A nonautonomous chain rule in $W^{1,p}$ and BV

Luigi Ambrosio[*], Graziano Crasta [†], Virginia De Cicco[‡], Guido De Philippis[§]

November 3, 2011

1 Introduction

The aim of this paper is to prove a generalization of the following chain rule formula in BV and in Sobolev spaces. Let $F\colon \mathbb{R}^h \to \mathbb{R}$ be a C^1 function with bounded gradient. It is well-known that, if $u \in BV_{\rm loc}(\mathbb{R}^n;\mathbb{R}^h)$, then the composite function $v(x) := F(u(x))$ belongs to $BV_{\rm loc}(\mathbb{R}^n)$ and the following chain rule formula holds in the sense of measures:

(i) (diffuse part) $\widetilde{D}v = \nabla F(\tilde{u})\widetilde{D}u$;

(ii) (jump part) $D^j v = [F(u^+) - F(u^-)]\nu_{J_u}\mathscr{H}^{n-1}\llcorner J_u$,

where $\widetilde{D}u$, $\tilde{u}$, J_u, ν_{J_u} and $u^\pm$ denote respectively the diffuse part of the measure Du, the precise representative of u, the jump set of u, the normal vector to J_u and the one-sided traces of u (see [2, Thm. 3.96]). The C^1 regularity of F can be omitted (requiring F to be only Lipschitz continuous), but in this case since the image of u might be a low-dimensional set the gradient ∇F appearing in (i) should be properly understood, see [1] and [8]; moreover, an analogous result holds true also in the vectorial case $F\colon \mathbb{R}^h \to \mathbb{R}^p$.

In recent years this formula has been generalized in order to deal with an explicit non-smooth dependence of F from the space variable x, especially in view to applications to semicontinuity results for integral functional (see [4, 5]) and to hyperbolic systems of conservation laws (see [3]).

The result proved in this paper goes in this direction: given $F\colon \mathbb{R}^n \times \mathbb{R}^h \to \mathbb{R}$, with $F(x,\cdot)$ of class $C^1(\mathbb{R}^h)$ for almost every $x \in \mathbb{R}^n$ and $F(\cdot,z) \in BV_{\rm loc}(\mathbb{R}^n)$ for all $z \in \mathbb{R}^h$, we establish the validity of a chain rule formula for the composite function $v(x) = F(x,u(x))$ with $u \in BV_{\rm loc}(\mathbb{R}^n;\mathbb{R}^h)$. We believe that, in our general setting, this formula can be a useful tool in the analysis of the problems mentioned above, in cases where the dependence from the space

[*]Scuola Normale Superiore, p.za dei Cavalieri 7, I-56126 Pisa, Italy, e–mail: l.ambrosio@sns.it

[†]Dipartimento di Matematica G. Castelnuovo, Univ. di Roma La Sapienza, P.le A. Moro 2, I-00185 Roma, Italy, e-mail: crasta@mat.uniroma1.it

[‡]Dipartimento di Scienze di Base e Applicate per l'Ingegneria, Univ. di Roma La Sapienza, Via A. Scarpa 10, I-00185 Roma, Italy, e-mail: virginia.decicco@sbai.uniroma1.it

[§]Scuola Normale Superiore, p.za dei Cavalieri 7, I-56126 Pisa, Italy, e–mail: guido.dephilippis@sns.it

variables is of BV type. In addition, our result provides also a chain rule in the case when u and $F(\cdot, z)$ are Sobolev, see (13).

The scalar case $h = 1$ has been considered in [4] and [5] in the case of a dependence of F with respect to x respectively of Sobolev type and of BV type.

The one-dimensional case $n = 1$ has been studied in [3], where a very explicit formula has been obtained at the price of some additional assumptions (see Remark 3.6 for a detailed comparison with the set of assumptions of the present paper).

In this paper we prove a chain rule formula in the general case $n \geq 1$ and $h \geq 1$ under some structural assumptions that now we briefly describe. According to the classical case mentioned at the beginning, the first additional assumption is a uniform bound on $\nabla_z F(x, z)$ (see (H1) below). Concerning the x-derivative, we need to require the existence of a Radon measure σ bounding from above all measures $|D_x F(\cdot, z)|$, uniformly with respect to $z \in \mathbb{R}^h$ (see assumption (H4) below). With these two bounds we can prove that for any $u \in BV_{\rm loc}$ the composite function $v(x) = F(x, u(x))$ belongs to $BV_{\rm loc}$ (see Lemma 3.7 and Remark 3.8). Moreover, we can show the existence of a countably $\mathscr{H}^{n-1}$-rectifiable set $\mathcal{N}_F$, independent of u and containing the jump set of $F(\cdot, z)$ for every $z \in \mathbb{R}^h$, such that the jump set of v is contained in $\mathcal{N}_F \cup J_u$.

On the other hand, in order to prove the validity of the chain rule formula we require that F satisfies other structural assumptions related to the uniform continuous dependence of $\nabla_z F$ and $\widetilde{D}_x F$ with respect to z (see assumptions (H2) and (H3) below). All these assumptions are enough in order to prove the following chain rule formula:

(i) (diffuse part) $|Dv| \ll \sigma + |Du|$ and, for any Radon measure μ such that $\sigma + |Du| \ll \mu$, it holds

$$\frac{d\widetilde{D}v}{d\mu} = \frac{d\widetilde{D}_x F(\cdot, \tilde{u}(x))}{d\mu} + \nabla_z \tilde{F}(x, \tilde{u}(x)) \frac{d\widetilde{D}u}{d\mu}.$$

(ii) (jump part) $D^j v = \big(F^+(x, u^+(x)) - F^-(x, u^-(x)\big)\nu_{\mathcal{N}_F \cup J_u} \mathscr{H}^{n-1} \llcorner (\mathcal{N}_F \cup J_u)$ in the sense of measures.

In particular, when u and $F(\cdot, z)$ belong to a Sobolev space, we obtain that the composite function v belongs to the same Sobolev space and the following formula holds:

$$\nabla v(x) = \nabla_x F(x, u(x)) + \nabla_z F(x, u(x)) \nabla u(x) \qquad \mathscr{L}^n\text{-a.e. in } \mathbb{R}^n \tag{1}$$

(see [6] for a analogous result in the more general context of vector fields with L^1 divergence).

In order to prove (i) we use a blow-up argument as in the proof of Theorem 3.96 in [2], which allows to treat at the same time all higher dimensional cases $n \geq 1$ and $h \geq 1$ (it this respect we recall that the approach based on convolutions does not seem to work in the case $h > 1$).

The explicit non-smooth dependence of F with respect to x gives rise to several major technical difficulties. It is then crucial to firstly investigate some fine properties of the function F (see Section 4). For example, we show that $\mathscr{H}^{n-1}$-a.e. in $\mathcal{N}_F$ the one-sided traces $F^{\pm}(\cdot, z)$ exist for all $z \in \mathbb{R}^h$ (see Proposition 4.2).

2 Notation and preliminary results

In this section we recall our main notation and preliminary facts on BV functions. A general reference is Chapter 3 of [2], and occasionally we will give more precise references therein.

We denote by $\mathscr{L}^n$ the Lebesgue measure in $\mathbb{R}^n$ and by $\mathscr{H}^k$ the k-dimensional Hausdorff measure. The restriction of $\mathscr{H}^k$ to a set A will be denoted by $\mathscr{H}^k \llcorner A$, so that $\mathscr{H}^k \llcorner A(B) = \mathscr{H}^k(A \cap B)$. By $⨍_A$ we mean averaged integral on a set A. By Radon measure we mean a nonnegative Borel measure finite on compact sets.

We say that $f \in L^1(\mathbb{R}^n)$ belongs to $BV(\mathbb{R}^n)$ if its derivative in the sense of distributions is representable by a vector-valued measure $Df = (D_1 f, \ldots, D_n f)$ whose total variation $|Df|$ is finite, i.e.

$$\int_{\mathbb{R}^n} f \frac{\partial \phi}{\partial x_i} dx = -\int \phi \, D_i f \qquad \forall \phi \in C_c^\infty(\mathbb{R}^n),\ i = 1, \ldots, n$$

and $|Df|(\mathbb{R}^n) < \infty$. The $BV_{\rm loc}$ definition is analogous, requiring that $|Df|$ is a Radon measure in $\mathbb{R}^n$.

Approximate continuity and jump points. We say that $x \in \mathbb{R}^n$ is an approximate continuity point of f if, for some $z \in \mathbb{R}$, it holds

$$\lim_{r \downarrow 0} ⨍_{B_r(x)} |f(y) - z| dy = 0.$$

The number z is uniquely determined at approximate continuity points and denoted by $\tilde{f}(x)$, the so-called approximate limit of f at x. The complement of the set of approximate continuity points, the so-called singular set of f, will be denoted by S_f.

Analogously, we say that x is a jump point of f, and we write $x \in J_f$, if there exists a unit vector $\nu \in \mathbf{S}^{n-1}$ and f^+, $f^- \in \mathbb{R}$ satisfying $f^+ \neq f^-$ and

$$\lim_{r \downarrow 0} ⨍_{B^\pm(x,r)} |f(y) - f^\pm| dy = 0,$$

where $B^\pm(x,r) := \{y \in B_r(x) : \pm\langle y - x, \nu\rangle \geq 0\}$ are the two half balls determined by ν. At points $x \in J_f$ the triplet (f^+, f^-, ν) is uniquely determined up to a permutation of (f^+, f^-) and a change of sign of ν; for this reason, with a slight abuse of notation, we do not emphasize the ν dependence of $f^\pm$ and $B^\pm(x,r)$. Since we impose $f^+ \neq f^-$, it is clear that $J_f \subset S_f$. Moreover

$$J_f \subset \Big\{x : \limsup_{r \downarrow 0} \frac{|Df|(B_r(x))}{\omega_{n-1} r^{n-1}} > 0\Big\} \tag{2}$$

and

$$\mathscr{H}^{n-1}\Big(\Big\{x : \limsup_{r \downarrow 0} \frac{|Df|(B_r(x))}{\omega_{n-1} r^{n-1}} > 0\Big\} \setminus J_f\Big) = 0, \tag{3}$$

see the proof of [2, Lemma 3.75].

Recall that a set Σ is said to be countably $\mathscr{H}^{n-1}$ rectifiable if $\mathscr{H}^{n-1}$-almost all of Σ can be covered by a sequence of C^1 hypersurfaces. For any $BV_{\rm loc}$ function f, the set S_f is countably $\mathscr{H}^{n-1}$ rectifiable and $\mathscr{H}^{n-1}(S_f \setminus J_f) = 0$.

Decomposition of the distributional derivative.

For any oriented and countably $\mathscr{H}^{n-1}$-rectifiable $\Sigma \subset \mathbb{R}^n$ we have

$$Df \llcorner \Sigma = (f^+ - f^-)\nu_\Sigma \mathscr{H}^{n-1} \llcorner \Sigma. \tag{4}$$

For any $f \in BV(\mathbb{R}^n)$, we can decompose Df as the sum of a diffuse part, that we shall denote $\widetilde{D}f$, and a jump part, that we shall denote by $D^j f$. The diffuse part is characterized by the property that $|\widetilde{D}f|(B) = 0$ whenever $\mathscr{H}^{n-1}(B)$ is finite, while the jump part is concentrated on a set σ-finite with respect to $\mathscr{H}^{n-1}$. The diffuse part can be then split as

$$\widetilde{D}f = D^a f + D^c f$$

where $D^a f$ is the absolutely continuous part with respect to the Lebesgue measure, while $D^c f$ is the so-called Cantor part. The density of Df with respect to $\mathscr{L}^n$ can be represented as follows

$$D^a f = \nabla f \, d\mathscr{L}^n, \tag{5}$$

where ∇f is the approximate gradient of f, i.e. the only vector such that

$$\lim_{r\downarrow 0} \frac{1}{r^{n+1}} \int_{B_r(x)} |f(y) - f(x) - \nabla f(x) \cdot (y - x)| = 0 \qquad \text{for almost every } x, \tag{6}$$

see [2, Proposition 3.71 and Theorem 3.83].

The jump part can be easily computed by taking $\Sigma = J_f$ (or, equivalently, S_f) in (4), namely

$$D^j f = Df \llcorner J_f = (f^+ - f^-)\nu_{J_f} \mathscr{H}^{n-1} \llcorner J_f. \tag{7}$$

All these concepts and results extend, mostly arguing component by component, to vector-valued BV functions, see [2] for details.

3 The chain rule

Let $F\colon \mathbb{R}^n \times \mathbb{R}^h \to \mathbb{R}$ be satisfying:

(a) $x \mapsto F(x, z)$ belongs to $BV_{\rm loc}(\mathbb{R}^n)$ for all $z \in \mathbb{R}^h$;

(b) $z \mapsto F(x, z)$ is continuously differentiable in $\mathbb{R}^h$ for almost every $x \in \mathbb{R}^n$.

We will use the notation $\nabla_z F(x, z)$ to denote the (continuous) gradient of $z \mapsto F(x, z)$ and $D_x F(\cdot, z)$ to denote the distributional gradient of $x \mapsto F(x, z)$. We will use the notation C_F to denote a Lebesgue negligible set of points such that $F(x, \cdot)$ is C^1 for all $x \in \mathbb{R}^n \setminus C_F$.

We assume throughout this paper that F satisfies, besides (a) and (b), the following *structural assumptions*:

(H1) For some constant M, $|\nabla_z F(x, z)| \leq M$ for all $x \in \mathbb{R}^n \setminus C_F$ and $z \in \mathbb{R}^h$.

(H2) For any compact set $H \subset \mathbb{R}^h$ there exists a modulus of continuity $\tilde{\omega}_H$ independent of x such that

$$|\nabla_z F(x,z) - \nabla_z F(x,z')| \le \tilde{\omega}_H(|z-z'|)$$

for all z, $z' \in H$ and $x \in \mathbb{R}^n \setminus C_F$.

(H3) For any compact set $H \subset \mathbb{R}^h$ there exist a positive Radon measure λ_H and a modulus of continuity ω_H such that

$$|\widetilde{D}_x F(\cdot,z)(A) - \widetilde{D}_x F(\cdot,z')(A)| \le \omega_H(|z-z'|)\lambda_H(A)$$

for all z, $z' \in H$ and $A \subset \mathbb{R}^n$ Borel.

(H4) The measure

$$\sigma := \bigvee_{z \in \mathbb{R}^h} |D_x F(\cdot,z)|, \tag{8}$$

(where $\bigvee$ denotes the least upper bound in the space of nonnegative Borel measures) is finite on compact sets, i.e. it is a Radon measure.

In connection with (H4), notice that an equivalent formulation of it would be to require the existence of a Radon measure θ bounding from above all measures $|D_x F(\cdot,z)|$. Taking the least possible θ has some advantages, as the following remark shows.

Remark 3.1. The measure σ in (8) vanishes on every $\mathscr{H}^{n-1}$-negligible and on every purely $(n-1)$-unrectifiable set (i.e. a set B such that $\mathscr{H}^{n-1}(B \cap \Gamma) = 0$ whenever Γ is countably $\mathscr{H}^{n-1}$-rectifiable). Indeed, these properties are valid for all the measures $|D_x F(\cdot,z)|$ thanks to the coarea formula, see for instance [2, Theorem 3.40].

We can now canonically build a countably $\mathscr{H}^{n-1}$-rectifiable set B_σ containing all jump sets of $F(\cdot,z)$ as follows. Indeed, we define

$$B_\sigma = \big\{x : \limsup_{r\downarrow 0} \frac{\sigma(B_r(x))}{\omega_{n-1}r^{n-1}} > 0\big\}. \tag{9}$$

Writing B_σ as the union of the sets

$$\big\{x \in B_k(0) : \limsup_{r\downarrow 0} \frac{\sigma(B_r(x))}{\omega_{n-1}r^{n-1}} > \frac{1}{k}\big\} \qquad k \ge 1$$

it is immediate to check that B_σ is σ-finite with respect to $\mathscr{H}^{n-1}$. Now, according to [7, Page 252], we can split B_σ as a disjoint union $B_\sigma = L \cup U$ with L countably $\mathscr{H}^{n-1}$-rectifiable and U purely $(n-1)$-unrectifiable. In order to prove the rectifiability of B_σ, we are thus led to show that $\mathscr{H}^{n-1}(U) = 0$. To see this notice that

$$\mathscr{H}^{n-1}\left(U \cap \left\{\limsup_{r\downarrow 0} \frac{\sigma(B_r(x))}{\omega_{n-1}r^{n-1}} \ge \varepsilon\right\}\right) \le \frac{1}{\varepsilon}\sigma(U) = 0$$

thanks to Remark 3.1. By (2) and the inequality $\sigma \ge |D_x F(\cdot,z)|$ we know that $B_\sigma \supset J_{F(\cdot,z)}$ for all $z \in \mathbb{R}^h$.

The main result of the paper is the following chain rule.

Theorem 3.2. *Let F be satisfying (a), (b), (H1)-(H2)-(H3)-(H4) above. Then there exists a countably $\mathscr{H}^{n-1}$-rectifiable set $\mathcal{N}_F$ such that, for any function $u \in BV_{\rm loc}(\mathbb{R}^n;\mathbb{R}^h)$, the function $v(x) := F(x,u(x))$ belongs to $BV_{\rm loc}(\mathbb{R}^n)$ and the following chain rule holds:*

(i) (diffuse part) $|Dv| \ll \sigma + |Du|$ and, for any Radon measure μ such that $\sigma + |Du| \ll \mu$, it holds

$$\frac{d\widetilde{D}v}{d\mu} = \frac{d\widetilde{D}_x F(\cdot,\tilde{u}(x))}{d\mu} + \nabla_z \tilde{F}(x,\tilde{u}(x))\frac{d\widetilde{D}u}{d\mu} \qquad \mu\text{-a.e. in } \mathbb{R}^n. \tag{10}$$

(ii) (jump part) $J_v \subset \mathcal{N}_F \cup J_u$ and, denoting by $u^\pm(x)$ and $F^\pm(x,z)$ the one-sided traces of u and $F(\cdot,z)$ induced by a suitable orientation of $\mathcal{N}_F \cup J_u$, it holds

$$D^j v = \big(F^+(x,u^+(x)) - F^-(x,u^-(x))\big)\nu_{\mathcal{N}_F\cup J_u}\mathscr{H}^{n-1} \llcorner (\mathcal{N}_F \cup J_u) \tag{11}$$

in the sense of measures.

Moreover for a.e. x the map $y \mapsto F(y,u(x))$ is approximately differentiable at x and

$$\nabla v(x) = \nabla_x F(x,u(x)) + \nabla_z F(x,u(x))\nabla u(x) \qquad \mathscr{L}^n\text{-a.e. in } \mathbb{R}^n\,. \tag{12}$$

Here (and in the sequel) the expression

$$\frac{d\widetilde{D}_x F(\cdot,\tilde{u}(x))}{d\mu}$$

means the pointwise density of the measure $\widetilde{D}_x F(\cdot,z)$ with respect to μ, computed choosing $z = \tilde{u}(x)$ (notice that the composition is Borel measurable thanks to the Scorza-Dragoni Theorem and Lemma 3.9 below). Analogously, the expression $\tilde{F}(x,z)$ is well defined at points x such that $x \notin S_{F(\cdot,z)}$ and we will prove that, $\nabla_z \tilde{F}(x,z)$ is well defined for all z out of a countably $\mathscr{H}^{n-1}$-rectifiable set of points x.

Remark 3.3. It is an easy exercise of measure theory to see that (10) holds for *any* Radon measure μ such that $\sigma + |Du| \ll \mu$ if and only if it holds for *one* such measure. For this reason, we shall prove the formula with $\mu = \sigma + M|Du|$.

In the Sobolev framework the following chain rule holds (see [6] for a similar result relative to vector fields with controlled divergence.

Corollary 3.4. *Let F be satisfying (b), (H1), (H2) above and the following conditions:*

(a)' The function $x \mapsto F(x,z)$ belongs to $W^{1,1}_{loc}(\mathbb{R}^n)$ for all $z \in \mathbb{R}^h$;

(H3)' For any compact set $H \subset \mathbb{R}^h$ there exist a positive function $g_H \in L^1_{\rm loc}(\mathbb{R}^n)$ and a modulus of continuity ω_H such that

$$|\nabla_x F(x,z) - \nabla_x F(x,z')| \le \omega_H(|z-z'|)g_H(x)$$

for all z, $z' \in H$ and for a.e. $x \in \mathbb{R}^n$;

(H4)' There exists a positive function $g \in L^1_{\rm loc}(\mathbb{R}^n)$ such that

$$|\nabla_x F(x,z)| \le g(x)$$

for all $z \in \mathbb{R}^h$ and for a.e. $x \in \mathbb{R}^n$.

Then for any function $u \in W^{1,1}_{\rm loc}(\mathbb{R}^n;\mathbb{R}^h)$ the function $v(x) := F(x,u(x))$ belongs to $W^{1,1}_{\rm loc}(\mathbb{R}^n)$ and

$$\nabla v(x) = \nabla_x F(x,u(x)) + \nabla_z F(x,u(x))\nabla u(x) \qquad \mathscr{L}^n\text{-a.e. in } \mathbb{R}^n. \tag{13}$$

Remark 3.5 (Locally bounded functions in domains Ω). We stated the results for globally defined functions u, but obviously it extends to the case when the domain is an open set Ω. In addition, if u is locally bounded in Ω, then Theorem 3.2 holds true replacing all the assumptions with their local counterpart, we show for instance how (H1) and (H4) should be modified:

(H1-loc) For any pair of compact sets $K \subset \Omega$ and $H \subset \mathbb{R}^h$ there exists a constant $M_{K,H}$ such that $|\nabla_z F(x,z)| \le M_{K,H}$ for every $x \in K \setminus C_F$ and $z \in H$.

(H4-loc) For any compact set $H \subset \mathbb{R}^h$ the measure

$$\sigma_H := \bigvee_{z\in H} |D_x F(\cdot,z)| \tag{14}$$

is a Radon measure in Ω.

In this local case we can define $\mathcal{N}_F := \cup_{j\ge 1}\mathcal{N}_j$, where

$$\mathcal{N}_j := \left\{x : \limsup_{r\downarrow 0} \frac{\sigma_j(B_r(x))}{\omega_{n-1} r^{n-1}} > 0\right\}, \qquad \sigma_j := \bigvee_{z\in B_j(0)} |D_x F(\cdot,z)|,$$

and, in the any subdomain $\Omega' \Subset \Omega$ where $|u| \le j$, the measure σ in Theorem 3.2(i) should be replaced by σ_j.

Remark 3.6. In [3] the one-dimensional analogous (i.e. for $n = 1$) of Theorem 3.2 has been proved under the following structural assumptions:

(A1) There exists a countable set $\mathcal{N}_F$ containing $S_{F(\cdot,z)}$ for every $z \in \mathbb{R}^h$.

(A2) For every compact set $H \subset \mathbb{R}^h$ there exists a Radon measure λ_H such that

$$|D_x F(\cdot,z)(A) - D_x F(\cdot,z')(A)| \le |z-z'|\lambda_H(A)$$

for all $z, z' \in \mathbb{R}^h$ and every Borel set $A \subset \mathbb{R}$.

(A3) For every compact set $H \subset \mathbb{R}^h$ there exists a constant M_H such that $|\nabla_z F(x,z)| \le M_H$ for every $x \in \mathbb{R} \setminus \mathcal{N}_F$ and every $z \in H$.

(A4) $x \mapsto \nabla_z F(x,z)$ belongs to $BV(\mathbb{R})$ for every $z \in \mathbb{R}^h$.

(A5) There exists a positive finite Cantor measure λ (i.e., a measure whose diffuse part is orthogonal to the Lebesgue measure) such that $D^c_x F(\cdot, z) \ll \lambda$ for every $z \in \mathbb{R}^h$.

It is apparent that the new assumption (H4) allows to drop both assumptions (A1) (see the construction of the set B_σ in (9)) and (A5). Moreover, assumption (H3) involves only the diffuse part of the measure $|D_x F(\cdot, z)|$, and so it is weaker than (A2). The other assumptions are almost equivalent (even the "local" assumption (A3) is equivalent to (H1) since a BV function of the real line is locally bounded).

We first prove that under assumptions (H1) and (H4), the composite function v belongs to $BV_{\rm loc}(\mathbb{R}^n)$.

Lemma 3.7. *The function v belongs to $BV_{\rm loc}(\mathbb{R}^n)$ and $|Dv| \le \sigma + M|Du|$.*

Proof. It is obvious that $v \in L^1_{\rm loc}(\mathbb{R}^n)$, so we only have to check that $|Dv|(K) < \infty$ for any compact set K. In order to do this choose a standard family of mollifiers ϱ_ε and define

$$F_\varepsilon(x, z) = F(\cdot, z) * \varrho_\epsilon(x) = \int \varrho_\varepsilon(x - x')F(x', z)dx'$$

which is a C^1 function satisfying $\sup_{(x,z)} |\nabla_z F_\varepsilon(x, z)| \le M$ and set $v_\varepsilon(x) = F_\varepsilon(x, u(x))$. If $u \in C^1$ the standard chain rule and the inequality

$$\sup_z |\nabla_x F_\varepsilon(x, z)| \le \sigma * \varrho_\epsilon(x)$$

give, for any compact set K,

$$\begin{aligned} \int_K |\nabla v_\epsilon|\, dx &\le \int_K |\nabla_x F_\varepsilon(x, u(x))|\, dx + \int_K |\nabla_z F_\varepsilon(x, u(x))||\nabla u(x)|\, dx \\ &\le \int_K \sigma * \varrho_\epsilon(x)\, dx + M \int_K |\nabla u|\, dx. \end{aligned}$$

By approximation the same inequality holds, now with $|Du|(K)$ in place of $\int_K |\nabla u|\, dx$, if $u \in BV_{\rm loc}$. Eventually we can use the arbitrariness of K to get $|Dv_\varepsilon| \le \sigma * \varrho_\epsilon + M|Du|$.

The only thing to check is that $v_\varepsilon \to v$ in $L^1_{\rm loc}(\mathbb{R}^n)$, to do this it suffices to check the pointwise convergence since

$$|v_\varepsilon(x)| \le |F(x, 0)| * \varrho_\varepsilon + M|u| * \varrho_\varepsilon(x)$$

and the right-hand side is convergent in $L^1_{\rm loc}$. We know that for fixed $z \in \mathbb{R}^h$ it holds that

$$\lim_{\varepsilon \to 0} F_\varepsilon(x, z) = F(x, z)$$

for every $x \in A_z$ where A_z is the set of Lebesgue points of $x \to F(x, z)$. To prove the almost everywhere convergence of $F_\varepsilon(x, u(x)) \to F(x, u(x))$ it is thus enough to show that

$$\mathscr{L}^n\Big(\bigcup_{z \in \mathbb{R}^h} (\mathbb{R}^n \setminus A_z)\Big) = 0.$$

To see this let us choose a countable dense set $U \subset \mathbb{R}^h$; we claim that

$$\bigcap_{z\in U} A_z = \bigcap_{z\in\mathbb{R}^h} A_z.$$

Indeed, if $x \in \bigcap\limits_{z\in U} A_z$, $z \in \mathbb{R}^h$ and $z_k \in U$ converges to z, then for all $0 < r < R$ it holds

$$\begin{aligned}
&\Big| \fint_{B_R(x)} F(y,z)\,dy - \fint_{B_r(x)} F(y,z)\,dy\Big| \le \fint_{B_R(x)} |F(y,z) - F(y,z_k)|\,dy \\
&+ \Big| \fint_{B_R(x)} F(y,z_k)\,dy - \fint_{B_r(x)} F(y,z_k)\,dy\Big| + \fint_{B_r(x)} |F(y,z) - F(y,z_k)|\,dy \\
&\le 2M|z_k - z| + \Big| \fint_{B_R(x)} F(y,z_k)\,dy - \fint_{B_r(x)} F(y,z_k)\,dy\Big|.
\end{aligned}$$

This proves that the averages $\fint_{B_r(x)} F(y,z)\,dy$ are Cauchy as $r \downarrow 0$ for every $x \in \bigcap\limits_{z\in U} A_z$ and $z \in \mathbb{R}^h$. Denoting by $\tilde F$ the limit, in the same way one can prove that

$$\fint_{B_r(x)} |F(y,z) - \tilde F(x,z)|\,dy \to 0.$$

The claim now immediately follows. □

Remark 3.8. Even in the Sobolev case the assumption (H1) is needed in order to have $v \in W^{1,1}_{\mathrm{loc}}$. For instance, if $F(x,z) = f(x)z$, with $f \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ not locally bounded, the composite function $v(x) = f(x)u(x)$ may not be locally integrable, unless u is locally bounded. But, even in that case, the term $f\nabla u$ might be not locally integrable.

Lemma 3.9. *Any Radon measure μ in $\mathbb{R}^n$ is concentrated on a Borel set A with the following property:*

$$\frac{d\widetilde{D}_x F(\cdot,z)}{d\mu}(y) = \lim_{r\downarrow 0} \frac{\widetilde{D}_x F(\cdot,z)(B_r(y))}{\mu(B_r(y))} \tag{15}$$

exists for every $y \in A$, $z \in \mathbb{R}^h$ and is Borel-measurable in y and continuous in z, more precisely it holds

$$\left| \frac{d\widetilde{D}_x F(\cdot,z_1)}{d\mu}(y) - \frac{d\widetilde{D}_x F(\cdot,z_2)}{d\mu}(y) \right| \le \omega_H(|z_1 - z_2|)\frac{d\lambda_k}{d\mu}(y)$$

for any z_1, $z_2 \in H$, $H \subset \mathbb{R}^h$ compact.

Proof. Let us fix a compact $H \subset \mathbb{R}^h$, we will show that statement of the lemma holds for any $z \in H$ and $y \in A_H$ with $\mu(\mathbb{R}^n \setminus A_H) = 0$. Then, writing $\mathbb{R}^h$ as a countable union of compact sets H_k, we obtain the thesis with $A := \bigcap_k A_{H_k}$. In the sequel H will be fixed and when referring to the measure λ_H and the modulus ω_H appearing in (H3) we will drop the subscript.

Let $U \subset H$ be a countable dense set, then there exists a set A with $\mu(\mathbb{R}^n \setminus A) = 0$ such that for every $z \in U$ and every $y \in A$ the following limit exists

$$f(y,z) := \lim_{r \downarrow 0} \frac{\widetilde{D}_x F(\cdot, z)(B_r(y))}{\mu(B_r(y))}.$$

Moreover, possibly removing from A a μ-negligible set and using Besicovitch differentiation theorem, we can assume that for every $y \in A$ there exists the limit $\lim_r \lambda(B_r(y))/\mu(B_r(y))$ and coincides with a fixed version of the Radon-Nikodým derivative $d\lambda/d\mu$. We claim that any $z \in H$ has the required property: indeed, for any such point y we have

$$|f(y,z_1) - f(y,z_2)| \leq \omega(|z_1 - z_2|) \frac{d\lambda}{d\mu}(y)$$

for every $z_1, z_2 \in U$. Let choose now $z \in H$ and $z_k \in U$, $z_k \to z$, and define $f(y,z) = \lim_k f(y,z_k)$, which exists and does not depend on (z_k). Then

$$\begin{aligned} &\left| \frac{\widetilde{D}_x F(\cdot, z)(B_r(y))}{\mu(B_r(y))} - f(y,z) \right| \\ &\leq \left| \frac{\widetilde{D}_x F(\cdot, z_k)(B_r(y))}{\mu(B_r(y))} - f(y,z_k) \right| + \omega(|z_k - z|) \frac{\lambda(B_r(y))}{\mu(B_r(y))} + |f(y,z_k) - f(y,z)|. \end{aligned}$$

Passing to the limit first as $r \downarrow 0$ and then as $k \to \infty$ we obtain the thesis. Borel measurability in x and continuity in z easily follow. □

4 Fine properties of F

The proof of the following lemma, concerning differentiability of integrals depending on a parameter, is a direct consequence of the dominated convergence theorem.

Lemma 4.1. *If $A \subset \mathbb{R}^n$ is a bounded measurable set, the function*

$$z \mapsto m_A(z) := \fint_A F(x,z)dx$$

is continuously differentiable in $\mathbb{R}^h$ and with bounded gradient given by

$$\nabla_z m_A(z) = \fint_A \nabla_z F(x,z)dx.$$

In particular $|m_A(z) - m_A(z')| \leq M|z - z'|$ and, if $\tilde{\omega}_H$ is as in (H2), then

$$|\nabla_z m_A(z) - \nabla_z m_A(z')| \leq \tilde{\omega}_H(|z - z'|) \qquad \text{for all } z,\, z' \in H. \tag{16}$$

Proposition 4.2. *The following two statements hold:*

(i) There exists a $\mathscr{H}^{n-1}$-negligible set A such that, defining $B = B_\sigma \cup A$, for all $x \in \mathbb{R}^n \setminus B$ the function $F(\cdot, z)$ is approximately continuous at x for every $z \in \mathbb{R}^h$ and the function $z \mapsto \tilde{F}(x, z)$ is C^1 with bounded derivative.

(ii) Let Σ be a countably $\mathscr{H}^{n-1}$-rectifiable set oriented by ν_Σ. Then, for $\mathscr{H}^{n-1}$-a.e. $x \in \Sigma$ the one-sided limits $F^+(x, z)$ and $F^-(x, z)$ defined by

$$\lim_{r\downarrow 0} \fint_{B_r^\pm(x)} |F(y, z) - F^\pm(x, z)| dy = 0$$

exist for all $z \in \mathbb{R}^h$ and $z \mapsto F^\pm(x, z)$ is C^1 with bounded derivative.

Proof. (i) Choose a countable dense set U in $\mathbb{R}^h$ and

$$A := \bigcup_{z\in U} S_{F(\cdot,z)} \setminus J_{F(\cdot,z)}.$$

This set is clearly $\mathscr{H}^{n-1}$-negligible and, since B_σ contains all jump sets of $F(\cdot, z)$, $B = A \cup B_\sigma$ contains all sets $S_{F(\cdot,z)}$, $z \in U$, so that all points in $\mathbb{R}^n \setminus B$ are approximate continuity points of $F(\cdot, z)$, $z \in U$. Using assumption (H1) and a density argument as in Lemma 3.9 we obtain that all functions $F(\cdot, z)$ have approximate limits at all points in $\mathbb{R}^n \setminus B$. Since Lemma 4.1 gives

$$\nabla_z \fint_{B_r(x)} F(y, z) dy = \fint_{B_r(x)} \nabla_z F(y, z) dy$$

we can pass to the limit as $r \downarrow 0$ and use (16) to obtain that $\nabla_z \tilde{F}(x, z)$ exists and is continuous.
(ii) Arguing as in the proof of (i), we can find a $\mathscr{H}^{n-1}$-negligible set $A \subset \Sigma$ such that for every $x \in \Sigma \setminus A$ the limits

$$F^\pm(x, z) := \lim_{r\downarrow 0} \fint_{B_r^\pm(x)} F(y, z) dy$$

exist for every $z \in \mathbb{R}^h$ and

$$\lim_{r\downarrow 0} \fint_{B^\pm(x)} |F(y, z) - F^\pm(x, z)| dy = 0 \qquad \forall z \in \mathbb{R}^h.$$

In addition, $z \mapsto F^\pm(x, z)$ is continuously differentiable in $\mathbb{R}^h$. ◻

Thanks to the previous proposition we know that the Borel map $x \mapsto \nabla_z \tilde{F}(x, z)$ is well defined for every $x \in \mathbb{R}^n \setminus B$ and $z \in \mathbb{R}^h$. Since B is σ-finite with respect to $\mathscr{H}^{n-1}$, $\nabla_z \tilde{F}(x, \tilde{f}(x))$ is well defined for $|\widetilde{D}f|$-almost every point x for every BV function f.

In the next lemma we provide a more precise expression for the derivative of $z \mapsto \tilde{F}(x, z)$.

Lemma 4.3. *Let $x \notin \cup_z S_{F(\cdot,z)}$. Then, for all $z \in \mathbb{R}^h$ the function $y \mapsto \nabla_z F(y, z)$ is approximately continuous at x and*

$$\nabla_z \tilde{F}(x, z) = \widetilde{\nabla_z F}(\cdot, z)(x),$$

where $\widetilde{\nabla_z F}(\cdot, z)(x)$ *is the approximate limit at* x *of* $y \mapsto \nabla_z F(y,z)$. *In addition, for all* $z \in \mathbb{R}^h$ *the functions*

$$G^r(y) := \nabla_z F(x+ry, z)$$

weak converge in* $L^\infty(B_1(0))$ *to* $\nabla_z \tilde{F}(x,z)$ *as* $r \downarrow 0$.

Proof. Let x a point where

$$\fint_{B_r(x)} F(y,z)dz \to \tilde{F}(x,z) \tag{17}$$

for every z. Fix a direction $e \in S^{n-1}$, $z_0 \in \mathbb{R}^h$ and consider for $h \in (0,1]$

$$g_r(h) := \fint_{B_r(x)} \frac{F(y, z_0 + he) - F(y, z_0)}{h} dy.$$

By the mean value theorem and hypothesis (H2) one has

$$|g_r(h_1) - g_r(h_2)| \leq \tilde{\omega}_H(|h_1 - h_2|),$$

where H is a compact neighbourhood of z_0. This and (17) imply that

$$\lim_{r\downarrow 0} g_r(h) = \frac{\tilde{F}(x, z_0 + he) - \tilde{F}(x, z_0)}{h}$$

uniformly in h, so exchanging the limits as $h \downarrow 0$ and $r \downarrow 0$ one gets

$$\begin{aligned}
\langle \widetilde{\nabla_z F}(\cdot, z_0)(x), e\rangle &= \lim_{r\downarrow 0}\lim_{h\downarrow 0} g_r(h) = \lim_{h\downarrow 0}\lim_{r\downarrow 0} g_r(h) \\
&= \lim_{h\downarrow 0} \frac{\tilde{F}(x, z_0 + he) - \tilde{F}(x, z_0)}{h} = \langle \nabla_z \tilde{F}(x, z_0), e\rangle.
\end{aligned}$$

The second part of the claim is obvious since the family of functions $\{G^r(y)\}$ is uniformly bounded and pointwise convergent. □

In case $\mu = \mathscr{L}^n$, Lemma 3.9 can be refined in the following way:

Lemma 4.4. *There exists a Lebesgue negligible set* A *such that for any* $x \in A^c$ *and every* $z \in \mathbb{R}^h$ *it holds*

$$\nabla_x F(x,z) = \lim_{r\downarrow 0} \frac{D_x F(\cdot, z)(B_r(x))}{\mathscr{L}^n(B_r(x))}, \tag{18}$$

where $\nabla_x F(x,z)$ *is the approximate gradient at* x *of* $y \mapsto F(y,z)$. *Moreover* $\nabla_x F(x,z)$ *is Borel measurable in* x *and continuous in* z.

Proof. By an easy approximation argument we can assume that z varies in a compact set K. Thanks to Lemma 3.9, we know that the limit in (18) exists for every $z \in \mathbb{R}^h$ and $x \in B$, where $\mathscr{L}^n(B^c) = 0$. Moreover, if we call such limit $L(x,z)$ we have for any $x \in B$, z_1, $z_2 \in K$ it holds

$$|L(x, z_1) - L(x, z_2)| \leq \omega_K(|z_1 - z_2|) \frac{d\lambda_K}{d\mathscr{L}^n}(x).$$

Possibly removing a negligible set we can also assume that for every x in B $F(\cdot,z)$ is approximately continuous for any z (see Lemma 4.2), $d\lambda_K/d\mathscr{L}^n$ exists and that (since clearly $\sigma \llcorner B_\sigma \perp \mathscr{L}^n$)

$$\lim_{r\downarrow 0}\frac{\sigma \llcorner B_\sigma(B_r(x))}{r^n}=0. \tag{19}$$

Let us now choose a countable dense set $H\subset\mathbb{R}^h$; thanks to [2, Theorem 3.83] we know that there exists a Borel set $\tilde B$ with Lebesgue negligible complement such that for any $x\in\tilde B$ and $z\in H$

$$\nabla_x F(x,z)=L(x,z)=\lim_{r\downarrow 0}\frac{\widetilde{D}_x F(\cdot,z)(B_r(x))}{\mathscr{L}^n(B_r(x))}$$

and

$$\lim_{r\downarrow 0}\frac{1}{r^{n+1}}\int_{B_r(x)}|F(y,z)-F(x,z)-\nabla_x F(x,z)\cdot(y-x)|=0.$$

Choosing $x\in\tilde B\cap B$, $z\in\mathbb{R}^h$ and a sequence $(z_k)\subset H$ converging to z, we have

$$\begin{aligned}
&\frac{1}{r^{n+1}}\int_{B_r(x)}|F(y,z)-\tilde F(x,z)-L(x,z)\cdot(y-x)|\\
&\le \frac{1}{r^{n+1}}\int_{B_r(x)}|F(y,z_k)-\tilde F(x,z_k)-\nabla_x F(x,z_k)\cdot(y-x)|\\
&+\frac{1}{r^{n+1}}\int_{B_r(x)}|F(y,z)-F(y,z_k)-\tilde F(x,z)+\tilde F(x,z_k)-(\nabla_x F(x,z_k)-L(x,z))\cdot(y-x)|\\
&\le o_r(1)+\sup_{t\in(0,1)}\frac{|D_xF(\cdot,z)-D_xF(\cdot,z_k)|(B_{tr}(x))}{(tr)^n}+\omega_n|L(x,z)-\nabla_x F(x,z_k)|\\
&\le o_r(1)+\sup_{t\in(0,1)}\left(\frac{|D_xF(\cdot,z)-D_xF(\cdot,z_k)|(B_{tr}(x)\cap(\mathbb{R}^n\setminus B_\sigma))}{(tr)^n}+\frac{2\sigma\llcorner B_\sigma(B_{tr}(x))}{(tr)^n}\right)\\
&+\omega_n|L(x,z)-\nabla_x F(x,z_k)|\\
&\le o_r(1)+\sup_{t\in(0,1)}\frac{|\widetilde{D}_xF(\cdot,z)-\widetilde{D}_xF(\cdot,z_k)|(B_{tr}(x))}{(tr)^n}+\omega_n|L(x,z)-\nabla_x F(x,z_k)|\\
&\le o_r(1)+\omega(|z-z_k|)\sup_{t\in(0,1)}\frac{\lambda_K(B_{tr}(x))}{(tr)^n}+\omega_n|L(x,z)-\nabla_x F(x,z_k)|
\end{aligned}$$

where in the second inequality we applied [2, Lemma 3.81] to the function $F(\cdot,z)-F(\cdot,z_k)$, in the fourth the fact that $D_xF(\cdot,z)\llcorner(\mathbb{R}^n\setminus B_\sigma)\le\widetilde{D}_xF(\cdot,z)$ for any z and equation (19) and finally in the last one hypothesis (H3). Passing to the limit first in r and then in k we obtain the thesis. □

5 Proof of the chain rule

In this section we prove Theorem 3.2.

Define $\mu = \sigma + M|Du|$ and

$$J := \big\{x : \limsup_{r\downarrow 0} \frac{\mu(B_r(x))}{\omega_{n-1} r^{n-1}} > 0\big\}.$$

Notice that J is σ-finite with respect to $\mathscr{H}^{n-1}$ and that

$$\big\{x : \limsup_{r\downarrow 0} \frac{|Du|(B_r(x))}{\omega_{n-1} r^{n-1}} > 0\big\} \cup B_\sigma = J,$$

where B_σ is defined in (9). By (2) and (3) we also get

$$J \supset B_\sigma \cup J_u \qquad \text{and} \qquad \mathscr{H}^{n-1}\big(J \setminus (B_\sigma \cup J_u)\big) = 0. \tag{20}$$

By Proposition 4.2 and (2) we can add to J a $\mathscr{H}^{n-1}$-negligible set A to obtain a set $\tilde{J} = J \cup A$ satisfying

$$\lim_{r\downarrow 0} \fint_{B_r(x)} |F(y,z) - \tilde{F}(x,z)|dy = 0 \quad \forall z \in \mathbb{R}^h,$$
$$\lim_{r\downarrow 0} \fint_{B_r(x)} |u(y) - \tilde{u}(x)|dy = 0$$

for all $x \in \mathbb{R}^n \setminus \tilde{J}$. We claim that v is approximately continuous at all points in $\mathbb{R}^n \setminus \tilde{J}$ with precise representative $\tilde{v}(x) = \tilde{F}(x, \tilde{u}(x))$. To see this, compute

$$\begin{aligned}
&\limsup_{r\downarrow 0} \fint_{B_r(x)} |F(y,u(y)) - \tilde{F}(x,\tilde{u}(x))|dy \\
&\leq \limsup_{r\downarrow 0} \fint_{B_r(x)} |F(y,u(y)) - F(y,\tilde{u}(x))|dy + \limsup_{r\downarrow 0} \fint_{B_r(x)} |F(y,\tilde{u}(x)) - \tilde{F}(x,\tilde{u}(x))|dy \\
&\leq \limsup_{r\downarrow 0} M \fint_{B_r(x)} |u(y) - \tilde{u}(x)|dy + \limsup_{r\downarrow 0} \fint_{B_r(x)} |F(y,\tilde{u}(x)) - \tilde{F}(x,\tilde{u}(x))|dy = 0.
\end{aligned}$$

In particular it holds that

$$\begin{aligned}
&Dv \llcorner (\mathbb{R}^n \setminus \tilde{J}) = Dv \llcorner (\mathbb{R}^n \setminus J) = \widetilde{D}v, \\
&Du \llcorner (\mathbb{R}^n \setminus \tilde{J}) = Du \llcorner (\mathbb{R}^n \setminus J) = \widetilde{D}u, \\
&D_x F(\cdot,z) \llcorner (\mathbb{R}^n \setminus \tilde{J}) = D_x F(\cdot,z) \llcorner (\mathbb{R}^n \setminus J) = \widetilde{D}_x F(\cdot,z) \quad \forall z \in \mathbb{R}^h.
\end{aligned} \tag{21}$$

Since $Dv \ll \mu$, in order to characterize $\widetilde{D}v$ it suffices to show that for μ-almost every $x_0 \in \mathbb{R}^n \setminus \tilde{J}$ it holds:

$$\frac{d\widetilde{D}v}{d\mu}(x_0) = \frac{d\widetilde{D}_x F(\cdot, \tilde{u}(x_0))}{d\mu}(x_0) + \nabla_z \tilde{F}(x_0, \tilde{u}(x_0)) \frac{d\widetilde{D}u}{d\mu}(x_0). \tag{22}$$

Choose a point $x_0 \in \mathbb{R}^n \setminus \tilde{J}$ such that (understanding densities in the pointwise sense as in (15)):

(i) there exists $d\widetilde{D}u/d\mu$ at x_0,

(ii) there exists $d\widetilde{D}v/d\mu$ at x_0,

(iii) there exists $d\widetilde{D}_x F(\cdot, z)/d\mu$ at x_0 for every $z \in \mathbb{R}^h$,

(iv) there exists $d\lambda_H/d\mu$ at x_0, with H compact neighborhood of $\tilde{u}(x_0)$,

(v) x_0 is a point of density 1 for $\mathbb{R}^n \setminus \tilde{J}$ with respect to μ,

(vi) $\mathrm{Tan}(\mu, x_0)$ contains a nonzero measure, i.e. there exist $r_h \downarrow 0$ such that the measures $\mu_{x_0,r_h}/\mu(B_{r_h}(x_0))$ (with $\mu_{x_0,r}(B) = \mu(x_0 + rB)$) weakly converge, in the duality with $C_c(B_1(0))$, to $\nu \neq 0$.

Notice that μ-almost every $x_0 \in \mathbb{R}^n \setminus \tilde{J}$ satisfies the previous assumptions. Indeed, (iii) is satisfied thanks to Lemma 3.9, while (v) follows from [2, Proposition 2.42].

Define for $y \in B_1(0)$

$$u^r(y) := \frac{u(x_0 + ry) - m_r(u)}{\mu(B_r(x_0))/r^{n-1}} \qquad \text{and} \qquad v^r(y) := \frac{v(x_0 + ry) - m_r^F(m_r(u))}{\mu(B_r(x_0))/r^{n-1}}$$

where

$$m_r(u) = \fint_{B_r(x_0)} u(x)dx \qquad \text{and} \qquad m_r^F(z) = \fint_{B_r(x_0)} F(x, z)dx.$$

We claim that v^r is relatively compact in $L^1(B_1(0))$. Namely, let us relate more precisely v^r to u^r, writing

$$\begin{aligned}
v^r(y) &= \frac{F(x_0 + ry, u(x_0 + ry)) - m_r^F(m_r(u))}{\mu(B_r(x_0))/r^{n-1}} \\
&= \frac{1}{\mu(B_r(x_0))/r^{n-1}} \Big\{ F(x_0 + ry, u(x_0 + ry)) - F(x_0 + ry, m_r(u))) \\
&\quad + F(x_0 + ry, m_r(u)) - m_r^F(m_r(u)) \Big\} \\
&= \frac{1}{\mu(B_r(x_0))/r^{n-1}} \Big\{ \nabla_z F(x_0 + ry, m_r(u))(u(x_0 + ry) - m_r(u)) + R(y) \\
&\quad + F(x_0 + ry, m_r(u)) - m_r^F(m_r(u)) \Big\} \\
&= \nabla_z F(x_0 + ry, m_r(u)) \frac{u(x_0 + ry) - m_r(u)}{\mu(B_r(x_0))/r^{n-1}} + \frac{F(x_0 + ry, m_r(u)) - m_r^F(m_r(u))}{\mu(B_r(x_0))/r^{n-1}} + \overline{R}(y)
\end{aligned} \tag{23}$$

with

$$R(y) = \big(u(x_0{+}ry){-}m_r(u)\big)\Big(\int_0^1 \nabla_z F\big(x_0{+}ry, m_r(u){+}t(u(x_0{+}ry){-}m_r(u)\big)dt{-}\nabla_z F(x_0{+}ry, m_r(u))\Big)$$

and $\overline{R} = Rr^{n-1}/\mu(B_r(x_0))$. By Poincaré inequality

$$u^r(y) := \frac{u(x_0+ry) - m_r(u)}{\mu(B_r(x_0))/r^{n-1}}$$

is bounded in $BV(B_1(0))$ and therefore relatively compact in the strong topology of $L^1(B_1(0))$. In addition

$$\nabla_z F(x_0+ry, m_r(u)) \stackrel{*}{\rightharpoonup} \nabla_z \tilde{F}(x_0, \tilde{u}(x_0)) \quad \text{in } L^\infty(B_1(0)) \tag{24}$$

thanks to (H2) and Lemma 4.3.

The second term in (23), namely

$$F^r(y) := \frac{F(x_0+ry, m_r(u)) - m_r^F(m_r(u))}{\mu(B_r(x_0))/r^{n-1}}$$

satisfies

$$|D_y F^r|(B_1(0)) \le \frac{|D_x F(\cdot, m_r(u))|(B_r(x_0))}{\mu(B_r(x_0))} \le \frac{\sigma(B_r(x_0))}{\mu(B_r(x_0))} \le 1$$

and

$$\int_{B_1(0)} F^r(y)dy = 0,$$

so again thanks to Poincaré inequality and to the compactness of the embedding of BV in L^1 it is relatively compact in the strong topology of $L^1(B_1(0))$. Finally

$$\overline{R}(y) = u^r(y)\Big(\int_0^1 \nabla_z F\big(x_0+ry, m_r(u)+t(u(x_0+ry)-m_r(u)\big)dt - \nabla_z F(x_0+ry, m_r(u)\big)\Big) \to 0 \tag{25}$$

in $L^1(B_1(0))$ thanks to (H2), proving the claim.

Now choose a sequence $r_h \downarrow 0$ such that

(a) $v^{r_h} \to v^0$ in $L^1(B_1(0))$,

(b) $u^{r_h} \to u^0 \in BV(B_1(0))$ in $L^1(B_1(0))$,

(c) $F^{r_h} \to F^0 \in BV(B_1(0))$ in $L^1(B_1(0))$,

(d) the measures $\mu_{x_0,r_h}/\mu(B(x_0,r_h))$ weakly converge in the duality with $C_c(B_1(0))$ to $\nu \neq 0$.

Thanks to (23), (24) and (25) we have

$$v^0(y) = F^0(y) + \nabla_z \tilde{F}(x_0, \tilde{u}(x_0))u^0(y)$$

and hence for all $t \in [0,1]$

$$Dv^0(B_t(0)) = DF^0(B_t(0)) + \nabla_z \tilde{F}(x_0, \tilde{u}(x_0))Du^0(B_t(0)). \tag{26}$$

We now compute the terms $DF^0(B_t(0))$ and $Du^0(B_t(0))$ in the previous equality. Clearly $DF^{r_h} \overset{*}{\rightharpoonup} DF^0$, now choose $t \in (0,1]$ such that $\nu(B_t(0)) \neq 0$ and $\nu(\partial B_t(0)) = 0$, then we have (writing in short $D_xF(z)$ for $D_xF(\cdot, z)$)

$$\begin{aligned} DF^0(B_t(0)) &= \lim_{h\to\infty} DF^{r_h}(B_t(0)) \\ &= \lim_{h\to\infty} \frac{D_xF(m_{r_h}(u))(B_{tr_h}(x_0))}{\mu(B_{tr_h}(x_0))} \frac{\mu(B_{tr_h}(x_0))}{\mu(B_{r_h}(x_0))} \\ &= \lim_{h\to\infty} \frac{\big[D_xF(m_{r_h}(u)) - D_xF(\tilde{u}(x_0)) + D_xF(\tilde{u}(x_0))\big](B_{tr_h}(x_0))}{\mu(B_{tr_h}(x_0))} \frac{\mu(B_{tr_h}(x_0))}{\mu(B_{r_h}(x_0))} \\ &= \Big(\lim_{h\to\infty} I(r_h) + \frac{dD_xF(\tilde{u}(x_0))}{d\mu}(x_0)\Big)\nu(B_t(0)) = \frac{dD_xF(\cdot,\tilde{u}(x_0))}{d\mu}(x_0)\nu(B_t(0)), \end{aligned} \tag{27}$$

since thanks to equation (21), (v) and (H3)

$$\begin{aligned} \limsup_{r\downarrow 0} |I(r)| &= \limsup_{r\downarrow 0} \frac{|D_xF(m_r(u))(B_r(x_0)) - D_xF(\tilde{u}(x_0))(B_r(x_0))|}{\mu(B_r(x_0))} \\ &\limsup_{r\downarrow 0} \left(\frac{|\widetilde{D}_xF(m_r(u))(B_r(x_0)) - \widetilde{D}_xF(\tilde{u}(x_0))(B_r(x_0))|}{\mu(B_r(x_0))} + \frac{2\mu(B_r(x_0)\cap \tilde{J})}{\mu(B_r(x_0))} \right) \\ &\leq \frac{d\lambda_H}{d\mu}(x_0) \limsup_{r\downarrow 0} \omega_H(|m_r(u) - \tilde{u}(x_0)|) = 0. \end{aligned}$$

A similar (and simpler) calculation gives

$$Du^0(B_t(0)) = \frac{dDu^0}{d\mu}(x_0)\nu(B_t(0)), \qquad Dv^0(B_t(0)) = \frac{dDv^0}{d\mu}(x_0)\nu(B_t(0)) \tag{28}$$

for all $t \in (0,1]$ such that $\nu(\partial B_t(0)) = 0$. Comparing (27) and (28) with (26) we obtain (22) and hence statement (i).

We now prove statement (ii), notice that (4), (20) and the rectifiability of B_σ imply

$$\begin{aligned} D^j v = D^j v \llcorner J_v = D^j v \llcorner J = D^j v \llcorner (J_u \cup B_\sigma) \\ = (v^+(x) - v^-(x))\nu_{J_u\cup B_\sigma}\mathscr{H}^{n-1} \llcorner J_u \cup B_\sigma. \end{aligned}$$

So we have only to check that

$$v^\pm = F^\pm(x, u^\pm(x))$$

$\mathscr{H}^{n-1}$-almost everywhere in $J_u \cup B_\sigma$. To do this, recall that, thanks to Proposition 4.2(ii), we have for $\mathscr{H}^{n-1}$-almost every $x \in J_u \cup B_\sigma$

$$\lim_{r\to 0} \fint_{B_r^\pm(x)} |F(y,z) - F^\pm(x,z)|dy = 0$$

for every $z \in \mathbb{R}^h$ and that the same is true for u and v. Choose any such x, then

$$\begin{aligned} v^\pm(x) &= \lim_{r\to 0} \fint_{B_r^\pm(x)} F(y,u(y))dy \\ &= \lim_{r\to 0} \fint_{B_r^\pm(x)} F(y,u^\pm(x))dy + \lim_{r\to 0} \fint_{B_r^\pm(x)} F(y,u(y)) - F(y,u^\pm(x))dy \\ &= F^\pm(x,u^\pm(x)), \end{aligned}$$

since

$$\lim_{r\to 0} \fint_{B_r^\pm(x)} |F(y,u(y)) - F(y,u^\pm(x))|dy \le \lim_{r\to 0} M \fint_{B_r^\pm(x)} |u(y)-u^\pm(x)|dy = 0.$$

So, part (b) of the theorem follows with $\mathcal{N}_F = B_\sigma$.

Finally we prove (12). Choosing $\mu = \sigma + M|Du|$ in (10), multiplying both sides by $d\mu/d\mathscr{L}^n$ and recalling that if $\nu \ll \mu$ then

$$\frac{d\nu}{d\mathscr{L}^n} = \frac{d\nu}{d\mu}\frac{d\mu}{d\mathscr{L}^n},$$

we get

$$\frac{dDv}{d\mathscr{L}^n} = \frac{D_x F(\cdot,\tilde{u}(x))}{d\mathscr{L}^n} + \nabla_z \tilde{F}(x,\tilde{u}(x))\frac{dDu}{d\mathscr{L}^n}.$$

Using equation (5) and Lemma 4.4 we thus obtain:

$$D^a v = \nabla v \, d\mathscr{L}^n = \nabla_x F(x,\tilde{u}(x))\, d\mathscr{L}^n + \nabla_z \tilde{F}(x,\tilde{u}(x))\nabla u(x)\, d\mathscr{L}^n$$

where ∇v, $\nabla_x F(x,z)$ and ∇u are the approximate gradients of the maps v, $F(\cdot,z)$ and u (notice again that the composition $\nabla_x F(x,\tilde{u}(x))$ is well defined thanks to the Scorza Dragoni Theorem). □

References

[1] L. Ambrosio and G. Dal Maso, *A general chain rule for distributional derivatives*, Proc. Amer. Math. Soc. **108** (1990), no. 3, 691–702.

[2] L. Ambrosio, N. Fusco, and D. Pallara, Functions of bounded variation and free discontinuity problems, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.

[3] G. Crasta and V. De Cicco, *A chain rule formula in the space BV and applications to conservation laws*, SIAM J. Math. Anal. **43** (2011), no. 1, 430–456.

[4] V. De Cicco, N. Fusco, and A. Verde, *On L^1-lower semicontinuity in BV*, J. Convex Anal. **12** (2005), no. 1, 173–185.

[5] V. De Cicco, N. Fusco, and A. Verde, *A chain rule formula in BV and application to lower semicontinuity*, Calc. Var. Partial Differential Equations **28** (2007), no. 4, 427–447.

[6] V. De Cicco, and G. Leoni, *A chain rule in* $L^1(div;\Omega)$ *and its applications to lower semicontinuity*, Calc. Var. Partial Differential Equations **19** (2004), no. 1, 23–51.

[7] H. Federer, Geometric measure theory, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.

[8] G. Leoni and M. Morini, *Necessary and sufficient conditions for the chain rule in* $W^{1,1}_{\mathrm{loc}}(\mathbb{R}^N;\mathbb{R}^d)$ *and* $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^N;\mathbb{R}^d)$, J. Eur. Math. Soc. (JEMS) **9** (2007), no. 2, 219–252.