Convex duality and uniqueness for BV-minimizers

Lisa Beck[*] Thomas Schmidt [†‡]

June 6, 2013

Abstract

We are concerned with the Dirichlet minimization problem for variational integrals with linear growth. In the existence theory for minimizers, one commonly considers a generalized formulation in the space of functions of bounded variation, while on the other hand the dual problem, in the classical sense of convex analysis, is a maximization problem in the space of divergence-free bounded vector fields. In this paper, we characterize extremals of the generalized and the dual problem by pointwise extremality relations, and we discuss uniqueness issues for both kinds of solutions. Our approach is sufficiently general to cover arbitrary dimensions, non-smooth integrands, and unbounded, irregular domains.

MSC (2010): 49K20, 49N15 (primary); 26B25, 26B30, 46G10 (secondary)

Contents

1 Introduction

Throughout this paper we fix two positive integers n and N and a non-empty, open set Ω (which is not necessarily smooth or bounded) in $\mathbb{R}^n$, and we investigate variational integrals of the type

$$F[w] := \int_\Omega f(\,\cdot\,,\nabla w)\,\mathrm{d}x \qquad \text{for } w\colon \Omega \to \mathbb{R}^N\,,$$

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement GeMeThnES n° 246923.

[*]Hausdorff Center for Mathematics, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany. Email address: `lisa.beck@hcm.uni-bonn.de`.

[†]Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, 56126 Pisa, Italy. Email address: `thomas.schmidt@sns.it`.

[‡]Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland. Email address: `thomas.schmidt@math.uzh.ch`.

with a given Borel measurable integrand $f\colon \Omega \times \mathbb{R}^{Nn} \to \mathbb{R}$. We consider the problem to

$$\text{minimize } F \text{ in the Dirichlet class } \mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N) := u_0|_\Omega + \mathrm{W}^{1,1}_0(\Omega,\mathbb{R}^N)\,, \tag{P}$$

where we permanently assume — for later convenience — that the boundary values are prescribed with the help of a globally defined function $u_0 \in \mathrm{W}^{1,1}(\mathbb{R}^n,\mathbb{R}^N)$. As our main assumption on the integrand f we impose a linear growth condition[1]

$$|f(x,z)| \le \Psi(x) + L|z| \qquad \text{for all } (x,z) \in \Omega \times \mathbb{R}^{Nn} \tag{Lin}$$

with a $[0,\infty)$-valued function $\Psi \in \mathrm{L}^1(\Omega)$ and a constant $L \in [0,\infty)$. This condition ensures that F is finite on $\mathrm{W}^{1,1}(\Omega,\mathbb{R}^N)$, and in particular that the infimum $\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F$ of the problem (P) cannot take the value ∞, while the value $-\infty$ remains possible. However, even if reasonable extra assumptions on f are made and the infimum is finite, it is not necessarily attained, in other words F need not have a minimizer in $\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$. For this reason one commonly considers a generalized formulation of (P) in the space $\mathrm{BV}(\Omega,\mathbb{R}^N)$ of functions of bounded variation. Postponing the introduction of this BV-formulation and the appropriate concept of generalized minimizers to Section 3.2, for the moment let us just point out that generalized minimizers in $\mathrm{BV}(\Omega,\mathbb{R}^N)$ exist significantly more often than minimizers in $\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$ itself, but they may have worse uniqueness properties.

In this paper are concerned with the interplay between the generalized BV-formulation of (P) and the dual problem in the sense of convex analysis, which is extensively discussed, for instance, in the monograph [19]. The latter problem involves the conjugate function $f^*\colon \mathbb{R}^n \times \mathbb{R}^{Nn} \to \mathbb{R} \cup \{\infty\}$ of f (with respect to the z-variable), which is given by $f^*(x,z^*) := \sup_{z\in\mathbb{R}^{Nn}} \big[z^*\cdot z - f(x,z)\big]$, and in fact, when we set[2]

$$\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn}) := \{\tau \in \mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn}) \,:\, \operatorname{div}\tau \equiv 0 \text{ in the sense of distributions on } \Omega\}\,,$$
$$R_{u_0}[\tau] := \int_\Omega \big[\tau\cdot\nabla u_0 - f^*(\,\cdot\,,\tau)\big]\,\mathrm{d}x \qquad \text{for } \tau \in \mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})\,,$$

the dual problem is to

$$\text{maximize } R_{u_0} \text{ in } \mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})\,. \tag{P*}$$

We briefly mention that, in many applications, the dual problem can be seen as a maximization problem for a physically relevant quantity, called the stress tensor; see [38, 24, 35, 37, 19], for instance. Here we will not further discuss this aspect, but we rather explain another classical way to understand the relationship between (P) and (P*): first, by the definition of the conjugate function one has $f(x,z) \ge z^*\cdot z - f^*(x,z^*)$ for all $x \in \Omega$ and $z,z^* \in \mathbb{R}^{Nn}$, and thus on the one hand one gets

$$\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F \ge \inf_{w\in\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} \left[\sup_{\tau\in\mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})} \int_\Omega \big[\tau\cdot\nabla w - f^*(\,\cdot\,,\tau)\big]\,\mathrm{d}x\right]. \tag{1.1}$$

For convex f one moreover has $f(x,z) = f^{**}(x,z) := \sup_{z^*\in\mathbb{R}^{Nn}} \big[z^*\cdot z - f^*(x,z^*)\big]$ (compare Section 3.3), and thus we expect equality in (1.1), an anticipation which will eventually turn out to be true[3]. On the other hand one also has

$$\sup_{\tau\in\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})} R_{u_0} = \sup_{\tau\in\mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})} \left[\inf_{w\in\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} \int_\Omega \big[\tau\cdot\nabla w - f^*(\,\cdot\,,\tau)\big]\,\mathrm{d}x\right], \tag{1.2}$$

[1] We find it worth remarking that, when Ω is bounded and f is independent of x and convex in z, (Lin) reduces to the requirement $f(z) \le L(1+|z|)$. In particular, in this case a lower bound of the type $f(z) \ge -L(1+|z|)$ is an automatic consequence of convexity.

[2] Here, we multiply matrices from $\mathbb{R}^{Nn}$ in the sense of the Hilbert-Schmidt product, and the distributional divergence is understood as the adjoint of the gradient operator with respect to this inner product.

[3] Indeed, equality in (1.1) can be inferred from the following arguments and Theorem 1.1 below or alternatively from [19, Chapter IX.2].

since by partial integration the infimum on the right-hand side equals $R_{u_0}[\tau]$ in the case $\operatorname{div}\tau \equiv 0$ (while it equals $-\infty$ otherwise). All in all, we can read off that (P) and (P*) differ essentially by the priority of the inf- and the sup-operation on the right-hand sides of (1.1) and (1.2). Moreover, the inequality $\inf\big[\sup\dots\big] \geq \sup\big[\inf\dots\big]$ between these right-hand sides is obvious, so that by the preceding elementary arguments we have in fact shown

$$\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F \geq \sup_{\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})} R_{u_0}\,. \tag{1.3}$$

Actually, one can even prove that the interchange of inf and sup does not change the resulting value at all so that equality holds in (1.3); this is a classical result on the duality correspondence between (P) and (P*), which is detailed, for instance, in [19], and which we restate in our setting as follows.

Theorem 1.1 (duality formula and existence of a dual solution)**.** *Assume that f satisfies* (Lin) *and that $f(x,\cdot)\colon \mathbb{R}^{Nn}\to\mathbb{R}$ is a convex function for $\mathscr{L}^n$-a. e. $x\in\Omega$. Then the infimum in* (P) *equals the supremum in* (P*)*, that is*

$$\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F = \sup_{\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})} R_{u_0} \in [-\infty,\infty)\,. \tag{1.4}$$

Moreover, whenever the common value is not $-\infty$, then the problem (P*) *has a solution, that is, the supremum in* (1.4) *is in fact a maximum.*

As already said above, Theorem 1.1 is a special case of more general results in [19]. However, we want to stress that the proof given there does not only require the abstract duality theory [19, Chapter III] on the level of functionals, but also the representation [19, Chapter IX.2] of the (bi-)dual problem in terms of the (bi-)conjugate, which is somewhat less elementary and relies also on a measurable selection theorem. As a side benefit, our methods yield an alternative proof of Theorem 1.1, which will be provided in Section 5. Though our approach relies on the same basic tools, we believe that it has a slight advantage over the more classical strategy: in the special case that f is C^1 in z, all measurable selection issues drop out of our argument (compare Remark 5.3), while a similar simplification of the reasoning in [19] does not seem obvious.

It is well known that the duality formula of Theorem 1.1 leads to characterizations of extremality in terms of pointwise relations. Let us address this point in detail:

Corollary 1.2 (extremality relations for minimizers in $\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$)**.** *Assume that f satisfies* (Lin) *and that $f(x,\cdot)\colon \mathbb{R}^{Nn}\to\mathbb{R}$ is a convex function for $\mathscr{L}^n$-a. e. $x\in\Omega$. Then, for $u\in \mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$ and $\sigma\in \mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$, the following four conditions are* equivalent:

$$u \textit{ solves } \mathrm{(P)} \textit{ and } \sigma \textit{ solves } \mathrm{(P^*)}\,, \tag{1.5}$$

$$f(\,\cdot\,,\nabla u) = \sigma\cdot\nabla u - f^*(\,\cdot\,,\sigma) \qquad \mathscr{L}^n\text{-a. e. on } \Omega\,, \tag{1.6}$$

$$\sigma \in \partial_z f(\,\cdot\,,\nabla u) \qquad \mathscr{L}^n\text{-a. e. on } \Omega\,, \tag{1.7}$$

$$\nabla u \in \partial_{z^*} f^*(\,\cdot\,,\sigma) \qquad \mathscr{L}^n\text{-a. e. on } \Omega\,. \tag{1.8}$$

Here, $\partial_z f$ and $\partial_{z^*} f^*$ denote the subdifferentials — as specified in Definition 3.3 below — of f and f^* with respect to the second variable.

Proof. By the definition of the conjugate function, $f(\,\cdot\,,\nabla u)\geq \sigma\cdot\nabla u - f^*(\,\cdot\,,\sigma)$ holds $\mathscr{L}^n$-a. e. on Ω. Thus, (1.6) is equivalent to the integral identity

$$\int_\Omega f(\,\cdot\,,\nabla u)\,\mathrm{d}x = \int_\Omega \sigma\cdot\nabla u\,\mathrm{d}x - \int_\Omega f^*(\,\cdot\,,\sigma)\,\mathrm{d}x\,. \tag{1.9}$$

As we are assuming $u\in \mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$ and $\sigma\in\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$, the first integral on the right-hand side remains unchanged if we replace u with u_0, and thus (1.9) just means $F[u]=R_{u_0}[\sigma]$. By Theorem 1.1, the last equality characterizes the extremality properties in (1.5), and hence we have established the equivalence of (1.5) and (1.6).

With the help of (3.3) (which is essentially the definition of the subdifferential) and the equality $f^{**}=f$, (1.6) can be rewritten in the two equivalent forms given in (1.7) and (1.8). $\square$

We emphasize however that the preceding reasoning makes essential use of the assumption $u \in \mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$, while the existence theory for (P) yields generalized minimizers in $\mathrm{BV}(\Omega,\mathbb{R}^N)$ rather than minimizers in $\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$. In the present paper, we close this gap and extend more specific results of Bildhauer & Fuchs [12, 9, 13, 10, 11], which are in turn inspired by previous ideas of Seregin [35, 36, 37]: indeed, by our Theorem 2.1, the extremality relations (1.6), (1.7), (1.8) remain true for *all* extremals $u \in \mathrm{BV}(\Omega,\mathbb{R}^N)$ and $\sigma \in \mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$ *provided that* ∇u *is understood as (the density of) the absolutely continuous part of the gradient measure* $\mathrm{D}u$. This result answers a question raised in [11, Remark 2.30], and it recovers and extends, in an elegant way, uniqueness results for u and σ in certain situations, covering in particular the singular integrals of [7] which have originally motivated our investigation of the duality correspondence. Additionally, in Theorem 2.1, we also provide a new extremality relation for the singular part of $\mathrm{D}u$, which is strongly connected with previous ideas of Anzellotti [2, 4] and Kohn & Temam [25, 26]. Combined with (1.6), (1.7), (1.8), this extra relation completely characterizes extremals in our BV-setup.

2 Statement of the results

Postponing the discussion of non-convex cases to Appendix A, we now state our results under — as we believe — quite general and sharp assumptions on the domain Ω and the integrand f.

Specifically, for the open set Ω, we do *not* require boundedness (allowing, at the cost of some technical complications in Sections 4 and 5, such natural domains as the whole space or a half-space), but we only impose the mild boundary regularity hypothesis

$$\mathbb{1}_\Omega \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^n) \text{ and } |\mathrm{D}\mathbb{1}_\Omega| = \mathscr{H}^{n-1} \llcorner \partial\Omega\,. \tag{\textbf{Per}}$$

This condition, introduced in [34], will only be relevant in connection with the strict approximation result of Lemma 4.3, and it can be rephrased by saying that Ω is a set of locally finite perimeter in $\mathbb{R}^n$ such that its topological boundary differs from its reduced boundary only by a set of zero $\mathscr{H}^{n-1}$-measure. In particular, (Per) implies that $\partial\Omega$ is $\mathscr{H}^{n-1}$-σ-finite and has zero $\mathscr{L}^n$-measure.

Furthermore, for the integrand f, we rely, in addition to (Lin), on the following continuity hypothesis:

$$\begin{gathered} f(x,\,\cdot\,)\colon \mathbb{R}^{Nn} \to \mathbb{R} \text{ is a continuous function for } \mathscr{L}^n\text{-a. e. } x \in \Omega \\ \text{and the limit } \lim_{\substack{\widetilde{x}\to x\\ \widetilde{z}\to z\\ t\searrow 0}} t f(\widetilde{x},\widetilde{z}/t) \text{ exists in } \mathbb{R} \text{ for all } (x,z) \in \overline{\Omega} \times (\mathbb{R}^{Nn}\setminus\{0\})\,. \end{gathered} \tag{\textbf{Con}}$$

Assumption (Con) is only needed in order to apply Theorem 4.1, a version of the Reshetnyak continuity result. The first part of (Con) is commonly phrased by saying that f is a Carathéodory function, and the second part of (Con) can be reformulated as the requirement that $(x,z) \mapsto (1{-}|z|)f(x,z/(1{-}|z|))$ extends from $\Omega{\times}\mathrm{B}^{Nn}_1$ to a function on $(\Omega{\times}\mathrm{B}^{Nn}_1) \cup \big(\overline{\Omega}{\times}\partial\mathrm{B}^{Nn}_1\big)$ which is continuous at all points of $\overline{\Omega}{\times}\partial\mathrm{B}^{Nn}_1$. Furthermore, for both illustration and later usage, we record that (Con) implies the following continuity condition in x:

$$\begin{gathered} \text{For all } x_0 \in \overline{\Omega} \text{ and } \varepsilon > 0 \text{ there exists a } \delta > 0 \text{ such that for } x,\widetilde{x} \in \Omega \text{ and } z \in \mathbb{R}^{Nn} \text{ we have} \\ |x{-}x_0| + |\widetilde{x}{-}x_0| + |z|^{-1} < \delta \implies |f(\widetilde{x},z) - f(x,z)| < \varepsilon|z|\,. \end{gathered} \tag{2.1}$$

We point out that, under the hypothesis that f is convex in z with (Lin), the conditions (Con) and (2.1) are even equivalent. In particular, (2.1) is trivially satisfied in the case of an x-independent integrand f, and hence, in this case, (Con) follows already from convexity and (Lin).

Now we are ready to state our main result.

Theorem 2.1 (extremality relations for generalized minimizers)**.** *Assume that* Ω *satisfies* (Per)*, that* f *satisfies* (Lin) *and* (Con)*, and that* $f(x,\,\cdot\,)\colon \mathbb{R}^{Nn}\to\mathbb{R}$ *is a convex function for* $\mathscr{L}^n$*-a. e.* $x\in\Omega$*. Then, for* $u\in \mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)$ *and* $\sigma\in \mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$ *we have the following* equivalence: *u is a generalized minimizer of* (P) *and* σ *a solution of* (P*)*,* if and only if *any of the relations*

$$f(\,\cdot\,,\nabla u)=\sigma\cdot\nabla u-f^*(\,\cdot\,,\sigma)\qquad \mathscr{L}^n\text{-a. e. on }\Omega\,, \tag{2.2}$$

$$\sigma\in\partial_z f(\,\cdot\,,\nabla u)\qquad \mathscr{L}^n\text{-a. e. on }\Omega\,, \tag{2.3}$$

$$\nabla u\in\partial_{z^*} f^*(\,\cdot\,,\sigma)\qquad \mathscr{L}^n\text{-a. e. on }\Omega \tag{2.4}$$

holds for ∇u*, and, at the same time,* $\mathrm{D}^{\mathrm{s}}u$ *satisfies*

$$f^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\mathrm{D}^{\mathrm{s}}u}{\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|}\Big)=\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]^{\mathrm{s}}}{\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|}\qquad |\mathrm{D}^{\mathrm{s}}u|\text{-a. e. on }\overline{\Omega}\,. \tag{2.5}$$

Here, ∇u *denotes the density of the absolutely continuous part in the Lebesgue decomposition* $\mathrm{D}u=(\nabla u)\mathscr{L}^n+\mathrm{D}^{\mathrm{s}}u$ *of the gradient measure* $\mathrm{D}u$ *with respect to* $\mathscr{L}^n$*, and we refer to Sections 3.2 and 7 for all further details of our terminology.*

In order to illustrate the meaning of this statement for a non-smooth f, let us consider, in the very simple case $n{=}N{=}1$ (then $\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$ consists of constant functions σ, thus $[\![\sigma\cdot\mathrm{D}u]\!]^{\mathrm{s}}=\sigma\mathrm{D}^{\mathrm{s}}u$), the integrand $f(x,z)=|z|$ (for which we have $\partial_z f(x,0)=[-1,1]$ and $f^\infty(x,z)=|z|$). In this situation, (2.3) and (2.5) show that $u\in\mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)$ and $\sigma\in\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$ are extremals of (P) and (P*) if and only if one of the following three possibilities occurs: either $\sigma\equiv 1$ and $\mathrm{D}u\geq 0$, or $\sigma\equiv -1$ and $\mathrm{D}u\leq 0$, or $\sigma\in(-1,1)$ and $\mathrm{D}u\equiv 0$; compare also Remark 7.5. We further observe that, in this example, (2.2), (2.3), (2.4) do also hold for every constant $\sigma\in[-1,1]$ and every non-monotone pure-jump function u. Therefore, the additional relation (2.5) is indeed inevitable in the characterization of BV-extremals.

First, in Section 5, we give a partial proof of Theorem 2.1, which establishes the relations (2.2), (2.3), (2.4) for all extremals $u\in\mathrm{BV}\big(\overline{\Omega},\mathbb{R}^N\big)$ and $\sigma\in\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$. The method employed there is based on approximations and Ekeland's variational principle, it yields Theorem 1.1 as a byproduct, it remains comparably elementary, and it provides precisely the assertion which is relevant for the following applications. The relation (2.5) requires more refined measure-theoretic concepts, in particular a pairing $[\![\sigma\cdot\mathrm{D}u]\!]$ of gradient measures $\mathrm{D}u$ and $\mathrm{L}^\infty_{\mathrm{div}}$-functions σ in the spirit of Anzellotti [2] (compare Definition 7.1). Under additional regularity assumptions on f and Ω, these tools have been employed in Anzellotti's subsequent work [4], and the last relation (2.5) follows from the validity of (2.3) and [4, Theorem 1.3]. In our less regular setting, however, we establish (2.5) and the full equivalence of Theorem 2.1 only in Section 7 — by a second approach which relies on $|\mathrm{D}^{\mathrm{s}}u|$-a. e. properties of $[\![\sigma\cdot\mathrm{D}u]\!]$ and which now utilizes Theorem 1.1 as a prerequisite.

We stress that, as already indicated, Theorem 2.1 is not the first duality result in the BV-context: for instance, under more stringent assumptions on f — namely x-independence, strict convexity, and C^2-regularity with a bound for $\nabla^2 f$ — Bildhauer & Fuchs [12, 10, 11] have proved some regularity properties of the dual solution and the existence of *at least one* generalized minimizer which satisfies the extremality relations (2.2), (2.3), (2.4), while Bildhauer [9] has established uniqueness of the dual solution under the same hypothesis. One advantage of our Theorem 2.1 is that it recovers the latter uniqueness result as a direct corollary and under the sole additional hypothesis that f is C^1 in the z-variable:

Corollary 2.2 (uniqueness of σ)**.** *Assume that* Ω *satisfies* (Per)*, that* f *satisfies* (Lin) *and* (Con)*, and that* $f(x,\,\cdot\,)\colon\mathbb{R}^{Nn}\to\mathbb{R}$ *is a convex* C^1*-function for* $\mathscr{L}^n$*-a. e.* $x\in\Omega$*. If a generalized minimizer* $u\in\mathrm{BV}(\Omega,\mathbb{R}^N)$ *of* (P) *exists, then the dual problem* (P*) *has a unique solution.*

Proof. By the convexity and differentiability assumptions, we are in the case of a single-valued subdifferential $\partial_z f(x,z)=\{\nabla_z f(x,z)\}$, and then (2.3) determines σ. ☐

We remark that one may also view — as done in [9] — the uniqueness result as an outcome of a strict convexity property of the dual problem, as in fact it follows from the following two observations: on the one hand, the differentiability assumption of the corollary implies that $f^*(x,\,\cdot\,)$ is strictly convex on[4] $\operatorname{Im}\partial_z f(x,\,\cdot\,)$, while on the other hand (2.3) shows that all dual solutions take values in the latter sets.

Clearly, if σ is a dual solution, then $\sigma(x)$ cannot lie in the region where $f^*(x,\cdot)$ is infinite. In the next statement we impose a slightly stronger condition, namely that $\sigma(x)$ is even (locally uniformly in x) bounded away from this region. Following quite closely a strategy from [37, 13], which is also described in [11, Section A.3], we then obtain:

Theorem 2.3 ($\mathrm{W}^{1,1}$-regularity of generalized minimizers via a duality criterion)**.** *Assume that* Ω *satisfies* (Per)*, that* f *satisfies* (Lin) *and* (Con)*, and that* $f(x,\cdot)\colon \mathbb{R}^{Nn}\to\mathbb{R}$ *is a convex function for* $\mathscr{L}^n$*-a. e.* $x\in\Omega$*. If the dual problem* (P*) *has a solution* $\sigma\in \mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$ *such that*

$$f^*(x,\,\cdot\,)\le\Lambda(x) \text{ holds on } \mathrm{B}_{\varepsilon(x)}(\sigma(x)) \tag{2.6}$$

for $\mathscr{L}^n$*-a. e.* $x\in\Omega$*, with a continuous function* $\varepsilon\colon\Omega\to(0,\infty)$ *and some* $\Lambda\in\mathrm{L}^1_{\mathrm{loc}}(\Omega)$*, then every generalized minimizer* $u\in\mathrm{BV}(\Omega,\mathbb{R}^N)$ *of* (P) *is in* $\mathrm{W}^{1,1}(\Omega,\mathbb{R}^N)$*.*

Theorem 2.3 will be established in Section 6.

Finally, we turn to the case that f is even strictly convex in z. Then, a minimizer in $\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$ is necessarily unique, while for generalized minimizers $u\in\mathrm{BV}(\Omega,\mathbb{R}^N)$ only the absolutely continuous part $(\nabla u)\mathscr{L}^n$ of their gradient is uniquely determined. Clearly, the latter assertion trivially implies uniqueness of the full gradient $\mathrm{D}u$ whenever one can prove $\mathrm{W}^{1,1}$-regularity for all generalized minimizers. While in [8] we have treated a borderline case of this regularity problem, we here discuss less subtle situations where it can be resolved — in a simpler and more elegant way — via Theorem 2.3. In particular, this happens in the following corollaries, which provide two simple sufficient criteria for (2.6), namely continuity of a dual solution and C^1-regularity for *one* generalized minimizer.

Corollary 2.4 (continuity of σ implies uniqueness of $\mathrm{D}u$)**.** *Assume that* Ω *satisfies* (Per)*, that* f *is continuous on* $\Omega\times\mathbb{R}^{Nn}$ *with* (Lin) *and* (Con)*, and that* $f(x,\cdot)\colon\mathbb{R}^{Nn}\to\mathbb{R}$ *is a strictly convex function for all* $x\in\Omega$*. If the dual problem* (P*) *has a* continuous *solution* $\sigma\in\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$*, then all generalized minimizers* $u,v\in\mathrm{BV}(\Omega,\mathbb{R}^N)$ *of* (P) *are in* $\mathrm{W}^{1,\infty}_{\mathrm{loc}}(\Omega,\mathbb{R}^N)$*, and we have* $\mathrm{D}u{=}(\nabla u)\mathscr{L}^n{=}(\nabla v)\mathscr{L}^n{=}\mathrm{D}v$*.*

The derivation of Corollary 2.4 from Theorems 2.1 and 2.3 will be implemented at the end of Section 6.

Combining the above results we finally infer:

Corollary 2.5 (uniqueness up to constants of a C^1 generalized minimizer)**.** *Assume that* Ω *is connected and satisfies* (Per)*, that* f *is continuous on* $\Omega\times\mathbb{R}^{Nn}$ *with* (Lin) *and* (Con)*, that* f *is* C^1 *in* z *with continuous gradient* $\nabla_z f$ *on* $\Omega\times\mathbb{R}^{Nn}$*, and that* $f(x,\,\cdot\,)\colon\mathbb{R}^{Nn}\to\mathbb{R}$ *is a strictly convex function for all* $x\in\Omega$*. If there exists* one *generalized minimizer* $u\in\mathrm{BV}(\Omega,\mathbb{R}^N)$ *of* (P) *which is in* C^1 *in the interior of* Ω*, then for* every *generalized minimizer* v *of* (P) *there is a constant* $c\in\mathbb{R}^N$ *such that* $\mathscr{L}^n$*-a. e. on* Ω *we have* $v=u{+}c$*.*

Proof. By Theorem 2.1, the existence of the C^1 generalized minimizer u and the continuity of $\nabla_z f$ give the continuous solution $\sigma:=\nabla_z f(\,\cdot\,,\nabla u)$ of the dual problem (P*). Hence, the assumptions of Corollary 2.4 are satisfied, and the claim follows from Corollary 2.4 via the connectedness of Ω and the constancy theorem. ☐

We remark that, when considering generalized minimizers, uniqueness up to constants, as stated in the last corollary, can only be improved to full uniqueness in quite specific situations. For the area

[4] By this strict convexity assertion we mean precisely that $f^*(x,\,\cdot\,)$ is not affine on any line segment with both endpoints in $\operatorname{Im}\partial_z f(x,\,\cdot\,)$; compare Proposition 3.8. Notice however that in the generality of our setup $f^*(x,\,\cdot\,)$ can be finite and non-strictly convex somewhere outside $\operatorname{Im}\partial_z f(x,\,\cdot\,)$.

integrand $f(x,z) = \sqrt{1+|z|^2}$ in codimension $N = 1$, for instance, Miranda's boundary continuity result [29, 30] yields full uniqueness in case of a continuous boundary datum on a Lipschitz domain, while the examples of Santi [33] and Baldo & Modica [6] show that uniqueness up to constants is optimal for general data. For a detailed discussion of such non-uniqueness phenomena, we refer also to [8] and [19, Chapter V.2].

3 Preliminaries

3.1 Some notation

Though we mostly stick to standard notations, we briefly comment on a few of them.

We write $\mathbb{1}_A \colon \mathbb{R}^n \to \mathbb{R}$ for the characteristic function of a subset A of $\mathbb{R}^n$, and by $\mathrm{B}_R := \{x \in \mathbb{R}^n : |x| < R\}$ we denote the open ball with radius R in $\mathbb{R}^n$. Regarding measures, we only work with (possibly signed or vector-valued) Radon measures on subsets of $\mathbb{R}^n$, which are often given as weighted measures $f\mu$, with weight function f and non-negative base measure μ, or as restrictions $\mu \llcorner A := (\mathbb{1}_A)\mu$ of μ to a set A. For a Radon measure ν, we write ν^{a} and ν^{s} for the absolutely continuous and the singular part in its Lebesgue decomposition with respect to the Lebesgue measure $\mathscr{L}^n$. Further, if ν is absolutely continuous with respect to a non-negative Radon measure μ, then $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ stands for the Radon-Nikodym density of ν with respect to μ, so that we have $\nu = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mu$. Moreover, $\|w\|_{p;\Omega}$ is the L^p-norm, taken on a subset Ω of $\mathbb{R}^n$ with respect to $\mathscr{L}^n$ (and the Euclidean norm on the finite-dimensional target of w). Finally, the space $\mathrm{BV}(\Omega, \mathbb{R}^N)$ of functions of bounded variation is defined as the collection of all functions in $\mathrm{L}^1(\Omega, \mathbb{R}^N)$ whose distributional derivative is represented by a finite $\mathbb{R}^{Nn}$-valued Radon measure. All further terminology for BV-functions follows closely the one of the book [1] — up to a few additional conventions that are explained in the following subsection.

3.2 The Dirichlet problem in BV

Recalling that $u_0 \in \mathrm{W}^{1,1}(\mathbb{R}^n, \mathbb{R}^N)$ is fixed, for every $w \in \mathrm{BV}(\Omega, \mathbb{R}^N)$ we set

$$\overline{w}(x) := \begin{cases} w(x) & \text{for } x \in \Omega \\ u_0(x) & \text{for } x \in \mathbb{R}^n \setminus \Omega \end{cases},$$

and we introduce the class

$$\mathrm{BV}_{u_0}\big(\overline{\Omega}, \mathbb{R}^N\big) := \{w \in \mathrm{BV}(\Omega, \mathbb{R}^N) \,:\, \overline{w} \in \mathrm{BV}(\mathbb{R}^n, \mathbb{R}^N)\}.$$

We stress that, if Ω has a bounded Lipschitz boundary, then [1, Corollary 3.89] implies $\mathrm{BV}_{u_0}\big(\overline{\Omega}, \mathbb{R}^N\big) = \mathrm{BV}(\Omega, \mathbb{R}^N)$. For less regular Ω (for instance, in the presence of sharp external cusps), $\mathrm{BV}_{u_0}\big(\overline{\Omega}, \mathbb{R}^N\big)$ can be strictly smaller than $\mathrm{BV}(\Omega, \mathbb{R}^N)$, but it still contains $\mathrm{W}^{1,1}_{u_0}(\Omega, \mathbb{R}^N)$ and is in particular non-empty. When considering functions $w \in \mathrm{BV}_{u_0}\big(\overline{\Omega}, \mathbb{R}^N\big)$, we will understand in the following that the derivative $\mathrm{D}w$ extends to a measure on $\overline{\Omega}$ which is given by $\mathrm{D}w(B) := \mathrm{D}\overline{w}(B)$ for all Borel subsets B of $\overline{\Omega}$. We denote by $\mathrm{D}^{\mathrm{a}}w := (\mathrm{D}w)^{\mathrm{a}}$ and $\mathrm{D}^{\mathrm{s}}w := (\mathrm{D}w)^{\mathrm{s}}$, respectively, the absolutely continuous and the singular part of this measure (with respect to $\mathscr{L}^n$), and we write ∇w for the density of $\mathrm{D}^{\mathrm{a}}w$ so that in fact we have the Lebesgue decomposition

$$\mathrm{D}w = (\nabla w)\mathscr{L}^n + \mathrm{D}^{\mathrm{s}}w\,.$$

For any $f \colon \Omega \times \mathbb{R}^m \to \mathbb{R}$ with (Lin), we define the recession function $f^\infty \colon \overline{\Omega} \times \mathbb{R}^{Nn} \to \mathbb{R}$ by

$$f^\infty(x,z) := \liminf_{\substack{\widetilde{x} \to x \\ \widetilde{z} \to z \\ t \searrow 0}} t f(\widetilde{x}, \widetilde{z}/t) \qquad \text{for } (x,z) \in \overline{\Omega} \times (\mathbb{R}^{Nn} \setminus \{0\}) \tag{3.1}$$

and $f^\infty(x,0) := 0$ for $x \in \overline{\Omega}$. We observe that f^∞ is lower semicontinuous[5] in (x,z) and positively 1-homogeneous in z, and that (Lin) implies the bound $|f^\infty(x,z)| \le L|z|$. Moreover, we record that the second part of our assumption (Con) just means that the lower limit in (3.1) is indeed a limit, and that in this case f^∞ is automatically continuous in (x,z).

Finally, we are in the position to extend F in a natural way from $\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$ to a class of BV-functions: indeed, following an idea of [23], for $w \in \mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)$ we set

$$\overline{F}_{u_0}[w] := \int_\Omega f(\,\cdot\,,\nabla w)\,\mathrm{d}x + \int_{\overline{\Omega}} f^\infty\Big(\,\cdot\,,\frac{\mathrm{dD^s}w}{\mathrm{d}|\mathrm{D^s}w|}\Big)\,\mathrm{d}|\mathrm{D^s}w|\,,$$

where we have involved u_0 in order to extend $\mathrm{D}w$ to $\overline{\Omega}$ as explained above. The functional $\overline{F}_{u_0}$ will play a crucial role in the present paper, and in particular it is used to specify the notion of generalized minimizers as follows.

Definition 3.1 (generalized minimizer)**.** *Suppose that f fulfills* (Lin)*. A function $u \in \mathrm{BV}(\Omega,\mathbb{R}^N)$ is called a generalized minimizer of the Dirichlet problem* (P) *if we have $u \in \mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)$ and*

$$\overline{F}_{u_0}[u] \le \overline{F}_{u_0}[w] \qquad \textit{for all } w \in \mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)\,.$$

We highlight the two main features of this notion, which have originally been observed in [21, Section 2] under slightly stronger assumptions: first, by Reshetnyak's semicontinuity theorem, generalized minimizers do always exist if f is convex in z with (Lin), linearly coercive[6], and lower semicontinuous in (x,z); second, under our assumptions (Lin), (Per), (Con) it follows from Theorem 4.1 and Lemma 4.3 below that the generalization preserves the infimum value of (P) in the sense of

$$\inf_{\mathrm{BV}_{u_0}(\overline{\Omega},\mathbb{R}^N)} \overline{F}_{u_0} = \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F\,. \tag{3.2}$$

3.3 Convex duality

In this subsection we recall some basic facts from convex analysis; compare, for instance, [19] and [11, Chapter 2].

Background definitions. As the conjugate function f^* in our main results can take the value ∞, it is convenient to provide the following statements for extended real-valued functions $h\colon \mathbb{R}^m \to \mathbb{R}\cup\{-\infty,\infty\}$ in arbitrary dimension $m \in \mathbb{N}$. We define the effective domain $\operatorname{dom} h$ of such an h as

$$\operatorname{dom} h := \{z \in \mathbb{R}^m \,:\, h(z) < \infty\}\,,$$

and by $\operatorname{int}(\operatorname{dom} h)$ we denote the topological interior of $\operatorname{dom} h$.

Definition 3.2 (conjugate function)**.** *Consider an arbitrary function $h\colon \mathbb{R}^m \to \mathbb{R}\cup\{-\infty,\infty\}$. Then its conjugate function $h^*\colon \mathbb{R}^m \to \mathbb{R}\cup\{-\infty,\infty\}$ is given by*

$$h^*(z^*) := \sup_{\xi\in\mathbb{R}^m} \big[z^*\cdot\xi - h(\xi)\big] \qquad \textit{for all } z^* \in \mathbb{R}^m\,.$$

The conjugate function of h^* is called the bi-conjugate function and is denoted by h^{**}.

[5] In the literature one can find several variations of (3.1), which do all coincide in the case of (Con), but may otherwise differ. The main advantage of the variant which we have singled out here is the general validity of the semicontinuity property, which in turn seems favorable in order to gain the existence of minimizers.

[6] Here, we call f linearly coercive if there exist an affine function ℓ and a positive ε such that $f(x,z) \ge \ell(z) + \varepsilon|z|$ holds for all $(x,z) \in \Omega\times\mathbb{R}^{Nn}$.

In addition, we record that, when $h \not\equiv \infty$ is not constantly infinite, then h^* has values in $\mathbb{R} \cup \{\infty\}$ (while for $h \equiv \infty$ we evidently have $h^* \equiv -\infty$). Moreover, being a pointwise supremum of affine functions, h^* is always convex and lower semicontinuous. Finally, for $h^* \not\equiv \infty$ (or equivalently for $h^{**} \not\equiv -\infty$), it is well-known that the bi-conjugate h^{**} coincides with the lower semicontinuous, convex envelope of h, that means that it is the largest convex function $\mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ which is nowhere larger than h; see [19, Proposition I.4.1].

Definition 3.3 (subdifferentiability and subgradients)**.** *For a function $h \colon \mathbb{R}^m \to \mathbb{R} \cup \{-\infty, \infty\}$, one defines the subdifferential $\partial h(z)$ of h at a point $z \in \mathbb{R}^m$ as the set of all $z^* \in \mathbb{R}^m$ with*

$$h(\xi) \geq h(z) + z^* \cdot (\xi - z) \qquad \text{for all } \xi \in \mathbb{R}^m .$$

One says that h is subdifferentiable at $z \in \mathbb{R}^m$ if $\partial h(z)$ is non-empty, and then one calls the elements of $\partial h(z)$ the subgradients of h at z. The collection of all subgradients of h is $\operatorname{Im} \partial h := \bigcup_{z \in \mathbb{R}^m} \partial h(z)$.

Clearly, $\partial h(z)$ is always convex and closed in $\mathbb{R}^m$, and the existence of a classical gradient $\nabla h(z)$ implies that $\partial h(z)$ is either empty or the singleton $\{\nabla h(z)\}$.

Excluding the value $-\infty$, from now on we specialize to functions $h \colon \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$. Then, for $h \not\equiv \infty$, we record a useful characterization of subgradients: there holds $z^* \in \partial h(z)$ if and only if the above supremum in the definition of the conjugate function is attained for the vector $\xi = z$, in other words

$$z^* \in \partial h(z) \iff h(z) + h^*(z^*) = z^* \cdot z \,. \tag{3.3}$$

From (3.3) we read off that one has

$$\operatorname{Im} \partial h \subset \operatorname{dom} h^* , \tag{3.4}$$

and, for convex and lower semicontinuous h, one can additionally show $\operatorname{dom} h^* \subset \overline{\operatorname{Im} \partial h}$. Though we will not need the latter inclusion let us briefly remark that it can be obtained by applying, for $z^* \in \operatorname{dom} h^*$, the Ekeland type result [19, Theorem I.6.2] to a maximizing sequence for $z \mapsto z^* \cdot z - h(z)$.

Subdifferentials of convex functions. Next we turn, specifically, to convex $h \colon \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$. In this case, it is well known that h is locally Lipschitz continuous and subdifferentiable on $\operatorname{int}(\operatorname{dom} h)$; see for instance [19, Corollary I.2.4, Proposition I.5.2]. Moreover, for $z \in \operatorname{int}(\operatorname{dom} h)$, the quantity $|\partial h(z)| := \sup_{z^* \in \partial h(z)} |z^*|$ (with the convention $|\partial h(z)| = 0$ for $\partial h(z) = \emptyset$) is bounded by the Lipschitz constant of h on an arbitrarily small neighborhood of z, and in particular $\partial h(z)$ is a bounded and thus — as we already observed its closedness — compact set. If, in addition to convexity, h is also lower semicontinuous, then h is even subdifferentiable on all of $\operatorname{dom} h$, and by the above interpretation of h^{**} as the lower semicontinuous, convex envelope of h, we necessarily have $h^{**} = h$ on $\mathbb{R}^m$.

In the following we recall some more statements involving subdifferentials of convex functions, but for convenience and completeness we now sketch the proofs in our setting.

Lemma 3.4 (continuity of the subdifferential)**.** *Suppose that $h \colon \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ is convex and that z_k converges to z in $\operatorname{int}(\operatorname{dom} h)$. Then for every choice of $z_k^* \in \partial h(z_k)$ the sequence $(z_k^*)_{k \in \mathbb{N}}$ is bounded in $\mathbb{R}^m$ and all its cluster points are contained in $\partial h(z)$.*

Proof. In view of the preceding observations, $|\partial h(z_k)|$ is bounded for $k \to \infty$ by the Lipschitz constant on a neighborhood of z, so that also z_k^* remains bounded. In addition, the inclusion of the cluster points in $\partial h(z)$ follows straightforwardly from Definition 3.3 and the convergence $\lim_{k \to \infty} h(z_k) = h(z)$. □

Lemma 3.5 (one-sided directional derivatives give subgradients)**.** *Consider a convex function $h \colon \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$, $z \in \operatorname{int}(\operatorname{dom} h)$, and $v \in \mathbb{R}^m$. Then we have*

$$\lim_{s \searrow 0} \frac{h(z{+}sv) - h(z)}{s} = z_v^* \cdot v \qquad \text{for some } z_v^* \in \partial h(z) ,$$

and in particular the limit exists.

Proof. By the convexity inequality, the quantity $\frac{h(z+sv)-h(z)}{s}$ is increasing in s for $0<s\ll 1$, so that its limit for $s\searrow 0$ exists as asserted. By the definition of the subdifferential, we can moreover bound the same quantity from above by $z^*(s)\cdot v$ with arbitrarily chosen $z^*(s)\in\partial h(z{+}sv)$ and from below by $z^*\cdot v$ with any $z^*\in\partial h(z)$. Involving Lemma 3.4, we choose $z_v^*\in\partial h(z)$ as a cluster point of the $z^*(s)$ for $s\searrow 0$ and deduce the claimed equality. □

Lemma 3.6 (criterion for subgradients)**.** *Suppose that $h\colon\mathbb{R}^m\to\mathbb{R}\cup\{\infty\}$ is lower semicontinuous. If some $z_0^*\in\mathbb{R}^m$ satisfies $\lim_{|z|\to\infty}[h(z)-z_0^*\cdot z]=\infty$, then we have $z_0^*\in\operatorname{Im}\partial h$.*

Proof. We can assume $h\not\equiv\infty$. Then, direct minimization gives a minimum point z_0 of $z\mapsto h(z)-z_0^*\cdot z$ in $\mathbb{R}^m$. We infer $h(z_0)+z_0^*\cdot(z{-}z_0)\le h(z)$ for all $z\in\mathbb{R}^m$, and thus we get $z_0^*\in\partial h(z_0)$. □

Proposition 3.7 (convexity and openness of $\operatorname{Im}\partial h$)**.** *Suppose that $h\colon\mathbb{R}^m\to\mathbb{R}\cup\{\infty\}$ is convex and lower semicontinuous, and assume that h is even strictly convex on $\operatorname{dom}h$. Whenever we have $z_0^*\in\operatorname{Im}\partial h$, then there exist positive constants ε and M such that we have*

$$h(z)-z_0^*\cdot z\ge\varepsilon|z|-M\qquad\text{for all } z\in\mathbb{R}^m\,. \tag{3.5}$$

Furthermore, $\operatorname{Im}\partial h$ is convex and open in $\mathbb{R}^m$.

Proof. We can assume $h\not\equiv\infty$.

After subtraction of an affine function and translation, it suffices to prove the claim (3.5) only if $h(0)=0$ and $z_0^*=0\in\partial h(0)$ (otherwise we can take $z_0\in\mathbb{R}^m$ with $z_0^*\in\partial h(z_0)$ and then replace h with $z\mapsto h(z_0+z)-h(z_0)-z_0^*\cdot z$). Then h is non-negative, and the strict convexity implies that $\varepsilon:=\inf_{|z|=1}h(z)$ is positive. By the convexity inequality we get $h(z)\ge h(z/|z|)|z|\ge\varepsilon|z|$ whenever $|z|\ge 1$, and (3.5) follows.

Given any $z_0^*\in\operatorname{Im}\partial h$, as a consequence of (3.5) we have $\lim_{|z|\to\infty}[h(z)-z^*\cdot z]=\infty$ for all $z^*\in\mathbb{R}^m$ with $|z^*{-}z_0^*|<\varepsilon$. By Lemma 3.6, these z^* are contained in $\operatorname{Im}\partial h$, and it is proved that $\operatorname{Im}\partial h$ is open.

Finally, we show that for $z_1^*,z_2^*\in\operatorname{Im}\partial h$ also every convex combination z^* of z_1^* and z_2^* is contained in $\operatorname{Im}\partial h$. First, by (3.5) we have $h(z)-z_1^*\cdot z\ge\varepsilon_1|z|-M_1$ for all $z\in\mathbb{R}^m$ with positive constants ε_1 and M_1. In the case $(z_1^*{-}z^*)\cdot z\ge 0$ we moreover infer $h(z)-z^*\cdot z\ge\varepsilon_1|z|-M_1$. As z^* is a convex combination of z_1^* and z_2^*, in the remaining case $(z_1^*{-}z^*)\cdot z<0$ we have $(z_2^*{-}z^*)\cdot z\ge 0$, and we can apply a completely analogous reasoning with $(z_2^*,\varepsilon_2,M_2)$ in place of $(z_1^*,\varepsilon_1,M_1)$. All in all we get $h(z)-z^*\cdot z\ge\min\{\varepsilon_1,\varepsilon_2\}|z|-\max\{M_1,M_2\}$ for all $z\in\mathbb{R}^m$. Via Lemma 3.6 we deduce $z^*\in\operatorname{Im}\partial h$, and the proof of Proposition 3.7 is complete. □

Proposition 3.8 (∂h^* is the inverse of ∂h)**.** *If $h\colon\mathbb{R}^m\to\mathbb{R}\cup\{\infty\}$ is convex and lower semicontinuous, then for $z,z^*\in\mathbb{R}^m$ we have the equivalence*

$$z^*\in\partial h(z)\qquad\Longleftrightarrow\qquad z\in\partial h^*(z^*)\,, \tag{3.6}$$

and, in particular, h^* is subdifferentiable on $\operatorname{Im}\partial h$. Moreover, if $h\colon\mathbb{R}^m\to\mathbb{R}$ is strictly convex, then h^* is of class C^1 on the open set $\operatorname{Im}\partial h$.

Proof. Assuming $h\not\equiv\infty$, we start with the proof of the forward implication in (3.6). From (3.3) we infer that $z^*\in\partial h(z)$ implies $h(z)+h^*(z^*)=z^*\cdot z$, which we rewrite in turn, with the help of the above-mentioned equality $h^{**}=h$, as $h^*(z^*)+h^{**}(z)=z\cdot z^*$. When we involve (3.3) again, but now with $h^*\not\equiv\infty$ in place of h, the last equality implies $z\in\partial h^*(z^*)$. This proves the forward implication in (3.6), and the backward one follows, when we use $h^{**}=h$ once more and switch the roles of h and h^*.

Turning to the remaining claim, we first observe that the strict convexity implies $\partial h(z_1)\cap\partial h(z_2)=\emptyset$ whenever $z_1\ne z_2$ in $\mathbb{R}^m$. By the subdifferentiability of h and the reverse implication in (3.6), we deduce that in fact $\partial h^*(z^*)=\{g(z^*)\}$ is a singleton for all $z^*\in\operatorname{Im}\partial h$. By Proposition 3.7 and (3.4), $\operatorname{Im}\partial h$ is open and contained in $\operatorname{int}(\operatorname{dom}h^*)$, and then Lemma 3.4 gives continuity of g on $\operatorname{Im}\partial h$, while Lemma 3.5 identifies g as the classical derivative of h^*. □

Finally, we come to the more general case of the x-dependent integrand $f\colon \Omega\times\mathbb{R}^m\to\mathbb{R}$, where for the remainder of this paper we permanently fix[7] $m:=Nn$. In this connection we always consider conjugate functions and subgradients with respect to the second variable, and we use the terminology

$$\operatorname{Im}(x,\partial_z f):=\bigcup_{x\in\Omega}\big[\{x\}\times\operatorname{Im}\partial_z f(x,\,\cdot\,)\big]=\{(x,z^*)\in\Omega\times\mathbb{R}^m\,:\,z^*\in\partial_z f(x,z)\text{ for some }z\in\mathbb{R}^m\}\,.$$

Lemma 3.9 (openness of $\operatorname{Im}(x,\partial_z f)$)**.** *Suppose that f satisfies* (Con) *and that $f(x,\,\cdot\,)\colon\mathbb{R}^m\to\mathbb{R}$ is a strictly convex function for all $x\in\Omega$. Then $\operatorname{Im}(x,\partial_z f)$ is open in $\Omega\times\mathbb{R}^m$.*

Proof. We consider $(x_0,z_0^*)\in\operatorname{Im}(x,\partial_z f)$, that is $x_0\in\Omega$ and $z_0^*\in\operatorname{Im}\partial_z f(x_0,\,\cdot\,)$. Then Proposition 3.7 gives positive constants ε and M such that we have $f(x_0,z)-z_0^*\cdot z\geq\varepsilon|z|-M$ for all $z\in\mathbb{R}^m$. Via (2.1) we find some $\delta>0$ such that for all $(x,z)\in\Omega\times\mathbb{R}^m$ with $|x{-}x_0|+|z|^{-1}<\delta$ we have $|f(x,z)-f(x_0,z)|\leq\frac13\varepsilon|z|$. When we additionally consider an arbitrary $z^*\in\mathbb{R}^m$ with $|z^*{-}z_0^*|<\frac13\varepsilon$, then we get

$$f(x,z)-z^*\cdot z>f(x_0,z)-z_0^*\cdot z-\tfrac23\varepsilon|z|\geq\tfrac13\varepsilon|z|-M\,.$$

With the help of Lemma 3.6 we deduce that all $(x,z^*)\in\Omega\times\mathbb{R}^m$ with $|x{-}x_0|<\delta$ and $|z^*{-}z_0^*|<\frac13\varepsilon$ are contained in $\operatorname{Im}(x,\partial_z f)$. Thus, the latter set is open. $\square$

Lemma 3.10 (joint continuity of f^*)**.** *Suppose that f is continuous on $\Omega\times\mathbb{R}^m$ with* (Con) *and that $f(x,\,\cdot\,)\colon\mathbb{R}^m\to\mathbb{R}$ is a strictly convex function for all $x\in\Omega$. Then f^* is continuous on $\operatorname{Im}(x,\partial_z f)$.*

Proof. By the continuity assumption on f and the definition of the conjugate, f^* is a supremum of continuous functions and is thus lower semicontinuous on all of $\Omega\times\mathbb{R}^m$. Now we argue indirectly and assume that upper semicontinuity of f^* fails at a point $(x_0,z_0^*)\in\operatorname{Im}(x,\partial_z f)$. After subtraction of a linear function from f, we assume $z_0^*=0$, which means $0\in\operatorname{Im}\partial_z f(x_0,\,\cdot\,)$, and then we find a sequence $(x_k,z_k^*)_{k\in\mathbb{N}}$ in $\operatorname{Im}(x,\partial_z f)$ with

$$\lim_{k\to\infty}(x_k,z_k^*)=(x_0,0)\qquad\text{and}\qquad\liminf_{k\to\infty}f^*(x_k,z_k^*)>f^*(x_0,0)\,.\tag{3.7}$$

By the definition of the conjugate, we choose a sequence $(z_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^m$ with

$$\liminf_{k\to\infty}f^*(x_k,z_k^*)\leq\liminf_{k\to\infty}\Big[|z_k^*||z_k|-f(x_k,z_k)\Big]\,.\tag{3.8}$$

If we now had $\limsup_{k\to\infty}|z_k|=\infty$, then (2.1) and Proposition 3.7, the latter applied to $f(x_0,\,\cdot\,)$, would yield $f(x_k,z_k)\geq f(x_0,z_k)-\frac12\varepsilon|z_k|\geq\frac12\varepsilon|z_k|-M$ for infinitely many k and fixed positive constants ε and M, so that the right-hand side of (3.8) would equal $-\infty$. As the left-hand side of (3.8) is bounded from below by (3.7), this cannot happen and we must have $\sup_{k\in\mathbb{N}}|z_k|<\infty$. With this information and the continuity of f we can now bound the right-hand side of (3.8) from above by $\sup_{z\in\mathbb{R}^m}\big[-f(x_0,z)\big]=f^*(x_0,0)$, which contradicts (3.7) and completes the proof. $\square$

Let us briefly remark that the strict convexity assumptions in the last two lemmas cannot be dropped. This is shown already for $n=1$, every Ω which contains 0, and arbitrary $z_0\in\mathbb{R}^m\setminus\{0\}$, by the simple example $f(x,z)=|xz{-}z_0|$, for which we have $\operatorname{Im}(x,\partial_z f)=\{(x,z^*)\in\Omega\times\mathbb{R}^m\,:\,|z^*|\leq|x|\}$, $f^*(x,0)=0$ for $x\neq0$, and $f^*(0,0)=-|z_0|$. Similarly, when we take $n=m=1$, $0\in\Omega$, and $f(x,z)=|xz{-}1|-z+\sqrt{1{+}z^2}$, then we get $f^*(x,0)=(1/x)-\sqrt{1{+}(1/x)^2}$ for $0<x\leq1$ and $f^*(0,0)=-1$; hence, even under the strict convexity assumption, one cannot strengthen the conclusion of Lemma 3.10 to continuity of f^* on all of $\{f^*<\infty\}$.

[7] Nevertheless, we sometimes write $\mathbb{R}^{Nn}$ and sometimes $\mathbb{R}^m$ depending on whether the matrix structure of $z\in\mathbb{R}^{Nn}$ is relevant at the respective stage or not.

4 Reshetnyak continuity and strict approximation

Next we state a refined version of Reshetnyak's continuity theorem [31] which requires only the assumptions (Lin) and (Con) for the integrand f. In particular, these hypotheses comprise the Carathéodory property for f and joint continuity of f^∞ in (x,z), but we emphasize that they do not imply continuity of f itself in x, which is imposed in more common versions [21, 17, 28] of the result. Indeed, the dropping of the latter continuity assumption in x seems quite natural, but to our knowledge it has been carried out only recently by Kristensen & Rindler [27]. Here, we take their corresponding statements as a starting point and then generalize the result to our setup with possibly unbounded Ω.

Theorem 4.1 (Reshetnyak continuity)**.** *Suppose that $\partial\Omega$ has zero $\mathscr{L}^n$-measure and that $f\colon \Omega\times\mathbb{R}^m \to \mathbb{R}$ satisfies* (Con) *and* (Lin) *with some positive $\Psi \in \mathrm{L}^1(\Omega)$ which is bounded away from 0 on every bounded subset of Ω. Assume moreover that $(\mu_k)_{k\in\mathbb{N}}$ weak-$*$-converges*[8] *to μ in the space of finite $\mathbb{R}^m$-valued Radon measures on $\overline{\Omega}$. If there holds*

$$\lim_{k\to\infty} |(\Psi\mathscr{L}^n,\mu_k)|\big(\overline{\Omega}\big) = |(\Psi\mathscr{L}^n,\mu)|\big(\overline{\Omega}\big)$$

for the $\mathbb{R}^{m+1}$-valued measures $(\Psi\mathscr{L}^n,\mu_k)$ and $(\Psi\mathscr{L}^n,\mu)$, then we also have

$$\lim_{k\to\infty}\left[\int_\Omega f\Big(\,\cdot\,,\frac{\mathrm{d}\mu_k^{\mathrm{a}}}{\mathrm{d}\mathscr{L}^n}\Big)\,\mathrm{d}x + \int_{\overline{\Omega}} f^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\mu_k^{\mathrm{s}}}{\mathrm{d}|\mu_k^{\mathrm{s}}|}\Big)\,\mathrm{d}|\mu_k^{\mathrm{s}}|\right] = \int_\Omega f\Big(\,\cdot\,,\frac{\mathrm{d}\mu^{\mathrm{a}}}{\mathrm{d}\mathscr{L}^n}\Big)\,\mathrm{d}x + \int_{\overline{\Omega}} f^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\mu^{\mathrm{s}}}{\mathrm{d}|\mu^{\mathrm{s}}|}\Big)\,\mathrm{d}|\mu^{\mathrm{s}}|\,.$$

Proof. We fix $\varepsilon>0$, and — observing that the second condition in (4.1) is satisfied for all but countably many R — we find a radius R with

$$|(\Psi\mathscr{L}^n,\mu)|\big(\overline{\Omega}\setminus \mathrm{B}_R\big)\le\varepsilon \quad\text{and}\quad |(\Psi\mathscr{L}^n,\mu)|(\overline{\Omega}\cap\partial\mathrm{B}_R)=0\,. \tag{4.1}$$

Then it follows that we have

$$\lim_{k\to\infty} |(\Psi\mathscr{L}^n,\mu_k)|\big(\overline{\Omega\cap \mathrm{B}_R}\big) = |(\Psi\mathscr{L}^n,\mu)|\big(\overline{\Omega\cap \mathrm{B}_R}\big) \quad\text{and}\quad \limsup_{k\to\infty} |(\Psi\mathscr{L}^n,\mu_k)|\big(\overline{\Omega}\setminus\mathrm{B}_R\big)\le\varepsilon\,. \tag{4.2}$$

Next, exploiting the strict positivity assumption on Ψ, we fix continuous and positive functions Ψ_l on $\overline{\Omega\cap\mathrm{B}_R}$ such that Ψ/Ψ_l converges to 1 in $\mathrm{L}^1(\Omega\cap\mathrm{B}_R)$ for $l\to\infty$ (for instance, one can choose the Ψ_l as mollifications). Then we introduce the auxiliary functions $h_l\colon \overline{\Omega\cap\mathrm{B}_R}\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}$ by

$$h_l(x,t,z) := \sqrt{\big[t/\Psi_l(x)\big]^2+|z|^2}\,,$$

and we record that h_l is positively 1-homogeneous in (t,z) and continuous in all variables with $0\le h_l(x,t,z)\le \big[1{+}1/\inf_{\overline{\Omega\cap\mathrm{B}_R}}\Psi_l\big]|(t,z)|$ (where $\overline{\Omega\cap\mathrm{B}_R}$ is compact and thus $\inf_{\overline{\Omega\cap\mathrm{B}_R}}\Psi_l$ is positive). By Reshetnyak's continuity theorem, as stated in [1, Theorem 2.39][9], we thus obtain

$$\lim_{k\to\infty}\int_{\overline{\Omega\cap\mathrm{B}_R}} h_l\bigg(\,\cdot\,,\frac{\mathrm{d}(\Psi\mathscr{L}^n,\mu_k)}{\mathrm{d}|(\Psi\mathscr{L}^n,\mu_k)|}\bigg)\,\mathrm{d}|(\Psi\mathscr{L}^n,\mu_k)| = \int_{\overline{\Omega\cap\mathrm{B}_R}} h_l\bigg(\,\cdot\,,\frac{\mathrm{d}(\Psi\mathscr{L}^n,\mu)}{\mathrm{d}|(\Psi\mathscr{L}^n,\mu)|}\bigg)\,\mathrm{d}|(\Psi\mathscr{L}^n,\mu)| \tag{4.3}$$

for all $l\in\mathbb{N}$. Relying on the 1-homogeneity of h in (t,z), we can split the integrals and use $h_l(x,\Psi(x),z)=\sqrt{[\Psi(x)/\Psi_l(x)]^2+|z|^2}$ and $h(x,0,z)=|z|$ to infer

$$\lim_{k\to\infty}|([\Psi/\Psi_l]\mathscr{L}^n,\mu_k)|\big(\overline{\Omega\cap\mathrm{B}_R}\big) = |([\Psi/\Psi_l]\mathscr{L}^n,\mu)|\big(\overline{\Omega\cap\mathrm{B}_R}\big)\,,$$

[8] In our terminology this convergence means precisely $\lim_{k\to\infty}\int_{\overline{\Omega}}\varphi\,\mathrm{d}\mu_k = \int_{\overline{\Omega}}\varphi\,\mathrm{d}\mu_k$ for every continuous function $\varphi\colon\overline{\Omega}\to\mathbb{R}^m$ with compact (or, here equivalently, with bounded) support.

[9] We remark that the domain of integration in (4.3) is not open as required in [1, Theorem 2.39]. Nevertheless, we can easily deduce (4.3) from the statement of [1] when we extend Ψ_l (and thus h_l) continuously and the measures μ_k and μ by 0 to an open neighborhood of the compactum $\overline{\Omega\cap\mathrm{B}_R}$.

still for all $l \in \mathbb{N}$. Recalling the $\mathrm{L}^1(\Omega)$-convergence of Ψ/Ψ_l and the assumption $\mathscr{L}^n(\partial\Omega) = 0$, we further deduce

$$\lim_{k\to\infty} |(\mathscr{L}^n, \mu_k)|\big(\overline{\Omega \cap \mathrm{B}_R}\big) = |(\mathscr{L}^n, \mu)|\big(\overline{\Omega \cap \mathrm{B}_R}\big)\,.$$

In the following we borrow some results and terminology from [27]. By [27, Proposition 1], the assumed weak-$*$-convergence and the last equality imply also the weak-$*$-convergence in $\mathbf{Y}(\Omega \cap \mathrm{B}_R, \mathbb{R}^N)$ of the elementary Young measures generated by μ_k to the one generated by μ. To exploit this convergence, we introduce, for arbitrary $M \geq 0$ and the constant L from (Lin), the truncated integrand f_M, defined as

$$f_M(x,z) := \begin{cases} -M - 2L|z| & \text{if } f(x,z) < -M - 2L|z| \\ f(x,z) & \text{if } |f(x,z)| \leq M + 2L|z| \\ M + 2L|z| & \text{if } f(x,z) > M + 2L|z| \end{cases}.$$

Then we have $|f_M(x,z)| \leq M+2L|z|$, i. e. the linear growth condition required in [27]. With the help of (Lin) we see that $(f_M)^\infty = f^\infty$ holds and that (Con) carries over from f to f_M. Therefore, the restriction[10] of f_M to $\overline{\Omega \cap \mathrm{B}_R} \times \mathbb{R}^m$ is a representation integrand in the sense of [27, Section 2.4], and we are in the position to apply [27, Proposition 2], which yields

$$\begin{aligned} &\lim_{k\to\infty} \bigg[\int_{\Omega\cap\mathrm{B}_R} f_M\Big(\,\cdot\,, \frac{\mathrm{d}\mu_k^{\mathrm{a}}}{\mathrm{d}\mathscr{L}^n}\Big)\,\mathrm{d}x + \int_{\overline{\Omega\cap\mathrm{B}_R}} f^\infty\Big(\,\cdot\,, \frac{\mathrm{d}\mu_k^{\mathrm{s}}}{\mathrm{d}|\mu_k^{\mathrm{s}}|}\Big)\,\mathrm{d}|\mu_k^{\mathrm{s}}| \bigg] \\ &\qquad = \int_{\Omega\cap\mathrm{B}_R} f_M\Big(\,\cdot\,, \frac{\mathrm{d}\mu^{\mathrm{a}}}{\mathrm{d}\mathscr{L}^n}\Big)\,\mathrm{d}x + \int_{\overline{\Omega\cap\mathrm{B}_R}} f^\infty\Big(\,\cdot\,, \frac{\mathrm{d}\mu^{\mathrm{s}}}{\mathrm{d}|\mu^{\mathrm{s}}|}\Big)\,\mathrm{d}|\mu^{\mathrm{s}}|\,. \end{aligned}$$

When we observe $|f(x,z) - f_M(x,z)| \leq (\Psi(x) - M)_+$ and send $M \to \infty$, we can replace f_M with f in the last equality. Hence, we get the claimed convergence, but initially with $\Omega \cap \mathrm{B}_R$ in place of Ω. Then we employ (Lin), (4.1), and (4.2) in order to control the corresponding integrals over $\overline{\Omega} \setminus \overline{\Omega \cap \mathrm{B}_R} \subset \overline{\Omega} \setminus \mathrm{B}_R$ so that we get the claim up to an error of at most $2(1+L)\varepsilon$. Taking into account the arbitrariness of ε, the proof is complete. □

Remark 4.2. *In fact, a version of Theorem 4.1 holds true for arbitrary open sets Ω (even when $\mathscr{L}^n(\partial\Omega) > 0$); to formulate this version, one needs to define f and Ψ, require the respective assumptions, and take all integrals* on the closure $\overline{\Omega}$ of Ω.

Starting from a given $u \in \mathrm{BV}_{u_0}\big(\overline{\Omega}, \mathbb{R}^N\big)$, we next construct convenient strict approximations $(w_k)_{k\in\mathbb{N}}$ such that in particular the preceding theorem applies for the convergence of the gradient measures. Basically, the existence of such approximations is classical, but a detailed proof of the here relevant version which covers bounded Lipschitz domains Ω seems to have been written down only in [11, Lemma B.2]. Moreover, the generality of bounded Ω with (Per) has been reached in [34]. In the following we continue to work under the hypothesis (Per), which is used only implicitly by quoting the relevant constructions from [34], and we generalize the strict approximation result [34, Theorem 1.2] to possibly unbounded Ω. Moreover — as a slight but decisive extra feature in the spirit of [5, Lemma 5.1] — we achieve the almost-everywhere convergence (4.5) for the absolutely continuous parts of the gradients.

Finally, we state the approximation lemma, which heavily uses the convention of Section 3.2 that gradient measures of functions on Ω are extended to $\overline{\Omega}$ with the aid of the fixed u_0.

Lemma 4.3 (strict and almost-everywhere approximation in BV)**.** *Suppose that Ω satisfies* (Per) *and consider an arbitrary non-negative $\Psi \in \mathrm{L}^1(\Omega)$. For every $u \in \mathrm{BV}_{u_0}\big(\overline{\Omega}, \mathbb{R}^N\big)$ there exists a sequence*

[10] In our setting, $(f_M)^\infty = f^\infty$ is (and needs to be) defined on $\overline{\Omega}\times\mathbb{R}^m$, while f_M itself is initially only given on $\Omega\times\mathbb{R}^m$. In view of $\mathscr{L}^n(\partial\Omega) = 0$, we can however assume that f_M extends suitably to $\overline{\Omega}\times\mathbb{R}^m \supset \overline{\Omega \cap \mathrm{B}_R}\times\mathbb{R}^m$; indeed, we can take $f_M(x,z) := f^\infty(x,z)$ for $x \in \partial\Omega$.

$(w_k)_{k\in\mathbb{N}}$ *in* $u_0|_\Omega + \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega,\mathbb{R}^N) \subset \mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$ *such that* $(w_k)_{k\in\mathbb{N}}$ *converges to* u *in* $\mathrm{L}^1(\Omega,\mathbb{R}^N)$ *and* $(\mathrm{D}w_k)_{k\in\mathbb{N}}$ *weak-∗-converges to* $\mathrm{D}u$ *as finite* $\mathbb{R}^{Nn}$*-valued measures on* $\overline{\Omega}$ *with*

$$\lim_{k\to\infty} |(\Psi\mathscr{L}^n, \mathrm{D}w_k)|\big(\overline{\Omega}\big) = |(\Psi\mathscr{L}^n, \mathrm{D}u)|\big(\overline{\Omega}\big)\,, \tag{4.4}$$

$$\nabla w_k \to \nabla u \quad \mathscr{L}^n\text{-a. e. in } \Omega\,. \tag{4.5}$$

Proof. For bounded Ω, $\Psi \equiv 1$, and without (4.5), the statement is just a reformulation of [34, Theorem 1.2]. We will now show how the general statement can be deduced.

Step 1. We prove Lemma 4.3 under the stated assumptions on Ω and Ψ, but at first without (4.5). To this end, we choose, for every $k \in \mathbb{N}$, a radius $r_k \in (k, k{+}1)$ with the following three properties: $\partial\Omega$ intersects $\partial\mathrm{B}_{r_k}$ in a set of zero $\mathscr{H}^{n-1}$-measure, the two one-sided traces of the BV-function $\overline{u}$ from Section 3.2 coincide $\mathscr{H}^{n-1}$-a. e. on $\partial\mathrm{B}_{r_k}$ with its Lebesgue representative (and hence with each other), and we have

$$\int_{\partial\mathrm{B}_{r_k}} |\overline{u}{-}u_0|\,\mathrm{d}\mathscr{H}^{n-1} \le \int_{\mathbb{R}^n\setminus\mathrm{B}_k} |\overline{u}{-}u_0|\,\mathrm{d}x\,. \tag{4.6}$$

This choice is possible, as (Per) and the Federer-Vol'pert theorem (see [1, Theorem 3.77, Theorem 3.78, Remark 3.79]) imply that all but countably many radii have the first two desired properties, while the following Fubini type argument guarantees the validity of the last property for the radii in a subset of $(k, k{+}1)$ with positive $\mathscr{L}^1$-measure: indeed, denoting by S the set of radii for which (4.6) fails, we have

$$\mathscr{L}^1(S)\int_{\mathbb{R}^n\setminus\mathrm{B}_k} |\overline{u}{-}u_0|\,\mathrm{d}x < \int_S\int_{\partial\mathrm{B}_r} |\overline{u}{-}u_0|\,\mathrm{d}\mathscr{H}^{n-1}\,\mathrm{d}r \le \int_{\mathbb{R}^n\setminus\mathrm{B}_k} |\overline{u}{-}u_0|\,\mathrm{d}x\,,$$

hence we get $\mathscr{L}^1(S) < 1$, and the complement of S has positive $\mathscr{L}^1$-measure. Once r_k is chosen, it is not difficult to verify that also

$$\Omega_k := \Omega \cap \mathrm{B}_{r_k}$$

satisfies the condition (Per). We now set $u_k := u|_{\Omega_k}$ and write $\overline{u_k}$ for the extension of u_k to $\mathbb{R}^n$ by the values of u_0. From (4.6) and [1, Theorem 3.84] we infer $\overline{u_k} = \mathbb{1}_{\mathrm{B}_{r_k}}\overline{u} + \mathbb{1}_{\mathbb{R}^n\setminus\mathrm{B}_{r_k}} u_0 \in \mathrm{BV}(\mathbb{R}^n,\mathbb{R}^N)$ and

$$|\mathrm{D}\overline{u_k}| \llcorner \overline{\Omega_k} = |\mathrm{D}\overline{u}| \llcorner \big(\overline{\Omega}\cap\mathrm{B}_{r_k}\big) + |u_0{-}\overline{u}|\mathscr{H}^{n-1} \llcorner \big(\overline{\Omega}\cap\partial\mathrm{B}_{r_k}\big)\,. \tag{4.7}$$

In particular, we have $u_k \in \mathrm{BV}_{u_0}\big(\overline{\Omega_k},\mathbb{R}^N\big)$, and, analogous to the convention of Section 3.2, we use $\mathrm{D}\overline{u_k}$ in order to extend $\mathrm{D}u_k$ to a measure on $\overline{\Omega_k}$. Applying [34, Theorem 1.2] to u_k on the bounded set Ω_k, we find a sequence $(\widetilde{w}_{k,\ell})_{\ell\in\mathbb{N}}$ in $u_0|_{\Omega_k} + \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega_k,\mathbb{R}^N)$ with

$$\lim_{\ell\to\infty} \|\widetilde{w}_{k,\ell}{-}u_k\|_{1;\Omega_k} = 0\,,$$

$$\lim_{\ell\to\infty} |(\mathscr{L}^n, \mathrm{D}\widetilde{w}_{k,\ell})|\big(\overline{\Omega_k}\big) = |(\mathscr{L}^n, \mathrm{D}u_k)|\big(\overline{\Omega_k}\big)\,.$$

It follows that a subsequence of $(\mathrm{D}\widetilde{w}_{k,\ell})_{\ell\in\mathbb{N}}$ weak-∗-converges in the sense of measures on $\overline{\Omega_k}$, and the limit measure must be $\mathrm{D}u_k$. Next we choose M_k large enough that $\Psi_k := \min\{\Psi, M_k\}$ satisfies $\|\Psi_k - \Psi\|_{1;\Omega} \le \frac{1}{k}$. Then we make use of Theorem 4.1, applied with Ω_k in place of Ω, the constant 1 in place of Ψ, and the integrand[11] $(x,z) \mapsto |(\Psi_k(x), z)|$ in place of f. Consequently, we can take $\ell(k)$ large enough that

$$\|\widetilde{w}_{k,\ell(k)}{-}u_k\|_{1;\Omega_k} \le \frac{1}{k}\,,$$

$$|(\Psi_k\mathscr{L}^n, \mathrm{D}\widetilde{w}_{k,\ell(k)})|\big(\overline{\Omega_k}\big) \le |(\Psi_k\mathscr{L}^n, \mathrm{D}u_k)|\big(\overline{\Omega_k}\big) + \frac{1}{k}\,,$$

[11] The approximations Ψ_k are needed, since the integrand $(x,z) \mapsto |(\Psi(x), z)|$ does not satisfy the relevant assumption (Con) in Theorem 4.1 if Ψ is unbounded on Ω_k.

and via the choice of Ψ_k, (4.7), and (4.6) we also get

$$
\begin{aligned}
|(\Psi\mathscr{L}^n, \mathrm{D}\widetilde{w}_{k,\ell(k)})|\big(\overline{\Omega_k}\big) &\le |(\Psi\mathscr{L}^n, \mathrm{D}u)|\big(\overline{\Omega} \cap \mathrm{B}_{r_k}\big) + \int_{\partial \mathrm{B}_{r_k}} |u_0 - \overline{u}| \,\mathrm{d}\mathscr{H}^{n-1} + \frac{2}{k} \\
&\le |(\Psi\mathscr{L}^n, \mathrm{D}u)|\big(\overline{\Omega}\big) + \int_{\mathbb{R}^n \setminus \mathrm{B}_k} |u_0 - \overline{u}| \,\mathrm{d}x + \frac{2}{k} \,.
\end{aligned}
$$

Now we introduce $w_k \in u_0\big|_\Omega + \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega, \mathbb{R}^N)$ as the extension of $\widetilde{w}_{k,\ell(k)}$ from Ω_k to Ω via the values of u_0. Then the preceding estimates readily yield

$$
\|w_k - u\|_{1;\Omega} \le \frac{1}{k} + \|u_0 - u\|_{1;\Omega \setminus \mathrm{B}_k}
$$

and

$$
\begin{aligned}
|(\Psi\mathscr{L}^n, \mathrm{D}w_k)|\big(\overline{\Omega}\big) &\le |(\Psi\mathscr{L}^n, \mathrm{D}\widetilde{w}_{k,\ell(k)})|\big(\overline{\Omega_k}\big) + \int_{\overline{\Omega} \setminus \mathrm{B}_k} |(\Psi, \nabla u_0)| \,\mathrm{d}x \\
&\le |(\Psi\mathscr{L}^n, \mathrm{D}u)|\big(\overline{\Omega}\big) + \frac{2}{k} + \int_{\mathbb{R}^n \setminus \mathrm{B}_k} \big[|(\Psi, \nabla u_0)| + |u_0 - \overline{u}|\big] \,\mathrm{d}x \,.
\end{aligned}
$$

We infer that $(w_k)_{k\in\mathbb{N}}$ converges to u in $\mathrm{L}^1(\Omega, \mathbb{R}^N)$ with

$$
\limsup_{k\to\infty} |(\Psi\mathscr{L}^n, \mathrm{D}w_k)|\big(\overline{\Omega}\big) \le |(\Psi\mathscr{L}^n, \mathrm{D}u)|\big(\overline{\Omega}\big) \,,
$$

and this in turn implies — via a standard argument with subsequences — that $(\mathrm{D}w_k)_{k\in\mathbb{N}}$ weak-$*$-converges to $\mathrm{D}u$ in the sense of measures on $\overline{\Omega}$. By the lower semicontinuity of the total variation, we arrive at (4.4).

Step 2. We finally establish the full statement of Lemma 4.3 including (4.5). To this end we now denote

$$
\Omega_k := \big\{x \in \Omega \,:\, \mathrm{dist}(x, \partial\Omega) > k^{-1}\big\} \,,
$$

and we first consider mollifications $u_k \in \mathrm{W}^{1,1}(\Omega_{2k}, \mathbb{R}^N)$ of u such that we have[12]

$$
\begin{gathered}
\|u_k - u\|_{1;\Omega_{2k}} \le k^{-2} \qquad \text{for all } k \in \mathbb{N} \,, \\
\nabla u_k \to \nabla u \qquad \mathscr{L}^n\text{-a. e. in } \Omega \,, \\
\limsup_{k\to\infty} |(\Psi\mathscr{L}^n, \mathrm{D}u_k)|(\Omega_{2k}) \le |(\Psi\mathscr{L}^n, \mathrm{D}u)|(\Omega) \,.
\end{gathered}
$$

Moreover, by the preceding Step 1 we can also find a sequence $(\widetilde{w}_k)_{k\in\mathbb{N}}$ in $u_0\big|_\Omega + \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega, \mathbb{R}^N)$ such that we have

$$
\begin{gathered}
\|\widetilde{w}_k - u\|_{1;\Omega} \le k^{-2} \,, \\
|(\Psi\mathscr{L}^n, \mathrm{D}\widetilde{w}_k)|\big(\overline{\Omega}\big) \le |(\Psi\mathscr{L}^n, \mathrm{D}u)|\big(\overline{\Omega}\big) + \frac{1}{k}
\end{gathered}
$$

for all $k \in \mathbb{N}$. We record that $(\mathrm{D}\widetilde{w}_k)_{k\in\mathbb{N}}$ weak-$*$-converges to $\mathrm{D}u$ in the sense of measures on $\overline{\Omega}$. Now, for all $k \in \mathbb{N}$, we choose cut-off functions $\eta_k \in \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega)$ which satisfy $\mathbb{1}_{\Omega_k} \le \eta_k \le \mathbb{1}_{\Omega_{2k}}$ and $|\nabla \eta_k| \le 4k$ on Ω. Introducing $w_k := \eta_k u_k + (1-\eta_k)\widetilde{w}_k \in u_0\big|_\Omega + \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega, \mathbb{R}^N)$ we observe that $(w_k)_{k\in\mathbb{N}}$ converges to u in $\mathrm{L}^1(\Omega, \mathbb{R}^N)$ and that (4.5) is valid. Then for fixed $\ell \in \mathbb{N}$ and $k \ge \ell$ we find

$$
\begin{aligned}
|(\Psi\mathscr{L}^n, \mathrm{D}w_k)|\big(\overline{\Omega}\big) &\le |\eta_k(\Psi\mathscr{L}^n, \mathrm{D}u_k)|\big(\overline{\Omega}\big) + |(1-\eta_k)(\Psi\mathscr{L}^n, \mathrm{D}\widetilde{w}_k)|\big(\overline{\Omega}\big) + \int_\Omega |(u_k - \widetilde{w}_k) \otimes \nabla\eta_k| \,\mathrm{d}x \\
&\le |(\Psi\mathscr{L}^n, \mathrm{D}u_k)|(\Omega_{2k}) + |(\Psi\mathscr{L}^n, \mathrm{D}\widetilde{w}_k)|\big(\overline{\Omega} \setminus \Omega_\ell\big) + 8k^{-1} \,,
\end{aligned}
$$

[12] In connection with the a. e. convergence observe that the ∇u_k are mollifications of the measure $\mathrm{D}u = (\nabla u)\mathscr{L}^n + \mathrm{D}^{\mathrm{s}}u$; the mollifications of ∇u converge $\mathscr{L}^n$-a. e. to ∇u, while the mollifications of $\mathrm{D}^{\mathrm{s}}u$ converge $\mathscr{L}^n$-a. e. to 0.

where we estimated the last term via the fast convergences of u_k and $\widetilde{w}_k$ and the bound for $|\nabla\eta_k|$. Splitting $|(\Psi\mathscr{L}^n,\mathrm{D}\widetilde{w}_k)|\big(\overline{\Omega}\setminus\Omega_\ell\big)=|(\Psi\mathscr{L}^n,\mathrm{D}\widetilde{w}_k)|\big(\overline{\Omega}\big)-|(\Psi\mathscr{L}^n,\mathrm{D}\widetilde{w}_k)|(\Omega_\ell)$ we now send first k to ∞ and use the lower semicontinuity of the total variation on the open Ω_ℓ to arrive at

$$\limsup_{k\to\infty}|(\Psi\mathscr{L}^n,\mathrm{D}w_k)|\big(\overline{\Omega}\big)\leq|(\Psi\mathscr{L}^n,\mathrm{D}u)|(\Omega)+|(\Psi\mathscr{L}^n,\mathrm{D}u)|\big(\overline{\Omega}\setminus\Omega_\ell\big)\,.$$

Then we pass also ℓ to ∞, and we conclude

$$\limsup_{k\to\infty}|(\Psi\mathscr{L}^n,\mathrm{D}w_k)|\big(\overline{\Omega}\big)\leq|(\Psi\mathscr{L}^n,\mathrm{D}u)|\big(\overline{\Omega}\big)\,.$$

As usual, we deduce that $(\mathrm{D}w_k)_{k\in\mathbb{N}}$ weak-$*$-converges to $\mathrm{D}u$ in the sense of measures on $\overline{\Omega}$, and the lower semicontinuity of the total variation gives (4.4). ☐

Remark 4.4. *If Ω, f, and Ψ satisfy the assumptions of Theorem 4.1 and Lemma 4.3, then the approximations of Lemma 4.3 have the following property, which we record for later usage: whenever for a Borel set B in $\mathbb{R}^n$ we have $(\mathscr{L}^n{+}|\mathrm{D}u|)\big(\overline{\Omega}\cap\partial B\big)=0$, then there holds*

$$\lim_{k\to\infty}\int_{\Omega\cap B}f(\,\cdot\,,\nabla w_k)\,\mathrm{d}x=\int_{\Omega\cap B}f(\,\cdot\,,\nabla u)\,\mathrm{d}x+\int_{\overline{\Omega}\cap B}f^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\mathrm{D}^{\mathrm{s}}u}{\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|}\Big)\,\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|\,,\tag{4.8}$$

where, in the case that f is positively 1-homogeneous in its second variable, the right-hand side simplifies to $\int_{\overline{\Omega}\cap B}f\big(\,\cdot\,,\frac{\mathrm{dD}u}{\mathrm{d}|\mathrm{D}u|}\big)\,\mathrm{d}|\mathrm{D}u|$. Indeed, in order to prove (4.8), one uses the semicontinuity of the total variation on the relatively open sets $\overline{\Omega}\cap\operatorname{int}B$ and $\overline{\Omega}\setminus\overline{B}$ to deduce the convergence $\lim_{k\to\infty}|(\Psi\mathscr{L}^n,\mathrm{D}w_k)|\big(\overline{\Omega\cap\operatorname{int}B}\big)=|(\Psi\mathscr{L}^n,\mathrm{D}u)|\big(\overline{\Omega\cap\operatorname{int}B}\big)$ (compare with [20, Theorem 1.9.1]); then one concludes by Theorem 4.1, with $\Omega\cap\operatorname{int}B$ in place of Ω.

5 The duality formula and the extremality relation for ∇u

We start with a separation lemma from functional analysis which we have chosen to state for all $p\in(1,\infty)$. We will use this lemma only for $p=2$ (and alternatively we could use it for every other fixed choice of $p>1$), but we want to emphasize that the statement does not carry over to the case $p=1$ which will cause slight technical complications later on.

Lemma 5.1 (separation lemma)**.** *Consider $\delta>0$, $p\in(1,\infty)$, a convex set C in $\mathrm{L}^\infty(\Omega,\mathbb{R}^m)$, and a closed subspace S of $\mathrm{L}^p(\Omega,\mathbb{R}^m)$ such that $\|\Phi\|_{1;\Omega}\leq M\|\Phi\|_{p;\Omega}$ holds for all $\Phi\in S$ and a constant M. If for every $\Phi\in S$ there is some $\tau_\Phi\in C$ with*

$$\int_\Omega\tau_\Phi\cdot\Phi\,\mathrm{d}x<\delta\|\Phi\|_{p;\Omega}\,,$$

then there also exists some $\tau\in C$ with

$$\int_\Omega\tau\cdot\Phi\,\mathrm{d}x<\delta\|\Phi\|_{p;\Omega}\qquad\text{for all }\Phi\in S\,.$$

Proof. In view of the assumed inequality $\|\Phi\|_{1;\Omega}\leq M\|\Phi\|_{p;\Omega}$, the specification $\langle R\tau,\Phi\rangle:=\int_\Omega\tau\cdot\Phi\,\mathrm{d}x$ defines a continuous linear operator $R\colon\mathrm{L}^\infty(\Omega,\mathbb{R}^m)\to S^*$. We now prove the claimed implication by a contradiction argument. Indeed, if the conclusion were wrong, we would have $\|R\tau\|_{S^*}\geq\delta$ for every $\tau\in C$. By the Hahn-Banach separation theorem (see for instance [19, Corollary I.1.1]) we could then separate the convex set $R(C)$ from the open ball with radius δ and center 0 in S^*, meaning that we would have $\big\langle F,R\tau\big\rangle\geq\delta$ for all $\tau\in C$ and some $F\in S^{**}$ with $\|F\|_{S^{**}}=1$. As we are assuming $1<p<\infty$, the space $\mathrm{L}^p(\Omega,\mathbb{R}^m)$ and its closed subspace S are reflexive, and F would coincide with the evaluation on some $\Phi\in S$ such that $\|\Phi\|_{p;\Omega}=1$. Hence, we would get

$$\int_\Omega\tau\cdot\Phi\,\mathrm{d}x=\big\langle R\tau,\Phi\big\rangle=\big\langle F,R\tau\big\rangle\geq\delta=\delta\|\Phi\|_{p;\Omega}\qquad\text{for all }\tau\in C\,.$$

Clearly, the existence of such a Φ would contradict our premise, and the lemma is proved. ☐

The next lemma, based on Ekeland's variational principle, is crucial for our approach.

Proposition 5.2 (approximative solutions)**.** *Assume that f satisfies* (Lin) *and that $f(x,\,\cdot\,)\colon \mathbb{R}^m \to \mathbb{R}$ is a convex function for $\mathscr{L}^n$-a. e. $x \in \Omega$. Furthermore, consider $\varepsilon, \chi \in (0,\infty)$ and a closed subspace S of $\mathrm{L}^2(\Omega,\mathbb{R}^m)$ such that $\|\Phi\|_{1;\Omega} \le M\|\Phi\|_{2;\Omega}$ holds for all $\Phi \in S$ and a constant M. Then, for every $\widehat{w} \in \mathrm{L}^1(\Omega,\mathbb{R}^m)$ with*

$$\int_\Omega f(\,\cdot\,,\widehat{w})\,\mathrm{d}x \le \inf_{\Theta\in\widehat{w}+S}\int_\Omega f(\,\cdot\,,\Theta)\,\mathrm{d}x + \varepsilon$$

there exist approximative solutions $\widehat{v} \in \widehat{w}+S$ and $\tau \in \mathrm{L}^\infty(\Omega,\mathbb{R}^m)$ such that we have

$$\int_\Omega f(\,\cdot\,,\widehat{v})\,\mathrm{d}x \le \inf_{\Theta\in\widehat{w}+S}\int_\Omega f(\,\cdot\,,\Theta)\,\mathrm{d}x + 2\varepsilon\,, \tag{5.1}$$

$$\|\widehat{v}-\widehat{w}\|_{2;\Omega} \le \chi\,, \tag{5.2}$$

$$\tau(x) \in \partial_z f(x,\widehat{v}(x)) \text{ for } \mathscr{L}^n\text{-a. e. } x\in\Omega\,, \tag{5.3}$$

$$\int_\Omega \tau\cdot\Phi\,\mathrm{d}x < \frac{2\varepsilon}{\chi}\|\Phi\|_{2;\Omega} \text{ for all } \Phi\in S\,. \tag{5.4}$$

Proof. By the convexity assumption, $f(x,\,\cdot\,)$ is a continuous function for $\mathscr{L}^n$-a. e. $x\in\Omega$, and we infer with the help of (Lin) and the assumed inequality $\|\Phi\|_{1;\Omega} \le M\|\Phi\|_{2;\Omega}$ that $\Theta \mapsto \int_\Omega f(\,\cdot\,,\Theta)\,\mathrm{d}x$ is finite and continuous on the complete metric space $(\widehat{w}+S, \|\cdot\|_{2;\Omega})$. An application of Ekeland's variational principle [18, Theorem 1.1] thus yields a function $\widehat{v}\in\widehat{w}+S$ with

$$\begin{aligned}\|\widehat{v}-\widehat{w}\|_{2;\Omega} &\le \chi\,,\\ \int_\Omega f(\,\cdot\,,\widehat{v})\,\mathrm{d}x &\le \int_\Omega f(\,\cdot\,,\Theta)\,\mathrm{d}x + \frac{\varepsilon}{\chi}\|\Theta-\widehat{v}\|_{2;\Omega} \qquad \text{for all } \Theta\in\widehat{w}+S\,.\end{aligned} \tag{5.5}$$

In particular, we get $\int_\Omega f(\,\cdot\,,\widehat{v})\,\mathrm{d}x \le \int_\Omega f(\,\cdot\,,\widehat{w})\,\mathrm{d}x + \varepsilon \le \inf_{\Theta\in\widehat{w}+S}\int_\Omega f(\,\cdot\,,\Theta)\,\mathrm{d}x + 2\varepsilon$, and thus (5.1) and (5.2) are verified. When we test (5.5) with $\Theta = \widehat{v} - s\Phi$, where $s>0$ and $\Phi\in S$ are arbitrary, we deduce

$$-\int_\Omega \frac{f(\,\cdot\,,\widehat{v}-s\Phi) - f(\,\cdot\,,\widehat{v})}{s}\,\mathrm{d}x \le \frac{\varepsilon}{\chi}\|\Phi\|_{2;\Omega}\,. \tag{5.6}$$

For $\mathscr{L}^n$-a. e. $x\in\Omega$, we now use Lemma 3.5 to find some $\tau_\Phi(x)$ with

$$\begin{aligned}\tau_\Phi(x) &\in \partial_z f(x,\widehat{v}(x))\,,\\ -\tau_\Phi(x)\cdot\Phi(x) &= \lim_{s\searrow 0}\frac{f(x,\widehat{v}(x)-s\Phi(x)) - f(x,\widehat{v}(x))}{s}\,.\end{aligned} \tag{5.7}$$

We immediately observe from (5.7) that $\tau_\Phi\cdot\Phi$ is Lebesgue measurable, while on the other hand it is not evident that τ_Φ itself is measurable. We claim however that one can modify the τ_Φ so that they become Lebesgue measurable, while (5.7) still holds for $\mathscr{L}^n$-a. e. $x\in\Omega$. Indeed, let us briefly sketch how this last claim can be justified using the theory of measurable multifunctions as described in [32]: first, by [32, Corollary 2X][13] the multifunction $\Gamma\colon\Omega\to\mathbb{R}^m$ with $\Gamma(x) := \partial_z f(x,\widehat{v}(x))$ is closed-valued and Lebesgue measurable — in one of the equivalent senses of [32, Proposition 1A]. Similarly, also $\Upsilon_\Phi(x) := \{z^*\in\mathbb{R}^m : z^*\cdot\Phi(x) = \tau_\Phi(x)\cdot\Phi(x)\}$ defines a closed-valued Lebesgue measurable multifunction $\Upsilon_\Phi\colon\Omega\to\mathbb{R}^m$ (this follows from the measurability of $\tau_\Phi\cdot\Phi$ and can be easily verified with the help of [32, Corollary 1.D]). By [32, Theorem 1.M] also the pointwise intersection $\Gamma\cap\Upsilon_\Phi$ is closed-valued and Lebesgue measurable. Moreover, the existence of the above τ_Φ shows that the values of $\Gamma\cap\Upsilon_\Phi$ are non-empty. Hence, by [32, Theorem 1.C] we can choose a Lebesgue measurable selection $\widetilde{\tau}_\Phi\colon\Omega\to\mathbb{R}^m$ with $\widetilde{\tau}_\Phi(x)\in\Gamma(x)\cap\Upsilon_\Phi(x)$ for $\mathscr{L}^n$-a. e. $x\in\Omega$. By the definitions of Γ and Υ_Φ,

[13]Notice also that, as our integrands f are always Borel measurable, the normality assumption in [32, Corollary 2X] is satisfied as a consequence of [32, Theorem 2F].

the last inclusion shows that (5.7) still holds with $\widetilde{\tau}_\Phi$ in place of τ_Φ. A posteriori we can thus assume that the τ_Φ themselves are all measurable with (5.7). From (Lin) and [22, Lemma 5.2], we infer that f is Lipschitz continuous in z, uniformly in x, so that $|\partial_z f|$ is bounded on $\Omega \times \mathbb{R}^{Nn}$. Hence, we read off from (5.7) that we can see τ_Φ as an element of $\mathrm{L}^\infty(\Omega,\mathbb{R}^m)$, and dominated convergence in (5.6) gives

$$\int_\Omega \tau_\Phi \cdot \Phi \,\mathrm{d}x < \frac{2\varepsilon}{\chi}\|\Phi\|_{2;\Omega}\,.$$

We are thus in the position to apply Lemma 5.1 with $p=2$, the convex set

$$\Big\{\vartheta \in \mathrm{L}^\infty(\Omega,\mathbb{R}^m)\,:\,\vartheta(x)\in\partial_z f(x,\widehat{v}(x)) \text{ for } \mathscr{L}^n\text{-a. e. } x\in\Omega\Big\}\,,$$

and the closed subspace S of $\mathrm{L}^2(\Omega,\mathbb{R}^m)$. The lemma then gives a function $\tau \in \mathrm{L}^\infty(\Omega,\mathbb{R}^m)$ with $\tau(x)\in\partial_z f(x,\widehat{v}(x))$ for $\mathscr{L}^n$-a. e. $x\in\Omega$ and with

$$\int_\Omega \tau\cdot\Phi\,\mathrm{d}x < \frac{2\varepsilon}{\chi}\|\Phi\|_{2;\Omega} \qquad \text{for all } \Phi\in S\,.$$

Thus we have established (5.3) and (5.4), and the proof of the proposition is complete. □

Remark 5.3. *If f is of class C^1 in z, then the proof of Proposition 5.2 simplifies considerably. Indeed, neither measurable selections nor Lemma 5.1 are needed in this situation, as the manifest choice $\tau := \nabla_z f(\,\cdot\,,\widehat{v})$ satisfies* (5.7) *for all $\Phi\in S$.*

Next we turn to the proof of Theorem 1.1, in which the existence of the approximative solutions of Proposition 5.2 will be exploited in order to apply the following simple lemma.

Lemma 5.4. *Assume that f satisfies* (Lin)*. If for some sequences $(v_k)_{k\in\mathbb{N}}$ in $\mathrm{W}^{1,1}(\Omega,\mathbb{R}^N)$ and $(\tau_k)_{k\in\mathbb{N}}$ in $\mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})$ we have*

$$\begin{aligned} \tau_k(x) &\in \partial_z f(x,\nabla v_k(x)) \qquad \textit{for } \mathscr{L}^n\textit{-a. e. } x\in\Omega\,,\\ \limsup_{k\to\infty} &\int_\Omega \tau_k\cdot(\nabla u_0-\nabla v_k)\,\mathrm{d}x \ge 0\,, \end{aligned} \tag{5.8}$$

and if $(\tau_k)_{k\in\mathbb{N}}$ weak-$*$-converges in $\mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})$ to a limit $\sigma\in\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$, then we have

$$R_{u_0}[\sigma] \ge \liminf_{k\to\infty} F[v_k]\,.$$

Proof. We record that f^* is convex and lower semicontinuous in its second variable, and (Lin) gives the lower bound $f^*(x,z^*)\ge -\Psi(x)$ with the L^1-function Ψ. In this situation, [16, Theorem 3.20][14] guarantees upper semicontinuity of $\vartheta\mapsto -\int_\Omega f^*(\,\cdot\,,\vartheta)\,\mathrm{d}x$ with respect to weak-$*$-convergence in $\mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})$, and thus we get

$$R_{u_0}[\sigma] = \int_\Omega \big[\sigma\cdot\nabla u_0 - f^*(\,\cdot\,,\sigma)\big]\,\mathrm{d}x \ge \limsup_{k\to\infty}\int_\Omega\big[\tau_k\cdot\nabla u_0 - f^*(\,\cdot\,,\tau_k)\big]\,\mathrm{d}x\,. \tag{5.9}$$

From the first part of (5.8) and (3.3) we deduce $f(\,\cdot\,,\nabla v_k)+f^*(\,\cdot\,,\tau_k)=\tau_k\cdot\nabla v_k$. With the help of this equality we can rewrite (5.9) as

$$R_{u_0}[\sigma]\ge\limsup_{k\to\infty}\left[\int_\Omega f(\,\cdot\,,\nabla v_k)\,\mathrm{d}x+\int_\Omega\tau_k\cdot(\nabla u_0-\nabla v_k)\,\mathrm{d}x\right],$$

and the claim follows via the second part of (5.8) . □

[14] Actually, [16, Theorem 3.20] is not directly formulated for case of weak-$*$-convergence in L^∞, but it implies the required statement (compare [16, Remark 3.25]). This follows easily from the fact that weak-$*$-convergence in L^∞ comprises — at least on subsets of finite measure — weak convergence in L^p for all $p<\infty$.

We remark that, in Lemma 5.4, neither $v_k \in \mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$ nor $\operatorname{div}\tau_k \equiv 0$ is assumed, and thus the functions v_k and τ_k need not be admissible competitors in (P) and (P*), respectively. Nevertheless, when applying the lemma in the following, we will utilize Proposition 5.2 to choose at least the v_k admissible. In this way we now provide a

Proof of Theorem 1.1. By (Lin), $\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F = \infty$ cannot happen, and if we have $\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F = -\infty$, the claim follows from (1.3). Thus, we now assume that the infimum is finite. From the continuity of f in z and from (Lin), we get that F is continuous in $\mathrm{W}^{1,1}(\Omega,\mathbb{R}^N)$, and thus we can find a sequence $(w_k)_{k\in\mathbb{N}}$ in the dense subset $u_0|_\Omega + \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega,\mathbb{R}^N)$ of $\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$ such that we have $\lim_{k\to\infty} F[w_k] = \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F$. Since each $w_k - u_0$ vanishes near the boundary of Ω, we can moreover choose an increasing sequence $(G_k)_{k\in\mathbb{N}}$ of bounded open subsets of Ω with $\bigcup_{k=1}^\infty G_k = \Omega$ and such that $w_k = u_0$ holds on $\Omega \setminus G_k$. In addition, we take a null sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ in $(0,\infty)$ with $F[w_k] \le \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F + \varepsilon_k$, and we set

$$
\begin{aligned}
\chi_k &:= 1 + \|\nabla w_k - \nabla u_0\|_{2;\Omega}\,,\\
\mathrm{W}^{1,2}_{0;G_k}(\Omega,\mathbb{R}^N) &:= \{\varphi \in \mathrm{W}^{1,2}_0(\Omega,\mathbb{R}^N) \,:\, \varphi \equiv 0 \text{ on } \Omega\setminus G_k\}\,,\\
S_k &:= \{\nabla\varphi \,:\, \varphi \in \mathrm{W}^{1,2}_{0;G_k}(\Omega,\mathbb{R}^N)\}\,.
\end{aligned}
$$

With the help of Poincaré's inequality and weak compactness, it follows that S_k is a closed subspace of $\mathrm{L}^2(\Omega,\mathbb{R}^N)$. Furthermore, we have $\|\Phi\|_{1;\Omega} \le \sqrt{\mathscr{L}^n(G_k)}\,\|\Phi\|_{2;\Omega}$ for all $\Phi \in S_k$ and

$$
\int_\Omega f(\,\cdot\,,\nabla w_k) \le \inf_{\Theta\in\nabla w_k + S_k} \int_\Omega f(\,\cdot\,,\Theta)\,\mathrm{d}x + \varepsilon_k\,.
$$

For each fixed $k \in \mathbb{N}$, we can thus apply Proposition 5.2 with the constants ε_k, χ_k, the subspace S_k, and and the L^1-function ∇w_k. Consequently, we find $\widehat{v}_k \in \nabla w_k + S_k$, which we can directly write as $\widehat{v}_k = \nabla v_k$ with $v_k \in u_0|_\Omega + \mathrm{W}^{1,2}_{0;G_k}(\Omega,\mathbb{R}^N) \subset \mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$, and $\tau_k \in \mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})$ such that

$$
\begin{aligned}
\|\nabla v_k - \nabla w_k\|_{2;\Omega} &\le \chi_k\,,\\
\tau_k(x) \in \partial_z f(x,\nabla v_k(x)) &\qquad \text{for } \mathscr{L}^n\text{-a. e. } x\in\Omega\,,\\
\int_\Omega \tau_k\cdot\nabla\varphi\,\mathrm{d}x < \frac{2\varepsilon_k}{\chi_k}\|\nabla\varphi\|_{2;\Omega} &\qquad \text{for all } \varphi \in \mathrm{W}^{1,2}_{0;G_k}(\Omega,\mathbb{R}^N)\,.
\end{aligned}
$$

As $|\partial_z f|$ is bounded on $\Omega\times\mathbb{R}^{Nn}$ via (Lin) and [22, Lemma 5.2], $(\tau_k)_{k\in\mathbb{N}}$ is a bounded sequence in $\mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})$. Possibly passing to a subsequence, we can assume that $(\tau_k)_{k\in\mathbb{N}}$ weak-$*$-converges to a limit σ in $\mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})$. Since every $\varphi \in \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega,\mathbb{R}^N)$ is for $k \gg 1$ in $\mathrm{W}^{1,2}_{0;G_k}(\Omega,\mathbb{R}^N)$ and ε_k/χ_k tends to 0, we easily infer $\sigma \in \mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^N)$. Recalling the precise choice of the χ_k we moreover have

$$
\int_\Omega \tau_k\cdot(\nabla v_k - \nabla u_0)\,\mathrm{d}x \le \frac{2\varepsilon_k}{\chi_k}\Big[\|\nabla v_k - \nabla w_k\|_{2;\Omega} + \|\nabla w_k - \nabla u_0\|_{2;\Omega}\Big] \le 4\varepsilon_k \underset{k\to\infty}{\longrightarrow} 0\,,
$$

so that all assumptions of Lemma 5.4 are available. By the latter lemma, we thus conclude

$$
R_{u_0}[\sigma] \ge \liminf_{k\to\infty} F[v_k] \ge \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F\,.
$$

Taking (1.3) into account, we see that σ solves the dual problem (P*), and that we have in fact equality in the last estimate. ☐

Similar in spirit and also based on Proposition 5.2, but choosing as the starting sequence $(w_k)_{k\in\mathbb{N}}$ the refined approximations of Lemma 4.3, we next establish a part of Theorem 2.1.

Proof that (2.2)*,* (2.3)*,* (2.4) *hold for all extremals u and σ, in the situation of Theorem 2.1.* Possibly passing from Ψ to $\max\{\Psi, \Phi|_{\Omega}\}$ with some positive and continuous $\Phi \in \mathrm{L}^1(\mathbb{R}^n)$, we assume in the following that the function $\Psi \in \mathrm{L}^1(\Omega)$ in assumption (Lin) is positive and bounded away from 0 on bounded subsets of Ω. Starting from a given generalized minimizer u for F in $\mathrm{W}^{1,1}_{u_0}(\Omega, \mathbb{R}^N)$ we then work with the sequence $(w_k)_{k\in\mathbb{N}}$ of Lemma 4.3, and in view of the assumption just made we can also apply Theorem 4.1. From this theorem and the minimizing property of u we deduce

$$\lim_{k\to\infty} F[w_k] = \lim_{k\to\infty} \int_\Omega f(\nabla w_k)\,\mathrm{d}x = \int_\Omega f(\nabla u)\,\mathrm{d}x + \int_{\overline{\Omega}} f^\infty\Big(\frac{\mathrm{d}\mathrm{D}^{\mathrm{s}}u}{\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|}\Big)\,\mathrm{d}|\mathrm{D}^{\mathrm{s}}u| = \overline{F}_{u_0}[u] \le \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F\,.$$

Now we take a null sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ in $(0,\infty)$ with $F[w_k] \le \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F + \varepsilon_k$, and we also fix a Borel representative of a solution $\sigma \in \mathrm{L}^\infty_{\mathrm{div}}(\Omega, \mathbb{R}^N)$ of the dual problem (P*). Setting $f_\sigma(x,z) := f(x,z) - \sigma(x)z$, we observe that f_σ is Borel measurable, convex in z, and satisfies (Lin) (possibly with $L+\|\sigma\|_{\infty;\Omega}$ in place of L). Using $\operatorname{div}\sigma \equiv 0$ in the first step and the duality formula of Theorem 1.1 in the last one, we get

$$\begin{aligned}\int_\Omega f_\sigma(\,\cdot\,, \nabla w_k)\,\mathrm{d}x &= F[w_k] - \int_\Omega \sigma\cdot\nabla u_0\,\mathrm{d}x\\ &\le \varepsilon_k + \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F - \int_\Omega \sigma\cdot\nabla u_0\,\mathrm{d}x = \varepsilon_k - \int_\Omega f^*(\,\cdot\,,\sigma)\,\mathrm{d}x\,. \end{aligned} \tag{5.10}$$

Moreover, by the definition of the conjugate function we have

$$-f^*(x,\sigma(x)) \le f(x,z) - \sigma(x)\cdot z = f_\sigma(x,z) \qquad \text{for all } (x,z) \in \Omega\times\mathbb{R}^{Nn}$$

so that in fact there holds

$$-\int_\Omega f^*(\,\cdot\,,\sigma)\,\mathrm{d}x \le \int_\Omega f_\sigma(\,\cdot\,,\Theta)\,\mathrm{d}x \qquad \text{for all } \Theta \in \mathrm{L}^1(\Omega,\mathbb{R}^N)\,. \tag{5.11}$$

Next, we choose an increasing sequence $(G_k)_{k\in\mathbb{N}}$ of bounded open subsets of Ω with $\bigcup_{k=1}^\infty G_k = \Omega$. When we introduce the closed subspace

$$\mathrm{L}^2_{G_k}(\Omega,\mathbb{R}^{Nn}) := \{\Phi \in \mathrm{L}^2(\Omega)\,:\,\Phi \equiv 0 \text{ on } \Omega\setminus G_k\}$$

of $\mathrm{L}^2(\Omega,\mathbb{R}^{Nn})$, we can combine (5.10) and (5.11) to obtain in particular

$$\int_\Omega f_\sigma(\,\cdot\,,\nabla w_k)\,\mathrm{d}x \le \inf_{\Theta\in\nabla w_k + \mathrm{L}^2_{G_k}(\Omega,\mathbb{R}^{Nn})} \int_\Omega f_\sigma(\,\cdot\,,\Theta)\,\mathrm{d}x + \varepsilon_k\,.$$

Similar to the above proof of Theorem 1.1, we apply Proposition 5.2 — but this time with f_σ in place of f and $\sqrt{\varepsilon_k}$ in place of χ — to find for each $k\in\mathbb{N}$ some $\widehat{v}_k \in \nabla w_k + \mathrm{L}^2_{G_k}(\Omega,\mathbb{R}^N)$ and $\tau_k \in \mathrm{L}^\infty(\Omega,\mathbb{R}^{Nn})$ with

$$\|\widehat{v}_k - \nabla w_k\|_{2;\Omega} \le \sqrt{\varepsilon_k}\,, \tag{5.12}$$

$$\tau_k(x) \in \partial_z f_\sigma(x,\widehat{v}_k(x)) \qquad \text{for } \mathscr{L}^n\text{-a. e. } x\in\Omega\,, \tag{5.13}$$

$$\int_\Omega \tau_k\cdot\Phi\,\mathrm{d}x < 2\sqrt{\varepsilon_k}\,\|\Phi\|_{2;\Omega} \qquad \text{for all } \Phi\in\mathrm{L}^2_{G_k}(\Omega,\mathbb{R}^N)\,. \tag{5.14}$$

From (5.14) we conclude that $(\tau_k)_{k\in\mathbb{N}}$ converges to 0 in $\mathrm{L}^2_{\mathrm{loc}}(\Omega,\mathbb{R}^N)$, and by (5.12) $(\widehat{v}_k - \nabla w_k)_{k\in\mathbb{N}}$ converges to 0 in $\mathrm{L}^2(\Omega,\mathbb{R}^N)$. Passing to a subsequence we can assume that these convergences hold also $\mathscr{L}^n$-a. e. on Ω, and — taking into account the extra information of (4.5) — it follows that $(\widehat{v}_k)_{k\in\mathbb{N}}$ converges $\mathscr{L}^n$-a. e. to ∇u. Using these convergences and (5.13), and applying Lemma 3.4 pointwisely, we infer

$$0 \in \partial_z f_\sigma(x,\nabla u(x)) \qquad \text{for } \mathscr{L}^n\text{-a. e. } x\in\Omega\,.$$

Recalling the definition of f_σ, we have $\partial_z f_\sigma(x,\nabla u(x)) = \partial_z f(x,\nabla u(x)) - \sigma(x)$, and hence we finally arrive at (2.3). With the help of (3.3) and the equality $f^{**} = f$, we see that (2.2) and (2.4) hold as well. □

6 Proofs of the regularity and uniqueness statements

In this section, we start with an elementary bound for the subdifferential, and then we proceed with the proofs of Theorem 2.3 and Corollary 2.4, which essentially follow previous arguments of [37, 13].

Lemma 6.1. *Consider $h\colon \mathbb{R}^m \to \mathbb{R} \cup \{\infty, -\infty\}$ and suppose that $|h|$ is bounded by a constant Λ on a subset A of $\mathbb{R}^m$. Then, for all interior points z of A, we have*

$$|\partial h(z)| \le \frac{2\Lambda}{\operatorname{dist}(z, \mathbb{R}^m \setminus A)}\,.$$

Proof. If the lemma were false, for some interior point z of A and some $z^* \in \partial h(z)$ we would have $|z^*| > 2\Lambda/\operatorname{dist}(z, \mathbb{R}^m \setminus A)$. We could then choose a $\xi \in \mathbb{R}^m$ which points in the same direction as z^* and satisfies $2\Lambda/|z^*| < |\xi| < \operatorname{dist}(z, \mathbb{R}^m \setminus A)$. Consequently, we would get $z{+}\xi \in A$ and

$$h(z{+}\xi) - h(z) \ge z^* \cdot \xi = |z^*||\xi| > 2\Lambda\,,$$

which clearly contradicts our assumptions. □

Proof of Theorem 2.3. We consider an arbitrary generalized minimizer $u \in \mathrm{BV}(\Omega, \mathbb{R}^N)$ of (P) and the approximations $w_k \in u_0|_\Omega + \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega, \mathbb{R}^N)$ of Lemma 4.3, for which in particular $(\mathrm{D}w_k)_{k\in\mathbb{N}}$ weak-$*$-converges to $\mathrm{D}u$ in the sense of measures. Exactly as in the proof of Theorem 2.1 we use Theorem 4.1 to deduce $\lim_{k\to\infty} F[w_k] \le \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F$, and by Theorem 1.1 we conclude

$$\lim_{k\to\infty} F[w_k] \le \int_\Omega \sigma \cdot \nabla u_0 \,\mathrm{d}x - \int_\Omega f^*(\,\cdot\,, \sigma)\,\mathrm{d}x \tag{6.1}$$

for the fixed dual solution σ of Theorem 2.3. Now we consider $t > 0$ and a test function $\varphi \in \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega, \mathbb{R}^{Nn})$. Then, exploiting the definition of f^* and recalling $\operatorname{div} \sigma \equiv 0$, we infer

$$\begin{aligned} F[w_k] &\ge \int_\Omega (\sigma + t\varphi) \cdot \nabla w_k \,\mathrm{d}x - \int_\Omega f^*(\,\cdot\,, \sigma + t\varphi)\,\mathrm{d}x \\ &= \int_\Omega \sigma \cdot \nabla u_0 \,\mathrm{d}x + t \int_\Omega \varphi \cdot \nabla w_k \,\mathrm{d}x - \int_\Omega f^*(\,\cdot\,, \sigma + t\varphi)\,\mathrm{d}x\,. \end{aligned}$$

Using (6.1), the weak-$*$-convergence of the $\mathrm{D}w_k$, and the continuity of φ, we pass to the limit $k \to \infty$ in the last inequality, and we find

$$\int_\Omega \sigma \cdot \nabla u_0 \,\mathrm{d}x - \int_\Omega f^*(\,\cdot\,, \sigma)\,\mathrm{d}x \ge \int_\Omega \sigma \cdot \nabla u_0 \,\mathrm{d}x + t \int_\Omega \varphi \cdot \mathrm{d}\mathrm{D}u - \int_\Omega f^*(\,\cdot\,, \sigma + t\varphi)\,\mathrm{d}x\,.$$

The integrals involving u_0 cancel out, and we arrive at

$$\int_\Omega \varphi \cdot \mathrm{d}\mathrm{D}u \le \int_\Omega \frac{f^*(\,\cdot\,, \sigma + t\varphi) - f^*(\,\cdot\,, \sigma)}{t}\,\mathrm{d}x\,. \tag{6.2}$$

Reasoning for $\mathscr{L}^n$-a. e. $x \in \Omega$, we next notice that by (Lin) and (2.6) we have $-\Psi(x) \le f^*(x, \,\cdot\,) \le \Lambda(x)$ on $\mathrm{B}_{\varepsilon(x)}(\sigma(x))$, so that by Lemma 6.1 we get $|\partial_{z^*} f^*(x, \,\cdot\,)| \le 4\max\{\Lambda(x), \Psi(x)\}/\varepsilon(x)$ on $\mathrm{B}_{\varepsilon(x)/2}(\sigma(x))$. For every fixed $\varphi \in \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega, \mathbb{R}^{Nn})$, we know that ε is bounded from below on $\operatorname{spt}\varphi$, and thus we can take $t > 0$ sufficiently small that $\sigma(x) + t\varphi(x)$ is contained in $\mathrm{B}_{\varepsilon(x)/2}(\sigma(x))$ for all $x \in \Omega$. For such t we have the control

$$\frac{f^*(\,\cdot\,, \sigma + t\varphi) - f^*(\,\cdot\,, \sigma)}{t} \le |\partial_{z^*} f^*(\,\cdot\,, \sigma + t\varphi)|\,|\varphi| \le \frac{4\max\{\Lambda, \Psi\}}{\varepsilon}|\varphi| \qquad \mathscr{L}^n\text{-a. e. on } \Omega \tag{6.3}$$

in terms of the function $\max\{\Lambda, \Psi\}/\varepsilon \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. When we combine the estimates (6.2) and (6.3) and exploit the arbitrariness of φ, we find

$$|\mathrm{D}u| \le \frac{4\max\{\Lambda, \Psi\}}{\varepsilon}\mathscr{L}^n\,.$$

As $u \in \mathrm{BV}(\Omega, \mathbb{R}^N)$ already means $u \in \mathrm{L}^1(\Omega, \mathbb{R}^N)$ and $|\mathrm{D}u|(\Omega) < \infty$, we have thus shown $u \in \mathrm{W}^{1,1}(\Omega, \mathbb{R}^N)$. $\square$

Proof of Corollary 2.4. We first argue that (2.6) is valid: By Lemmas 3.9 and 3.10, the strict convexity assumption on f implies that f^* is continuous on the open subset $\mathrm{Im}(x, \partial_z f)$ of $\Omega \times \mathbb{R}^{Nn}$. Moreover, by the continuity of σ and by Theorem 2.1, for every compact subset K of Ω, the graph Σ_K of $\sigma|_K$ is a compact subset of $\mathrm{Im}(x, \partial_z f)$. Consequently, for some $\varepsilon_K > 0$ the ε_K-neighborhood $\mathcal{N}_{\varepsilon_K}(\Sigma_K)$ of Σ_K is still relatively compact in $\mathrm{Im}(x, \partial_z f)$, and $|f^*|$ is bounded on $\mathcal{N}_{\varepsilon_K}(\Sigma_K)$ by some constant Λ_K. As K is an arbitrary compact subset of Ω, it is now easy to obtain (2.6) with continuous positive functions ε and Λ. Consequently, the conclusions of Theorem 2.3 are available, and every generalized minimizer u of (P) is in $\mathrm{W}^{1,1}(\Omega, \mathbb{R}^N)$. Taking Proposition 3.8 into account, the strict convexity also gives that f^* is C^1 in z^* on $\mathrm{Im}(x, \partial_z f)$ with single-valued subdifferential $\partial_{z^*} f^*(x, z^*) = \{\nabla_{z^*} f^*(x, z^*)\}$, and Lemma 6.1 yields the bound $|\nabla_{z^*} f^*| \le 4\Lambda_K/\varepsilon_K$ on $\mathcal{N}_{\varepsilon_K/2}(\Sigma_K)$. Therefore, (2.4) determines ∇u and bounds it in $\mathrm{L}^\infty_{\mathrm{loc}}(\Omega, \mathbb{R}^{Nn})$. $\square$

7 A pairing of $\boldsymbol{\sigma}$ and $\mathbf{D}\boldsymbol{u}$ and the extremality relation for $\mathbf{D}^{\mathrm{s}}\boldsymbol{u}$

In this section we prove the full statement of Theorem 2.1. To this end, we follow ideas of Anzellotti [2] (compare also [14, 26]), and we introduce, for $u \in \mathrm{BV}_{u_0}(\overline{\Omega}, \mathbb{R}^N)$ and $\sigma \in \mathrm{L}^\infty_{\mathrm{div}}(\Omega, \mathbb{R}^{Nn})$, a pairing of the gradient measure $\mathrm{D}u$ and the possibly discontinuous function σ. Indeed, imposing the assumption (Per) on Ω, we use the approximations from Lemma 4.3 and integration by parts in order to handle an up-to-the-boundary version of Anzellotti's pairing. In the first place, these tools allow to show continuity of the linear functional

$$\mathrm{C}^\infty_{\mathrm{cpt}}(\mathbb{R}^n) \to \mathbb{R},\ \varphi \mapsto \int_\Omega \varphi\sigma \cdot \nabla u_0 \,\mathrm{d}x - \int_\Omega \sigma \cdot ((u{-}u_0) \otimes \nabla\varphi) \,\mathrm{d}x$$

(and its extension to $\mathrm{C}^0_{\mathrm{cpt}}(\mathbb{R}^n)$) in the sup-norm. In view of the Riesz representation theorem for continuous linear functionals on $\mathrm{C}^0_{\mathrm{cpt}}$, we can then give the following variant of [2, Definition 1.4, Theorem 1.5].

Definition 7.1 (up-to-the-boundary pairing of $\mathrm{D}u$ and σ)**.** *Suppose that Ω satisfies* (Per). *For every $u \in \mathrm{BV}_{u_0}(\overline{\Omega}, \mathbb{R}^N)$ and $\sigma \in \mathrm{L}^\infty_{\mathrm{div}}(\Omega, \mathbb{R}^{Nn})$ we define $[\![\sigma \cdot \mathrm{D}u]\!]$ as the uniquely determined signed Radon measure on $\overline{\Omega}$ such that*

$$\int_{\overline{\Omega}} \varphi \,\mathrm{d}[\![\sigma \cdot \mathrm{D}u]\!] = \int_\Omega \varphi\sigma \cdot \nabla u_0 \,\mathrm{d}x - \int_\Omega \sigma \cdot ((u{-}u_0) \otimes \nabla\varphi) \,\mathrm{d}x \qquad \textit{holds for all } \varphi \in \mathrm{C}^\infty_{\mathrm{cpt}}(\mathbb{R}^n)\,.$$

We stress that the up-to-the boundary feature in this definition lies in the fact that only $\varphi \in \mathrm{C}^\infty_{\mathrm{cpt}}(\mathbb{R}^n)$, but *not* $\mathrm{spt}\,\varphi \subset \Omega$, is required; as a result, Definition 7.1 incorporates the possible deviation of u from the boundary values prescribed by u_0, or, in other words, it takes into account the measure $\mathrm{D}u \llcorner \partial\Omega$.

However, when $\mathrm{spt}\,\varphi \subset \Omega$ holds, then integration by parts and standard approximation of u_0 give $\int_{\overline{\Omega}} \varphi \,\mathrm{d}[\![\sigma \cdot \mathrm{D}u]\!] = -\int_\Omega \sigma \cdot (u \otimes \nabla\varphi) \,\mathrm{d}x$ for all $\varphi \in \mathrm{C}^\infty_{\mathrm{cpt}}(\Omega)$, so that our pairing $[\![\sigma \cdot \mathrm{D}u]\!]$ coincides on Ω with Anzellotti's original one. Therefore, from [2, Theorem 2.4] we can deduce the representation

$$[\![\sigma \cdot \mathrm{D}u]\!]^{\mathrm{a}} = (\sigma \cdot \nabla u)\mathscr{L}^n \tag{7.1}$$

of the absolutely continuous part of $[\![\sigma \cdot \mathrm{D}u]\!]$. Approximation, based on Lemma 4.3 with $\Psi \equiv 0$, also yields (compare with [2, Theorem 1.5, Corollary 1.6])

$$|[\![\sigma \cdot \mathrm{D}u]\!]| \le \|\sigma\|_{\infty;\Omega} |\mathrm{D}u| \tag{7.2}$$

as an inequality of measures on $\overline{\Omega}$, and hence the existence of the density $\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}$ follows. In addition, the usage of test functions φ with $\varphi\equiv 1$ on large balls, gives the equality

$$[\![\sigma\cdot\mathrm{D}u]\!]\big(\overline{\Omega}\big)=\int_\Omega \sigma\cdot\nabla u_0\,\mathrm{d}x\,. \tag{7.3}$$

Finally, we have the following statements, which are crucial for our purposes:

Theorem 7.2 ($|\mathrm{D}u|$-a. e. density control on $[\![\sigma\cdot\mathrm{D}u]\!]$)**.** *Suppose that* Ω *satisfies* (Per). *Consider* $u\in\mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)$ *and* $\sigma\in\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$*, and a common Lebesgue point*[15] x_0 *of* $\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}$ *and* $\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}$ *with respect to* $|\mathrm{D}u|$ *in* $\overline{\Omega}$*. If* $\sigma\in K$ *holds* $\mathscr{L}^n$*-a. e. on a neighborhood of* x_0 *in* Ω*, for some closed convex set* K *in* $\mathbb{R}^{Nn}$*, then there exists a* $\sigma_0\in K$ *with*

$$\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}(x_0)=\sigma_0\cdot\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}(x_0)\,. \tag{7.4}$$

Corollary 7.3. *Suppose that* Ω *satisfies* (Per) *and that* (Lin) *and* (Con) *hold for* f*. Then, for all* $u\in\mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)$ *and* $\sigma\in\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$ *such that* $f^*(\,\cdot\,,\sigma)<\infty$ *holds* $\mathscr{L}^n$*-a. e. on* Ω*, we have*

$$\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}\le f^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}\Big)\qquad |\mathrm{D}u|\text{-a. e. on }\overline{\Omega}\,. \tag{7.5}$$

Before turning to the proofs of Theorem 7.2 and Corollary 7.3, let us highlight their most decisive feature: indeed, for $\mathscr{L}^n$-a. e. $x_0\in\overline{\Omega}$ with $\nabla u(x_0)\neq 0$, one can directly read off from (7.1) that

$$\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}(x_0)=\sigma(x_0)\cdot\frac{\nabla u(x_0)}{|\nabla u(x_0)|}=\sigma(x_0)\cdot\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}(x_0)$$

holds, so that the validity of (7.4) with the 'concrete' value $\sigma_0=\sigma(x_0)$ is obvious for $|\mathrm{D}^{\mathrm{a}}u|$-a. e. x_0. However, the crucial point of Theorem 7.2 — which will enable us to deal with the extremality relation (2.5) for $\mathrm{D}^{\mathrm{s}}u$ — is that it gives (7.4) not only for $|\mathrm{D}^{\mathrm{a}}u|$-a. e. x_0, but also for $|\mathrm{D}^{\mathrm{s}}u|$-a. e. x_0. We believe that it is possible to deduce the latter assertion — which is clearly more subtle, as σ cannot be evaluated $|\mathrm{D}^{\mathrm{s}}u|$-a. e. — from an adaption of Anzellotti's aureate representation formula [3, Theorem 3.6] for $\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}$ (stated for $N=1$, see also [4, Fact 1.1] and [14, Proposition 1.6]). However, the adaption to our case of an up-to-a-non-smooth-boundary pairing would require a considerable effort, and we prefer to follow a more elementary line of argument. Our approach yields a less precise information about $\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}$, which however still suffices for our purposes:

Proof of Theorem 7.2. We assume $0\in K$ (otherwise we fix some $z_0^*\in K$, and in view of $[\![(\sigma-z_0^*)\cdot\mathrm{D}u]\!]=[\![\sigma\cdot\mathrm{D}u]\!]-z_0^*\cdot\mathrm{D}u$, we can pass from K to $\{z^*-z_0^*\,:\,z^*\in K\}$ and from σ to $\sigma-z_0^*$), and we work with approximations w_k of Lemma 4.3, corresponding to an arbitrarily fixed, positive $\Psi\in\mathrm{L}^1(\Omega)$. Using the L^1-convergence of the w_k and integration by parts in Definition 7.1, we get

$$\int_{\overline{\Omega}}\varphi\,\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]=\lim_{k\to\infty}\int_\Omega\varphi\sigma\cdot\nabla w_k\,\mathrm{d}x\qquad\text{for all }\varphi\in\mathrm{C}^\infty_{\mathrm{cpt}}(\mathbb{R}^n)\,.$$

Approximating the characteristic functions of balls with the φ and keeping (7.2) in mind, this implies, in a standard way,

$$[\![\sigma\cdot\mathrm{D}u]\!]\big(\overline{\Omega}\cap\mathrm{B}_R(x_0)\big)=\lim_{k\to\infty}\int_{\Omega\cap\mathrm{B}_R(x_0)}\sigma\cdot\nabla w_k\,\mathrm{d}x\qquad\text{whenever }|\mathrm{D}u|\big(\overline{\Omega}\cap\partial\mathrm{B}_R(x_0)\big)=0\,. \tag{7.6}$$

[15] We call $x_0\in\operatorname{spt}\mu$ a Lebesgue point of a μ-measurable function $G\colon\overline{\Omega}\to\mathbb{R}^m$ with respect to a non-negative Radon measure μ in $\overline{\Omega}$ if there exists a $z_0\in\mathbb{R}^m$ with

$$\lim_{R\searrow 0}\frac{1}{\mu(\overline{\Omega}\cap\mathrm{B}_R(x_0))}\int_{\overline{\Omega}\cap\mathrm{B}_R(x_0)}|G-z_0|\,\mathrm{d}\mu=0\,.$$

For such points, the value z_0 is uniquely determined, is called the Lebesgue value of G at x_0, and is denoted by $G(x_0)$.

Here, the last requirement is fulfilled for all but countably many R, and we tacitly understand in the following that it is met by all radii in our computations. We now introduce the Lebesgue value

$$v := \frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}(x_0) \qquad \text{with } |v| = 1\,.$$

Writing $\mathrm{p}_{v^\perp} : \mathbb{R}^{Nn} \to \mathbb{R}^{Nn}$ for the orthogonal projection on the orthogonal complement of v, we then split

$$\sigma \cdot \nabla w_k = (\sigma \cdot v)(v \cdot \nabla w_k)_+ - (\sigma \cdot v)(v \cdot \nabla w_k)_- + \mathrm{p}_{v^\perp}(\sigma) \cdot \mathrm{p}_{v^\perp}(\nabla w_k)$$

and estimate the resulting terms on the right-hand side of (7.6) separately. For one term, we use Remark 4.4, with the integrand $(x,z) \mapsto |\mathrm{p}_{v^\perp}(z)|$, and the inequality $|\mathrm{p}_{v^\perp}(z)| = |\mathrm{p}_{v^\perp}(z-v)| \le |z-v|$ to get

$$\begin{aligned}
\limsup_{k\to\infty} \left| \int_{\Omega\cap \mathrm{B}_R(x_0)} \mathrm{p}_{v^\perp}(\sigma) \cdot \mathrm{p}_{v^\perp}(\nabla w_k)\,\mathrm{d}x \right| &\le \|\sigma\|_{\infty;\Omega} \lim_{k\to\infty} \int_{\Omega\cap \mathrm{B}_R(x_0)} |\mathrm{p}_{v^\perp}(\nabla w_k)|\,\mathrm{d}x \\
&= \|\sigma\|_{\infty;\Omega} \int_{\overline{\Omega}\cap \mathrm{B}_R(x_0)} \left|\mathrm{p}_{v^\perp}\Big(\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}\Big)\right| \mathrm{d}|\mathrm{D}u| \qquad (7.7)\\
&\le \|\sigma\|_{\infty;\Omega} \int_{\overline{\Omega}\cap \mathrm{B}_R(x_0)} \left|\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|} - v\right| \mathrm{d}|\mathrm{D}u|\,.
\end{aligned}$$

Arguing analogously with $(x,z) \mapsto (v \cdot z)_-$ and $(v\cdot z)_- \le |z-v|$, we also get

$$\begin{aligned}
\limsup_{k\to\infty} \left| \int_{\Omega\cap \mathrm{B}_R(x_0)} (\sigma\cdot v)(v \cdot \nabla w_k)_-\,\mathrm{d}x \right| &\le \|\sigma\|_{\infty;\Omega} \lim_{k\to\infty} \int_{\Omega\cap \mathrm{B}_R(x_0)} (v\cdot\nabla w_k)_-\,\mathrm{d}x \\
&= \|\sigma\|_{\infty;\Omega} \int_{\overline{\Omega}\cap \mathrm{B}_R(x_0)} \Big(v \cdot \frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}\Big)_- \mathrm{d}|\mathrm{D}u| \qquad (7.8)\\
&\le \|\sigma\|_{\infty;\Omega} \int_{\overline{\Omega}\cap \mathrm{B}_R(x_0)} \left|\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|} - v\right| \mathrm{d}|\mathrm{D}u|\,.
\end{aligned}$$

In order to treat the remaining term, we set $M := \max\{z^* \cdot v \,:\, z^* \in K\,, |z^*| \le \|\sigma\|_{\infty;\Omega}\}$, and we get $0 \le M \le \|\sigma\|_{\infty;\Omega}$ (as we supposed $0 \in K$). Now we take R sufficiently small that inclusion $\sigma \in K$ holds in $\Omega \cap \mathrm{B}_R(x_0)$. Then, Remark 4.4, applied with $(x,z) \mapsto (v\cdot z)_+$, and the inequality $(v\cdot z)_+ = v\cdot z + (v\cdot z)_- \le v \cdot z + |z-v|$ give

$$\begin{aligned}
\limsup_{k\to\infty} \int_{\Omega\cap \mathrm{B}_R(x_0)} (\sigma\cdot v)(v \cdot \nabla w_k)_+\,\mathrm{d}x &\le M \lim_{k\to\infty} \int_{\Omega\cap \mathrm{B}_R(x_0)} (v\cdot\nabla w_k)_+\,\mathrm{d}x \\
&= M \int_{\overline{\Omega}\cap \mathrm{B}_R(x_0)} \Big(v \cdot \frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}\Big)_+ \mathrm{d}|\mathrm{D}u| \qquad (7.9)\\
&\le M \int_{\overline{\Omega}\cap \mathrm{B}_R(x_0)} \left[v \cdot \frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|} + \left|\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|} - v\right|\right] \mathrm{d}|\mathrm{D}u|\,.
\end{aligned}$$

Collecting the estimates (7.6), (7.7), (7.8), (7.9), we arrive at

$$[\![\sigma \cdot \mathrm{D}u]\!]\big(\overline{\Omega} \cap \mathrm{B}_R(x_0)\big) \le Mv \cdot \mathrm{D}u\big(\overline{\Omega}\cap \mathrm{B}_R(x_0)\big) + 3\|\sigma\|_{\infty;\Omega} \int_{\overline{\Omega}\cap \mathrm{B}_R(x_0)} \left|\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|} - v\right| \mathrm{d}|\mathrm{D}u|\,.$$

Now we divide on both sides by $|\mathrm{D}u|\big(\overline{\Omega}\cap \mathrm{B}_R(x_0)\big)$ and take the limit for $R \searrow 0$. Recalling that x_0 is a Lebesgue point of $\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}$ with Lebesgue value v and also a Lebesgue point of $\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}$ (in particular $x_0 \in \operatorname{spt}|\mathrm{D}u|$, so that for $0 < R \ll 1$ we are not dividing by 0), we obtain

$$\frac{\mathrm{d}[\![\sigma \cdot \mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}(x_0) \le Mv \cdot v\,.$$

Recalling $|v|=1$ and the choice of M, this implies

$$\frac{\mathrm{d}[\![\sigma\cdot \mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}(x_0)=\sigma_M\cdot v$$

for some $\sigma_M\in K$. Using, as a substitute for (7.9), a very similar estimate from below, we can also find a $\sigma_m\in K$ with

$$\frac{\mathrm{d}[\![\sigma\cdot \mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}(x_0)\geq\sigma_m\cdot v\,,$$

and, together, the two last inequalities show that (7.4) holds, when we take $\sigma_0\in K$ as a suitable convex combination of σ_M and σ_m. $\square$

In order to deduce the statement of Corollary 7.3, the following simple continuity lemma will be useful to cope with the x-dependence of the integrand f.

Lemma 7.4. *Suppose that $g\colon \overline{\Omega}\times\mathbb{R}^m\to\mathbb{R}$ is continuous. Then, for every $(x_0,z_0)\in\overline{\Omega}\times\mathbb{R}^m$ and every $\varepsilon>0$, there exists a $\delta>0$ such that we have*

$$\partial_z g(x,z_0)\subset\mathcal{N}_\varepsilon(\partial_z g(x_0,z_0))\qquad\text{for all } x\in\overline{\Omega} \text{ with } |x{-}x_0|<\delta\,.$$

Here, we used $\mathcal{N}_\varepsilon(\,\cdot\,)$ for the ε-neighborhood of a set.

Proof. We may assume $z_0=0$. In order to prove the lemma by contradiction, we now suppose that the claim fails for some $x_0\in\overline{\Omega}$ and some $\varepsilon>0$. Then we can find a sequence $(x_k)_{k\in\mathbb{N}}$ in Ω, converging to x, and a sequence $(z_k^*)_{k\in\mathbb{N}}$ in $\mathbb{R}^m$ such that $z_k^*\in\partial_z g(x_k,0)$ and $\operatorname{dist}(z_k^*,\partial_z g(x_0,0))\geq\varepsilon$ hold for all $k\in\mathbb{N}$. As $g(x_k,0)$ and $g(x_k,z_k^*/|z_k^*|)$ remain bounded for $k\to\infty$, the estimate

$$g(x_k,0)+|z_k^*|=g(x_k,0)+z_k^*\cdot z_k^*/|z_k^*|\leq g(x_k,z_k^*/|z_k^*|)$$

gives boundedness of $(z_k^*)_{k\in\mathbb{N}}$, hence a subsequence $(z_{k_l}^*)_{l\in\mathbb{N}}$ converges to a limit $z_0^*\in\mathbb{R}^m$. For the limit, we have on the one hand $\operatorname{dist}(z_0^*,\partial_z g(x_0,0))\geq\varepsilon$, while on the other hand we infer

$$g(x_0,z)=\lim_{l\to\infty}g(x_{k_l},z)\geq\lim_{l\to\infty}g(x_{k_l},0)+z_{k_l}^*\cdot z=g(x_0,0)+z_0^*\cdot z$$

for all $z\in\mathbb{R}^m$, so that we get $z_0^*\in\partial_z g(x_0,0)$. This contradiction ends the proof of the lemma. $\square$

Proof of Corollary 7.3. It suffices to show that (7.5) holds at every common Lebesgue point x_0 of $\frac{\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}$ and $\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}$ with respect to $|\mathrm{D}u|$ in $\overline{\Omega}$. To see this, we first record that, by the definition of the conjugate function, we have $tf(x,z/t)\geq\sigma(x)\cdot z-tf^*(x,\sigma(x))$ for all $(x,z)\in\Omega\times\mathbb{R}^{Nn}$ and all $t>0$. Sending t to 0 and recalling that in view of (Con) the lower limit in (3.1) is in fact a limit, we infer

$$f^\infty(x,z)\geq\sigma(x)\cdot z\qquad\text{for all } z\in\mathbb{R}^{Nn}$$

whenever $x\in\Omega$ such that $f^*(x,\sigma(x))$ is finite. By assumption, the last finiteness requirement is available, and we thus have $\sigma(x)\in\partial_z f^\infty(x,0)$, for $\mathscr{L}^n$-a. e. $x\in\Omega$. For an arbitrary $\varepsilon>0$, we now apply Lemma 7.4 to f^∞ (which under (Lin) and (Con) is jointly continuous in (x,z)), and we infer that $\sigma(x)\in\mathcal{N}_\varepsilon(\partial_z f^\infty(x_0,0))$ holds for $\mathscr{L}^n$-a. e. x in a neighborhood of x_0 in Ω. At this stage we employ Theorem 7.2 with the closure of the convex set $\mathcal{N}_\varepsilon(\partial_z f^\infty(x_0,0))$ in place of K, and we infer

$$\frac{\mathrm{d}[\![\sigma\cdot \mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}(x_0)=\sigma_0\cdot\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}(x_0)\qquad\text{for some } \sigma_0\in\overline{\mathcal{N}_\varepsilon(\partial_z f^\infty(x_0,0))}\,.$$

As a consequence, we can find a subgradient $\sigma_*\in\partial_z f^\infty(x_0,0)$ with $|\sigma_*{-}\sigma_0|\leq\varepsilon$, and we get

$$\frac{\mathrm{d}[\![\sigma\cdot \mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}(x_0)=\sigma_0\cdot\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}(x_0)\leq\sigma_*\cdot\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}(x_0)+\varepsilon\Big|\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}(x_0)\Big|\leq f_\infty\Big(x_0,\frac{\mathrm{d}\mathrm{D}u}{\mathrm{d}|\mathrm{D}u|}(x_0)\Big)+\varepsilon\,.$$

Sending ε to 0, the proof is complete. $\square$

Building on Theorem 1.1 and Corollary 7.3, we can provide a short

Proof of Theorem 2.1. As in the proof of Corollary 1.2, for all $u \in \mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)$ and $\sigma \in \mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})$, it follows from the definition of the conjugate function that

$$f(\,\cdot\,,\nabla u) \geq \sigma\cdot\nabla u - f^*(\,\cdot\,,\sigma) \qquad \text{holds } \mathscr{L}^n\text{-a. e. on } \Omega\,. \tag{7.10}$$

Turning to the singular part, we first record that $\frac{\mathrm{dD^s}u}{\mathrm{d}|\mathrm{D^s}u|}=\frac{\mathrm{dD}u}{\mathrm{d}|\mathrm{D}u|}$ holds $|\mathrm{D^s}u|$-a. e. on $\overline{\Omega}$. In addition, (7.2) implies $|[\![\sigma\cdot\mathrm{D}u]\!]^{\mathrm{s}}| \leq \|\sigma\|_{\infty;\Omega}|\mathrm{D^s}u|$, and therefore the density $\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]^{\mathrm{s}}}{\mathrm{d}|\mathrm{D^s}u|}$ is well-defined and $|\mathrm{D^s}u|$-a. e. equal to $\frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|}$. With these observations at hand, Corollary 7.3 shows that

$$f^\infty\Big(\,\cdot\,,\frac{\mathrm{dD^s}u}{\mathrm{d}|\mathrm{D^s}u|}\Big) \geq \frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]^{\mathrm{s}}}{\mathrm{d}|\mathrm{D^s}u|} \qquad \text{holds } |\mathrm{D^s}u|\text{-a. e. on } \Omega \tag{7.11}$$

whenever $f^*(\,\cdot\,,\sigma)<\infty$ is valid $\mathscr{L}^n$-a. e. on Ω.

After these initial remarks we now proceed with the proof of the claimed equivalence. First, from Theorem 1.1, (3.2), and the definition of R_{u_0}, we infer that extremality of u and σ means nothing but

$$\overline{F}_{u_0}[u] = \int_\Omega \big[\sigma\cdot\nabla u_0 - f^*(\,\cdot\,,\sigma)\big]\,\mathrm{d}x\,.$$

When we write out the left-hand side and make use of (7.3) and (7.1) on the right-hand side, this equality becomes

$$\int_\Omega f(\,\cdot\,,\nabla u)\,\mathrm{d}x + \int_{\overline{\Omega}} f^\infty\Big(\,\cdot\,,\frac{\mathrm{dD^s}u}{\mathrm{d}|\mathrm{D^s}u|}\Big)\,\mathrm{d}|\mathrm{D^s}u| = \int_\Omega \big[\sigma\cdot\nabla u - f^*(\,\cdot\,,\sigma)\big]\,\mathrm{d}x + [\![\sigma\cdot\mathrm{D}u]\!]^{\mathrm{s}}\big(\overline{\Omega}\big)\,. \tag{7.12}$$

As we have the pointwise estimates (7.10) and (7.11) for the integrands, (7.12) holds if and only if equality occurs in these estimates, or, in other words, if and only if (2.2) and (2.5) hold (where we have also exploited that (7.12) implies the finiteness condition on $f^*(\,\cdot\,,\sigma)$ which is needed for (7.11)). Hence, we have shown that extremality of u and σ is equivalent to the combination of (2.2) and (2.5). In view of (3.3) and $f^{**}=f$, we can also use (2.3) or (2.4) as a substitute for (2.2), and the proof is complete. $\square$

Remark 7.5 (extremality relations in the 1-homogeneous case)**.** *If, in the situation of Theorem 2.1, $f(x,\,\cdot\,)\colon \mathbb{R}^{Nn}\to\mathbb{R}$ is positively 1-homogeneous for $\mathscr{L}^n$-a. e. $x\in\Omega$, then the extremality relations can be restated in the alternative form*

$$f\Big(\,\cdot\,,\frac{\mathrm{dD}u}{\mathrm{d}|\mathrm{D}u|}\Big) = \frac{[\![\sigma\cdot\mathrm{D}u]\!]}{\mathrm{d}|\mathrm{D}u|} \qquad |\mathrm{D}u|\text{-a. e. on } \overline{\Omega} \qquad \text{and} \qquad f^*(\,\cdot\,,\sigma)\equiv 0 \qquad \mathscr{L}^n\text{-a. e. on } \Omega\,.$$

This follows from (7.1) via the observations that f^∞ equals f and that f^* takes only the values 0 and ∞. We also refer to [15] a further analysis of the 1-homogeneous case.

A Relaxation and non-convex problems

In this section, we restrict ourselves to bounded Ω and Ψ (so that we can quote suitable auxiliary results from the literature), and we point out that a weakening of the convexity assumptions on f is possible in Theorem 1.1, in Theorem 2.1 and consequently in Corollary 2.2, and in Theorem 2.3 (while the strict convexity in Corollaries 2.4 and 2.5 seems inevitable). It should however be noted that, under these weaker assumptions, no general existence results for (P) can be expected; hence, the practicability of the following general results is in fact limited to more specific situations.

To describe the new set of assumptions, we utilize quasiconvex functions in the sense of [16, Definition 5.1 (ii)], and we recall that the quasiconvex envelope $\mathrm{Q}f\colon \Omega\times\mathbb{R}^{Nn}\to[-\infty,\infty)$ of f (with respect to the z-variable) is defined at $(x,z)\in\Omega\times\mathbb{R}^{Nn}$ by

$$\mathrm{Q}f(x,z) := \sup\big\{g(z) \,:\, g\colon \mathbb{R}^{Nn}\to\mathbb{R} \text{ is quasiconvex with } g\le f(x,\,\cdot\,) \text{ on } \mathbb{R}^{Nn}\big\}$$

(with the usual convention $\sup\emptyset=-\infty$). Furthermore, for a Carathéodory[16] function $f\colon \Omega\times\mathbb{R}^{Nn}\to\mathbb{R}$ with

$$-L(1+|z|)\le \mathrm{Q}f(x,z)\le f(x,z)\le L(1+|z|)\,, \tag{A.1}$$

also $\mathrm{Q}f$ has the Carathéodory property, and setting $\mathrm{Q}F[w]:=\int_\Omega \mathrm{Q}f(\,\cdot\,,\nabla w)\,\mathrm{d}x$ we will rely one the well-known relaxation formula

$$\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} \mathrm{Q}F = \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F\,, \tag{A.2}$$

which can be inferred, for instance, from [16, Proposition 9.5, Theorem 9.8]. If we now require, as the decisive hypothesis of this appendix, that $\mathrm{Q}f$ is convex in the z-variable, then we can apply the preceding results with $\mathrm{Q}f$ in place of f, and — as will be clarified in the following — in view of (A.2) we can hope to come up with the same conclusions. Here, the convexity assumption on $\mathrm{Q}f$ is equivalent to the equality $\mathrm{Q}f=f^{**}$ and is much weaker than the analogous assumption for f itself: indeed, convexity of $\mathrm{Q}f$ in z holds generally true for a large class of rotationally symmetric integrands [16, Theorem 6.30], and most importantly it is *tautologically satisfied in the cases $N=1$ and $n=1$*, where quasiconvexity reduces to convexity. Thus, in the following we accept the convexity requirement for $\mathrm{Q}f$ as a reasonable hypothesis in order to state:

Corollary A.1. *For bounded Ω, the conclusions of Theorem 1.1 remains true if we solely impose the hypotheses that f is a Carathéodory function with* (A.1) *and that $\mathrm{Q}f(x,\,\cdot\,)$ is convex for $\mathscr{L}^n$-a. e. $x\in\Omega$.*

Corollary A.2. *For bounded Ω, the forward implication of Theorem 2.1 and all assertions of Theorems 2.3 remain true if we solely impose the following conditions on the integrand: f is a Carathéodory function with* (A.1)*, $\mathrm{Q}f$ is convex for $\mathscr{L}^n$-a. e. $x\in\Omega$, and* (Con) *holds for $\mathrm{Q}f$ in place of f.*

The remaining deficit in these statements lies in the fact that further hypotheses — most notably the validity of (Con) for $\mathrm{Q}f$ and less severely the requirement (A.1) — are formulated in terms of $\mathrm{Q}f$ rather than f. While in general it does not seem easy to overcome this point and to provide good criteria in terms of f itself, we stress that the problem automatically disappears in the case of an x-independent integrand f: indeed, when we assume convexity of $\mathrm{Q}f$ and the growth condition $f(z)\le L(1+z)$ and exclude the trivial situation $\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F=-\infty$ (which in this case happens if and only if $\mathrm{Q}f\equiv-\infty$), then (Con) for $\mathrm{Q}f$ and (A.1) are automatically satisfied.

Proof of Corollaries A.1 and A.2. From the assumption $\mathrm{Q}f=f^{**}$ and the general equality $f^{***}=f^*$ (which in turn follows from the convexity and lower semicontinuity of f^*; compare the beginning of Section 3.3) we infer $(\mathrm{Q}f)^*=f^*$. Consequently, f^*, R_{u_0}, and the solutions of the dual problem (P*) are completely invariant under passage from f to $\mathrm{Q}f$.

Clearly, under the assumptions stated in Corollary A.1 we can apply Theorem 1.1 with $\mathrm{Q}f$ in place of f, and and keeping the above invariance in mind we infer the equality

$$\inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} \mathrm{Q}F = \sup_{\mathrm{L}^\infty_{\mathrm{div}}(\Omega,\mathbb{R}^{Nn})} R_{u_0}$$

and the existence of a dual solution. Involving (A.2) it follows that the claims of Theorem 1.1 hold in the generality of Corollary A.1.

[16]Indeed, it suffices for both the relaxation formula and our purposes in this section if f is not Carathéodory, but only Borel measurable; compare [16, Remark 9.9 (ii)]. Nevertheless, we have decided to work with the Carathéodory property, as it is commonly postulated in the statement of the relaxation formula.

Turning to Corollary A.2, let us first show that every generalized minimizer u of F is also a generalized minimizer of $\mathrm{Q}F$ (with respect to the same u_0) with

$$\mathrm{Q}f(\,\cdot\,,\nabla u) = f(\,\cdot\,,\nabla u) \qquad \mathscr{L}^n\text{-a. e. on } \Omega\,, \tag{A.3}$$

$$(\mathrm{Q}f)^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\mathrm{D}^{\mathrm{s}}u}{\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|}\Big) = f^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\mathrm{D}^{\mathrm{s}}u}{\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|}\Big) \qquad |\mathrm{D}^{\mathrm{s}}u|\text{-a. e. on } \overline{\Omega}\,. \tag{A.4}$$

To this end, we observe — with the specifications of Section 3.2 — that, if $u \in \mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big)$ minimizes $\overline{F}_{u_0}$, then we also have

$$\overline{\mathrm{Q}F}_{u_0}[u] \le \overline{F}_{u_0}[u] = \inf_{\mathrm{BV}_{u_0}(\overline{\Omega},\mathbb{R}^N)} \overline{F}_{u_0} \le \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} F = \inf_{\mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)} \mathrm{Q}F = \inf_{\mathrm{BV}_{u_0}(\overline{\Omega},\mathbb{R}^N)} \overline{\mathrm{Q}F}_{u_0}\,.$$

Here, the first inequality follows from $\mathrm{Q}f \le f$ and $(\mathrm{Q}f)^\infty \le f^\infty$, the second one from $\mathrm{BV}_{u_0}\big(\overline{\Omega},\mathbb{R}^N\big) \supset \mathrm{W}^{1,1}_{u_0}(\Omega,\mathbb{R}^N)$, and the equalities result from the minimality of u, (A.2), and (3.2) (which in turn exploits the assumptions (Lin) and (Con) for $\mathrm{Q}f$). All in all, this reasoning shows that u minimizes $\overline{\mathrm{Q}F}_{u_0}$; in particular, the first inequality is in fact an equality, which in turn results in (A.3) and (A.4).

At this stage, we apply Theorem 2.1 with $\mathrm{Q}f$ in place of f, and we deduce

$$\mathrm{Q}f(\,\cdot\,,\nabla u) = \sigma\cdot\nabla u - f^*(\,\cdot\,,\sigma)\,, \qquad \mathscr{L}^n\text{-a. e. on } \Omega$$

$$(\mathrm{Q}f)^\infty\Big(\,\cdot\,,\frac{\mathrm{d}\mathrm{D}^{\mathrm{s}}u}{\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|}\Big) = \frac{\mathrm{d}[\![\sigma\cdot\mathrm{D}u]\!]^{\mathrm{s}}}{\mathrm{d}|\mathrm{D}^{\mathrm{s}}u|} \qquad |\mathrm{D}^{\mathrm{s}}u|\text{-a. e. on } \overline{\Omega}$$

for all generalized minimizers u of $\mathrm{Q}F$ and all dual solutions σ, where we have used $(\mathrm{Q}f)^* = f^*$ once more. By the preceding argument, the last equalities hold in particular for generalized minimizers u of F, and (2.4) follows once we recall $\mathrm{Q}f = f^{**}$ and (3.3). Furthermore, we can use (A.3) to deduce also (2.2), then (2.3) follows again via (3.3), and via (A.4) we also obtain (2.5). Thus, we have established the claimed generalization of Theorem 2.1.

Finally, the same arguments also suffice to generalize the implication of Theorem 2.3. □

References

[1] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, Oxford, 2000.

[2] G. Anzellotti, *Pairings between measures and bounded functions and compensated compactness*, Ann. Mat. Pura Appl., IV. Ser. **135** (1983), 293–318.

[3] G. Anzellotti, *Traces of bounded vectorfields and the divergence theorem*, Preprint (1983), 43 pages.

[4] G. Anzellotti, *On the minima of functionals with linear growth*, Rend. Sem. Mat. Univ. Padova **75** (1986), 91–110.

[5] G. Anzellotti, BV *solutions of quasilinear PDEs in divergence form*, Commun. Partial Differ. Equations **12** (1987), 77–122.

[6] S. Baldo and L. Modica, *Non uniqueness of minimal graphs*, Indiana Univ. Math. J. **40** (1991), 975–983.

[7] L. Beck and T. Schmidt, *Interior gradient regularity for* BV*-minimizers of singular variational problems*, Preprint (2013), 20 pages.

[8] L. Beck and T. Schmidt, *On the Dirichlet problem for variational integrals in BV*, J. Reine Angew. Math. **674** (2013), 113–194.

[9] M. Bildhauer, *A uniqueness theorem for the dual problem associated to a variational problem with linear growth*, J. Math. Sci., New York **115** (2003), 2747–2752; translation from Zap. Nauchn. Semin. POMI **271** (2000), 83–91.

[10] M. Bildhauer, *A note on degenerate variational problems with linear growth*, Z. Anal. Anwend. **20** (2001), 589–598.

[11] M. Bildhauer, *Convex variational problems. Linear, nearly linear and anisotropic growth conditions*, Springer Lecture Notes in Mathematics **1818**, Berlin, 2003.

[12] M. Bildhauer and M. Fuchs, *Regularity for dual solutions and for weak cluster points of minimizing sequences of variational problems with linear growth*, J. Math. Sci., New York **109** (2002), 1835–1850; translation from Zap. Nauchn. Semin. POMI **259** (1999), 46–66.

[13] M. Bildhauer and M. Fuchs, *On a class of variational integrals with linear growth satisfying the condition of μ-ellipticity*, Rend. Mat. Appl., VII. Ser. **22** (2002), 249–274.

[14] G. Bouchitté and G. Dal Maso, *Integral representation and relaxation of convex local functionals on $BV(\Omega)$*, Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser. **20** (1993), 483–533.

[15] A. Chambolle, M. Goldman, and M. Novaga, *Fine properties of the subdifferential for a class of* 1*-homogeneous functionals*, Preprint (2013), 16 pages.

[16] B. Dacorogna, *Direct methods in the calculus of variations. 2nd ed.*, Springer, New York, 2008.

[17] S. Delladio, *Lower semicontinuity and continuity of functions of measures with respect to the strict convergence*, Proc. R. Soc. Edinb., Sect. A, Math. **119** (1991), 265–278.

[18] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. **47** (1974), 324–353.

[19] I. Ekeland and R. Témam, *Convex analysis and variational problems*, Society for Industrial and Applied Mathematics, Philadelphia, 1999.

[20] L.C. Evans and R.F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, 1992.

[21] M. Giaquinta, G. Modica, and J. Souček, *Functionals with linear growth in the calculus of variations*, Commentat. Math. Univ. Carol. **20** (1979), 143–156, 157–172.

[22] E. Giusti, *Direct Methods in the Calculus of Variations*, World Scientific Publishing, Singapore, 2003.

[23] C. Goffman and J. Serrin, *Sublinear functions of measures and variational integrals*, Duke Math. J. **31** (1964), 159–178.

[24] R. Hardt and D. Kinderlehrer, *Elastic plastic deformation*, Appl. Math. Optimization **10** (1983), 203–246.

[25] R.V. Kohn and R. Temam, *Principes variationnels duaux et théorème de l'énergie dans le modèle de plasticité de Hencky*, C. R. Acad. Sci., Paris, Sér. I **294** (1982), 205–208.

[26] R.V. Kohn and R. Temam, *Dual spaces of stresses and strains, with applications to Hencky plasticity*, Appl. Math. Optimization **10** (1983), 1–35.

[27] J. Kristensen and F. Rindler, *Characterization of generalized gradient Young measures generated by sequences in* $\mathrm{W}^{1,1}$ *and* BV, Arch. Ration. Mech. Anal. **197** (2010), 539–598; correction in Arch. Ration. Mech. Anal. **203** (2012), 693–700.

[28] J. Kristensen and F. Rindler, *Relaxation of signed integral functionals in* BV, Calc. Var. Partial Differ. Equ. **37** (2010), 29–62.

[29] M. Miranda, *Un principio di massimo forte per le frontiere minimali e una sua applicazione alla risoluzione del problema al contorno per l'equazione delle superfici di area minima*, Rend. Sem. Mat. Univ. Padova **45** (1971), 355–366.

[30] M. Miranda, *Dirichlet problem with L^1 data for the non-homogeneous minimal surface equation*, Indiana Univ. Math. J. **24** (1974), 227–241.

[31] Y.G. Reshetnyak, *Weak convergence of completely additive vector functions on a set*, Sib. Math. J. **9** (1968), 1039–1045.

[32] T. Rockafellar, *Integral functionals, normal integrands and measurable selections*, Nonlinear operators and the calculus of variations, Summer School, Brussels, 1975, Springer Lecture Notes in Mathematics **543**, Berlin, 1976.

[33] E. Santi, *Sul problema al contorno per l'equazione delle superfici di area minima su domini limitati qualunque*, Ann. Univ. Ferrara, N. Ser., Sez. VII **17** (1972), 13–26.

[34] T. Schmidt, *Strict interior approximation of sets of finite perimeter and functions of bounded variation*, Preprint (2013), 14 pages, available at http://cvgmt.sns.it/paper/2152/.

[35] G. Seregin, *Variation-difference schemes for problems in the mechanics of ideally elastoplastic media*, U.S.S.R. Comput. Math. Math. Phys. **25** (1985), 153–165; translation from Zh. Vychisl. Mat. Mat. Fiz. **25** (1985), 237–253.

[36] G. Seregin, *Differential properties of solutions of variational problems for functionals of linear growth*, J. Sov. Math. **64** (1993), 1256–1277; translation from Probl. Mat. Anal. **11** (1990), 51–79.

[37] G. Seregin, *Two-dimensional variational problems of the theory of plasticity*, Izv. Math. **60** (1996), 179–216; translation from Izv. Ross. Akad. Nauk, Ser. Mat. **60** (1996), 175–210.

[38] R. Temam and G. Strang, *Duality and relaxation in the variational problems of plasticity*, J. Mec. **19** (1980), 493–527.