PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 108, Number 3, March 1990

A GENERAL CHAIN RULE FOR DISTRIBUTIONAL DERIVATIVES

L. AMBROSIO AND G. DAL MASO

(Communicated by Barbara L. Keyfitz)

ABSTRACT. We prove a general chain rule for the distribution derivatives of the composite function $v(x) = f(u(x))$, where $u: \mathbf{R}^n \to \mathbf{R}^m$ has bounded variation and $f: \mathbf{R}^m \to \mathbf{R}^k$ is Lipschitz continuous.

INTRODUCTION

The aim of the present paper is to prove a chain rule for the distributional derivative of the composite function $v(x) = f(u(x))$, where $u: \Omega \to \mathbf{R}^m$ has bounded variation in the open set $\Omega \subset \mathbf{R}^n$ and $f: \mathbf{R}^m \to \mathbf{R}^k$ is uniformly Lipschitz continuous. Under these hypotheses it is easy to prove that the function v has locally bounded variation in Ω, hence its distributional derivative Dv is a Radon measure in Ω with values in the vector space $\mathscr{L}_{n,m}$ of all linear maps from $\mathbf{R}^n$ to $\mathbf{R}^m$. The problem is to give an explicit formula for Dv in terms of the gradient ∇f of f and of the distributional derivative Du.

To illustrate our formula, we begin with the simpler case, studied by A. I. Vol′pert, where f is continuously differentiable. Let us denote by S_u the set of all jump points of u, defined as the set of all $x \in \Omega$ where the approximate limit $\tilde{u}(x)$ does not exist at x. Then the following identities hold in the sense of measures (see [19] and [20]):

$$Dv = \nabla f(\tilde{u}) \cdot Du \quad \text{on } \Omega \backslash S_u, \tag{0.1}$$

and

$$Dv = (f(u^+) - f(u^-)) \otimes \nu_u \cdot \mathscr{H}_{n-1} \quad \text{on } S_u, \tag{0.2}$$

where ν_u denotes the measure theoretical unit normal to S_u, u^+, u^- are the approximate limits of u from both sides of S_u, and $\mathscr{H}_{n-1}$ denotes the $(n-1)$-dimensional Hausdorff measure.

In this paper we prove that (0.2) remains valid when f is only Lipschitz continuous. The main difficulty in this case lies in the extension of the chain

Received by the editors May 15, 1988, and in revised form November 21, 1988.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49F20; Secondary 26B30, 26B40, 46E35, 46G05.

0002-9939/90 $1.00 + $.25 per page

rule (0.1). In fact it may happen that the function f is nowhere differentiable on the range of u. To overcome this difficulty, for every $x \in \Omega \setminus S_u$ we introduce the tangent space

$$T_x^u = \left\{ y \in \mathbf{R}^m : y = \tilde{u}(x) + \left\langle \frac{Du}{|Du|}(x), z \right\rangle \quad \text{for some } z \in \mathbf{R}^n \right\},$$

where $Du/|Du|$ denotes the Radon-Nikodym derivative of the $\mathscr{L}_{n,m}$-valued measure Du with respect to its variation $|Du|$. We prove that for $|Du|$-almost every $x \in \Omega \setminus S_u$ the restriction of f to T_x^u is differentiable at $\tilde{u}(x)$ and that the identity

$$Dv = \nabla(f|_{T_x^u})(\tilde{u}) \cdot Du \quad \text{on } \Omega \setminus S_u$$

holds in the sense of measures.

When u is a scalar function (i.e., $m = 1$), from the previous result we deduce easily that f is differentiable at $\tilde{u}(x)$ for $|Du|$-almost every $x \in \Omega \setminus S_u$ and that the usual chain rule (0.1) holds. For a different proof of this result we refer to [7].

When u is scalar and belongs to a Sobolev space $W^{1,1}(\Omega)$, the chain rule (0.1) is well known when f is continuously differentiable except for a finite number of points (see [18]). In the general case of a Lipschitz continuous function f, the chain rule was established (without proof) by G. Stampacchia in [17]. It can also be obtained from an unpublished result by J. Serrin (see [14]). Two different proofs of this formula can be found in the literature (see [14] and [4]).

When u is vector-valued and belongs to the Sobolev space $W^{1,1}(\Omega; \mathbf{R}^m)$, our result implies that for almost every $x \in \Omega$ the restriction of f to the affine space

$$T_x^u = \{ y \in \mathbf{R}^m : y = u(x) + \langle \nabla u(x), z \rangle \quad \text{for some } z \in \mathbf{R}^n \}$$

is differentiable at $u(x)$ and that

$$\nabla v = \nabla(f|_{T_x^u})(u) \cdot \nabla u \quad \text{a.e. in } \Omega.$$

Compare this result with the chain rule for tracks studied in [14].

1. Notation and basic results about functions of bounded variation

Let $\Omega \subset \mathbf{R}^n$ be an open set; by $\mathbf{B}(\Omega)$ we denote the σ-algebra of Borel sets $B \subset \Omega$, by $|B|$ the Borel-Lebesgue n-dimensional measure, and by $\mathscr{H}_{n-1}(B)$ the Hausdorff $(n-1)$-dimensional measure of any Borel set $B \subset \mathbf{R}^n$. The vector space of linear mappings $L: \mathbf{R}^n \to \mathbf{R}^m$ will be denoted by $\mathscr{L}_{n,m}$, and it will be endowed with the Hilbert-Schmidt norm

$$|L| = \sqrt{\sum_{i=1}^{n} |L(w_i)|^2}$$

where $w_1, \dots, w_n$ is any orthonormal basis of $\mathbf{R}^n$ (the definition is independent of the choice of the basis). If $L \in \mathscr{L}_{n,m}$, $z \in \mathbf{R}^n$, we often denote $L(z)$ by $\langle L, z\rangle$. For every pair of vectors $a \in \mathbf{R}^m$, $b \in \mathbf{R}^n$, the tensor product $a \otimes b \in \mathscr{L}_{n,m}$ is canonically defined by

$$\langle a \otimes b, p\rangle = \langle b, p\rangle a \qquad \forall p \in \mathbf{R}^n.$$

where $\langle \cdot, \cdot \rangle$ denotes scalar product in $\mathbf{R}^n$.

Let $(V, |\cdot|)$ be a finite dimensional vector space. If $\sigma: \mathbf{B}(\Omega) \to V$ is any measure, by $|\sigma|$ we denote its total variation, defined for every $B \in \mathbf{B}(\Omega)$ by

$$|\sigma|(B) = \sup\left\{\sum_{i=1}^{\infty} |\sigma(B_i)| : B = \bigcup_{i=1}^{\infty} B_i,\ B_i \in \mathbf{B}(\Omega), B_i \text{ mutually disjoint}\right\}.$$

All measures we shall deal with in this paper are measures of finite total variation. If $\mu : \mathbf{B}(\Omega) \to [0, +\infty[$ is a finite measure and $h: \Omega \to V$ is a Borel function such that $\int_\Omega |h|\, d\mu < +\infty$, we denote by $h \cdot \mu$ the vector measure defined by

$$h \cdot \mu(B) = \int_B h\, d\mu \qquad \forall B \in \mathbf{B}(\Omega).$$

If $\sigma : \mathbf{B}(\Omega) \to V$ is a measure such that $|\sigma|(\Omega) < +\infty$, by the Radon-Nikodym theorem the absolutely continuous part of σ with respect to μ is representable as $h \cdot \mu$ for some Borel function $h : \Omega \to V$ whose values are determined μ-almost everywhere. We denote such a function h by σ/μ. If $\mu : \mathbf{B}(\Omega) \to \mathscr{L}_{n,m}$ is a measure and $z \in \mathbf{R}^n$, we denote by $\langle \mu, z\rangle$ the scalar measure defined by $\langle \mu, z\rangle(B) = \langle \mu(B), z\rangle$.

We denote by $BV(\Omega; \mathbf{R}^m)$ the space of functions $u \in L^1(\Omega; \mathbf{R}^m)$ whose distributional derivative is representable as a measure of finite variation. For the main properties of functions of bounded variation we refer to [11], [12], [15], [19], [20]. For every function $u \in BV(\Omega; \mathbf{R}^m)$ we denote by $Du : \mathbf{B}(\Omega) \to \mathscr{L}_{n,m}$ the distributional derivative of u, characterized by the property

$$\int_\Omega \sum_{i=1}^m u^{(i)} \operatorname{div} g_i\, dx = -\int_\Omega \sum_{i=1}^m \left\langle \left\langle \frac{Du}{|Du|}, g_i \right\rangle, e_i \right\rangle d|Du|$$

for every $g \in C_0^1(\Omega; \mathbf{R}^{nm})$, $g = (g_1, \dots, g_m)$, where $e_1, \dots, e_m$ is the canonical basis of $\mathbf{R}^m$. For every open set $A \subset \Omega$, the above formula implies

$$(1.1) \qquad |Du|(A) = \sup\left\{\int_\Omega \sum_{i=1}^m u^{(i)} \operatorname{div} g_i\, dx : g \in C_0^1(A; \mathbf{R}^{nm}), |g| \le 1\right\},$$

where $g = (g_1, \dots, g_m)$. By Riesz's theorem, a function $u \in L^1(\Omega; \mathbf{R}^m)$ belongs to $BV(\Omega; \mathbf{R}^m)$ if and only if the quantity $|Du|(\Omega)$ defined by (1.1) is finite, and one can see immediately that $u \to |Du|(A)$ is lower semicontinuous with respect to the $L^1_{\text{loc}}(A; \mathbf{R}^m)$ convergence for every open set $A \subset \Omega$. By using mollifiers, it can be easily proved that

$$|Du|(A) = \int_A |\nabla u|\, dx \tag{1.2}$$

whenever u is locally Lipschitz continuous in A. By an approximation theorem first proved in the case $m = 1$ by Anzellotti and Giaquinta in [3] and later extended to vector functions by Ambrosio, Mortola, and Tortorelli (see [2, Proposition 4.2]), for every function $u \in BV(A;\mathbf{R}^m)$ it is possible to find a sequence $(u_h) \subset C^1(A;\mathbf{R}^m)$ such that

$$\lim_{h\to+\infty} \int_A |u_h - u|\, dx = 0, \qquad \lim_{h\to+\infty} |Du_h|(A) = |Du|(A). \tag{1.3}$$

For every function $u \in BV(\Omega;\mathbf{R}^m)$ we denote by S_u the set of points where u has not an approximate limit in the sense of [11, 2.9.12], i.e. $x \in \Omega\backslash S_u$ if and only if

$$\exists \tilde{u}(x) \in \mathbf{R}^m : \forall \varepsilon > 0 \ \lim_{\rho\to 0^+} \frac{|\{y \in B_\rho(x) : |u(y) - \tilde{u}(x)| > \varepsilon\}|}{\rho^n} = 0, \tag{1.4}$$

where $B_\rho(x)$ is the open ball centered at x with radius ρ. It can be proved (see [19, Theorem 15.2], [11, 3.2.29]) that S_u can be covered, up to $\mathscr{H}_{n-1}$-negligible sets, by a sequence of hypersurfaces of class 1, and $\tilde{u} : \Omega\backslash S_u \to \mathbf{R}^m$ is a Borel function equal to u almost everywhere [11, 2.9.13]. We split the distributional derivative Du into two parts $\widetilde{D}u$, Ju, setting

$$\widetilde{D}u(B) = Du(B\backslash S_u), \qquad Ju(B) = Du(B \cap S_u) \tag{1.5}$$

for every Borel set $B \subset \Omega$. By [19, Theorem 9.2] and [11, 3.2.26], in $\mathscr{H}_{n-1}$ almost every $x \in S_u$ it is possible to find u^+, $u^- \in \mathbf{R}^m$ and a versor $\nu_u \in \mathbf{R}^n$ such that

$$\lim_{\rho\to 0^+} \frac{|\{y \in B_\rho(x) : \langle y - x, \nu_u\rangle > 0, |u(y) - u^+| > \varepsilon\}|}{\rho^n} = 0, \tag{1.6}$$

and

$$\lim_{\rho\to 0^+} \frac{|\{y \in B_\rho(x) : \langle y - x, \nu_u\rangle < 0, |u(y) - u^-| > \varepsilon\}|}{\rho^n} = 0 \tag{1.7}$$

for every $\varepsilon > 0$. The triplet (u^+, u^-, ν_u) is uniquely determined up to an interchange of u^+, u^- and to a change of sign of ν_u. Moreover, (see [19, Theorem 15.1])

$$Ju(B) = \int_{B\cap S_u} (u^+ - u^-) \otimes \nu_u \, d\mathscr{H}_{n-1} \qquad \forall B \in \mathbf{B}(\Omega). \tag{1.8}$$

We recall also that Fleming-Rishel coarea formula implies (see, for instance, [1])

$$|\widetilde{D}u|(B) = 0 \quad \forall B \in \mathbf{B}(\Omega) \text{ with } \mathscr{H}_{n-1}(B) < +\infty. \tag{1.9}$$

In the proof of our theorem the following two results play a fundamental role. The first one, proved in [1], allows us to describe the distributional derivative of a function of bounded variation by means of the derivatives of the one-dimensional sections. The second one (see, for instance, [13, Appendix A]) is concerned with differentiation of measures on the real line.

To state the theorem below, we first need to introduce some new notation. Let $\Omega_1 \subset \mathbf{R}^p$, $\Omega_2 \subset \mathbf{R}^q$ be open sets, and let μ be a positive finite measure in Ω_1. Let σ_x be a mapping which assigns to each $x \in \Omega_1$ an $\mathbf{R}^r$-valued Radon measure in Ω_2 in such a way that $x \to \sigma_x(A)$ is a Borel mapping for every open set $A \subset \Omega_2$ and $\int_{\Omega_1} |\sigma_x|(\Omega_2)\, d\mu(x) < +\infty$. Under these assumptions, we can define an $\mathbf{R}^r$-valued measure in the product space $\Omega_1 \times \Omega_2$, which we denote by $\int_{\Omega_1} \sigma_x\, d\mu(x)$, characterized by the property

$$\int_{\Omega_1} \sigma_x\, d\mu(x)(A \times B) = \int_A \sigma_x(B)\, d\mu(x) \qquad \forall A \in \mathbf{B}(\Omega_1),\ \forall B \in \mathbf{B}(\Omega_2). \tag{1.10}$$

Moreover, by (1.10) one gets by approximation

$$\int_{\Omega_1 \times \Omega_2} h\, d\left(\int_{\Omega_1} \sigma_x\, d\mu(x)\right) = \int_{\Omega_1} \int_{\Omega_2} h(x, y)\, d\sigma_x(y)\, d\mu(x)$$

for every bounded Borel function $h: \Omega_1 \times \Omega_2 \to \mathbf{R}$. We remark that

$$\int_{\Omega_1} \sigma_x\, d\mu(x) = \int_{\Omega_1} \sigma'_x\, d\mu(x) \Leftrightarrow \sigma_x = \sigma'_x \qquad \mu\text{-almost everywhere} \tag{1.11}$$

and, using Aumann's selection theorem (see [5, Theorem III.30]), it is also possible to prove that

$$\left|\int_{\Omega_1} \sigma_x\, d\mu(x)\right| = \int_{\Omega_1} |\sigma_x|\, d\mu(x). \tag{1.12}$$

Theorem 1.1. *Let* $u \in BV(\mathbf{R}^n; \mathbf{R}^m)$, $\nu \in \mathbf{R}^n$, $|\nu| = 1$. *Let* $\pi_\nu \subset \mathbf{R}^n$, *be defined by*

$$\pi_\nu = \{y \in \mathbf{R}^n : \langle y, \nu\rangle = 0\},$$

and, for every $y \in \pi_\nu$, *let* $u_y : \mathbf{R} \to \mathbf{R}^m$ *be defined by*

$$u_y(t) = u(y + t\nu) \qquad \forall t \in \mathbf{R}.$$

Then for $\mathscr{H}_{n-1}$*-almost every* $y \in \pi_\nu$ *the function* u_y *has bounded variation and*

$$\begin{aligned} \langle \widetilde{D}u, \nu\rangle &= \int_{\pi_\nu} \widetilde{D}u_y\, d\mathscr{H}_{n-1}(y), \\ \langle Ju, \nu\rangle &= \int_{\pi_\nu} Ju_y\, d\mathscr{H}_{n-1}(y). \end{aligned} \tag{1.13}$$

In addition,

$$S_{u_y} = \{t \in \mathbf{R} : y + t\nu \in S_u\} \tag{1.14}$$

and

$$\tilde{u}_y(t) = \tilde{u}(y + t\nu) \qquad \forall t \in \mathbf{R}\backslash S_{u_y} \tag{1.15}$$

for $\mathscr{H}_{n-1}$*-almost every* $y \in \pi_\nu$.

Theorem 1.2. *Let μ, ν be Radon measures in $\mathbf{R}$, and assume that $\mu \geq 0$. Then the set*

$$S_\mu = \{t \in \mathbf{R} : \mu([t, s[) = 0 \text{ for some } s > t\}$$

belongs to $\mathbf{B}(\mathbf{R})$ and $\mu(S_\mu) = 0$. Moreover, the Borel functions

$$\Delta_-(t) = \liminf_{s \to t^+} \frac{\nu([t, s[)}{\mu([t, s[)}, \qquad \Delta_+(t) = \limsup_{s \to t^+} \frac{\nu([t, s[)}{\mu([t, s[)}$$

are equal μ-almost everywhere on $\mathbf{R} \backslash S_\mu$ and they are both versions of ν / μ.

2. Statement and proof of the main result

We can now prove the main theorem of this paper.

Theorem 2.1. *Let $\Omega \subset \mathbf{R}^n$ be an open set, let $u \in BV(\Omega; \mathbf{R}^m)$, and let*

$$T_x^u = \left\{ y \in \mathbf{R}^m : y = \tilde{u}(x) + \left\langle \frac{Du}{|Du|}(x), z \right\rangle \text{ for some } z \in \mathbf{R}^n \right\} \tag{2.1}$$

for every $x \in \Omega \backslash S_u$. Let $f: \mathbf{R}^m \to \mathbf{R}^k$ be a Lipschitz continuous function such that $f(0) = 0$, and let $v = f(u) : \Omega \to \mathbf{R}^k$. Then $v \in BV(\Omega; \mathbf{R}^k)$ and

$$Jv = (f(u^+) - f(u^-)) \otimes \nu_u \cdot \mathscr{H}_{n-1}|_{S_u}. \tag{2.2}$$

In addition, for $|\widetilde{D}u|$-almost every $x \in \Omega$ the restriction of the function f to T_x^u is differentiable at $\tilde{u}(x)$ and

$$\widetilde{D}v = \nabla(f|_{T_x^u})(\tilde{u}) \frac{\widetilde{D}u}{|\widetilde{D}u|} \cdot |\widetilde{D}u|. \tag{2.3}$$

Before proving the theorem, we state without proof three elementary remarks which will be useful in the sequel.

Remark 2.1. Let $\omega :]0, +\infty[\to]0, +\infty[$ be a continuous function such that $\omega(t) \to 0$ as $t \to 0$. Then

$$\lim_{h \to 0^+} g(\omega(h)) = L \Leftrightarrow \lim_{h \to 0^+} g(h) = L$$

for any function $g :]0, +\infty[\to \mathbf{R}$.

Remark 2.2. Let $g : \mathbf{R}^n \to \mathbf{R}$ be a Lipschitz continuous function and assume that

$$L(z) = \lim_{h \to 0^+} \frac{g(hz) - g(0)}{h}$$

exists for every $z \in \mathbf{Q}^n$ and that L is a linear function of z. Then g is differentiable at 0.

Remark 2.3. Let $A : \mathbf{R}^n \to \mathbf{R}^m$ be a linear function, and let $f : \mathbf{R}^m \to \mathbf{R}$ be a function. Then the restriction of f to the range of A is differentiable at 0 if and only if $f(A) : \mathbf{R}^n \to \mathbf{R}$ is differentiable at 0 and

$$\nabla(f|_{\mathrm{Im}(A)})(0)A = \nabla(f(A))(0).$$

Proof of Theorem 2.1. We begin by showing that $v \in BV(\Omega;\mathbf{R}^k)$ and

$$|Dv|(B) \leq K|Du|(B) \qquad \forall B \in \mathbf{B}(\Omega), \tag{2.4}$$

where $K > 0$ is the Lipschitz constant of f. By (1.2) and by the approximation result quoted in §1, it is possible to find a sequence $(u_h) \subset C^1(\Omega;\mathbf{R}^m)$ converging to u in $L^1(\Omega;\mathbf{R}^m)$ and such that

$$\lim_{h\to+\infty} \int_\Omega |\nabla u_h|\, dx = |Du|(\Omega).$$

The functions $v_h = f(u_h)$ are locally Lipschitz continuous in Ω, and the definition of differential implies that $|\nabla v_h| \leq K|\nabla u_h|$ almost everywhere in Ω. The lower semicontinuity of the total variation and (1.2) yield

$$\begin{aligned} |Dv|(\Omega) \leq \liminf_{h\to+\infty} |Dv_h|(\Omega) &= \liminf_{h\to+\infty} \int_\Omega |\nabla v_h|\, dx \\ &\leq K \liminf_{h\to+\infty} \int_\Omega |\nabla u_h|\, dx = K|Du|(\Omega). \end{aligned}$$

Since $f(0) = 0$, we have also

$$\int_\Omega |v|\, dx \leq K \int_\Omega |u|\, dx;$$

therefore $u \in BV(\Omega;\mathbf{R}^k)$. Repeating the same argument for every open set $A \subset \Omega$, we get (2.4) for every $B \in \mathbf{B}(\Omega)$, because $|Dv|$, $|Du|$ are Radon measures. To prove (2.2), first we observe that

$$S_v \subset S_u, \qquad \tilde{v}(x) = f(\tilde{u}(x)) \qquad \forall x \in \Omega\backslash S_u. \tag{2.5}$$

In fact, for every $\varepsilon > 0$ we have

$$\{y \in B_\rho(x) : |v(y) - f(\tilde{u}(x))| > \varepsilon\} \subset \{y \in B_\rho(x) \colon |u(y) - \tilde{u}(x)| > \varepsilon/K\},$$

hence

$$\lim_{\rho\to0^+} \frac{|\{y \in B_\rho(x) : |v(y) - f(\tilde{u}(x))| > \varepsilon\}|}{\rho^n} = 0$$

whenever $x \in \Omega\backslash S_u$. By a similar argument, if $x \in S_u$ is a point such that there exists a triplet (u^+, u^-, ν_u) satisfying (1.6), (1.7), then

$$(v^+(x) - v^-(x)) \otimes \nu_v = (f(u^+(x)) - f(u^-(x))) \otimes \nu_u \quad \text{if } x \in S_v$$

and $f(u^-(x)) = f(u^+(x))$ if $x \in S_u\backslash S_v$. Hence, by (1.8) we get

$$\begin{aligned} Jv(B) = \int_{B\cap S_v} (v^+ - v^-) \otimes \nu_v\, d\mathscr{H}_{n-1} &= \int_{B\cap S_v} (f(u^+) - f(u^-)) \otimes \nu_u\, d\mathscr{H}_{n-1} \\ &= \int_{B\cap S_u} (f(u^+) - f(u^-)) \otimes \nu_u\, d\mathscr{H}_{n-1} \end{aligned}$$

and (2.2) is proved.

To prove (2.3), it is not restrictive to assume that $k = 1$. Moreover, to simplify our notation, from now on we shall assume that $\Omega = \mathbf{R}^n$. The proof of (2.3) is divided into two steps. In the first step we prove the statement in the one-dimensional case ($n = 1$), using Theorem 1.2. In the second step we achieve the general result using Theorem 1.1.

Step 1. Assume that $n = 1$. Since S_u is at most countable, (1.9) yields that $|\widetilde{D}v|(S_u \setminus S_v) = 0$, so that (2.4) and (2.5) imply that $Dv = \widetilde{D}v + Jv$ is the Radon-Nikodym decomposition of Dv in absolutely continuous and singular part with respect to $|\widetilde{D}u|$. By Theorem 1.2, we have

$$\frac{\widetilde{D}v}{|\widetilde{D}u|}(t) = \lim_{s\to t^+} \frac{Dv([t,s[)}{|\widetilde{D}u|([t,s[)}, \qquad \frac{\widetilde{D}u}{|\widetilde{D}u|}(t) = \lim_{s\to t^+} \frac{Du([t,s[)}{|\widetilde{D}u|([t,s[)} \tag{2.6}$$

$|\widetilde{D}u|$-almost everywhere in $\mathbf{R}$. It is well known (see, for instance, [11, 2.5.16]) that every one-dimensional function of bounded variation w has a unique left continuous representative, i.e., a function $\hat{w}$ such that $\hat{w} = w$ almost everywhere and $\lim_{s\to t^-} \hat{w}(s) = \hat{w}(t)$ for every $t \in \mathbf{R}$. These conditions imply

$$\hat{u}(t) = Du(]-\infty, t[), \qquad \hat{v}(t) = Dv(]-\infty, t[) \qquad \forall t \in \mathbf{R} \tag{2.7}$$

and

$$\hat{v}(t) = f(\hat{u}(t)) \qquad \forall t \in \mathbf{R}. \tag{2.8}$$

Let $t \in \mathbf{R}$ be such that $|\widetilde{D}u|([t,s[) > 0$ for every $s > t$ and assume that the limits in (2.6) exist. By (2.7) and (2.8) we get

$$\begin{aligned} \frac{\hat{v}(s) - \hat{v}(t)}{|\widetilde{D}u|([t,s[)} &= \frac{f(\hat{u}(s)) - f(\hat{u}(t))}{|\widetilde{D}u|([t,s[)} \\ &= \frac{f(\hat{u}(s)) - f(\hat{u}(t) + \frac{\widetilde{D}u}{|\widetilde{D}u|}(t)|\widetilde{D}u|([t,s[))}{|\widetilde{D}u|([t,s[)} \\ &\quad + \frac{f(\hat{u}(t) + \frac{\widetilde{D}u}{|\widetilde{D}u|}(t)|\widetilde{D}u|([t,s[)) - f(\hat{u}(t))}{|\widetilde{D}u|([t,s[)} \end{aligned}$$

for every $s > t$. Using the Lipschitz condition on f we find

$$\begin{aligned} &\left| \frac{\hat{v}(s) - \hat{v}(t)}{|\widetilde{D}u|([t,s[)} - \frac{f(\hat{u}(t) + \frac{\widetilde{D}u}{|\widetilde{D}u|}(t)|\widetilde{D}u|([t,s[)) - f(\hat{u}(t))}{|\widetilde{D}u|([t,s[)} \right| \\ &\qquad \le K \left| \frac{\hat{u}(s) - \hat{u}(t)}{|\widetilde{D}u|([t,s[)} - \frac{\widetilde{D}u}{|\widetilde{D}u|}(t) \right|. \end{aligned}$$

By (1.9), the function $s \to |\widetilde{D}u|([t,s[)$ is continuous and converges to 0 as $s \downarrow t$. Therefore Remark 2.1 and the previous inequality imply

$$\frac{\widetilde{D}v}{|\widetilde{D}u|}(t) = \lim_{h\to 0^+} \frac{f(\hat{u}(t) + h\frac{\widetilde{D}u}{|\widetilde{D}u|}(t)) - f(\hat{u}(t))}{h} \qquad |\widetilde{D}u|\text{-a.e. in } \mathbf{R}.$$

By (2.7), $\hat{u}(x) = \tilde{u}(x)$ for every $x \in \mathbf{R}\setminus S_u$; moreover, applying the same argument to the functions $u'(t) = u(-t)$, $v'(t) = f(u'(t)) = v(-t)$, we get

$$\frac{\widetilde{D}v}{|\widetilde{D}u|}(t) = \lim_{h\to 0} \frac{f\left(\tilde{u}(t) + h\dfrac{\widetilde{D}u}{|\widetilde{D}u|}(t)\right) - f(\tilde{u}(t))}{h} \qquad |\widetilde{D}u|\text{-a.e. in } \mathbf{R}$$

and our statement is proved.

Step 2. Let us consider now the general case $n > 1$. Let $\nu \in \mathbf{R}^n$ be such that $|\nu| = 1$, and let $\pi_\nu = \{y \in \mathbf{R}^n : \langle y, \nu\rangle = 0\}$. In the following, we shall identify $\mathbf{R}^n$ with $\pi_\nu \times \mathbf{R}$, and we shall denote by y the variable ranging in π_ν and by t the variable ranging in $\mathbf{R}$. By the just proven one-dimensional result, and by Theorem 1.1, we get

$$\lim_{h\to 0} \frac{f\left(\tilde{u}(y+t\nu) + h\dfrac{\widetilde{D}u_y}{|\widetilde{D}u_y|}(t)\right) - f(\tilde{u}(y+t\nu))}{h} = \frac{\widetilde{D}v_y}{|\widetilde{D}u_y|}(t) \qquad |\widetilde{D}u_y|\text{-a.e. in } \mathbf{R}$$

for $\mathscr{H}_{n-1}$-almost every $y \in \pi_\nu$. We claim that

$$\frac{\langle \widetilde{D}u, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(y+t\nu) = \frac{\widetilde{D}u_y}{|\widetilde{D}u_y|}(t) \qquad |\widetilde{D}u_y|\text{-a.e. in } \mathbf{R} \tag{2.9}$$

for $\mathscr{H}_{n-1}$-almost every $y \in \pi_\nu$. In fact, by (1.12) and (1.13) we get

$$\int_{\pi_\nu} \frac{\widetilde{D}u_y}{|\widetilde{D}u_y|}\cdot |\widetilde{D}u_y|\, d\mathscr{H}_{n-1}(y) = \int_{\pi_\nu} \widetilde{D}u_y\, d\mathscr{H}_{n-1}(y)$$

$$= \langle \widetilde{D}u, \nu\rangle = \frac{\langle \widetilde{D}u, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}\cdot |\langle \widetilde{D}u, \nu\rangle| = \int_{\pi_\nu} \frac{\langle \widetilde{D}u, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(y+\cdot\nu)\cdot |\widetilde{D}u_y|\, d\mathscr{H}_{n-1}(y)$$

and (2.9) follows from (1.11). By the same argument it is possible to prove that

$$\frac{\langle \widetilde{D}v, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(y+t\nu) = \frac{\widetilde{D}v_y}{|\widetilde{D}u_y|}(t) \qquad |\widetilde{D}u_y|\text{-a.e. in } \mathbf{R} \tag{2.10}$$

for $\mathscr{H}_{n-1}$-almost every $y \in \pi_\nu$. By (2.9) and (2.10) we get

$$\lim_{h\to 0} \frac{f\left(\tilde{u}(y+t\nu) + h\dfrac{\langle \widetilde{D}u, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(y+t\nu)\right) - f(\tilde{u}(y+t\nu))}{h} = \frac{\langle \widetilde{D}v, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(y+t\nu)$$

$|\widetilde{D}u_y|$-a.e. in $\mathbf{R}$ for $\mathscr{H}_{n-1}$-almost every $y \in \pi_\nu$, and using again (1.12), (1.13) we get

$$\lim_{h\to 0} \frac{f\left(\tilde{u}(x) + h\dfrac{\langle \widetilde{D}u, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(x)\right) - f(\tilde{u}(x))}{h} = \frac{\langle \widetilde{D}v, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(x)$$

$|\langle \widetilde{D}u, \nu\rangle|$-a.e. in $\mathbf{R}^n$. Since the function $|\langle \widetilde{D}u, \nu\rangle|/|\widetilde{D}u|$ is strictly positive $|\langle \widetilde{D}u, \nu\rangle|$-almost everywhere, we obtain also

$$\lim_{h\to 0} \frac{f\left(\tilde{u}(x) + h\dfrac{|\langle \widetilde{D}u, \nu\rangle|}{|\widetilde{D}u|}(x)\dfrac{\langle \widetilde{D}u, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(x)\right) - f(\tilde{u}(x))}{h} = \frac{|\langle \widetilde{D}u, \nu\rangle|}{|\widetilde{D}u|}(x)\frac{\langle \widetilde{D}v, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|}(x) \tag{2.11}$$

$|\langle \widetilde{D}u, \nu\rangle|$-almost everywhere in $\mathbf{R}^n$. Finally, since

$$\frac{|\langle \widetilde{D}u, \nu\rangle|}{|\widetilde{D}u|}\frac{\langle \widetilde{D}u, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|} = \frac{\langle \widetilde{D}u, \nu\rangle}{|\widetilde{D}u|} = \left\langle \frac{\widetilde{D}u}{|\widetilde{D}u|}, \nu \right\rangle \qquad |\widetilde{D}u|\text{-a.e. in } \mathbf{R}^n$$

$$\frac{|\langle \widetilde{D}u, \nu\rangle|}{|\widetilde{D}u|}\frac{\langle \widetilde{D}v, \nu\rangle}{|\langle \widetilde{D}u, \nu\rangle|} = \frac{\langle \widetilde{D}v, \nu\rangle}{|\widetilde{D}u|} = \left\langle \frac{\widetilde{D}v}{|\widetilde{D}u|}, \nu \right\rangle \qquad |\widetilde{D}u|\text{-a.e. in } \mathbf{R}^n$$

and since both sides of (2.11) are zero $|\widetilde{D}u|$-almost everywhere on $|\langle \widetilde{D}u, \nu\rangle|$-negligible sets, we conclude that

$$\lim_{h\to 0} \frac{f\left(\tilde{u}(x) + h\left\langle \dfrac{\widetilde{D}u}{|\widetilde{D}u|}(x), \nu \right\rangle\right) - f(\tilde{u}(x))}{h} = \left\langle \frac{\widetilde{D}v}{|\widetilde{D}u|}(x), \nu \right\rangle$$

$|\widetilde{D}u|$-a.e. in $\mathbf{R}^n$. Since ν is arbitrary, by Remarks 2.2 and 2.3 the restriction of f to the affine space T_x^u is differentiable at $\tilde{u}(x)$ for $|\widetilde{D}u|$-almost every $x \in \mathbf{R}^n$ and (2.3) holds. Q.E.D.

3. Some corollaries

Formula (2.3) becomes simpler in some particular but important cases, i.e., when u is a scalar function $(m = 1)$ or when $u \in W^{1,p}(\Omega; \mathbf{R}^m)$ for some p, $1 \le p \le +\infty$.

Corollary 3.1. *Let $u \in BV(\Omega)$ and let $f: \mathbf{R} \to \mathbf{R}$ be a Lipschitz continuous function such that $f(0) = 0$. Then $v = f(u)$ belongs to $BV(\Omega)$ and*

$$Jv = (f(u^+) - f(u^-))\nu_u \cdot \mathscr{H}_{n-1}|_{S_u}.$$

In addition, for $|\widetilde{D}u|$-almost every $x \in \Omega$ the function f is differentiable at $\tilde{u}(x)$ and

$$\widetilde{D}v = \nabla f(\tilde{u}) \cdot \widetilde{D}u.$$

Proof. Since $\widetilde{D}u/|\widetilde{D}u| = \pm 1$ $|\widetilde{D}u|$-almost everywhere in Ω, the corollary is a straightforward consequence of Theorem 2.1. Q.E.D.

The last formula of the previous corollary can also be stated in the following form:

$$\widetilde{D}v = g(\tilde{u}) \cdot \widetilde{D}u$$

where $g\colon \mathbf{R} \to \mathbf{R}$ is any Borel function such that $g(t) = \nabla f(t)$ almost everywhere. In fact, by the Fleming-Rishel coarea formula, it is not difficult to desume that (see, for instance, [1])

$$|\widetilde{D}u|(\tilde{u}^{-1}(E)) = 0 \qquad \forall E \in \mathbf{B}(\mathbf{R}) \text{ with } |E| = 0.$$

In this form, when u belongs to a Sobolev space $W^{1,p}(\Omega)$, Corollary 3.1 has been proved by Marcus and Mizel in [14].

Corollary 3.2. *Let $p \in [1, +\infty]$, $u \in W^{1,p}(\Omega; \mathbf{R}^m)$, and let $f\colon \mathbf{R}^m \to \mathbf{R}^k$ be a Lipschitz continuous function such that $f(0) = 0$. Then $v = f(u)$ belongs to $W^{1,p}(\Omega; \mathbf{R}^k)$, for almost every $x \in \Omega$ the restriction of the function f to the affine space*

$$T_x^u = \{y \in \mathbf{R}^m \colon y = u(x) + \langle \nabla u(x), z\rangle \quad \text{for some } z \in \mathbf{R}^n\}$$

is differentiable at $u(x)$, and

$$\nabla v = \nabla(f|_{T_x^u})(u)\nabla u \quad \text{a.e. in } \Omega.$$

Proof. For functions $u \in W^{1,1}_{\text{loc}}(\Omega; \mathbf{R}^m)$ the set S_u is $\mathscr{H}_{n-1}$-negligible (see [11, 4.5.9(29)] and [19, Theorem 15.3]). By (2.2) and (2.4), $Jv = 0$ and Dv is absolutely continuous with respect to the Lebesgue measure. Since $\tilde{u} = u$ almost everywhere, and since (2.4) implies

$$|\nabla u(x)| = 0 \Rightarrow |\nabla v(x)| = 0 \quad \text{a.e. in } \Omega,$$

the statement of the corollary follows from (2.3). Q.E.D.

Acknowledgments

The authors wish to thank Prof. E. De Giorgi for having suggested to investigate the differentiability properties of a Lipschitz continuous function f on the tangent spaces T_x^u related to a function u of bounded variation.

References

1. L. Ambrosio, *A compactness theorem for a special class of functions of bounded variation* (to appear in Boll. Un. Mat. Ital.).
2. L. Ambrosio, S. Mortola, and V. M. Tortorelli, *Functionals with linear growth defined on vector valued BV functions* (to appear in J. Math. Pures Appl.).
3. G. Anzellotti and M. Giaquinta, *Funzioni BV e tracce*, Rend. Sem. Mat. Univ. Padova **60** (1978), 1–22.
4. L. Boccardo and F. Murat, *Remarques sur l'homogeneisation de certaines problèmes quasi-linéaires*, Portugal. Math. **41** (1982), 535–562.
5. A. P. Calderon and A. Zygmund, *On the differentiability of functions which are of bounded variation in Tonelli's sense*, Rev. Un. Mat. Argentina **20** (1960), 102–121.
6. C. Castaing and M. Valadier, *Convex analysis and measurable multifunctions*, Lecture Notes in Math., no. 580, Springer-Verlag, 1977.
7. G. Dal Maso, P. Le Floch, and F. Murat, (paper in preparation).
8. E. De Giorgi, *Su una teoria generale della misura $(r-1)$-dimensionale in uno spazio a r dimensioni*, Ann. Mat. Pura Appl. **36** (1954), 191–213.

9. ——, *Nuovi teoremi relativi alle misure $(r-1)$-dimensionali in uno spazio a r dimensioni*, Ricerche Mat. **4** (1955), 95–113.

10. E. De Giorgi, F. Colombini and L. C. Piccinini, *Frontiere orientate di misura minima e questioni collegate*, Quaderni, Scuola Norm. Sup., Pisa, 1972.

11. H. Federer, *Geometric measure theory*, Springer-Verlag, Berlin, 1969.

12. E. Giusti, *Minimal surfaces and functions of bounded variation*, Birkhäuser, Boston, 1984.

13. G. Letta, *Martingales et intégration stochastique*, Quaderni, Scuola Norm. Sup., Pisa, 1984.

14. M. Marcus and V. J. Mizel, *Absolute continuity on tracks and mappings of Sobolev spaces*, Arch. Rational Mech. Anal. **45**, (1972), 294–320.

15. M. Miranda, *Distribuzioni aventi derivate misure e insiemi di perimetro localmente finito*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **18** (1964), 27–56.

16. ——, *Superfici cartesiane generalizzate ed insiemi di perimetro localmente finito sui prodotti cartesiani*. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **18** (1964), 515–542.

17. G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble) **15** (1965), 189–258.

18. ——, *Équations elliptiques du second ordre à coefficients discontinus*, Les Presses de l'Université de Montréal, Montréal, 1966.

19. A. I. Vol′pert, *The spaces BV and quasi-linear equations*, Math. USSR-Sb. **2** (1967), 225–267.

20. A. I. Vol′pert and S. I. Hudjaev, *Analysis in classes of discontinuous functions and equations of mathematical physics*, Martinus Nijhoff Publisher, Dordrecht, 1985.

Dipartimento di Matematica, II Università di Roma, Tor Vergata, 00173 Roma, Italy

Sissa, Strada Costiera 11, 34014 Trieste, Italy