Calc. Var. 19, 23–51 (2004) **Calculus of Variations**
DOI: 10.1007/s00526-003-0192-2

Virginia De Cicco · Giovanni Leoni

A chain rule in L^1 (div; Ω) and its applications to lower semicontinuity

Received: 1 June 2002 / Accepted: 7 January 2003 /
Published online: 6 June 2003 –

Abstract. A chain rule in the space L^1 (div; Ω) is obtained under weak regularity conditions. This chain rule has important applications in the study of lower semicontinuity problems for general functionals of the form $\int_\Omega f(x,u,\nabla u)\,dx$ with respect to strong convergence in $L^1(\Omega)$. Classical results of Serrin and of De Giorgi, Buttazzo and Dal Maso are extended and generalized.

1. Introduction

It is well known that if $h:\mathbb{R}\to\mathbb{R}$ is a Lipschitz continuous function, then for every function $u\in W^{1,p}(\Omega)$, where $\Omega\subset\mathbb{R}^N$ is an open bounded set and $1\le p\le\infty$, the composition function $v=h\circ u:\Omega\to\mathbb{R}$ belongs to $W^{1,p}(\Omega)$. Moreover, if the function h is of class C^1, then it is easy to see that the classical chain rule formula holds, that is

$$\nabla v(x)=h'(u(x))\nabla u(x) \text{ for } \mathcal{L}^N \text{ a.e. } x\in\Omega. \tag{1.1}$$

If the function h is only Lipschitz, then the right side of (1.1) may not be defined since h' may not exist everywhere. However, by Rademacher's Theorem we know that the set $\mathcal{M}:=\{u\in\mathbb{R}:\ h'(u) \text{ does not exist}\}$ is $\mathcal{L}^1$-null. Thus

$$\nabla u(x)=0 \text{ for } \mathcal{L}^N \text{ a.e. } x\in u^{-1}(\mathcal{M}), \tag{1.2}$$

and in turn the right side of (1.1) is always well defined provided $h'(u(x))\nabla u(x)$ is interpreted to be zero whenever $\nabla u(x)=0$, irrespective of whether $h'(u(x))$ is defined. With this convention the chain rule formula (1.1) was first proved by Serrin [46] (see also [8, 10, 19] and [41]).

Property (1.2) is a consequence of the following crucial result:

Theorem 1.1 *Let E be a set of $\mathbb{R}$. Then the following are equivalent:*

(i) $\mathcal{H}^1(E)=0$;

Virginia De Cicco: Dipartimento di Metodi e Modelli, Matematici per le Scienze Applicate, Università di Roma "La Sapienza", Via Scarpa 16, I-00161 Rome, Italy

Giovanni Leoni: Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

(ii) for all $N \in \mathbb{N}$ and for any $u \in W^{1,1}_{\text{loc}}(\mathbb{R}^N)$

$$\nabla u\,(x) = 0 \textit{ for } \mathcal{L}^N \textit{ a.e. } x \in u^{-1}(E); \tag{1.3}$$

(iii) for all $N \in \mathbb{N}$ and for any $u \in BV_{\text{loc}}(\mathbb{R}^N)$

$$Du\,(x) = 0 \textit{ for } |Du| \textit{ a.e. } x \in \tilde{u}^{-1}(E) \cap L_u,$$

where $|Du|$, $\tilde{u}$ and L_u are, respectively, the total variation of the distributional derivative Du, a precise representative and the set of Lebesgue points of u.

The equivalence (i) $\Longleftrightarrow$ (ii) follows e.g. from a result of Serrin and Varberg [48] (see Theorem 2.6 below) in the case $N = 1$, while the general case $N > 1$ may be obtained from a slicing argument (see the work of Marcus and Mizel [41]). A different proof of the implication (i) $\Rightarrow$ (ii) may be found in the paper [10] of Boccardo and Murat (see also [8] for some recent extensions). The implication (i) $\Rightarrow$ (iii) is due to Dal Maso, Lefloch and Murat [19].

The situation is significantly more complicated in the vectorial case, namely when $h : \mathbb{R}^d \to \mathbb{R}^m$ is a Lipschitz continuous function with $d > 1$. In this case we still have that for every function $\mathbf{u} \in W^{1,p}\left(\Omega; \mathbb{R}^d\right)$ the composition function $\mathbf{v} = h \circ \mathbf{u}$ belongs to $W^{1,p}\left(\Omega; \mathbb{R}^m\right)$. However, if we define

$$\mathcal{M}_1 := \left\{\mathbf{u} \in \mathbb{R}^d : \nabla_{\mathbf{u}} h\,(\mathbf{u}) \text{ does not exist}\right\}, \tag{1.4}$$

(i.e. $\mathbf{u} \in \mathcal{M}_1$ if and only if at least one of the partial derivatives $\frac{\partial h_i}{\partial u_j}$ does not exist at $\mathbf{u}$), then again by Rademacher's Theorem we have that $\mathcal{L}^d\left(\mathcal{M}_1\right) = 0$, but *the analogous of property* (1.3) *is false for $\mathcal{L}^d$-null sets $E \subset \mathbb{R}^d$* as before. Hence in the vectorial chain rule

$$\nabla \mathbf{v}\,(x) = \nabla_{\mathbf{u}}\, h\,(\mathbf{u}\,(x))\, \nabla \mathbf{u}\,(x) \text{ for } \mathcal{L}^N \text{ a.e. } x \in \Omega \tag{1.5}$$

the expression $\nabla_{\mathbf{u}} h(\mathbf{u}(x)) \nabla \mathbf{u}(x)$ may not be defined. Indeed, let $d = 2$, $N = m = 1$, $\Omega = (0, 1)$ and consider the functions (cf. [41])

$$h\,(\mathbf{u}) := \max\{u_1, u_2\} \text{ for } \mathbf{u} = (u_1, u_2) \in \mathbb{R}^d$$

and

$$\mathbf{u}\,(x) := (x, x) \text{ for } x \in (0, 1)\,.$$

Then $v\,(x) := (h \circ \mathbf{u})\,(x) = x$ so that $v'\,(x) = 1$ while the right side of (1.5) is *nowhere defined.*

More recently Ambrosio and Dal Maso [3] have proved the following weaker form of the chain rule:

Theorem 1.2 *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set, $1 \le p \le \infty$, and let $h : \mathbb{R}^d \to \mathbb{R}^m$ be a Lipschitz continuous function. Then for every function $\mathbf{u} \in W^{1,p}\left(\Omega; \mathbb{R}^d\right)$ the composition function $\mathbf{v} = h \circ \mathbf{u}$ belongs to $W^{1,p}\left(\Omega; \mathbb{R}^m\right)$ and for $\mathcal{L}^N$ a.e. $x \in \Omega$ the restriction of the function h to the affine space*

$$T_x^{\mathbf{u}} := \left\{\mathbf{w} \in \mathbb{R}^d : \mathbf{w} = \mathbf{u}\,(x) + \nabla \mathbf{u}\,(x)\, \mathbf{z} \textit{ for some } \mathbf{z} \in \mathbb{R}^N\right\}$$

is differentiable at $\mathbf{u}\,(x)$ and

$$\nabla \mathbf{v}\,(x) = \nabla_{\mathbf{u}}(h|_{T_x^{\mathbf{u}}})\,(\mathbf{u}\,(x))\, \nabla \mathbf{u}\,(x)\,. \tag{1.6}$$

Theorem 1.2 leaves us with an important open problem: to establish under which additional conditions on the function h the right side of (1.6) coincides with right side of (1.5), in other words to *find necessary and sufficient conditions on h for the classical chain rule* (1.5) *to hold.* A first step in this direction is to understand for which sets $E \subset \mathbb{R}^d$ property (1.3) holds, namely under what hypotheses on E we have that for every $\mathbf{u} \in W^{1,p}\left(\Omega;\mathbb{R}^d\right)$ there holds

$$\nabla\mathbf{u}\,(x) = 0 \text{ for } \mathcal{L}^N \text{ a.e. } x \in \mathbf{u}^{-1}(E). \tag{1.7}$$

This problem was studied by Marcus and Mizel [41] who proved that property (1.7) holds if the set $E \subset \mathbb{R}^d$ has the *null intersection property,* that is

$$\mathcal{H}^1(E \cap \mathbf{w}(I)) = 0 \tag{1.8}$$

for any absolutely continuous function $\mathbf{w} : I \to \mathbb{R}^d$, where $I \subset \mathbb{R}$ is an interval. Property (1.8) turns out to be equivalent to the important notion of $\mathcal{H}^1$-*unrectifiability* (see e.g. [29]). We recall that a set E is *purely $\mathcal{H}^1$-unrectifiable* if

$$\mathcal{H}^1(E \cap \mathbf{w}(\mathbb{R})) = 0 \tag{1.9}$$

for any Lipschitz function $\mathbf{w} : \mathbb{R} \to \mathbb{R}^d$. Indeed we have the following result:

Theorem 1.3 *Let E be a set of $\mathbb{R}^d$. Then the following are equivalent:*

(i) E is purely $\mathcal{H}^1$-unrectifiable;
(ii) E has the null intersection property (1.8)*;*
(iii) $\mathcal{H}^1(E \cap \mathbf{w}(\mathbb{R})) = 0$ for any function $\mathbf{w} \in BV_{\mathrm{loc}}(\mathbb{R};\mathbb{R}^d) \cap C(\mathbb{R};\mathbb{R}^d)$;
(iv) for all $N \in \mathbb{N}$ and for any $\mathbf{u} \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^N;\mathbb{R}^d)$

$$\nabla\mathbf{u}\,(x) = 0 \textit{ for } \mathcal{L}^N \textit{ a.e. } x \in \mathbf{u}^{-1}(E); \tag{1.10}$$

(v) for all $N \in \mathbb{N}$ and for any $\mathbf{u} \in BV_{\mathrm{loc}}(\mathbb{R}^N;\mathbb{R}^d)$

$$\nabla\mathbf{u}\,(x) = 0 \textit{ for } \mathcal{L}^N \textit{ a.e. } x \in \tilde{\mathbf{u}}^{-1}(E),$$

where $\nabla\mathbf{u}$ is the Radon–Nikodym derivative of the distributional derivative $D\mathbf{u}$ of $\mathbf{u}$, with respect to the Lebesgue measure $\mathcal{L}^N$ and $\tilde{\mathbf{u}}$ is a precise representative of $\mathbf{u}$.

The equivalences (i) $\iff$ (ii) $\iff$ (iii) follow from a result of Choquet (cf. Theorem 16 in [17]); the implication (ii) $\Rightarrow$ (iv) is due to Marcus and Mizel (cf. the proof of Lemma 2.1 in [41]), while the remaining implications and equivalence are new in the vectorial case $d > 1$, to the best of our knowledge. We note, in particular, that the equivalence (i) $\iff$ (v) improves Proposition 3.92(a) in [4].

To prove the implications (i) $\Rightarrow$ (iv) and (i) $\Rightarrow$ (v) we combine the ideas of Marcus and Mizel (cf. the proof of Lemma 2.1 of [41]) with a Lusin type theorem, obtained by Liu [39] (see also the work of Acerbi and Fusco [1]) in the Sobolev setting and later extended by Ambrosio to functions of bounded variation (see Theorem 5.34 in [4]), which allows to approximate Sobolev functions and functions of bounded variation with Lipschitz functions. If one could find an extension of these approximation results for the Cantor part of the distributional derivative $D\mathbf{u}$ of a function $\mathbf{u}$ of bounded variation, then the analogous of Theorem 1.1(iii) would hold in the vectorial case.

The relevance of the equivalence (i) $\Longleftrightarrow$ (iv) lies in the fact that purely $\mathcal{H}^1$-unrectifiable sets may be characterized by the Structure Theorem of Besicovitch-Federer (see Theorem 2.1 below).

In view of Theorem 1.3 we have the following result:

Theorem 1.4 *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set, $1 \leq p \leq \infty$, and let $h : \mathbb{R}^d \to \mathbb{R}^m$ be a Lipschitz continuous function. Then a sufficient condition for the chain rule* (1.5) *to hold for every $\mathbf{u} \in W^{1,p}\left(\Omega; \mathbb{R}^d\right)$ is that the set*

$$\mathcal{M}_2 := \left\{\mathbf{u} \in \mathbb{R}^d : \ h \text{ is not differentiable at } \mathbf{u}\right\} \tag{1.11}$$

be purely $\mathcal{H}^1$-unrectifiable, while a necessary condition is that the set $\mathcal{M}_1$ defined in (1.4) *be purely $\mathcal{H}^1$-unrectifiable.*

We recall that in (1.5) the function $\nabla_{\mathbf{u}} h(\mathbf{u}(x))\, \nabla \mathbf{u}(x)$ is interpreted to be zero whenever $\nabla \mathbf{u}(x) = 0$, irrespective of whether $\nabla_{\mathbf{u}} h(\mathbf{u}(x))$ is defined. The proof of the sufficiency part follows exactly as in Theorem 2.1 in [41] (see also the proof of Theorem 1.5 below), while the necessity follows from the discussion above. Note that unlike the scalar case $d = 1$ the sets $\mathcal{M}_1$ and $\mathcal{M}_2$ do not coincide. It is actually possible to construct functions h for which $\nabla_{\mathbf{u}} h$ exist $\mathcal{L}^d$ a.e. in $\mathbb{R}^d$, but which are *nowhere differentiable* (see e.g. [13]).

Although the right side of (1.5) is formally well defined if $\mathbf{u}(x) \in \mathcal{M}_2 \setminus \mathcal{M}_1$, simple examples show that in general if $\mathbf{u}_0 \in \mathcal{M}_2 \setminus \mathcal{M}_1$ it is possible to construct a smooth curve $\mathbf{u} : (-1, 1) \to \mathbb{R}^d$ such that $\mathbf{u}\,(0) = \mathbf{u}_0$ and for which the chain rule fails at $x = 0$. Thus we are inclined to think that the $\mathcal{H}^1$-unrectifiability of $\mathcal{M}_2$ should also be a necessary condition for the chain rule to hold, but we have been unable to prove it.

We are now ready to state one of the main results of the paper.

Let us consider the space $L^1(\mathrm{div};\Omega) := \left\{u \in L^1(\Omega; \mathbb{R}^N) : \mathrm{div}\, u \in L^1(\Omega)\right\}$, where $\mathrm{div}\, v$ is the distributional divergence.

Theorem 1.5 (Chain rule in $L^1(\mathrm{div};\Omega)$) *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set and let $B : \Omega \times \mathbb{R}^d \to \mathbb{R}^N$ be a Borel function. Assume that there exist an $\mathcal{L}^N$-null set $\mathcal{N} \subset \Omega$ and a purely $\mathcal{H}^1$-unrectifiable set $\mathcal{M} \subset \mathbb{R}^d$ such that*

(i) for all $\mathbf{u} \in \mathbb{R}^d$ the function $B\,(\cdot, \mathbf{u}) \in L^1_{\mathrm{loc}}(\mathrm{div};\Omega)$;

(ii) for all $x \in \Omega \setminus \mathcal{N}$ the function $\mathrm{div}_x\, B(x, \cdot)$ is approximately continuous in $\mathbb{R}^d$;

(iii) for all $x \in \Omega \setminus \mathcal{N}$ the function $B\,(x, \cdot)$ is locally Lipschitz in $\mathbb{R}^d$ and differentiable in $\mathbb{R}^d \setminus \mathcal{M}$;

(iv) for every $\Omega' \times D \subset\subset \Omega \times \mathbb{R}^d$ there exist $g \in L^1(\Omega)$ and $L > 0$ such that

$$|B(x, \mathbf{u})| + |\mathrm{div}_x\, B(x, \mathbf{u})| \leq g(x)$$

for $\mathcal{L}^N$ a.e. $x \in \Omega'$ and for all $\mathbf{u} \in D$, and

$$|\nabla_{\mathbf{u}} B(x, \mathbf{u})| \leq L$$

for $\mathcal{L}^N$ a.e. $x \in \Omega'$ and for all $\mathbf{u} \in D \setminus \mathcal{M}$.

Then for every $\mathbf{u} \in W^{1,1}(\Omega;\mathbb{R}^d) \cap L^\infty_{\mathrm{loc}}(\Omega;\mathbb{R}^d)$ *the function* $\mathbf{v} : \Omega \to \mathbb{R}^N$, *defined by*

$$\mathbf{v}(x) := B(x, \mathbf{u}(x)) \quad x \in \Omega,$$

belongs to $L^1_{\mathrm{loc}}(\mathrm{div};\Omega)$ *and*

$$\mathrm{div}\, \mathbf{v}(x) = \mathrm{div}_x\, B(x, \mathbf{u}(x)) + \mathrm{tr}\,(\nabla_{\mathbf{u}} B(x, \mathbf{u}(x))\, \nabla \mathbf{u}(x))$$

for $\mathcal{L}^N$ *a.e.* $x \in \Omega$, *provided* $\nabla_{\mathbf{u}} B(x, \mathbf{u}(x))\, \nabla \mathbf{u}(x)$ *is interpreted to be zero whenever* $\nabla \mathbf{u}(x) = 0$, *irrespective of whether* $\nabla_{\mathbf{u}} B(x, \mathbf{u}(x))$ *is defined.*

Remark 1.6 It goes almost without saying that by strengthening the growth condition (iv) one may obtain $\mathbf{v} \in L^1(\mathrm{div};\Omega)$, or more generally $\mathbf{v} \in L^p(\mathrm{div};\Omega)$ (provided $\mathbf{u} \in W^{1,p}(\Omega;\mathbb{R}^d)$). We leave the details to the interested reader.

An important special case is given when B has the form

$$B(x, u, w) := \int_w^u b(x, s)\, ds, \tag{1.12}$$

where $b : \Omega \times \mathbb{R} \to \mathbb{R}^N$.

Corollary 1.7 *Let* $\Omega \subset \mathbb{R}^N$ *be an open bounded set and let* $b : \Omega \times \mathbb{R} \to \mathbb{R}^N$ *be a locally bounded Borel function. Assume that there exist an* $\mathcal{L}^N$*-null set* $\mathcal{N} \subset \Omega$ *and an* $\mathcal{L}^1$*-null set* $\mathcal{M} \subset \mathbb{R}$ *such that*

(i) for every $x \in \Omega \setminus \mathcal{N}$ *the function* $b(x, \cdot)$ *is approximately continuous in* $\mathbb{R} \setminus \mathcal{M}$;
(ii) $b(\cdot, u) \in L^1_{\mathrm{loc}}(\mathrm{div};\Omega)$ *for* $\mathcal{L}^1$ *a.e.* $u \in \mathbb{R}$ *with* $\mathrm{div}_x\, b \in L^1_{\mathrm{loc}}(\Omega \times \mathbb{R})$.

Then for every $u, w \in W^{1,1}(\Omega) \cap L^\infty_{\mathrm{loc}}(\Omega)$ *the function* $\mathbf{v} : \Omega \to \mathbb{R}^N$, *defined by*

$$\mathbf{v}(x) := \int_{w(x)}^{u(x)} b(x, s)\, ds,$$

belongs to $L^1_{\mathrm{loc}}(\mathrm{div};\Omega)$ *and for* $\mathcal{L}^N$ *a.e.* $x \in \Omega$

$$\mathrm{div}\, \mathbf{v}(x) = \int_{w(x)}^{u(x)} \mathrm{div}_x\, b(x, s)\, ds + b(x, u(x)) \cdot \nabla u(x) - b(x, w(x)) \cdot \nabla w(x).$$

Remark 1.8 We observe that the hypothesis (i) is verified in the following two cases:

(i) $b(x, \cdot)$ is continuous in $\mathbb{R} \setminus \mathcal{M}$ for every $x \in \Omega \setminus \mathcal{N}$;
(ii) $b(\cdot, u)$ is continuous in $\Omega \setminus \mathcal{N}$ for every $u \in \mathbb{R} \setminus \mathcal{M}$.

See Proposition 2.5 for more details.

Proposition 1.9 *Assume that* $b : \Omega \times \mathbb{R} \to \mathbb{R}^N$ *satisfies all the hypotheses of Corollary 1.7. Let* $\{u_n\}$ *be a sequence in* $W^{1,1}(\Omega)$ *converging to* $u \in W^{1,1}(\Omega)$ *with respect to the strong* $L^1(\Omega)$ *convergence, and such that* $\sup_n \|u_n\|_{L^\infty(\Omega)} < \infty$. *Then for every* $\phi \in C^1_c(\Omega)$

$$\lim_{n\to\infty} \int_\Omega b(x, u_n(x)) \cdot \nabla u_n(x) \phi(x)\, dx = \int_\Omega b(x, u(x)) \cdot \nabla u(x) \phi(x)\, dx. \tag{1.13}$$

Remark 1.10 The condition $\sup_n \|u_n\|_{L^\infty(\Omega)} < \infty$ may be removed if $b(x,u) = 0$ for $|u| \geq L$.

The structure of Proposition 1.9 is of course that of the celebrated Div-Curl Lemma of Murat and Tartar (see [49]). The situation here is simpler since the sequences $D_n := b(x, u_n(x))$ and $E_n := \nabla u_n(x)$ are closely related. At the same time here we have no uniform bounds on the L^1 norms of the $\nabla u_n(x)$.

The hypotheses that b is bounded and $b(\cdot,u) \in L^1_{\text{loc}}(\text{div};\Omega)$ may be significantly weakened to require b to be a summable function and $\text{div}_x\, b(\cdot,u)$ to be a measure (see Section 5 for more details). It would be interesting to find necessary and sufficient conditions for (1.13) to hold. Indeed the previous proposition turns out to be important in the study of lower semicontinuity properties for functionals of the form

$$F(u) := \int_\Omega (a(x,u) + b(x,u)\cdot\nabla u)^+ \, dx \quad \text{or}$$
$$F(u) := \int_\Omega |a(x,u) + b(x,u)\cdot\nabla u| \, dx.$$

Proposition 1.11 *Let b be a function satisfying all the hypotheses of Corollary 1.7. Let $a : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Borel function, such that $a(x,\cdot)$ is lower semicontinuous on $\mathbb{R}$. Assume also that for every $K \subset\subset \Omega \times \mathbb{R}$ there exists $h \in L^1(\Omega)$ such that*

$$a(x,u) \geq h(x) \tag{1.14}$$

for all $(x,u) \in K$. Then the functional $F(u) := \int_\Omega (a(x,u) + b(x,u)\cdot\nabla u)^+ \, dx$ is lower semicontinuous on $W^{1,1}(\Omega)$ with respect to the strong L^1 convergence.

Remark 1.12 We believe that the hypothesis that $a(x,\cdot)$ is lower semicontinuous on $\mathbb{R}$ is significantly stronger than necessary. Indeed, by taking $u_n(x) :\equiv c_n \in \mathbb{R}$, $u(x) :\equiv c \in \mathbb{R}$ with $c_n \to c$ we obtain that the correct necessary condition for lower semicontinuity on the function a is that $(a(x,\cdot))^+$ is lower semicontinuous on $\mathbb{R}$. It would be interesting to know if Proposition 1.11 continues to hold under this weaker condition.

Using the fact that $|u| = u^+ + (-u)^+$ we easily obtain from the previous proposition that

Proposition 1.13 *Let b be a bounded function satisfying all the hypotheses of Corollary 1.7. Let $a : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Borel function, such that $a(x,\cdot)$ is continuous on $\mathbb{R}$. Assume also that for every $K \subset\subset \Omega \times \mathbb{R}$ there exists $h \in L^1(\Omega)$ such that $|a(x,u)| \leq h(x)$ for all $(x,u) \in K$. Then the functional*

$$F(u) := \int_\Omega |a(x,u) + b(x,u)\cdot\nabla u| \, dx \tag{1.15}$$

is lower semicontinuous on $W^{1,1}(\Omega)$ with respect to the strong L^1 convergence.

The functional (1.15) has been studied by several authors. In particular Dal Maso [18] has shown that lower semicontinuity may fail, in general, if $b = b\,(x)$ is continuous, while Gori and Marcellini [36] have shown that the functional is lower semicontinuous if $b\,(\cdot, u)$ is Lipschitz, but fails to be lower semicontinuous if $b\,(\cdot, u)$ is only Holder continuous. This latter result has been extended by Gori, Maggi and Marcellini [37].

In the case when $a = a(x)$ and $b = b(x)$ the hypotheses of Proposition 1.13 reduce to

$$a \in L^1_{\mathrm{loc}}(\Omega) \quad \text{and} \quad b \in L^{\infty,1}_{\mathrm{loc}}(\mathrm{div};\Omega), \tag{1.16}$$

where the spaces $L^{p,q}_{\mathrm{loc}}(\mathrm{div};\Omega)$ are defined in Section 2. Note that neither a nor b needs to be continuous.

The conditions (1.16) are almost optimal. Indeed in [34] Gavioli (see also [8, 11], and [32]) has shown the following result.

Theorem 1.14 *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set, and let $f : \Omega \times \mathbb{R}^N \to [0, \infty)$ be a Borel integrand such that $f(x, \cdot)$ is convex in $\mathbb{R}^N$ for every $x \in \Omega$ and*

$$0 \le f(x, \xi) \le \psi(x) + C|\xi|$$

for every $x \in \Omega$ and $\xi \in \mathbb{R}^N$, for some function $\psi \in L^1(\Omega)$ and a positive constant C. Then the functional $F(u) := \int_\Omega f(x, \nabla u)\,dx$ is lower semicontinuous on $W^{1,1}(\Omega)$ with respect to the strong $L^1(\Omega)$ convergence if and only if there exist two sequences $\{a_n\} \subset L^1(\Omega)$, $\{b_n\} \subset L^\infty(\mathrm{div};\Omega)$ such that for $\mathcal{L}^N$ a.e. $x \in \Omega$ and for all $\xi \in \mathbb{R}^N$

$$f(x, \xi) = \sup_{n\in\mathbb{N}}\{a_n\,(x) + b_n(x) \cdot \xi\}.$$

We are now ready to state the main lower semicontinuity results of the paper. We consider first the case when f is continuous in the u variable.

Theorem 1.15 *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set, and let*

$$f : \Omega \times \mathbb{R} \times \mathbb{R}^N \to [0, \infty)$$

be a Borel function such that

(i) $f(x, u, \cdot)$ is convex in $\mathbb{R}^N$ for every $(x, u) \in \Omega \times \mathbb{R}$;
(ii) $f(x, \cdot, \xi)$ is continuous in $\mathbb{R}$ for all $(x, \xi) \in \Omega \times \mathbb{R}^N$;
(iii) f is locally bounded;
(iv) for $\mathcal{L}^1$ a.e. $u \in \mathbb{R}$ and for $\mathcal{L}^N$ a.e. $\xi \in \mathbb{R}^N$

$$\nabla_\xi f(\cdot, u, \xi) \in L^1_{\mathrm{loc}}\,(\mathrm{div};\Omega) \quad \textit{with}\ \mathrm{div}_x\, \nabla_\xi f \in L^1_{\mathrm{loc}}(\Omega \times \mathbb{R} \times \mathbb{R}^N).$$

Then the functional

$$F(u) := \int_\Omega f(x, u, \nabla u)\,dx$$

is lower semicontinuous on $W^{1,1}(\Omega)$ with respect to the strong $L^1(\Omega)$ convergence.

Theorem 1.15(ii) improves a classical result of Serrin [47] (see also [30] and [36]) where the function f was required to be continuous together with its derivatives $\nabla_x f$, $\nabla_\xi f$ and $\nabla_x (\nabla_\xi f)$.

The key tool in the proof of Theorem 1.15 is an approximation result for convex functions due to De Giorgi [23], who proved that every convex function $f : \mathbb{R}^N \to \mathbb{R}$ may be approximated by a sequence of affine functions of the form $A_\alpha + B_\alpha \cdot \xi$, where

$$A_\alpha := \int_{\mathbb{R}^N} f(z)\left((N+1)\,\alpha(z) + \nabla\alpha(z)\cdot z\right) dz \tag{1.17}$$
$$B_\alpha := -\int_{\mathbb{R}^N} f(z)\,\nabla\alpha(z)\,dz,$$

and the functions $\alpha \in C_c^1\left(\mathbb{R}^N\right)$ are nonnegative and $\int_{\mathbb{R}^N} \alpha(z)\,dz = 1$. The relevance of this approximation result lies in the fact that the coefficients A_α and B_α are given explicitly in terms of f and thus in the case when f depends also on x and u the properties of the coefficients A_α and B_α may be deduced from the hypotheses on f. This idea was first used in the recent paper of Gori and Marcellini [36] (see also [37]).

We remark that, when f depends explicitly also on x and u, the continuity of f in x or u (see hypothesis (ii) in the previous theorem and hypothesis (iii) in the next theorem) implies the continuity of A_α and B_α in x or u. On the contrary the lower semicontinuity of f does not seem to imply the lower semicontinuity of A_α; this is the condition on A_α which assures the lower semicontinuity of the approximating functionals (see Proposition 1.11, see also Remark 1.12).

Note that Theorem 1.15 cannot be extended to the vectorial case, see [14,15, 25,30] and [44]. This is actually clear from Proposition 1.9. Indeed if we want

$$\int_\Omega b(x,\mathbf{u}(x))\nabla\mathbf{u}(x)\phi(x)\,dx$$

to be continuous with respect to strong convergence in L^1 of admissible sequences $\{u_n\}$ without any assumption on $\{\nabla u_n\}$ it is clear that one must necessarily integrate by parts, but in the vectorial case to find a potential B such that $\nabla_{\mathbf{u}} B(x,\mathbf{u}) = b(x,\mathbf{u})$ we would need b to be curl free.

Next we consider the case when f is continuous in the x variable.

Theorem 1.16 *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set, and let*

$$f : \Omega \times \mathbb{R} \times \mathbb{R}^N \to [0,\infty)$$

be a Borel function such that

(i) $f(x,u,\cdot)$ is convex in $\mathbb{R}^N$ for every $(x,u) \in \Omega \times \mathbb{R}$;
(ii) $f(x,u,0) = 0$ for all $(x,u) \in \Omega \times \mathbb{R}$;
(iii) $f(\cdot,u,\xi)$ is continuous on Ω for all $(u,\xi) \in \mathbb{R} \times \mathbb{R}^N$;
(iv) for $\mathcal{L}^1$ a.e. $u \in \mathbb{R}$ and for $\mathcal{L}^N$ a.e. $\xi \in \mathbb{R}^N$

$$\nabla_\xi f(\cdot,u,\xi) \in L^1_{\mathrm{loc}}(\mathrm{div};\Omega) \quad \textit{with } \mathrm{div}_x \nabla_\xi f \in L^1_{\mathrm{loc}}(\Omega \times \mathbb{R} \times \mathbb{R}^N);$$

(v) for $\mathcal{L}^1$ a.e. $u \in \mathbb{R}$ and every compact set $K \subset \Omega \times \mathbb{R}^N$ there exists $h_u \in L^1_{\mathrm{loc}}(K)$ such that

$$f(x,u,\xi) \leq h_u(x)$$

for all $(x,\xi) \in K$.

Then the functional

$$F(u) := \int_\Omega f(x,u,\nabla u)\,dx$$

is lower semicontinuous on $W^{1,1}(\Omega)$ with respect to the strong $L^1(\Omega)$ convergence.

Theorem 1.16 was first proved De Giorgi, Buttazzo and Dal Maso [24] for integrands of the form $f = f(u,\xi)$ (see also [2,20,21,31]).

Remark 1.17 Condition (iv) in Theorems 1.15 and 1.16 may actually be replaced by the weaker condition that for $\mathcal{L}^1$ a.e. $u \in \mathbb{R}$ and for every $\varphi \in \mathcal{D}\left(\mathbb{R}^N\right)$

$$\langle \nabla_\xi f(x,u,\cdot), \varphi(\cdot)\rangle \in L^1_{\mathrm{loc}}(\mathrm{div};\Omega) \text{ with } \mathrm{div}_x \nabla_\xi f \in \mathcal{D}'\left(\mathbb{R}^N; L^1_{\mathrm{loc}}(\Omega\times\mathbb{R})\right),$$

where with this notation we mean that for all $A \subset\subset \mathbb{R}^N$, $\Omega' \subset\subset \Omega$, and $M > 0$ there exists $C > 0$ and an integer $k \in \mathbb{N}$ such that

$$\int_{-M}^{M}\int_{\Omega'} \left|\mathrm{div}_x \langle \nabla_\xi f(x,u,\cdot), \varphi(\cdot)\rangle_A\right| \,dx\,du \leq C\,\|\varphi\|_{C_c^k(A)}, \tag{1.18}$$

for all $\varphi \in C_c^\infty(A)$. Since

$$\begin{aligned}\langle \nabla_\xi f(x,u,\cdot), \varphi(\cdot)\rangle_A &= \int_A \nabla_\xi f(x,u,\xi)\,\varphi(\xi)\,d\xi \\ &= -\int_A f(x,u,\xi)\,\nabla_\xi\varphi(\xi)\,d\xi = -\langle f(x,u,\cdot), \nabla_\xi\varphi(\cdot)\rangle_A\end{aligned}$$

condition (1.18) becomes

$$\int_{-M}^{M}\int_{\Omega'} \left|\mathrm{div}_x \int_A f(x,u,\xi)\,\nabla_\xi\varphi(\xi)\,d\xi\right| \,dx\,du \leq C\,\|\varphi\|_{C_c^k(A)}. \tag{1.19}$$

Corollary 1.18 *Theorems 1.15 and 1.16 continue to hold if condition (iv) is replaced by the hypothesis that $f(\cdot,u,\xi) \in W^{1,1}_{\mathrm{loc}}(\Omega)$ for all $(u,\xi) \in \mathbb{R}\times\mathbb{R}^N$ with $\nabla_x f(\cdot,u,\cdot) \in L^1_{\mathrm{loc}}\left(\Omega\times\mathbb{R}^N;\mathbb{R}^N\right)$ for $\mathcal{L}^1$ a.e. $u \in \mathbb{R}$ and*

$$\mathrm{div}_\xi \nabla_x f \in L^1_{\mathrm{loc}}\left(\Omega\times\mathbb{R}; W^{-1,\infty}_{\mathrm{loc}}\left(\mathbb{R}^N\right)\right). \tag{1.20}$$

Condition (1.20) is implied by the simpler assumption that

$$\nabla_x f \in L^1_{\mathrm{loc}}\left(\Omega\times\mathbb{R}\times\mathbb{R}^N;\mathbb{R}^N\right).$$

Hence Corollary 1.18 improves the recent results of Gori and Marcellini in [36] and of Gori, Maggi and Marcellini [37] (see also the paper of Fusco, Giannetti and Verde [33] for an extension to the BV setting).

Further improvements may be found on Section 5 below.

Finally, we remark that an extension to the BV setting of our lower semicontinuity results is not straightforward , as different techniques are required to treat the jump part of the functional. This problem will be addressed in a forthcoming paper [22] by the first author in collaboration with Fusco and Verde.

Previous lower semicontinuity results in BV may be found in [18,20,21,31, 33], under various continuity or semicontinuity assumptions of the integrand f in the x variable, which are not needed in [22].

2. Preliminaries

In this section we collect some preliminary results which will be used in the sequel. We start with some notation. Here $\Omega \subset \mathbb{R}^N$ is an open bounded subset, $\mathcal{L}^k$ and $\mathcal{H}^k$ are, respectively, the k dimensional Lebesgue measure and the k dimensional Hausdorff measure in Euclidean spaces.

A set $E \subset \mathbb{R}^d$ is *purely* $\mathcal{H}^1$*-unrectifiable* if $\mathcal{H}^1(E \cap \mathbf{w}(\mathbb{R})) = 0$ for any Lipschitz function $\mathbf{w} : \mathbb{R} \to \mathbb{R}^d$.

Purely $\mathcal{H}^1$-unrectifiable sets with finite (or σ-finite) $\mathcal{H}^1$ measure may be characterized by virtue of the Structure Theorem of Besicovitch-Federer (see [9] for the case $N = 2$, [28] for the general case $N \geq 2$ and the recent simple proof of White [50])

Theorem 2.1 (Structure Theorem) *Let $E \subset \mathbb{R}^d$ be such that $0 < \mathcal{H}^1(E) < \infty$. Then E is purely $\mathcal{H}^1$-unrectifiable if and only if*

$$\mathcal{H}^1(\pi_L E) = 0 \text{ for } \mu_1 \text{ a.e. line } L \subset \mathbb{R}^d,$$

where $\pi_L : \mathbb{R}^d \to L$ is the orthogonal projection, and μ_1 is the Haar measure in the set of all lines in $\mathbb{R}^d$.

For more informations on the Haar measure we refer to [28] and [29].

Let $u : \mathbb{R}^N \to \mathbb{R}$ be Lebesgue measurable. We say that u is *approximately continuous* at $x_0 \in \mathbb{R}^N$ if for every $\varepsilon > 0$

$$\lim_{r \to 0^+} \frac{\mathcal{L}^N\left(B(x_0, r) \cap \left\{x \in \mathbb{R}^N : |u(x) - u(x_0)| \geq \varepsilon\right\}\right)}{\mathcal{L}^N(B(x_0, r))} = 0.$$

We say that $x_0 \in \mathbb{R}^N$ is a *Lebesgue point* of u if

$$\lim_{\varepsilon \to 0^+} \frac{1}{\mathcal{L}^N(B(x_0, \varepsilon))} \int_{B(x_0, \varepsilon)} |u(x) - u(x_0)| \, dx = 0.$$

The relation between points of approximate continuity and Lebesgue points is given by the following proposition:

Proposition 2.2 *Let $u : \mathbb{R}^N \to \mathbb{R}$ be Lebesgue measurable. If $x_0 \in \mathbb{R}^N$ is a Lebesgue point for u, then u is approximately continuous at x_0. Conversely, if $u : \mathbb{R}^N \to \mathbb{R}$ is locally bounded and is approximately continuous at x_0, then $x_0 \in \mathbb{R}^N$ is a Lebesgue point for u.*

In the sequel we will use often the following result.

Proposition 2.3 *Let $u : \mathbb{R}^N \to \mathbb{R}$ be Lebesgue measurable, let $\varphi_\varepsilon = \varphi_\varepsilon(x)$ be a standard mollifier and define*

$$u_\varepsilon(x) := \int_\Omega \varphi_\varepsilon(x-y)\, u(y)\, dy.$$

(i) If $x_0 \in \mathbb{R}^N$ is a Lebesgue point for u, then $u_\varepsilon(x_0) \to u(x_0)$ as $\varepsilon \to 0^+$;
(ii) if u is differentiable at $x_0 \in \mathbb{R}^N$, then $\nabla u_\varepsilon(x_0) \to \nabla u(x_0)$ as $\varepsilon \to 0^+$.

Remark 2.4 By (i) and (ii) it follows that $\nabla u_\varepsilon(x_0) \to \nabla u_\varepsilon(x_0)$ as $\varepsilon \to 0^+$ if either u is differentiable at x_0 or if ∇u exist $\mathcal{L}^N$ a.e in a neighborhood of x_0 and x_0 is a Lebesgue point for ∇u. These two conditions do not seem to be related in general. Thus it would be interesting to find a condition weaker than the previous two and which would still imply that $\nabla u_\varepsilon(x_0) \to \nabla u(x_0)$ as $\varepsilon \to 0^+$. This would allow to improve significantly Theorem 1.5.

Proposition 2.5 *Let E be a Lebesgue measurable subset of $\mathbb{R}^N$ and G a Borel subset of $\mathbb{R}^d$. Let $g : E \times G \to \mathbb{R}$ be a Borel function such that for $\mathcal{L}^N$ a.e. $x \in E$ the function $g(x,\cdot)$ is continuous on G. Then there exists an $\mathcal{L}^N$-null set $\mathcal{N} \subset \mathbb{R}^N$ such that for every $\mathbf{u} \in G$ the function $g(\cdot,\mathbf{u})$ is approximately continuous in $E \backslash \mathcal{N}$.*

Proof. It can be obtained as in [31] (cf. the proof of Theorem 1.3) by using Scorza-Dragoni Theorem. □

The following result may be found in the paper of Serrin and Varberg [48].

Theorem 2.6 *(i) If $u : \mathbb{R} \to \mathbb{R}$ has finite derivative on a set G, with $u'(x) = 0$ for $\mathcal{L}^1$ a.e. $x \in G$, then $\mathcal{L}^1(u(G)) = 0$;*
(ii) if $u : \mathbb{R} \to \mathbb{R}$ has derivative (finite or infinite) on a set G, with $\mathcal{L}^1(u(G)) = 0$, then $u'(x) = 0$ for $\mathcal{L}^1$ a.e. $x \in G$.

The next result is due to Tonelli and may be found e.g. in [45].

Theorem 2.7 *Let $I \subset R$ be an interval and consider a function $\mathbf{w} : I \to \mathbb{R}^d$ of bounded variation. Let s be a length function for $\mathbf{w}$. Then*

(i) $s'(x) = |\mathbf{w}'(x)|$ for $\mathcal{L}^1$ a.e. $x \in I$;
(ii) $\mathcal{H}^1(\mathbf{w}(E)) \le \mathcal{L}^1(s(E))$ for every measurable set $E \subset I$.

In the previous theorem a *length function* s for $\mathbf{w}$ is any real function $s : I \to \mathbb{R}$ such that $s(x) - s(x')$ is the total variation of $\mathbf{w}$ over $[x, x']$ for all $x, x' \in I$ with $x < x'$.

For every $1 \le p, q \le \infty$ consider the space

$$L^{p,q}(\mathrm{div};\Omega) := \left\{u \in L^p(\Omega;\mathbb{R}^N) : \mathrm{div}\, u \in L^q(\Omega)\right\},$$

where $\mathrm{div}\, v$ is the distributional divergence, namely for every $\phi \in C_c^\infty(\Omega)$ we have

$$\int_\Omega \phi(x)\ \mathrm{div}\, u(x)\, dx = -\int_\Omega \nabla\phi(x) \cdot u(x)\, dx.$$

If $p = q$ we use the notation $L^p(\mathrm{div};\Omega)$. The space $L^{p,q}(\mathrm{div};\Omega)$ has been studied by several authors. We refer e.g. to [5,6] and [38] for more information and a detailed bibliography. As usual we set

$$L^{p,q}_{\mathrm{loc}}(\mathrm{div};\Omega) := \{u \in L^{p,q}(\mathrm{div};\Omega') \text{ for every } \Omega' \subset\subset \Omega\}.$$

Finally, we present two approximation results due respectively to De Giorgi, Buttazzo and Dal Maso [24] and to De Giorgi [23].

Proposition 2.8 *Let $\mu : \mathcal{B}(X) \to [0,\infty]$ be a Radon measure defined on a locally compact Hausdorff space X. Consider a sequence of Borel measurable functions $u_n : X \to [0,\infty]$. Then*

$$\int_X \sup_n u_n \, d\mu = \sup_{k\in\mathbb{N}} \left\{ \sum_{i=1}^k \int_{A_i} u_i \, d\mu : \ A_i \subset X \text{ open and pairwise disjoint} \right\}.$$

Let $f : \mathbb{R}^N \to \mathbb{R}$ be a convex function and $\alpha \in C^1_c\left(\mathbb{R}^N\right)$ any function with $\alpha \geq 0$ and $\int_{\mathbb{R}^N} \alpha(z)\, dz = 1$. Define

$$\begin{aligned} A_\alpha &:= \int_{\mathbb{R}^N} f(z)\left((N+1)\,\alpha(z) + \nabla\alpha(z)\cdot z\right) dz \\ B_\alpha &:= -\int_{\mathbb{R}^N} f(z)\,\nabla\alpha(z)\, dz. \end{aligned} \tag{2.1}$$

Theorem 2.9 (De Giorgi) *Let f, α, A_α and B_α be as above. Then*

(i) $f(\xi) \geq A_\alpha + B_\alpha \cdot \xi$ *for all* $\xi \in \mathbb{R}^N$;
(ii) $f(\xi) = \sup_{\beta\in\mathcal{B}} \{A_\beta + B_\beta\cdot\xi\}$ *for all* $\xi \in \mathbb{R}^N$, *where*

$$\mathcal{B} := \left\{\beta : \ \beta(\cdot) := k^N \alpha\left(k(q-\cdot)\right),\ k\in\mathbb{N},\ q\in\mathbb{Q}^N\right\};$$

(iii) $f(\xi) = \lim_{j\to\infty} f_j(\xi)$ *for all* $\xi \in \mathbb{R}^N$, *where* $f_j(\xi) := \sup_{i\leq j}\{A_{\beta_i} + B_{\beta_i}\cdot\xi\}$ *for all* $\xi\in\mathbb{R}^N$, *with* $\{\beta_i\}$ *an ordering of the class* $\mathcal{B}$.

Remark 2.10 If $f(\xi) \geq 0$ for all $\xi \in \mathbb{R}^N$ or $f(\xi) \geq C(|\xi| - 1)$ for all $\xi \in \mathbb{R}^N$ and for some constant $C > 0$, then we may assume that the approximating functions f_j in (iii) satisfy the same properties. Indeed it suffices to replace f_j respectively with $\max\{f_j, 0\}$ and $\max\{f_j, C(|\cdot| - 1)\}$.

3. Proof of Theorems 1.3 and 1.5

In this Section we prove Theorems 1.3, 1.5 and their corollaries. We begin with Theorem 1.3. As we already mentioned in the introduction the equivalences (i) $\Longleftrightarrow$ (ii) $\Longleftrightarrow$ (iii) follow from a result of Choquet (cf. Theorem 16 in [17]), while the implication (ii) $\Rightarrow$ (iv) is due to Marcus and Mizel. For the convenience of the reader we give the full proof.

Proof of Theorem 1.3. We begin by showing that (i) (resp. (iii)) is equivalent to require

$$\mathcal{H}^1(E \cap \mathbf{w}(I)) = 0 \tag{3.1}$$

for any Lipschitz function $\mathbf{w} : I \to \mathbb{R}^d$ (resp. any $\mathbf{w} \in BV(I;\mathbb{R}^d) \cap C(I;\mathbb{R}^d)$), where I is any subinterval of $\mathbb{R}$. We only prove this for (i), the proof for (iii) being similar. Assume that (i) holds and for any Lipschitz function $\mathbf{w} : [a,b] \to \mathbb{R}^d$ define $\mathbf{w}(x) := \mathbf{w}(b)$ for all $x \geq b$ and $\mathbf{w}(x) := \mathbf{w}(a)$ for all $x \leq a$. Then $\mathbf{w} : \mathbb{R} \to \mathbb{R}^d$ is a Lipschitz function with $\mathbf{w}(\mathbb{R}) = \mathbf{w}([a,b])$. Hence by (i)

$$\mathcal{H}^1(E \cap \mathbf{w}([a,b])) = \mathcal{H}^1(E \cap \mathbf{w}(\mathbb{R})) = 0.$$

Conversely if (3.1) holds and $\mathbf{w} : \mathbb{R} \to \mathbb{R}^d$ is a Lipschitz function, then

$$\mathcal{H}^1(E \cap \mathbf{w}(\mathbb{R})) \leq \sum_{k=1}^{\infty} \mathcal{H}^1(E \cap \mathbf{w}([-k,k])) = 0.$$

Next we prove the equivalence (i)$\Leftrightarrow$(iii). It clearly suffices to show the implication (i)$\Rightarrow$(iii). The proof follows closely Choquet (cf. Theorem 16 in [17]). Let $\mathbf{w} \in BV(I;\mathbb{R}^d) \cap C(I;\mathbb{R}^d)$. Without loss of generality we may assume that $I = [0,1]$. Assume first that $\mathbf{w}$ is not constant on any subinterval of I. For each $x \in I$ let $s(x)$ be the length of the rectifiable curve $\mathbf{w}$ from 0 to x. Let $L := s(1)$. Let $h : [0,L] \to [0,1]$ be the inverse of the function $s : [0,1] \to [0,L]$ and define

$$\mathbf{u}(s) := \mathbf{w}(h(s)) \quad s \in [0,L].$$

Then for $0 \leq s_1 < s_2 < L$ we have

$$\left|\frac{\mathbf{u}(s_2) - \mathbf{u}(s_1)}{s_2 - s_1}\right| = \left|\frac{\mathbf{w}(h(s_2)) - \mathbf{w}(h(s_1))}{s_2 - s_1}\right| = \left|\frac{\mathbf{w}(x_2) - \mathbf{w}(x_1)}{s(x_2) - s(x_1)}\right| \leq 1$$

where $x_2 = h(s_2)$ and $x_1 = h(s_1)$. Hence $\mathbf{u} : [0,L] \to \mathbb{R}^d$ is a Lipschitz function and so by (3.1) we have

$$\mathcal{H}^1(E \cap \mathbf{u}([0,L])) = \mathcal{H}^1(E \cap \mathbf{w}(I)) = 0.$$

In the general case when $\mathbf{w}$ may be constant on subintervals of I and so the function $s(x)$ is not strictly increasing, it suffices to replace $s(x)$ with $s(x) + x$.

Clearly (iii) in the form (3.1) implies (ii).

Next we prove the equivalence (i)$\Leftrightarrow$(iv). Assume that (iv) holds. We claim that E is purely $\mathcal{H}^1$-unrectifiable. Indeed, assume by contradiction that there exists a Lipschitz function $\mathbf{w} : \mathbb{R} \to \mathbb{R}^d$ such that

$$\mathcal{H}^1(E \cap \mathbf{w}(\mathbb{R})) > 0. \tag{3.2}$$

By assumption $\mathbf{w}'(x) = 0$ for $\mathcal{L}^1$ a.e. $x \in \mathbf{w}^{-1}(E)$. Let s be the length function for $\mathbf{w}$. By Theorem 2.7 we have that $s'(x) = 0$ for $\mathcal{L}^1$ a.e. $x \in \mathbf{w}^{-1}(E)$. Let

$$G := \left\{x \in \mathbf{w}^{-1}(E) : s' \text{ is finite}\right\}.$$

By Theorem 2.6 we have that

$$\mathcal{L}^1\left(s\left(G\right)\right)=0. \tag{3.3}$$

Since s is Lipschitz it maps $\mathcal{L}^1$ null sets into $\mathcal{L}^1$ null sets. Hence

$$\mathcal{L}^1\left(s\left(\mathbf{w}^{-1}(E)\setminus G\right)\right)=0. \tag{3.4}$$

By Theorem 2.7, (3.3) and (3.4) we have

$$\mathcal{H}^1\left(\mathbf{w}\left(\mathbf{w}^{-1}(E)\right)\right)\le\mathcal{L}^1\left(s\left(\mathbf{w}^{-1}(E)\right)\right)=0$$

which is a contradiction.

Conversely let E be purely $\mathcal{H}^1$-unrectifiable and consider $\mathbf{u}\in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^N;\mathbb{R}^d)$. Assume first that $\mathbf{u}$ is Lipschitz continuous and, following Marcus and Mizel (cf the proof of Lemma 2.1 in [41]) let $L\subset\mathbb{R}^N$ be a line parallel to the x_i-axis, $i=1,\cdots,N$. Then

$$\mathcal{H}^1(E\cap\mathbf{u}(L))=0.$$

Let $M=\mathbf{u}^{-1}(E\cap\mathbf{u}(L))\cap L$. Then $u_j\left(M\right)$ is an $\mathcal{L}^1$-null set for all $j=1,\cdots,d$ since the projection of the $\mathcal{H}^1$-null set $E\cap\mathbf{u}(L)$ on any coordinate axis in $\mathbb{R}^d$ is $\mathcal{L}^1$-null. Hence by Theorem 2.6

$$\frac{\partial u_j\left(x\right)}{\partial x_i}=0 \text{ for } \mathcal{L}^1 \text{ a.e. } x\in M,\quad j=1,\cdots,d.$$

Since this is true for all lines $L\subset\mathbb{R}^N$ parallel to the x_i-axis it follows from Fubini's Theorem that

$$\nabla\mathbf{u}\left(x\right)=0 \text{ for } \mathcal{L}^N \text{ a.e. } x\in\mathbf{u}^{-1}(E).$$

We now remove the additional hypothesis that $\mathbf{u}$ is Lipschitz continuous. By Theorem 8 p. 208 in [35] for each $n\in\mathbb{N}$ we may find a Lipschitz function $\mathbf{u}_n:\mathbb{R}^N\to\mathbb{R}^d$ such that

$$\mathcal{L}^N(\{x\in\mathbb{R}^N:\mathbf{u}_n(x)\ne\mathbf{u}(x)\})\le 1/n. \tag{3.5}$$

Let $E_n=\{x\in\mathbb{R}^N:\mathbf{u}_n(x)\ne\mathbf{u}(x)\}$. By the previous part of the proof we have $\nabla\mathbf{u}_n\left(x\right)=0$ for $\mathcal{L}^N$ a.e. $x\in\mathbf{u}_n^{-1}(E)$. Hence $\nabla\mathbf{u}\left(x\right)=0$ for $\mathcal{L}^N$ a.e. $x\in\mathbf{u}^{-1}(E)\setminus E_n$ for each $n\in\mathbb{N}$, that is $\nabla\mathbf{u}\left(x\right)=0$ for $\mathcal{L}^N$ a.e. $x\in\mathbf{u}^{-1}(E)\setminus\bigcap_{n=1}^{\infty}E_n$. But since

$$\mathcal{L}^N\left(\bigcap_{n=1}^{\infty}E_n\right)\le\mathcal{L}^N\left(E_n\right)\le\frac{1}{n}\to 0$$

it follows that $\nabla\mathbf{u}\left(x\right)=0$ for $\mathcal{L}^N$ a.e. $x\in\mathbf{u}^{-1}(E)$ and the proof is complete.

The implication (i)⇒(v) follows exactly as in the Sobolev case, with the only difference that to obtain (3.5) we use Theorem 5.34 and Remark 3.93 in [4]. □

Next we prove Theorem 1.5.

Proof of Theorem 1.5. Fix $\phi\in C^1_c(\Omega)$ and let $\varphi_\varepsilon=\varphi_\varepsilon\left(x\right)$ and $\psi_\delta=\psi_\delta\left(\mathbf{u}\right)$ be standard mollifiers. In what follows, for a given function $G\left(x,\mathbf{u}\right)$, we will use the notation G_ε (resp. G_δ) to denote the convolution of $G(\cdot,\mathbf{u})$ (resp. $G(x,\cdot)$) with φ_ε (resp. ψ_δ). Define

$$B_{\varepsilon\delta}(x,\mathbf{w}) := \int_\Omega \varphi_\varepsilon(x-y)\int_{\mathbb{R}^d}\psi_\delta(\mathbf{w}-\mathbf{z})\;B(y,\mathbf{z})\;d\mathbf{z}\,dy$$

for $x\in\Omega'$ and $\mathbf{w}\in\mathbb{R}^d$, where $\operatorname{supp}\phi\subset\subset\Omega'\subset\subset\Omega$, and $0<\varepsilon<\operatorname{dist}(\Omega',\partial\Omega)$. Let $\mathbf{u}\in W^{1,1}(\Omega;\mathbb{R}^d)\cap L^\infty_{\mathrm{loc}}(\Omega;\mathbb{R}^d)$ and define

$$\mathbf{v}_{\varepsilon\delta}(x) := B_{\varepsilon\delta}(x,\mathbf{u}(x)).$$

As $B_{\varepsilon\delta}\in C^\infty\left(\overline{\Omega'}\times\mathbb{R}^d;\mathbb{R}^N\right)$ and $\mathbf{u}\in W^{1,1}(\Omega;\mathbb{R}^d)\cap L^\infty_{\mathrm{loc}}(\Omega;\mathbb{R}^d)$ we have that $\mathbf{v}_{\varepsilon\delta}\in L^1(\operatorname{div};\Omega')$ and

$$\begin{aligned}
&-\int_{\Omega'}\nabla\phi(x)\cdot B_{\varepsilon\delta}(x,\mathbf{u}(x))\;dx\\
&=-\int_{\Omega'}\nabla\phi(x)\cdot\mathbf{v}_{\varepsilon\delta}(x)\,dx\\
&=\int_{\Omega'}\phi(x)\;\operatorname{div}\mathbf{v}_{\varepsilon\delta}(x)\,dx \qquad (3.6)\\
&=\int_{\Omega'}\phi(x)\;\left[\operatorname{tr}\left(\nabla_{\mathbf{u}}B_{\varepsilon\delta}(x,\mathbf{u}(x))\nabla\mathbf{u}(x)\right)+\operatorname{div}_x B_{\varepsilon\delta}(x,\mathbf{u}(x))\right]\;dx.
\end{aligned}$$

Since $\nabla_{\mathbf{u}}B_\delta(x,\cdot)$ is continuous by Proposition 2.5 we may find an $\mathcal{L}^N$ null set $\mathcal{N}_1\subset\Omega'$ such that for each $\mathbf{w}\in\mathbb{R}^d$ the function $\nabla_{\mathbf{u}}B_\delta(\cdot,\mathbf{w})$ is approximately continuous in $\Omega'\setminus\mathcal{N}_1$. Since $\nabla_{\mathbf{u}}B_\delta$ is locally bounded by (iv) it follows that each $x\in\Omega'\setminus\mathcal{N}_1$ is a Lebesgue point for $\nabla_{\mathbf{u}}B_\delta(\cdot,\mathbf{w})$. In turn

$$\nabla_{\mathbf{u}}B_{\varepsilon\delta}(x,\mathbf{w})\to\nabla_{\mathbf{u}}B_\delta(x,\mathbf{w})$$

as $\varepsilon\to0^+$ for each $x\in\Omega'\setminus\mathcal{N}_1$ and $\mathbf{w}\in\mathbb{R}^d$. By (iv) and Lebesgue Dominated Convergence Theorem we have

$$\lim_{\varepsilon\to0^+}\int_{\Omega'}\phi\;\operatorname{tr}\left(\nabla_{\mathbf{u}}B_{\varepsilon\delta}(x,\mathbf{u})\nabla\mathbf{u}\right)\;dx=\int_{\Omega'}\phi\;\operatorname{tr}\left(\nabla_{\mathbf{u}}B_\delta(x,\mathbf{u})\nabla\mathbf{u}\right)\;dx. \quad (3.7)$$

By hypothesis (iv) there exists $g\in L^1(\Omega')$ such that

$$|B(x,\mathbf{w})|+|\operatorname{div}_x B(x,\mathbf{w})|\le g(x) \quad (3.8)$$

for $\mathcal{L}^N$ a.e. $x\in\Omega'$ and for all $\mathbf{w}\in B(0;2R)$, where $R>\|\mathbf{u}\|_{L^\infty(\Omega')}$. Let $\{\mathbf{w}_n\}\subset B(0;R)$, $n\ge1$, be dense in $B(0;R)$ and let Ω'_0 and Ω'_n be respectively the set of Lebesgue points of g and of $\operatorname{div}_x B_\delta(\cdot,\mathbf{w}_n)$. Define $\mathcal{N}_2:=\Omega'\setminus\bigcap_{n=0}^\infty\Omega'_n$. Clearly $\mathcal{L}^N(\mathcal{N}_2)=0$. We claim that we have that

$$\operatorname{div}_x B_{\varepsilon\delta}(x,\mathbf{w})\to\operatorname{div}_x B_\delta(x,\mathbf{w}) \quad (3.9)$$

as $\varepsilon\to0^+$ for all $x\in\Omega'\setminus\mathcal{N}_2$ and for all $\mathbf{w}\in B(0;R)$. Indeed since

$$\operatorname{div}_x B_{\varepsilon\delta}(x,\mathbf{w})=\left(\operatorname{div}_x B_\delta(x,\mathbf{w})\right)_\varepsilon$$

and since

$$\nabla_{\mathbf{u}} \left(\operatorname{div}_x B_\delta \left(x, \mathbf{w} \right) \right) = \int_{\mathbb{R}^d} \nabla_{\mathbf{u}} \psi_\delta \left(\mathbf{w} - \mathbf{z} \right) \operatorname{div}_x B \left(x, \mathbf{z} \right) \, d\mathbf{z}$$

it follows from (3.8) that

$$\left| \nabla_{\mathbf{u}} \left(\operatorname{div}_x B_\delta \left(x, \mathbf{w} \right) \right) \right| \leq g(x) C_\delta$$

for $\mathcal{L}^N$ a.e. $x \in \Omega'$ and for all $\mathbf{w} \in B\left(0; R\right)$. In turn

$$\left| \operatorname{div}_x B_\delta \left(x, \mathbf{w} \right) - \operatorname{div}_x B_\delta \left(x, \mathbf{w}_1 \right) \right| \leq g(x) C_\delta \left| \mathbf{w} - \mathbf{w}_1 \right| \tag{3.10}$$

for $\mathcal{L}^N$ a.e. $x \in \Omega'$ and for all $\mathbf{w}, \mathbf{w}_1 \in B\left(0; R\right)$. By taking $\mathcal{N}_2$ larger, if necessary, we may assume, without loss of generality that (3.10) holds for all $x \in \Omega' \setminus \mathcal{N}_2$ and for all $\mathbf{w}$, $\mathbf{w}_1 \in B\left(0; R\right)$. Fix $x_0 \in \Omega' \setminus \mathcal{N}_2$ and $\mathbf{w} \in B\left(0; R\right)$. Since the set $\{\mathbf{w}_n\} \subset B\left(0; R\right)$ is dense in $B\left(0; R\right)$ we may find a subsequence (not relabelled) of $\{\mathbf{w}_n\}$ such that $\mathbf{w}_n \to \mathbf{w}$. Then by (3.10)

$$\begin{aligned} &\left| \operatorname{div}_x B_\delta \left(x_0, \mathbf{w} \right) - \operatorname{div}_x B_\delta \left(x, \mathbf{w} \right) \right| \leq \left| \operatorname{div}_x B_\delta \left(x_0, \mathbf{w}_n \right) - \operatorname{div}_x B_\delta \left(x, \mathbf{w}_n \right) \right| \\ &\quad + \left| \operatorname{div}_x B_\delta \left(x, \mathbf{w} \right) - \operatorname{div}_x B_\delta \left(x, \mathbf{w}_n \right) \right| + \left| \operatorname{div}_x B_\delta \left(x_0, \mathbf{w} \right) - \operatorname{div}_x B_\delta \left(x_0, \mathbf{w}_n \right) \right| \\ &\quad \leq \left| \operatorname{div}_x B_\delta \left(x_0, \mathbf{w}_n \right) - \operatorname{div}_x B_\delta \left(x, \mathbf{w}_n \right) \right| + 2g(x) C_\delta \left| \mathbf{w} - \mathbf{w}_n \right|. \end{aligned}$$

By averaging over $B\left(x_0, r\right)$ and letting $r \to 0^+$ in the previous inequality we get

$$\begin{aligned} \limsup_{r \to 0^+} \frac{1}{\mathcal{L}^N \left(B \left(x_0, r \right) \right)} \int_{B(x_0, r)} \left| \operatorname{div}_x B_\delta \left(x_0, \mathbf{w} \right) - \operatorname{div}_x B_\delta \left(x, \mathbf{w} \right) \right| \, dx \\ \leq 2g(x_0) C_\delta \left| \mathbf{w} - \mathbf{w}_n \right|, \end{aligned}$$

where we have used the fact that x_0 is a Lebesgue point of g and of $\operatorname{div}_x B_\delta\left(\cdot, \mathbf{w}_n\right)$. Since $\mathbf{w}_n \to \mathbf{w}$ letting $n \to \infty$ in the previous inequality proves (3.9). By (3.8), (3.9) and by Lebesgue Dominated Convergence Theorem we get

$$\lim_{\varepsilon \to 0^+} \int_{\Omega'} \phi \, \operatorname{div}_x B_{\varepsilon\delta} \left(x, \mathbf{u} \right) \, dx = \int_{\Omega'} \phi \, \operatorname{div}_x B_\delta \left(x, \mathbf{u} \right) \, dx. \tag{3.11}$$

Similarly we have

$$\lim_{\varepsilon \to 0^+} - \int_{\Omega'} \nabla \phi \cdot B_{\varepsilon\delta} \left(x, \mathbf{u} \right) \, dx = - \int_{\Omega'} \nabla \phi \cdot B_\delta \left(x, \mathbf{u} \right) \, dx. \tag{3.12}$$

Letting $\varepsilon \to 0^+$ in (3.6) and using (3.7), (3.11) and (3.12) yields

$$\begin{aligned} &- \int_{\Omega'} \nabla \phi \cdot B_\delta \left(x, \mathbf{u} \right) \, dx = \\ &\qquad \int_{\Omega'} \phi \, \left(\operatorname{tr} \left(\nabla_{\mathbf{u}} B_\delta \left(x, \mathbf{u} \right) \nabla \mathbf{u} \right) + \operatorname{div}_x B_\delta \left(x, \mathbf{u} \right) \right) \, dx. \end{aligned} \tag{3.13}$$

Since $B\left(x, \cdot\right)$ is differentiable in $\mathbb{R}^d \setminus \mathcal{M}$ we have that

$$\nabla_{\mathbf{u}} B_\delta \left(x, \mathbf{w} \right) \to \nabla_{\mathbf{u}} B \left(x, \mathbf{w} \right)$$

as $\delta \to 0^+$ for each $x \in \Omega' \setminus \mathcal{N}$ and $\mathbf{w} \in \mathbb{R}^d \setminus \mathcal{M}$, while $\nabla \mathbf{u}(x) = 0$ for $\mathcal{L}^N$ a.e. $x \in \mathbf{u}^{-1}(\mathcal{M})$. Hence, again by Lebesgue Dominated Convergence Theorem

$$\lim_{\delta \to 0^+} \int_{\Omega'} \phi \, \operatorname{tr}\left(\nabla_{\mathbf{u}} B_\delta\left(x, \mathbf{u}\right) \nabla \mathbf{u}\right) \, dx = \int_{\Omega'} \phi \, \operatorname{tr}\left(\nabla_{\mathbf{u}} B\left(x, \mathbf{u}\right) \nabla \mathbf{u}\right) \, dx,$$

where we have used the fact that $B(x, \cdot)$ is locally Lipschitz.

For every $x \in \Omega \setminus \mathcal{N}$ the function $B(x, \cdot)$ is continuous in $\mathbb{R}^d$, while the function $\operatorname{div}_x B(x, \cdot)$ is approximately continuous in $\mathbb{R}^d$. Hence, since for $\mathcal{L}^N$ a.e. fixed $x \in \Omega'$ the functions $B(x, \cdot)$ and $\operatorname{div}_x B(x, \cdot)$ are bounded in $B(0; R)$ by virtue of (3.8), we have that

$$B_\delta\left(x, \mathbf{w}\right) \to B\left(x, \mathbf{w}\right), \quad \operatorname{div}_x B_\delta(x, \mathbf{w}) \to \operatorname{div}_x B(x, \mathbf{w})$$

as $\delta \to 0^+$ for $\mathcal{L}^N$ a.e. $x \in \Omega' \setminus \mathcal{N}$ and for all $\mathbf{w} \in B(0; R)$. In turn

$$\lim_{\delta \to 0^+} - \int_{\Omega'} \nabla \phi \cdot B_\delta\left(x, \mathbf{u}\right) \, dx = - \int_{\Omega'} \nabla \phi \cdot B\left(x, \mathbf{u}\right) \, dx,$$
$$\lim_{\delta \to 0^+} \int_{\Omega'} \phi \, \operatorname{div}_x B_\delta\left(x, \mathbf{u}\right) \, dx = \int_{\Omega'} \phi \, \operatorname{div}_x B\left(x, \mathbf{u}\right) \, dx.$$

In conclusion it follows from (3.13) that

$$- \int_{\Omega'} \nabla \phi \cdot B\left(x, \mathbf{u}\right) \, dx = \int_{\Omega'} \phi \, \left(\operatorname{tr}\left(\nabla_{\mathbf{u}} B\left(x, \mathbf{u}\right) \nabla \mathbf{u}\right) + \operatorname{div}_x B\left(x, \mathbf{u}\right)\right) \, dx$$

for every $\phi \in C_c^1(\Omega)$. This concludes the proof. □

Proof of Corollary 1.7. Without loss of generality we may take $w \equiv 0$. It is easy to show that the function $B(x, u) := \int_0^u b(x, s) \, ds$ satisfies all the hypotheses of the previous theorem. □

Proof of Proposition 1.9. To prove (1.13) it is enough to show that from any subsequence of $\{u_n\}$ we may extract a further subsequence for which (1.13) holds. Hence, without loss of generality, we may assume that $u_n \to u$ $\mathcal{L}^N$ a.e. in Ω. Let $\mathbf{v}_n : \Omega \to \mathbb{R}^N$ defined by $\mathbf{v}_n(x) := \int_{u(x)}^{u_n(x)} b(x, s) \, ds$. By Corollary 1.7 we have that $\mathbf{v}_n$ belongs to $L^1_{\mathrm{loc}}(\operatorname{div};\Omega)$ and

$$\operatorname{div} \mathbf{v}_n(x) = \int_{u(x)}^{u_n(x)} \operatorname{div}_x b(x, s) \, ds + b(x, u_n(x)) \cdot \nabla u_n(x) - b(x, u(x)) \cdot \nabla u(x).$$

Fix $\phi \in C_c^1(\Omega)$. Since $\int_\Omega \phi \, \operatorname{div} \mathbf{v}_n \, dx = - \int_\Omega \nabla \phi \cdot \mathbf{v}_n \, dx$, we get

$$\int_\Omega \phi b(x, u_n) \cdot \nabla u_n dx = - \int_\Omega \nabla \phi \cdot \mathbf{v}_n \, dx + \int_\Omega \phi \, b(x, u) \cdot \nabla u \, dx \qquad (3.14)$$
$$- \int_\Omega \phi\left(x\right) \left[\int_{u(x)}^{u_n(x)} \operatorname{div}_x b(x, s) \, ds \right] dx.$$

Write the third term on the right side of (3.14) as

$$\iint_{\operatorname{supp}\phi\times[-M,M]} \operatorname{sgn}\left(u_n\left(x\right)-u\left(x\right)\right)\chi_{D_n}\left(x,s\right)\ \phi(x)\operatorname{div}_x b(x,s)\,ds\,dx,$$

where $D_n \subset \operatorname{supp}\phi \times [-M,M]$ is the set of all pairs (x,s) such that s belongs to the segment of endpoints $u_n(x)$, $u(x)$ and $M := \sup_n \|u_n\|_{L^\infty(\Omega)}$. Since

$$\begin{aligned}&\left|\chi_{D_n}\left(x,s\right)\ \phi(x)\operatorname{div}_x b(x,s)\right|\\ &\le \|\phi\|_{L^\infty(\Omega)}\left|\operatorname{div}_x b(x,s)\right| \in L^1\left(\operatorname{supp}\phi\times[-M,M]\right),\end{aligned}$$

it follows by Lebesgue Dominated Convergence Theorem that

$$\lim_{n\to\infty} -\int_\Omega \phi(x)\left[\int_{u(x)}^{u_n(x)} \operatorname{div}_x b(x,s)\,ds\right]dx = 0. \tag{3.15}$$

Let

$$M_\phi := \operatorname{esssup}\left\{\left|b\left(x,u\right)\right| :\ x\in\operatorname{supp}\phi,\ u\in[-M,M]\right\},$$

Since

$$|\mathbf{v}_n(x)| \le M_\phi |u_n(x)-u(x)|$$

for all $x \in \operatorname{supp}\phi$, by letting $n \to \infty$ in (3.14) we obtain the desired result also by (3.15). □

4. Proof of the lower semicontinuity theorems

Proof of Proposition 1.11 Step 1: We begin by assuming that $a\left(x,u\right) = 0$ and $b(x,u) = 0$ for $|u| \ge L$ for some $L > 0$. It is not difficult to see that for every $u \in W^{1,1}\left(\Omega\right)$ we have

$$\begin{aligned}&\int_\Omega \left(a(x,u)+b(x,u)\cdot\nabla u\right)^+\,dx\\ &\quad= \sup\left\{\int_\Omega [a(x,u)+b(x,u)\cdot\nabla u]\phi\,dx :\ \phi\in C_c^1(\Omega),\ 0\le\phi\le 1\right\}\end{aligned}$$

Fix $\phi \in C_c^1(\Omega)$ with $0 \le \phi \le 1$ and let $K := \operatorname{supp}\phi \times [-L,L]$. By hypothesis (1.14) there exists $h_K \in L^1(\Omega)$ such that $a(x,u) \ge h_K(x)$ for all $(x,u) \in K$. Since $a(x,u_n(x))\,\phi(x) \ge h_K(x)\,\phi(x)$ for all $x \in \Omega$ and all $n \in \mathbb{N}$ we may apply Fatou's Lemma to obtain that

$$\liminf_{n\to\infty}\int_\Omega a(x,u_n)\,\phi\,dx \ge \int_\Omega a(x,u)\,\phi\,dx.$$

Using Proposition 1.9 and Remark 1.10 we obtain

$$\begin{aligned}\liminf_{n\to\infty} F(u_n) &\ge \liminf_{n\to\infty}\int_\Omega [a(x,u_n)+b(x,u_n)\cdot\nabla u_n]\phi\,dx\\ &\ge \int_\Omega [a(x,u)+b(x,u)\cdot\nabla u]\phi\,dx\end{aligned}$$

and by taking the supremum over all ϕ, we get the desired result.

Step 2: We now remove the extra assumption that $a(x,u) = 0$ and $b(x,u) = 0$ for $|u| \geq L$. For every $k \in \mathbb{N}$ let

$$\sigma_k(u) := \begin{cases} 1 & |u| \leq k-1, \\ -|u| + k & k-1 < |u| \leq k, \\ 0 & |u| > k, \end{cases} \tag{4.1}$$

Then

$$\liminf_{n\to\infty} F(u_n) \geq \liminf_{n\to\infty} \int_\Omega \sigma_k(u_n) \left(a(x,u_n) + b(x,u_n)\cdot\nabla u_n\right)^+ dx. \tag{4.2}$$

Since the functions $a_k(x,u) := \sigma_k(u)\, a(x,u)$, and $b_k(x,u) := \sigma_k(u)\, b(x,u)$ satisfy all the hypotheses of Step 1 from (4.2) we get

$$\begin{aligned} \liminf_{n\to\infty} F(u_n) &\geq \int_\Omega \sigma_k(u) \left(a(x,u) + b(x,u)\cdot\nabla u\right)^+ dx \\ &\geq \int_{\{x\in\Omega:\, |u(x)|\leq k-1\}} \left(a(x,u) + b(x,u)\cdot\nabla u\right)^+ dx. \end{aligned}$$

By Lebesgue Monotone Convergence Theorem, letting $k \to \infty$ yields the desired result. □

We are now ready to prove Theorem 1.15

Proof of Theorem 1.15. By Theorem 2.9 there exists a sequence $\{\alpha_k\} \subset C_c^\infty(\mathbb{R}^N)$, with $\alpha_k \geq 0$ and $\int_{\mathbb{R}^N} \alpha_k(\xi)\, d\xi = 1$ such that for every $(x,u,\xi) \in \Omega \times \mathbb{R} \times \mathbb{R}^N$ we have

$$f(x,u,\xi) = \sup_{k\in\mathbb{N}} \left(a_k(x,u) + b_k(x,u)\cdot\xi\right)^+, \tag{4.3}$$

where

$$\begin{aligned} a_k(x,u) &:= \int_{\mathbb{R}^N} f(x,u,\xi)\left((N+1)\alpha_k(\xi) + \nabla\alpha_k(\xi)\cdot\xi\right) d\xi \\ b_k(x,u) &:= -\int_{\mathbb{R}^N} f(x,u,\xi)\nabla\alpha_k(\xi)\, d\xi. \end{aligned} \tag{4.4}$$

By Proposition 2.8 we get that

$$F(u) = \sup_{j\in\mathbb{N}} \left\{ \sum_{k=1}^{j} \int_{A_k} \left(a_k(x,u) + b_k(x,u)\cdot\nabla u\right)^+ dx : A_k \subset \Omega \text{ open, disjoint} \right\}.$$

Thus to prove the lower semicontinuity of the functional $F(u)$ it suffices to prove lower semicontinuity of each functional

$$\int_{A_k} \left(a_k(x,u) + b_k(x,u)\cdot\nabla u\right)^+ dx. \tag{4.5}$$

We claim that the functions defined in (4.4) satisfy all the hypotheses of Proposition 1.11. Indeed, by (i)-(iii) it follows from Lebesgue Dominated Convergence

Theorem that the functions $a_k(x,\cdot)$ and $b_k(x,\cdot)$ are continuous. Moreover by (iii) the functions a_k and b_k are locally bounded.

Finally, since by Lemma 2.2 in [7] we have that for each fixed $(x,u)\in\Omega\times\mathbb{R}$

$$\sup_{\xi\in B(0;r)}|\nabla_\xi f(x,u,\xi)|\le\frac{\sqrt{N}}{r}\operatorname{osc}\left(f(x,u,\cdot);B\left(0;2r\right)\right),\tag{4.6}$$

it follows from (iii) that $\nabla_\xi f\in L^1_{\mathrm{loc}}(\Omega\times\mathbb{R}\times\mathbb{R}^N;\mathbb{R}^N)$. Fix $\phi\in C^1_c(\Omega)$, then by (iv), for $\mathcal{L}^1$ a.e. $u\in\mathbb{R}$

$$\begin{aligned}\int_\Omega\nabla\phi\left(x\right)\cdot b_k(x,u)\,dx&=-\int_\Omega\nabla\phi\left(x\right)\cdot\int_{\mathbb{R}^N}f(x,u,\xi)\nabla\alpha_k(\xi)\,d\xi\,dx\\&=\int_\Omega\nabla\phi\left(x\right)\cdot\int_{\mathbb{R}^N}\nabla_\xi f(x,u,\xi)\,\alpha_k(\xi)\,d\xi\,dx\\&=\int_{\mathbb{R}^N}\alpha_k(\xi)\int_\Omega\nabla\phi\left(x\right)\cdot\nabla_\xi f(x,u,\xi)\,dx\,d\xi\\&=-\int_{\mathbb{R}^N}\alpha_k(\xi)\int_\Omega\phi\left(x\right)\operatorname{div}_x\nabla_\xi f(x,u,\xi)\,dx\,d\xi\\&=-\int_\Omega\phi\left(x\right)\int_{\mathbb{R}^N}\alpha_k(\xi)\operatorname{div}_x\nabla_\xi f(x,u,\xi)\,d\xi\,dx.\end{aligned}\tag{4.7}$$

Hence

$$\operatorname{div}_x b_k\left(x,u\right)=\int_{\mathbb{R}^N}\alpha_k(\xi)\operatorname{div}_x\nabla_\xi f(x,u,\xi)\,d\xi$$

and $\operatorname{div}_x b_k\in L^1_{\mathrm{loc}}\left(\Omega\times\mathbb{R}\right)$ again by (iv). □

To prove Theorem 1.16 we begin with the following proposition.

Proposition 4.1 *Let*

$$a:\Omega\times\mathbb{R}\to\mathbb{R},\quad b:\Omega\times\mathbb{R}\to\mathbb{R}^N$$

be two Borel functions. Assume that there exists an $\mathcal{L}^1$-null set $\mathcal{M}\subset\mathbb{R}$ such that $a(\cdot,u)$ and $b(\cdot,u)$ are continuous on Ω for every $u\in\mathbb{R}\setminus\mathcal{M}$. Suppose also that $a(x,u)\le0$ for every $x\in\Omega$ and $u\in\mathbb{R}$ and that b satisfies condition (ii) of Corollary 1.7. Then the functional

$$F(u):=\int_\Omega\left(a(x,u)+b(x,u)\cdot\nabla u\right)^+\,dx$$

is lower semicontinuous on $W^{1,1}(\Omega)$ with respect to the strong $L^1(\Omega)$ convergence.

Proof. *Step 1:* Assume first that the functions a and b are bounded. As in Corollary 2.5 in [2], by Scorza-Dragoni theorem (see [26]) we may find an increasing sequence K_i of compact subsets of $\mathbb{R}$ such that

$$\mathcal{L}^1\left(\mathbb{R}\setminus E\right)=0,\ \text{where}\ E:=\bigcup_{i\in\mathbb{N}}K_i,\tag{4.8}$$