(x), u_*'(x)) + (0, w_1(x), w_2(x)) \tag{4.10}$$

is satisfied for a.e. $x \in (\alpha, \beta)$, *and the following transversality condition holds true:*

$$(0, \zeta(\alpha), -\zeta(\beta)) \in \underline{d}g_0(u_*(\alpha), u_*(\beta)) + (0, s_{10}, s_{20}) + \sum_{i=1}^{\ell_1} \lambda_i \big(\underline{d}g_i(u_*(\alpha), u_*(\beta)) + (0, s_{1i}, s_{2i})\big)$$

$$+ \sum_{j=\ell_1+1}^{\ell_2} \underline{\mu}_j \big(\underline{d}g_j(u_*(\alpha), u_*(\beta)) + (0, s_{1j}, s_{2j})\big) - \sum_{j=\ell_1+1}^{\ell_2} \overline{\mu}_j \big((0, r_{1j}, r_{2j}) + \overline{d}g_j(u_*(\alpha), u_*(\beta))\big). \quad (4.11)$$

Proof. Let us transform problem (4.7). To this end, recall that $u \in W^{1,p}((\alpha, \beta); \mathbb{R}^m))$ if and only if there exists $h \in L^p((\alpha, \beta); \mathbb{R}^m)$ such that $u(x) = u(\alpha) + \int_\alpha^x h(\tau)d\tau$ for a.e. $x \in (\alpha, \beta)$ (see, *e.g.* [54]). Therefore, the linear operator $\mathcal{T} \colon \mathbb{R}^m \times L^p((\alpha, \beta); \mathbb{R}^m) \to W^{1,p}((\alpha, \beta); \mathbb{R}^m))$ defined as $\mathcal{T}(\eta, h)(x) = \eta + \int_\alpha^x h(\omega) \, d\omega$ is a continuous one-to-one correspondence. Consequently, the pair $(u_*(\alpha), u_*')$ is a point of local minimum of the problem

$$\min_{(\eta,h) \in \mathbb{R}^m \times L^p((\alpha,\beta);\mathbb{R}^m))} \mathcal{J}_0(\eta, h) \quad \text{s.t.} \quad \mathcal{J}_i(\eta, h) \leq 0, \ i \in I, \quad \mathcal{J}_j(\eta, h) = 0, \ j \in J. \quad (4.12)$$

where $\mathcal{J}_0(\eta, h) = \mathcal{I}(\mathcal{T}(\eta, h))$ and $\mathcal{J}_i(\eta, h) = g_i(\eta, \mathcal{T}(\eta, h)(\beta))$, $i \in I \cup J$.

By our assumption the functions g_i, $i \in I \cup \{0\}$ are continuously codifferentiable at $(u_*(\alpha), u_*(\beta))$, while the functional $\int_\alpha^\beta f(x, u(x), u'(x)) \, dx$ is continuously codifferentiable by Theorem 3.3. Consequently, by Theorem 4.5 of [32] the functions $\mathcal{J}_i$, $i \in I \cup J \cup \{0\}$ are continuously codifferentiable at the point $(u_*(\alpha), u_*')$, the set

$$\left\{(A, x^*) \in \mathbb{R} \times \left(\mathbb{R}^m \times L^p((\alpha, \beta); \mathbb{R}^m)\right)^* \,\Big|\, A = \int_\alpha^\beta a(x) \, dx + a_0, \quad \forall(\eta, h) \in \mathbb{R}^m \times L^p((\alpha, \beta); \mathbb{R}^m) \right.$$

$$\langle x^*, (\eta, h) \rangle = \left\langle \int_\alpha^\beta v_1(x) \, dx + r_1 + r_2, \eta \right\rangle + \int_\alpha^\beta \left\langle \int_x^\beta v_1(\tau)d\tau + v_2(x) + r_2, h(x) \right\rangle dx,$$

$$\left. (a_0, r_1, r_2) \in \underline{d}g_0(u_*(\alpha), u_*(\beta)), \ (a(\cdot), v_1(\cdot), v_2(\cdot)) \text{ is a measurable selection of } \underline{d}_{u,\xi} f(\cdot, u_*(\cdot), u_*'(\cdot)) \right\} \quad (4.13)$$

is a hypodifferential of $\mathcal{J}_0$ at $(u_*(\alpha), u_*')$, while the set

$$\left\{(a, x^*) \in \mathbb{R} \times \left(\mathbb{R}^m \times L^p((\alpha, \beta); \mathbb{R}^m)\right)^* \,\Big|\, \langle x^*, (\eta, h) \rangle = \langle r_1 + r_2, \eta \rangle + \int_\alpha^\beta \langle r_2, h(x) \rangle dx \right.$$

$$\left. \forall(\eta, h) \in \mathbb{R}^m \times L^p((\alpha, \beta); \mathbb{R}^m), \quad (a, r_1, r_2) \in \underline{d}g_i(u_*(\alpha), u_*(\beta)) \right\} \quad (4.14)$$

is a hypodifferential of $\mathcal{J}_i$ at $(u_*(\alpha), u_*')$, $i \in I \cup J$. The hyperdifferentials $\overline{d}\mathcal{J}_i(u_*(\alpha), u_*')$ are defined in the same way. Thus, by Lemma 2.6 the functions $\mathcal{J}_i$, $i \in I \cup J \cup \{0\}$, are Hadamard quasidifferentiable at $(u_*(\alpha), u_*')$.

With the use of optimality conditions for nonsmooth mathematical programming problems in terms of quasidifferentials ([35], Cor. 4, Prop. 1, and Lem. 2) one obtains that if for some $x_j^* \in \underline{\partial}\mathcal{J}_j(u_*(\alpha), u_*')$, $j \in J$, and $y_i^* \in \overline{\partial}\mathcal{J}_i(u_*(\alpha), u_*')$, $i \in I \cup J$, one has

$$D_j \cap \text{cone}\{-D_k \mid k \in J \setminus \{j\}\} = \emptyset \ \forall j \in J, \quad \text{co}\left\{\underline{\partial}\mathcal{J}_i(u_*(\alpha), u_*') + y_i^* \mid i \in I(u_*)\right\} \cap \text{cone}\{-D_j \mid j \in J\} = \emptyset \quad (4.15)$$

(here $D_j = \{\underline{\partial}\mathcal{J}_j(u_*(\alpha), u_*') + y_j^*\} \cup \{-x_j^* - \overline{\partial}\mathcal{J}_j(u_*(\alpha), u_*')\}$ for any $j \in J$), then for any $y_0^* \in \overline{\partial}\mathcal{J}_0(u_*(\alpha), u_*')$ there exist $\lambda_i \geq 0$, $i \in I$, and $\underline{\mu}_j, \overline{\mu}_j \geq 0$, $j \in J$, such that $\lambda_i \mathcal{J}_i(u_*(\alpha), u_*') = 0$ for any $i \in I$ and

$$
\begin{aligned}
0 \in \underline{\partial}\mathcal{J}_0(u_*(\alpha), u_*') + y_0^* &+ \sum_{i \in I} \lambda_i \big(\underline{\partial}\mathcal{J}_j(u_*(\alpha), u_*') + y_i^*\big) \\
&+ \sum_{j \in J} \underline{\mu}_j \big(\underline{\partial}\mathcal{J}_j(u_*(\alpha), u_*') + y_j^*\big) - \sum_{j \in J} \overline{\mu}_j \big(x_j^* + \overline{\partial}\mathcal{J}_j(u_*(\alpha), u_*')\big).
\end{aligned}
\tag{4.16}
$$

Let us rewrite these optimality conditions in terms of the original problem (4.7).

For all $(s_1, s_2) \in \mathbb{R}^m \times \mathbb{R}^m$ define linear functional $\Theta(s_1, s_2)$ as follows:

$$
\langle \Theta(s_1, s_2), (\eta, h) \rangle = \langle s_{1i} + s_{2i}, \eta \rangle + \int_\alpha^\beta \langle s_{2i}, h(x) \rangle \mathrm{d}x \quad \forall (\eta, h) \in \mathbb{R}^m \times \mathrm{L}^p((\alpha, \beta); \mathbb{R}^m).
\tag{4.17}
$$

Fix any $(0, s_{1i}, s_{2i}) \in \overline{\mathrm{d}}g_i(u_*(\alpha), u_*(\beta))$, $i \in I \cup J$, and $(0, r_{1j}, r_{2j}) \in \underline{\mathrm{d}}g_j(u_*(\alpha), u_*(\beta))$, $j \in J$ satisfying (4.8) and (4.9), and put $y_i^* = \Theta(s_{1i}, s_{2i})$, $i \in I \cup J$, and $x_j^* = \Theta(r_{1j}, r_{2j})$, $j \in J$. Then $y_i^* \in \overline{\partial}\mathcal{J}_i(u_*(\alpha), u_*')$ for all $i \in I \cup J$ and $x_j^* \in \underline{\partial}\mathcal{J}_j(u_*(\alpha), u_*')$ for all $j \in J$ according to (4.14) and Lemma 2.6. Let us check that these functionals y_i^* and x_j^* satisfy constraint qualification (4.15).

Indeed, by virtue of (4.14) and Lemma 2.6 one has $\Theta(\underline{\partial}g_j(u_*(\alpha), u_*(\beta))) = \underline{\partial}\mathcal{J}_j(u_*(\alpha), u_*')$ and the same equality holds true for the superdifferentials. Hence taking into account the fact that Θ is a linear operator (see (4.17)) one obtains that $\Theta(C_j) = D_j$, $j \in J$, and $\Theta(\mathrm{cone}\{C_k \mid k \in J \setminus \{j\}\}) = \mathrm{cone}\{D_k \mid k \in J \setminus \{j\}\}$ for all $j \in J$. One can readily verify that Θ is an injective mapping (see (4.17)). Therefore (4.8) implies the first condition in (4.15). Similarly, (4.9) implies the second condition in (4.15).

Thus, constraint qualification (4.15) is satisfied. Consequently, with the use of (4.16), (4.13), (4.14), and Lemma 2.6 one gets that for all $(0, s_{10}, s_{20}) \in \overline{\mathrm{d}}g_0(u_*(\alpha), u_*(\beta))$ and for any measurable selection $(0, w_1(\cdot), w_2(\cdot))$ of the multifunction $\overline{\mathrm{d}}_{u,\xi} f(\cdot, u_*(\cdot), u_*'(\cdot))$ there exist $\lambda_i \geq 0$, $i \in I$, $\underline{\mu}_j, \overline{\mu}_j \geq 0$, $j \in J$, vectors $(0, r_{1i}, r_{2i}) \in \underline{\mathrm{d}}g_i(u_*(\alpha), u_*(\beta))$, $i \in I \cup \{0\}$, $(0, \xi_{1j}, \xi_{2j}) \in \underline{\mathrm{d}}g_j(u_*(\alpha), u_*(\beta))$, $j \in J$, and $(0, y_{1j}, y_{2j}) \in \overline{\mathrm{d}}g_j(u_*(\alpha), u_*(\beta))$, $j \in J$, and a measurable selection $(0, v_1(\cdot), v_2(\cdot))$ of the multifunction $\underline{\mathrm{d}}_{u,\xi} f(\cdot, u_*(\cdot), u_*'(\cdot))$ such that for all $i \in I$ one has $\lambda_i g_i(u_*(\alpha), u_*(\beta)) = 0$, for any $\eta \in \mathbb{R}^m$ one has

$$
\begin{aligned}
\Big\langle \int_\alpha^\beta (v_1(x) + w_1(x))\, \mathrm{d}x + r_{10} + r_{20} + s_{10} + s_{20} + \sum_{i \in I} \lambda_i \big(r_{1i} + r_{2i} + s_{1i} + s_{2i}\big) \\
+ \sum_{j \in J} \underline{\mu}_j \big(\xi_{1j} + \xi_{2j} + s_{1j} + s_{2j}\big) - \sum_{j \in J} \overline{\mu}_j \big(r_{1j} + r_{2j} + y_{1j} + y_{2j}\big), \eta \Big\rangle = 0,
\end{aligned}
\tag{4.18}
$$

and for any $h \in \mathrm{L}^p((\alpha, \beta); \mathbb{R}^m)$ one has

$$
\begin{aligned}
\int_\alpha^\beta \Big\langle \int_x^\beta (v_1(\tau) + w_1(\tau))\, \mathrm{d}\tau + v_2(x) + w_2(x) + r_{20} + s_{20} \\
+ \sum_{i \in I} \lambda_i \big(r_{2i} + s_{2i}\big) + \sum_{j \in J} \underline{\mu}_j \big(\xi_{2j} + s_{2j}\big) - \sum_{j \in J} \overline{\mu}_j \big(r_{2j} + y_{2j}\big), h(x) \Big\rangle \mathrm{d}x = 0.
\end{aligned}
\tag{4.19}
$$

Denote

$$
\zeta(x) = -\int_x^\beta (v_1(\tau) + w_1(\tau))\, \mathrm{d}\tau - r_{20} - s_{20} - \sum_{i \in I} \lambda_i \big(r_{2i} + s_{2i}\big) - \sum_{j \in J} \underline{\mu}_j \big(\xi_{2j} + s_{2j}\big) + \sum_{j \in J} \overline{\mu}_j \big(r_{2j} + y_{2j}\big)
$$

for any $x \in [\alpha, \beta]$. Then ζ is an absolutely continuous function such that

$$(0, \zeta'(x), \zeta(x)) = (0, v_1(x) + w_1(x), v_2(x) + w_2(x)) \quad \text{for a.e. } x \in (\alpha, \beta)$$

due to (4.19), $\zeta \in W^{1,p'}((\alpha, \beta); \mathbb{R}^m)$ due to the growth condition on the codifferential mapping $D_{u,\xi} f(\cdot)$ (see Def. 3.1), and

$$(0, \zeta(\alpha), -\zeta(\beta)) = (0, r_{10} + s_{10}, r_{20} + s_{20}) + \sum_{i \in I} \lambda_i (0, r_{1i} + s_{1i}, r_{2i} + s_{2i})$$

$$+ \sum_{j \in J} \underline{\mu}_j (0, \xi_{1j} + s_{1j}, \xi_{2j} + s_{2j}) - \sum_{j \in J} \overline{\mu}_j (0, r_{1j} + y_{1j}, r_{2j} + y_{2j})$$

due to (4.18). It remains to note that the first equality above is equivalent to (4.10), while the second one is equivalent to the transversality condition (4.11). □

Remark 4.4. (i) In the case when there are no equality constraints, the constraint qualification (4.8), (4.9) from the previous theorem takes an especially simple form. Namely, it is sufficient to suppose that for some $(0, s_{1i}, s_{2i}) \in \overline{\mathrm{d}} g_i(u_*(\alpha), u_*(\beta))$, $i \in I(u_*)$, one has $0 \notin \mathrm{co}\{\underline{\mathrm{d}} g_i(u_*(\alpha), u_*(\beta)) + (0, s_{1i}, s_{2i}) \mid i \in I(u_*)\}$. In the case when there are no inequality constraints and there is only one equality constraint, the constraint qualification (4.8), (4.9) also takes a very simple form. One has to suppose that $0 \notin \underline{\mathrm{d}} g_1(u_*(\alpha), u_*(\beta)) + (0, s_1, s_2)$ and $0 \notin (0, r_1, r_2) + \overline{\mathrm{d}} g_1(u_*(\alpha), u_*(\beta))$ for some $(0, s_1, s_2) \in \overline{\mathrm{d}} g_1(u_*(\alpha), u_*(\beta))$ and $(0, r_1, r_2) \in \underline{\mathrm{d}} g_1(u_*(\alpha), u_*(\beta))$.
(ii) It should be noted that there are *two* Lagrange multipliers $\underline{\mu}_j$ and $\overline{\mu}_j$ corresponding to each equality constraint $g_j(u(\alpha), u_j(\beta)) = 0$, which is a specific feature of optimality conditions for nonsmooth optimisation problems in terms of quasidifferentials. See [34, 35] for more details.

Let us also present an example illustrating optimality conditions for problem (4.7).

Example 4.5. Let $d = 1$, $m = 2$, $p = 2$, and $\Omega = (0, 1)$. Consider the following problem:

$$\min \mathcal{I}(u_1, u_2) = \int_0^1 \left(\frac{1}{2}((u_1'(x))^2 + (u_2'(x))^2) + |u_1(x)| + |u_2(x)| \right) \mathrm{d}x - u_1(1) + u_2(1) \tag{4.20}$$
$$\text{subject to} \quad |u_1(1)| - |u_2(1)| = 0, \quad u \in W^{1,2}((0, 1); \mathbb{R}^m).$$

In this case, $f(x, u, \xi) = |u_1| + |u_2| + 0.5(\xi_1)^2 + 0.5(\xi_2)^2$, $g_0(u(0), u(1)) = -u_1(1) + u_2(1)$, $I = \emptyset$, $J = \{1\}$, and $g_1(u(0), u(1)) = |u_1(1)| - |u_2(1)|$. Let us check whether the function $u_* = (0, 0)$ is an optimal solution of problem (4.20).

First we apply optimality conditions for problem (4.20) in terms of the Clarke subdifferential ([17], Thm. 4.4.1). Denote $L(u, \xi) = |u_1| + |u_2| + 0.5(\xi_1)^2 + 0.5(\xi_2)^2$. The Hamiltonian for problem (4.20) is defined as

$$H(u, p) = \sup_{\xi \in \mathbb{R}^2} (\langle p, \xi \rangle - L(u, \xi)) = -|u_1| - |u_2| + \frac{1}{2}(p_1^2 + p_2^2).$$

The Clarke subdifferentials of the Hamiltonian and the function $\ell(u(1)) = -u_1(1) + u_2(1)$ at the origin have the form:

$$\partial_{Cl} H(0, 0) = \mathrm{co}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\}, \quad \partial_{Cl} \ell(0) = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}.$$

Furthermore, one can verify that the Clarke normal cone $N_S(0)$ to the set $S = \{(y, z) \in \mathbb{R}^2 \mid |y| - |z| = 0\}$ at the origin (see [17], Sect. 2.4) is equal to $\mathbb{R}^2$. Therefore, for the function $p_*(x) \equiv 0$ one has

$$\begin{pmatrix} -p'_*(x) \\ u'_*(x) \end{pmatrix} \in \partial_{Cl} H(u_*(x), p_*(x)) \quad \forall x \in [0, 1], \quad p_*(0) = 0, \quad -p_*(1) \in \partial_{Cl} \ell(u_*(1)) + N_S(u_*(1)),$$

i.e. optimality conditions in terms of the Clarke subdifferential ([17], Thm. 4.4.1) are satisfied at u_*.

Next we check optimality conditions in terms of the limiting proximal subdifferential from [49]. Define $\ell(u(0), u(1)) = -u_1(1) + u_2(1)$, if $|u_1(1)| - |u_2(1)| = 0$, and $\ell(u(0), u(1)) = +\infty$, otherwise. Then problem (4.20) can be rewritten as the following generalised problem of Bolza:

$$\min \ell(u(0), u(1)) + \int_0^1 L(u(x), u'(x)) \, \mathrm{d}x, \quad u \in W^{1,1}((0, 1); \mathbb{R}^2). \tag{4.21}$$

One can readily verify that the limiting proximal subdifferentials, which we denote by ∂_p^∞, of the functions L and ℓ at the origin have the form:

$$\partial_p^\infty L(0, 0) = \mathrm{co} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\}, \quad \partial_p^\infty \ell(0, 0) = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} + \left\{ \begin{pmatrix} 0 \\ 0 \\ t \\ -t \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ t \\ t \end{pmatrix} \middle| t \in \mathbb{R} \right\}.$$

Therefore, for the function $p_*(x) \equiv 0$ one has

$$p'_*(x) \in \mathrm{co} \left\{ w \in \mathbb{R}^2 \mid (w, p_*(x)) \in \partial_p^\infty L(u_*(x), u'_*(x)) \right\} = \mathrm{co} \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\} \quad \forall x \in [0, 1]$$

and $(p_*(0), -p_*(1)) \in \partial_p^\infty \ell(u_*(0), u_*(1))$. Furthermore, one also has

$$L(u_*(x), y) = \frac{1}{2} \left((y_1)^2 + (y_2)^2 \right) \geq 0 = L(u_*(x), u'_*(x)) + \langle p_*(x), y - u'_*(x) \rangle \quad \forall y = (y_1, y_2) \in \mathbb{R}^2.$$

Thus, optimality conditions in terms of the limiting proximal subdifferential ([49], Thm. 1) are satisfied at u_* as well. In addition, one can check that the limiting proximal subdifferentials $\partial_p^\infty L(0, 0)$ and $\partial_p^\infty l(0, 0)$ coincide with the corresponding limiting Fréchet subdifferentials (which, in turn, coincide with the Mordukhovich basic subdifferentials by [65], Thm. 1.89), which implies that the optimality conditions in terms of the limiting Fréchet subdifferential ([51], Thm. 3.4) are satisfied at u_* as well.

Let us finally check optimality conditions in terms of codifferentials from Theorem 4.3. Applying the codifferential calculus [27] one obtains that

$$\underline{\mathrm{d}}_{u,\xi} f(x, u, \xi) = \mathrm{co} \left\{ \left((-1)^i u_1 - |u_1| + (-1)^j u_2 - |u_2|, (-1)^i, (-1)^j, \xi_1, \xi_2 \right) \in \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \middle| i, j \in \{1, 2\} \right\}$$

and $\overline{\mathrm{d}}_{u,\xi} f(x, u, \xi) = \{0\}$. One also gets that

$$\underline{\mathrm{d}} g_0(y, z) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\}, \quad \overline{\mathrm{d}} g_0(y, z) = \{0\}, \quad \underline{\mathrm{d}} g_1(y, z) = \mathrm{co} \left\{ \begin{pmatrix} \pm z_1 - |z_1| \\ 0 \\ 0 \\ \pm 1 \\ 0 \end{pmatrix} \right\}, \quad \overline{\mathrm{d}} g_1(y, z) = \mathrm{co} \left\{ \begin{pmatrix} \pm z_2 + |z_2| \\ 0 \\ 0 \\ 0 \\ \pm 1 \end{pmatrix} \right\},$$

for all $y = (y_1, y_2) \in \mathbb{R}^2$ and $z = (z_1, z_2) \in \mathbb{R}^2$. Thus, the integrand f satisfies the codifferentiability conditions of order $p = 2$, the functions g_0 and g_1 are continuously codifferentiable, and the sets $\underline{\mathrm{d}} g_1(u_*(0), u_*(1))$ and $\overline{\mathrm{d}} g_1(u_*(0), u_*(1))$ are polytopes. Furthermore, observe that for vectors $\eta_1 = (0, 0, 0, 1, 0) \in \underline{\mathrm{d}} g_1(u_*(0), u_*(1))$ and $\eta_2 = (0, 0, 0, 0, 1) \in \overline{\mathrm{d}} g_1(u_*(0), u_*(1))$ one has $0 \notin \underline{\mathrm{d}} g_1(u_*(0), u_*(1)) + \eta_2$ and $0 \notin \eta_1 + \overline{\mathrm{d}} g_1(u_*(0), u_*(1))$, *i.e.* the constraint qualification from Theorem 4.3 is satisfied at u_* (see Rem. 4.4).

Suppose that optimality conditions from Theorem 4.3 are satisfied at u_*. Then there exist an absolutely continuous function $\zeta \in \mathrm{W}^{1,2}((0,1); \mathbb{R}^2)$ and $\underline{\mu}_1, \overline{\mu}_1 \geq 0$ such that

$$(0, \zeta'(x), \zeta(x)) \in \underline{\mathrm{d}}_{u,\xi} f(x, u_*(x), u'_*(x)) = \mathrm{co}\left\{ \left(0, (-1)^i, (-1)^j, 0, 0\right) \in \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \,\Big|\, i, j \in \{1, 2\}\right\}$$

for a.e. $x \in (0, 1)$, and the transversality condition

$$(0, \zeta(0), -\zeta(1)) \in \nabla g_0(u_*(0), u_*(1)) + \underline{\mu}_1\big(\underline{\mathrm{d}} g_1(u_*(0), u_*(1)) + \eta_2\big) - \overline{\mu}_1\big(\eta_1 + \overline{\mathrm{d}} g_1(u_*(0), u_*(1))\big).$$

holds true. Therefore $\zeta(x) \equiv 0$ and the transversality condition takes the form

$$0 \in \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \underline{\mu}_1 \,\mathrm{co}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\} - \overline{\mu}_1 \,\mathrm{co}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}$$

or, equivalently,

$$-1 - \underline{\mu}_1 - \overline{\mu}_1 \leq 0 \leq -1 + \underline{\mu}_1 - \overline{\mu}_1, \quad 1 + \underline{\mu}_1 - \overline{\mu}_1 \leq 0 \leq 1 + \underline{\mu}_1 + \overline{\mu}_1.$$

The third inequality implies that $1 + \underline{\mu}_1 \leq \overline{\mu}_1$, while the second one yields $1 + \overline{\mu}_1 \leq \underline{\mu}_1$. Consequently, $2 + \underline{\mu}_1 \leq 1 + \overline{\mu}_1 \leq \underline{\mu}_1$, which is impossible. Hence optimality conditions from Theorem 4.3 are not satisfied at u_*, and one can conclude that this point is not an optimal solution of problem (4.20). Thus, optimality conditions in terms of codifferentials detect the non-optimality of u_*, while optimality conditions in terms of Clarke, limiting proximal, and limiting Fréchet subdifferentials fail to do so.

4.3. Problems with isoperimetric constraints

Let us now consider problems with isoperimetic inequality constraints. With the use of optimality conditions for general quasidifferentiable programming problems in Banach spaces [34, 35] one can derive optimality conditions for nonsmooth problems with both isoperimetric equality and inequality constraints. However, this approach requires the use of constraint qualifications (similar to the ones used in Thm. 4.3), whose reformulation in the case of isoperimetric constraints leads to very cumbersome assumptions, which we do not present here for the sake of shortness.

Consider isoperimetric problem of the form:

$$\min \mathcal{I}_0(u) = \int_\Omega f_0(x, u(x), \nabla u(x)) \, \mathrm{d}x$$
$$\text{subject to} \quad \mathcal{I}_i(u) = \int_\Omega f_i(x, u(x), \nabla u(x)) \, \mathrm{d}x \leq 0, \quad u \in u_0 + \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m). \tag{4.22}$$

Here $\Omega \subset \mathbb{R}^d$ is an open set, $f_i \colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to \mathbb{R}$, $f_i = f_i(x, u, \xi)$, are nonsmooth functions, $i \in I \cup \{0\}$, $I = \{1, \ldots, \ell\}$, and $u_0 \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ is a fixed function. Denote $I(u) = \{i \in I \mid \mathcal{I}_i(u) = 0\}$.

Our aim is to derive optimality conditions for problem (4.22) with the use of general optimality conditions in terms of quasidifferentials for inequality constrained nonsmooth optimisation problem ([35], Cor. 5). It should be noted that these optimality conditions were largely inspired by B.N. Pschenichny work [73] and are derived with the use of the standard trick, which goes back to Pschenichny, of reducing an inequality constrained optimisation problem to the problem of minimising the nonsmooth max-envelope of the objective function and constraints.

Theorem 4.6. *Let f_i, $i \in I \cup \{0\}$, satisfy the codifferentiability conditions of order $p \in [1, +\infty]$, and let either $1 < p \leq +\infty$ or the set Ω be bounded and have the segment property. Suppose also that u_* is a locally optimal solution of problem (4.22). Let finally $(0, w_{1i}(\cdot), w_{2i}(\cdot))$ be measurable selections of the multifunctions*

$\overline{\mathrm{d}}_{u,\xi} f_i(\cdot, u_*(\cdot), \nabla u_*(\cdot))$, $i \in I$, such that there does not exist $\zeta \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^{m \times d}; \mathrm{div})$ satisfying the following inclusion for a.e. $x \in \Omega$:

$$(0, \mathrm{div}(\zeta)(x), \zeta(x)) \in \mathrm{co}\left\{\underline{\mathrm{d}}_{u,\xi} f_i(x, u_*(x), \nabla u_*(x)) + (0, w_{1i}(x), w_{2i}(x)) \mid i \in I(u_*)\right\}. \tag{4.23}$$

Then for any measurable selection $(0, w_{10}(\cdot), w_{20}(\cdot))$ of the set-valued map $\overline{\mathrm{d}}_{u,\xi} f_0(\cdot, u_*(\cdot), \nabla u_*(\cdot))$ one can find $\lambda_i \geq 0$, $i \in I$, and $\zeta \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^{m \times d}; \mathrm{div})$ such that $\lambda_i \mathcal{I}_i(u_*) = 0$ for any $i \in I$, and for a.e. $x \in \Omega$ one has

$$(0, \mathrm{div}(\zeta)(x), \zeta(x)) \in \underline{\mathrm{d}}_{u,\xi} f_0(x, u_*(x), \nabla u_*(x)) + (0, w_{10}(x), w_{20}(x))$$
$$+ \sum_{i=1}^{\ell} \lambda_i \left(\underline{\mathrm{d}}_{u,\xi} f_i(x, u_*(x), \nabla u_*(x)) + (0, w_{1i}(x), w_{2i}(x))\right). \tag{4.24}$$

Proof. For any $h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m)$ define $\mathcal{J}_0(h) = \mathcal{I}_0(u_* + h)$, and $\mathcal{J}_i(h) = \mathcal{I}_i(u_* + h)$, $i \in I$. By Corollary 3.11 and Remark 3.12 the functions $\mathcal{J}_i$ are correctly defined and quasidifferentiable at $h = 0$. Moreover, the point $h = 0$ is a locally optimal solution of the problem

$$\min_{h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m)} \mathcal{J}_0(h) \quad \text{subject to} \quad \mathcal{J}_i(h) \leq 0, \quad i \in I,$$

since u_* is a locally optimal solution of problem (4.22). Hence by applying optimality conditions for quasidifferentiable programming problems with inequality constraints ([35], Cor. 5) one obtains that if $y_i^* \in \overline{\partial} \mathcal{J}_i(0)$, $i \in I$, are such that

$$0 \notin \mathrm{co}\{\underline{\partial} \mathcal{J}_i(0) + y_i^* \mid i \in I(u_*)\}, \tag{4.25}$$

then for any $y_0^* \in \overline{\partial} \mathcal{J}_0(0)$ one can find $\lambda_i \geq 0$, $i \in I$, such that $\lambda_i \mathcal{J}_i(0) = 0$ for any $i \in I$ and

$$0 \in \underline{\partial} \mathcal{J}_0(0) + y_0^* + \sum_{i \in I} \lambda_i \left(\underline{\partial} \mathcal{J}_i(0) + y_i^*\right). \tag{4.26}$$

Let us reformulate these optimality conditions in term of problem (4.22).

Fix any $(0, w_{1i}(\cdot), w_{2i}(\cdot))$, $i \in I$, satisfying the assumptions of the theorem. Define

$$\langle y_i^*, h \rangle = \int_{\Omega} \left(\langle w_{1i}(x), h(x) \rangle + \langle w_{2i}(x), \nabla h(x) \rangle\right) \mathrm{d}x \quad \forall h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m). \tag{4.27}$$

Then by Corollary 3.11 one has $y_i^* \in \overline{\partial} \mathcal{J}_i(0)$, $i \in I$. Let us check that constraint qualification (4.25) holds true. Indeed, arguing by reductio ad absurdum suppose that condition (4.25) is not satisfied. Then for any $i \in I$ there exist $x_i^* \in \underline{\partial} \mathcal{J}_i(0)$ and $\alpha_i \geq 0$ such that

$$\sum_{i=1}^{\ell} \alpha_i (x_i^* + y_i^*) = 0, \quad \sum_{i=1}^{\ell} \alpha_i = 1.$$

Hence with the use of Corollary 3.11 one obtains that for any $i \in I$ there exists a measurable selection $(0, v_{1i}(\cdot), v_{2i}(\cdot))$ of the multifunction $\underline{\mathrm{d}}_{u,\xi} f_i(\cdot, u_*(\cdot), \nabla u_*(\cdot))$ such that

$$\int_{\Omega} \left(\langle \omega_1(x), h(x) \rangle + \langle \omega_2(x), \nabla h(x) \rangle\right) \mathrm{d}x = 0 \quad \forall h \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^m), \tag{4.28}$$

where

$$\omega_1(x) = \sum_{i=1}^{m} \alpha_i(v_{1i}(x) + w_{1i}(x)), \quad \omega_2(x) = \sum_{i=1}^{m} \alpha_i(v_{12}(x) + w_{12}(x)).$$

Equality (4.28) implies that there exists the weak divergence of the function $\zeta = \omega_2$, and $\operatorname{div} \zeta = \omega_1$. By the growth condition on the codifferential mappings $D_{u,\xi} f_i(\cdot)$ (see Def. 3.1) one has $\omega_2 \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^{m \times d})$ and $\omega_1 \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^m)$. Thus, there exists $\zeta \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^{m \times d}; \operatorname{div})$ such that

$$(0, \operatorname{div}(\zeta)(x), \zeta(x)) = \sum_{i=1}^{\ell} \alpha_i\big((a_i(x), v_{1i}(x), v_{2i}(x)) + (0, w_{1i}(x), w_{2i}(x))\big)$$

for a.e. $x \in \Omega$, which contradicts (4.23). Therefore constraint qualification (4.25) holds true.

Choose any measurable selection $(0, w_{10}(\cdot), w_{20}(\cdot))$ of the set-valued mapping $\overline{\mathrm{d}}_{u,\xi} f_0(\cdot, u_*(\cdot), \nabla u_*(\cdot))$. Define linear functional y_0^* in the same way as in (4.27). Then by Corollary 3.11 one has $y_0 \in \overline{\partial} \mathcal{J}_0(0)$. Consequently, there exist $\lambda_i \geq 0$, $i \in I$, such that $\lambda_i \mathcal{J}_i(0) = \lambda_i \mathcal{I}_i(u_*) = 0$ for any $i \in I$ and (4.26) holds true. Now, arguing in the same way as in the proof of Theorem 4.1 one can readily verify that optimality condition (4.26) is equivalent to (4.24). □

Let us give an example illustrating optimality conditions for isoperimetric problems from the theorem above.

Example 4.7. Let $d = m = p = 1$ and $\Omega = (0, 1)$. Consider the following problem:

$$\min \mathcal{I}_0(u) = \int_0^1 \max\big\{-|u(x)|, -|u'(x)|\big\}\, \mathrm{d}x \tag{4.29}$$
$$\text{subject to } \mathcal{I}_1(u) = \int_0^1 u(x)\, \mathrm{d}x \leq 0, \quad u(0) = u(1) = 0, \quad u \in \mathrm{W}^{1,1}(0, 1).$$

In this case $f_0(x, u, \xi) = \max\{-|u|, -|\xi|\}$, $I = \{1\}$ and $f_1(x, u, \xi) = u$. Let us check whether optimality conditions are satisfied at $u_* \equiv 0$. It is easily seen that this function is not a locally optimal solution of problem (4.29), since for the function $u_\alpha(x) = \alpha x(x - 1)$ one obviously has $\mathcal{I}_0(u_\alpha) < 0$ and $\mathcal{I}_1(u_\alpha) = -\alpha/6 < 0$ for any $\alpha > 0$. In actuality, u_* is a point of unconstrained global maximum of $\mathcal{I}_0(u)$.

To the best of the author's knowledge, optimality conditions for nonsmooth variational problems with isoperimetric constraints have been obtained earlier only in ([6], Thm. 3.5.1). Let us verify whether these optimality conditions hold true at u_*. The limiting Fréchet subdifferential of the function $u \mapsto f_0(x, u, \xi)$ with respect to ξ (see [52]) at the point $(x, 0, 0)$, which we denote by $\partial_{F,u}^\infty f_0(x, \cdot)(0, 0)$, is equal to $[-1, 1]$. Therefore, for the function $p(\cdot) \equiv 0$ and for all $x \in [0, 1]$ one has

$$p'(x) \in \partial_{F,u}^\infty f_0(x, \cdot)(u_*(x), u_*'(x)), \quad p(x)u_*'(x) - f_0(x, u_*(x), u_*'(x)) = 0 = \max_{v \in \mathbb{R}}\big(p(x)v - f_0(x, u_*(x), v)\big),$$

that is, the optimality conditions ([6], Thm. 3.5.1) are satisfied for $p(\cdot) \equiv 0$, $\lambda = 1$, and $\gamma = 0$.

To apply other optimality condition to problem (4.29), one needs to transform this problem to an equivalent one without isoperimetric constraints. Such transformation can be done in many different ways. Following Example 4.5.4 of [17] we can reformulate problem (4.29) as the following Mayer problem with nonholonomic inequality constraints:

$$\min f(x(1)) = x_3(1) \quad \text{subject to} \quad x(0) = 0, \quad x_1(1) = 0, \quad x_2(1) \leq 0,$$
$$\varphi_1(x(t), \dot{x}(t)) = x_1(t) - \dot{x}_2(t) \leq 0, \quad \varphi_2(x(t), \dot{x}(t)) = \max\{-|x_1(t)|, -|\dot{x}_1(t)|\} - \dot{x}_3(t) \leq 0, \quad t \in [0, 1]. \tag{4.30}$$

Let us verify optimality conditions for this problem ([17], Cor. 4.5.1) at the point $x_* \equiv 0$, which corresponds to the point $u_* \equiv 0$ in problem (4.29). Indeed, the Clarke subdifferentials of the functions φ_i at the origin have the form:

$$\partial_{Cl}\varphi_1(0) = \{(1, 0, 0, 0, -1, 0)\}, \quad \partial_{Cl}\varphi_2(0) = \mathrm{co}\left\{\left(\pm 1, 0, 0, 0, 0, -1\right), \left(0, 0, 0, \pm 1, 0, -1\right)\right\}.$$

As is readily seen, the constraint qualifications from Corollary 4.5.1 of [17] is satisfied at x_*. Note also that the Clarke normal cone to the set $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 = 0, x_2 \le 0\}$ at the origin has the form $N_S(0) = \{(t, s, 0) \in \mathbb{R}^3 \mid t \in \mathbb{R}, s \ge 0\}$. Therefore, for $p(t) \equiv (0, 0, -1)$, $\lambda_1(t) \equiv 0$, $\lambda_2(t) \equiv 1$, and $\lambda_0 = 1$ one has

$$(\dot{p}(t), p(t)) \in \lambda_1(t)\varphi_1(x_*(t), \dot{x}_*(t)) + \lambda_2(t)\varphi_2(x_*(t), \dot{x}_*(t)) \quad \forall t \in [0, 1]$$

and $-p(1) \in \lambda_0 \partial_{Cl} f(x_*(1)) + N_S(x_*(1))$. Thus, optimality conditions for problem (4.30) in terms of the Clarke subdifferential ([17], Cor. 4.5.1) are satisfied at the point $x_*(\cdot) = 0$.

Problem (4.29) can also be rewritten as the following nonsmooth optimal control problem:

$$\min \ \mathcal{J}(x, u) = \int_0^1 \max\left\{-|x_1(t)|, -|u(t)|\right\} \mathrm{d}t \quad \text{subject to}$$
$$\dot{x}_1(t) = u(t), \quad \dot{x}_2(t) = x_1(t), \quad u(t) \in \mathbb{R}, \quad t \in [0, 1], \quad x_1(0) = x_1(1) = 0, \quad x_2(0) = 0, \quad x_2(1) \le 0.$$

One can verify that the pair $(x_*, u_*) \equiv (0, 0)$ satisfies various existing optimality conditions for this problem in terms of subdifferentials and normal cones [11, 12, 17, 19, 47, 51, 55, 66, 70–72, 81]. We leave the laborious task of verifying these conditions to the interested reader. Instead, let us check here whether optimality conditions from Theorem 4.6 are satisfied at u_*.

The function f_0 can be rewritten as

$$f_0(x, u, \xi) = \max\{|u|, |\xi|\} - |u| - |\xi| = \max\{u, -u, \xi, -\xi\} + \min\{u, -u\} + \min\{\xi, -\xi\}.$$

Hence with the use of the codifferential calculus [27] one gets that

$$\underline{\mathrm{d}}_{u,\xi} f_0(x, u, \xi) = \mathrm{co}\left\{\begin{pmatrix} \pm u - g(u,x) \\ \pm 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \pm \xi - g(u,x) \\ 0 \\ \pm 1 \end{pmatrix}\right\}, \overline{\mathrm{d}}_{u,\xi} f_0(x, u, \xi) = \mathrm{co}\left\{\begin{pmatrix} (-1)^i u + (-1)^j \xi + |u| + |\xi| \\ (-1)^i \\ (-1)^j \end{pmatrix} \,\middle|\, i, j \in \{1, 2\}\right\},$$

where $g(u, \xi) = \max\{|u|, |\xi|\}$, while $\underline{\mathrm{d}}_{u,\xi} f_1(x, u, \xi) = \{(0, 1, 0)\}$ and $\overline{\mathrm{d}}_{u,\xi} f_1(x, u, \xi) = \{0\}$. Therefore, as one can readily see, both functions f_0 and f_1 satisfy the codifferentiability conditions of order $p = 1$. The set $\Omega = (0, 1)$ is obviously bounded and has the segment-property. Moreover, if for some $\zeta \in \mathrm{L}^\infty((0, 1); \mathbb{R}, \mathrm{div}) = \mathrm{W}^{1,\infty}(0, 1)$ one has $(0, \zeta'(x), \zeta(x)) \in \underline{\mathrm{d}}_{u,\xi} f_1(x, u'_*(x), u_*(x)) = \{(0, 1, 0)\}$ for a.e. $x \in (0, 1)$, then $\zeta(x) = 0$, while $\zeta'(x) = 1$ for a.e. $x \in (0, 1)$, which is impossible. Thus, constraint qualification (4.23) holds true at u_*.

Suppose that optimality conditions from Theorem 4.6 are satisfied at u_*. Then for any measurable selection $(0, w_{10}(\cdot), w_{20}(\cdot))$ of the set-valued map $\overline{\mathrm{d}}_{u,\xi} f_0(\cdot, u_*(\cdot), u'_*(\cdot))$ there exist $\lambda_1 \ge 0$ and $\zeta \in \mathrm{W}^{1,\infty}(0, 1)$ such that for a.e. $x \in (0, 1)$ one has

$$(0, \zeta'(x), \zeta(x)) \in \underline{\mathrm{d}}_{u,\xi} f_0(x, u_*(x), u'_*(x)) + (0, w_{10}(x), w_{20}(x)) + \lambda_1 \underline{\mathrm{d}}_{u,\xi} f_1(x, u_*(x), u'_*(x)).$$

Define $z_0(x) = (0, w_{10}(x), w_{20}(x)) = (0, 1, 1) \in \overline{\mathrm{d}} f_0(x, 0, 0)$, if $x \in [0, 0.5]$ and $z_0(x) = (0, 1, -1) \in \overline{\mathrm{d}} f_0(x, 0, 0)$, if $x \in (0.5, 1]$. Then there exist $\lambda_1 \ge 0$ and $\zeta \in \mathrm{W}^{1,\infty}(0, 1)$ such that

$$\begin{pmatrix} 0 \\ \zeta'(x) \\ \zeta(x) \end{pmatrix} \in \mathrm{co}\left\{\begin{pmatrix} 0 \\ 2+\lambda_1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \lambda_1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1+\lambda_1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1+\lambda_1 \\ 2 \end{pmatrix}\right\} \quad \text{for a.e. } x \in [0, 0.5], \tag{4.31}$$

$$\begin{pmatrix} 0 \\ \zeta'(x) \\ \zeta(x) \end{pmatrix} \in \mathrm{co} \left\{ \begin{pmatrix} 0 \\ 2+\lambda_1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ \lambda_1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1+\lambda_1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1+\lambda_1 \\ -2 \end{pmatrix} \right\} \quad \text{for a.e. } x \in (0.5, 1]. \tag{4.32}$$

Consequently, $\zeta'(x) \geq 0$ for a.e. $x \in (0,1)$, $\zeta(x) \geq 0$ for a.e. $x \in (0, 0.5)$, and $\zeta(x) \leq 0$ for a.e. $x \in (0.5, 1)$. Redefining, if necessary, the function ζ on a set of measure zero one can suppose that ζ is Lipschitz continuous (see, *e.g.* [54], Thm. 7.17). Therefore, from the inequality $\zeta'(\cdot) \geq 0$ it follows that the function ζ is non-decreasing. Hence with the use of the inequalities $\zeta(x) \geq 0$ for a.e. $x \in (0, 0.5)$ and $\zeta(x) \leq 0$ for a.e. $x \in (0.5, 1)$ one obtains that $\zeta(x) \equiv 0$ and $\zeta'(x) \equiv 0$, which contradicts the fact that the zero vector does not belong to the right-hand sides of (4.31) and (4.32). Thus, optimality conditions from Theorem 4.6 are not satisfied at u_*, and once again optimality conditions in terms of codifferentials were able to detect the non-optimality of the point u_*, when subdifferential-based optimality conditions failed to do so.

5. Conclusions

In this paper we presented a general theory of first order necessary optimality conditions for nonsmooth multidimensional problems of the calculus of variations on arbitrary (not necessarily bounded) domains. This theory is based on the concepts of codifferentiability and quasidifferentiability of nonsmooth functions developed in the finite dimensional case by Demyanov, Rubinov, and Polyakova (see [26–28]). We proved that a nonsmooth integral functional defined on the Sobolev space is continuously codifferentiable and computed its codifferential and quasidifferential under the assumption that the integrand satisfies the codifferentiability conditions introduced in this paper. These conditions, in essence, mean that the integrand is continuously codifferentiable and satisfies, along with its codifferential, some natural growth conditions. In comparison with our previous paper [31], in this work we proved the codifferentiability of the integral functional without the assumption that the domain of integration is bounded and has the segment property (provided $p > 1$), demonstrated that the obscure and hard to verify assumption on *uniform codifferentiability with respect to the Sobolev space* is completely redundant (thus, giving a positive answer to the second question raised [31], Rem. 4.22), and proved the *continuous* codifferentiability of the integral functional for all $1 \leq p \leq +\infty$ (in [31] the continuity of the codifferential mapping was proved only in the case $p = +\infty$).

The explicit expressions for a codifferential and a quasidifferential of the integral functional obtained in this article allowed us to apply general necessary optimality conditions for constrained nonsmooth optimisation problems in Banach spaces in terms of quasidifferentials [34, 35] to easily obtain necessary optimality conditions for constrained nonsmooth problems of the calculus of variations, including problems with additional constraints at the boundary and problems with isoperimetric constraints. As is demonstrated by a series of simple examples, our optimality conditions are sometimes better than the existing ones in terms of various subdifferentials, since they are able to detect the non-optimality of a given point, when subdifferential-based optimality conditions fail to disqualify this point as non-optimal.

References

[1] R.A. Adams, Sobolev Spaces. Academic Press, New York (1975).

[2] J.-P. Aubin and H. Frankowska, Set-Valued Analysis. Birkhäuser, Boston (1990).

[3] L. Bayón, J.M. Grau, M.M. Ruiz and P.M. Suárez, Nonsmooth optimization of hydrothermal problems. *J. Comput. Appl. Math.* **192** (2006) 11–19.

[4] L. Bayón, J.M. Grau, M.M. Ruiz and P.M. Suárez, A constrained and non-smooth hydrothermal problem. *Appl. Math. Comput.* **209** (2009) 10–18.

[5] L. Bayón, J.M. Grau, M.M. Ruiz and P.M. Suárez, A hydrothermal problem with non-smooth Lagrangian. *J. Ind. Manag. Optim.* **10** (2014) 761–776.

[6] S. Bellaassali, *Contributions à l'optimisation multicritère*, Ph.D. thesis, Université de Bourgogne, Laboratoire Analyse Appliquée et Optimisation, Dijon, France (2003). Available at: https://tel.archives-ouvertes.fr/file/index/docid/46039/filename/tel-00004421.pdf.

[7] D.N. Bessis, Yu. S. Ledyaev and R.B. Vinter, Dualization of the Euler and Hamiltonian inclusions. *Nonlinear Anal.* **43** (2001) 861–882.

[8] V.I. Bogachev, Vol. I of Measure Theory. Springer-Verlag, Berlin, Heidelberg (2007).

[9] G. Bonfanti and A. Cellina, The validity of the Euler-Lagrange equation. *Discret. Contin. Dyn. Syst.* **28** (2010) 511–517.

[10] P. Bousquet, The Euler equation in the multiple integrals calculus of variations. *SIAM J. Control Optim.* **51** (2013) 1047–1062.

[11] F. Clarke, Necessary Conditions in Dynamic Optimization. American Mathematical Society, Providence, Rhode Island (2005).

[12] F. Clarke, Functional Analysis, Calculus of Variations and Optimal Control. Springer-Verlag, London (2013).

[13] F.H. Clarke, The Euler-Lagrange differential inclusion. *J. Differ. Equ.* **19** (1975) 80–90.

[14] F.H. Clarke, The generalized problem of Bolza. *SIAM J. Control Optim.* **14** (1976) 682–699.

[15] F.H. Clarke, Multiple integrals of Lipschitz functions in the calculus of variations. *Proc. Am. Math. Soc.* **64** (1977) 260–264.

[16] F.H. Clarke, The Erdmann condition and Hamiltonian inclusions in optimal control and the calculus of variations. *Can. J. Math.* **32** (1980) 494–509.

[17] F.H. Clarke, Optimization and Nonsmooth Analysis. Wiley–Interscience, New York (1983).

[18] F.H. Clarke, A decoupling principle in the calculus of variations. *J. Math. Anal. Appl.* **172** (1993) 92–105.

[19] F.H. Clarke and M.R. de Pinho, The nonsmooth maximum principle. *Control Cybern.* **38** (2009) 1151–1167.

[20] F.H. Clarke, Yu. S. Ledyaev, R.J. Stern and P.R. Wolenski, Nonsmooth Analysis and Control Theory. Springer-Verlag, New York (1998).

[21] G. Cupini, M. Guidorzi and C. Marcelli, Necessary conditions and non-existence results for autonomous nonconvex variational problems. *J. Differ. Equ.* **243** (2007) 329–348.

[22] B. Dacorogna, Direct Methods in the Calculus of Variations. Springer-Verlag, New York (2008).

[23] V.F. Demyanov, Continuous generalized gradients for nonsmooth functions, in Optimization, Parallel Processing and Applications, edited by A. Kurzhanski, K. Neumann and D. Pallaschke. Springer, Berlin, Heidelberg (1988) 24–27.

[24] V.F. Demyanov, On codifferentiable functions. *Vestn. Leningr. Univ., Math.* **2** (1988) 22–26.

[25] V.F. Demyanov, Smoothness of nonsmooth functions, in Nonsmooth Optimization and Related Topics, edited by F. Clarke, V. Demyanov and F. Giannesssi. Springer, Boston (1989) 79–88.

[26] V.F. Demyanov and L.C.W. Dixon, Quasidifferential Calculus. Springer, Berlin, Heidelberg (1986).

[27] V.F. Demyanov and A.M. Rubinov, Constructive Nonsmooth Analysis. Peter Lang, Frankfurt am Main (1995).

[28] V.F. Demyanov and A.M. Rubinov, Quasidifferentiability and Related Topics. Kluwer Academic Publishers, Dordrecht (2000).

[29] V.F. Demyanov, G.E. Stavroulakis, L.N. Polyakova and P.D. Panagiotopoulos, Quasidifferentiability and Nonsmooth Modelling in Mechanics, Engineering and Economics. Kluwer Academic Publishers, Dordrecht (1996).

[30] M.V. Dolgopolik, Codifferential calculus in normed spaces. *J. Math. Sci.* **173** (2011) 441–462.

[31] M.V. Dolgopolik, Nonsmooth problems of calculus of variations via codifferentiation. *ESAIM: COCV* **20** (2014) 1153–1180.

[32] M.V. Dolgopolik, Abstract convex approximations of nonsmooth functions. *Optim* **64** (2015) 1439–1469.

[33] M.V. Dolgopolik, A convergence analysis of the method of codifferential descent. *Comput. Optim. Appl.* **71** (2018) 879–913.

[34] M.V. Dolgopolik, Metric regularity of quasidifferentiable mappings and optimality conditions for nonsmooth mathematical programming problems. *Set-Valued Var. Anal.* **28** (2019) 427–449.

[35] M.V. Dolgopolik, A new constraint qualification and sharp optimality conditions for nonsmooth mathematical programming problems in terms of quasidifferentials. *SIAM J. Optim.* **30** (2020) 2603–2627.

[36] N. Dunford and J.T. Schwartz, Linear Operators, Part 1: General Theory. John Wiley & Sons, New Jersey (1958).

[37] I. Ekeland and R. Temam, Convex Analysis and Variational Problems. SIAM, Philadelphia (1999).

[38] G.B. Folland, Real Analysis. Modern Techniques and Their Applications. Interscience Publishers, New York (1984).

[39] Y. Gao, On the minimal quasidifferential in the one-dimensional case. *Soochow J. Math.* **24** (1998) 211–218.

[40] F. Giannessi, A common understanding or a common misunderstanding? *Numer. Funct. Anal. Optim.* **16** (1995) 1359–1363.

[41] J. Grzybowski, D. Pallaschke and R. Urbański, On the reduction of pairs of bounded closed convex sets. *Studia Math.* **189** (2008) 1–12.

[42] J. Grzybowski, D. Pallaschke and R. Urbański, On the amount of minimal pairs of convex sets. *Optim. Methods Softw.* **25** (2010) 89–96.

[43] J. Grzybowski and R. Urbański, Minimal pairs of bounded closed convex sets. *Studia Math.* **126** (1997) 95–99.

[44] J. Grzybowski and R. Urbański, Three criteria of minimality for pairs of compact convex sets. *Optim* **55** (2006) 569–576.

[45] M. Handschug, On equivalent quasidifferentials in the two-dimensional case. *Optim* **20** (1989) 37–43.

[46] A. Ioffe, Euler-Lagrange and Hamiltonian formalisms in dynamic optimization. *Trans. Am. Math. Soc.* **349** (1997) 2871–2900.

[47] A.D. Ioffe, On necessary conditions for a minimum. *J. Math. Sci.* **217** (2016) 751–772.

[48] A.D. Ioffe, On generalized Bolza problems and its application to dynamic optimization. *J. Optim. Theory Appl.* **182** (2019) 285–309.

[49] A.D. Ioffe and R.T. Rockafellar, The Euler and Weierstrass conditions for nonsmooth variational problems. *Calc. Var. Partial Differ. Equ.* **4** (1996) 59–87.

[50] A.D. Ioffe and V.M. Tihomirov, Theory of Extremal Problems. North-Holland Publishing Company, Amsterdam (1979).

[51] A. Jourani, Lagrangian and Hamiltonian necessary conditions for the generalized Bolza problem and applications. *J. Nonlinear Convex Anal.* **10** (2009) 437–454.

[52] A. Jourani and L. Thibault, Approximate subdifferential and metric regularity: the finite-dimensional case. *Math. Program.* **47** (1990) 203–218.

[53] L. Kuntz, A characterization of continuously codifferentiable functions and some consequences. *Optim* **22** (1991) 539–547.

[54] G. Leoni, A First Course in Sobolev spaces. American Mathematical Society, Providence, RI (2009).

[55] P.D. Loewen, Optimal Control via Nonsmooth Analysis. American Mathematical Society, Providence, Rhode Island (1993).

[56] P.D. Loewen and R.T. Rockafellar, The adjoint arc in nonsmooth optimization. *Trans. Am. Math. Soc.* **325** (1991) 39–72.

[57] P.D. Loewen and R.T. Rockafellar, Optimal control of unbounded differential inclusions. *SIAM J. Control Optim.* **32** (1994) 442–470.

[58] P.D. Loewen and R.T. Rockafellar, New necessary conditions for the generalized problem of Bolza. *SIAM J. Control Optim.* **34** (1996) 1496–1511.

[59] P.D. Loewen and R.T. Rockafellar, Bolza problem with general time constraints. *SIAM J. Control Optim.* **35** (1997) 2050–2069.

[60] C. Marcelli, Variational problems with nonconvex, noncoercive, highly discontinuous integrands: characterization and existence of minimizers. *SIAM J. Control Optim.* **40** (2002) 1473–1490.

[61] C. Marcelli, Necessary and sufficient conditions for optimality of nonconvex, noncoercive autonomous variational problems with constraints. *Trans. Am. Math. Soc.* **360** (2008) 5201–5227.

[62] C. Marcelli, E. Outkine and M. Sytchev, Remarks on necessary conditions for minimizers of one-dimensional variational problems. *Math. Prepr. Arch.* **2001** (2001) 1145–1163.

[63] B.S. Mordukhovich, Approximation Methods in Problems of Optimization and Control. Nauka, Moscow (1988). [in Russian].

[64] B.S. Mordukhovich, Discrete approximation and refined Euler-Lagrange conditions for nonconvex differential inclusions. *SIAM J. Control Optim.* **33** (1995) 882–915.

[65] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation I: Basic Theory. Springer-Verlag, Berling, Heidelberg (2006).

[66] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation II: Applications. Springer-Verlag, Berling, Heidelberg (2006).

[67] I.V. Orlov and A.V. Tsygankova, Multidimensional variational functionals with subsmooth integrands. *Eurasian Math. J.* **6** (2015) 54–75.

[68] D. Pallaschke and R. Urbański, Some criteria for the minimality of pairs of compact convex sets. *ZOR — Methods Models Oper. Res.* **37** (1993) 129–150.

[69] D. Pallaschke and R. Urbański, Quasidifferentiable calculus and minimal pairs of compact convex sets. *Schedae Informaticae* **21** (2012) 107–125.

[70] E.S. Polovinkin, Differential inclusions with unbounded right-hand side and necessary optimality conditions. *Proc. Stekov Inst. Math.* **291** (2015) 237–252.

[71] E.S. Polovinkin, Necessary optimality conditions for the Mayer problem with unbounded differential inclusion. *IFAC-PapersOnline* **51** (2018) 521–524.

[72] E.S. Polovinkin, Pontryagin's direct method for optimization problems with differential inclusions. *Proc. Stekov Inst. Math.* **304** (2019) 241–256.

[73] B.N. Pshenichnyi, Necessary Conditions for an Extremum. Marcel Dekker, New York (1971).

[74] R.T. Rockafellar, Conjugate convex functions in optimal control and the calculus of variations. *J. Math. Anal. Appl.* **32** (1970) 174–222.

[75] R.T. Rockafellar, Generalized Hamiltonian equations for convex problems of Lagrange. *Pac. J. Math.* **33** (1970) 411–427.

[76] R.T. Rockafellar, Existence and duality theorems for convex problems of Bolza. *Trans. Am. Math. Soc.* **159** (1971) 1–40.

[77] R.T. Rockafellar, Dualization of subgradient conditions for optimality. *Nonlinear Anal.* **20** (1993) 627–646.

[78] S. Scholtes, Minimal pairs of convex bodies in two dimensions *Mathematika* **39** (1992) 267–273.

[79] M.H.N. Skandari, A.V. Kamyad and S. Effati, Generalized Euler-Lagrange equation for nonsmooth calculus of variations. *Nonlinear Dyn.* **75** (2014) 85–100.

[80] R. Vinter and H. Zheng, The extended Euler-Lagrange condition for nonconvex variational problems. *SIAM J. Control Optim.* **35** (1997) 56–77.

[81] R.B. Vinter, Optimal Control. Birkhäuser, Boston (2000).

[82] A. Zaffaroni, Codifferentiable mappings with applications to vector optimality. *Pilska Studia Mathematica Bulgarica* **12** (1998) 255–266.

[83] A. Zaffaroni, Continuous approximations, codifferentiable functions and minimization methods, in Quasidifferentiability and related Topics, edited by V.F. Demyanov and A.M. Rubinov. Kluwer Academic Publishers, Dordrecht (2000) 361–391.