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 \( u: \mathbb{R}^n \rightarrow \mathbb{R}^m \) has bounded variation and \( f: \mathbb{R}^m \rightarrow \mathbb{R}^k \) 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 \( u: \Omega \rightarrow \mathbb{R}^m \) has bounded variation in the open set \( \Omega \subset \mathbb{R}^n \) and \( f: \mathbb{R}^m \rightarrow \mathbb{R}^k \) is uniformly Lipschitz continuous. Under these hypotheses it is easy to prove that the function \( v \) has locally bounded variation in \( \Omega \), hence its distributional derivative \( Dv \) is a Radon measure in \( \Omega \) with values in the vector space \( \mathcal{L}_{n,m} \) of all linear maps from \( \mathbb{R}^n \) to \( \mathbb{R}^m \). The problem is to give an explicit formula for \( Dv \) in terms of the gradient \( \nabla f \) 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]):

\[
(0.1) \quad Dv = \nabla f(\tilde{u}) \cdot Du \quad \text{on} \quad \Omega \setminus S_u,
\]

and

\[
(0.2) \quad Dv = (f(u^+) - f(u^-)) \otimes \nu_u \cdot \mathcal{H}^1 \quad \text{on} \quad S_u,
\]

where \( \nu_u \) denotes the measure theoretical unit normal to \( S_u \), \( u^+, \ u^- \) are the approximate limits of \( u \) from both sides of \( S_u \), and \( \mathcal{H}^1 \) 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
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

\[
T_x^u = \left\{ y \in \mathbb{R}^m : y = \tilde{u}(x) + \left( \frac{Du}{|Du|}(x), z \right) \forall z \in \mathbb{R}^n \right\},
\]

where \( \frac{Du}{|Du|} \) denotes the Radon-Nikodym derivative of the \( \mathcal{L}^m \)-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 \( |Dv| \)-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 \( W^{1,1}(\Omega; \mathbb{R}^m) \), our result implies that for almost every \( x \in \Omega \) the restriction of \( f \) to the affine space

\[
T_x^u = \left\{ y \in \mathbb{R}^m : y = u(x) + (\nabla u(x), z) \forall z \in \mathbb{R}^n \right\}
\]

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

\[
|L| = \sqrt{\sum_{i=1}^n |L(w_i)|^2}
\]
where \( w_1, \ldots, w_n \) is any orthonormal basis of \( \mathbb{R}^n \) (the definition is independent of the choice of the basis). If \( L \in \mathcal{L}_{n,m} \), \( z \in \mathbb{R}^n \), we often denote \( L(z) \) by \( \langle L, z \rangle \). For every pair of vectors \( a \in \mathbb{R}^m, b \in \mathbb{R}^n \), the tensor product \( a \otimes b \in \mathcal{L}_{n,m} \) is canonically defined by

\[
\langle a \otimes b, p \rangle = \langle b, p \rangle a \quad \forall p \in \mathbb{R}^n.
\]

where \( \langle \cdot, \cdot \rangle \) denotes scalar product in \( \mathbb{R}^n \).

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

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

All measures we shall deal with in this paper are measures of finite total variation. If \( \mu : \mathcal{B}(\Omega) \to [0, +\infty[ \) 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

\[
h \cdot \mu(B) = \int_B h \, d\mu \quad \forall B \in \mathcal{B}(\Omega).
\]

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

\[
\langle \mu, z \rangle(B) = \mu(B, z).
\]

We denote by \( BV(\Omega; \mathbb{R}^m) \) the space of functions \( u \in L^1(\Omega; \mathbb{R}^m) \) 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 \( u \in BV(\Omega; \mathbb{R}^m) \) we denote by \( Du : \mathcal{B}(\Omega) \to \mathcal{L}_{n,m} \) the distributional derivative of \( u \), characterized by the property

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

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

\[
(1.1) \quad |Du|(A) = \sup \left\{ \int_{\Omega} \sum_{i=1}^{m} u^{(i)} \text{div} g_i \, dx : g \in C_0^1(A; \mathbb{R}^m), |g| \leq 1 \right\},
\]

where \( g = (g_1, \ldots, g_m) \). By Riesz's theorem, a function \( u \in L^1(\Omega; \mathbb{R}^m) \) belongs to \( BV(\Omega; \mathbb{R}^m) \) 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 \( L^1_{loc}(A; \mathbb{R}^m) \) convergence for every open set \( A \subset \Omega \). By
using mollifiers, it can be easily proved that

\[
|Du|(A) = \int_A |\nabla u| \, dx
\]

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 \( u \in BV(A; \mathbb{R}^m) \) it is possible to find a sequence \( (u_h) \subset C^1(A; \mathbb{R}^m) \) such that

\[
\lim_{h \to +\infty} \int_A |u_h - u| \, dx = 0, \quad \lim_{h \to +\infty} |Du_h|(A) = |Du|(A).
\]

For every function \( u \in BV(\Omega; \mathbb{R}^m) \) 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. \( x \in \Omega \setminus S_u \) if and only if

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

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

\[
\tilde{D}u(B) = Du(B \setminus S_u), \quad J_u(B) = Du(B \cap S_u)
\]

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

\[
\lim_{\rho \to 0^+} \frac{|\{y \in B_\rho(x) : (y - x, \nu_u) > 0, |u(y) - u^+| > \varepsilon\}|}{\rho^n} = 0,
\]

and

\[
\lim_{\rho \to 0^+} \frac{|\{y \in B_\rho(x) : (y - x, \nu_u) < 0, |u(y) - u^-| > \varepsilon\}|}{\rho^n} = 0
\]

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

\[
J_u(B) = \int_{B \cap S_u} (u^+ - u^-) \otimes \nu_u \, d\mathcal{H}^{n-1} \quad \forall B \in \mathcal{B}(\Omega).
\]

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

\[
|\tilde{D}u|(B) = 0 \quad \forall B \in \mathcal{B}(\Omega) \text{ with } \mathcal{H}^{n-1}(B) < +\infty.
\]
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 $\Omega_1 \subset \mathbb{R}^p$, $\Omega_2 \subset \mathbb{R}^q$ be open sets, and let $\mu$ be a positive finite measure in $\Omega_1$. Let $\sigma_x$ be a mapping which assigns to each $x \in \Omega_1$ an $\mathbb{R}^r$-valued Radon measure in $\Omega_2$ 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 $\mathbb{R}^r$-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

\begin{equation}
(1.10) \quad \int_{\Omega_1} \sigma_x \, d\mu(x)(A \times B) = \int_A \sigma_x(B) \, d\mu(x) \quad \forall A \in \mathcal{B}(\Omega_1), \forall B \in \mathcal{B}(\Omega_2).
\end{equation}

Moreover, by (1.10) one gets by approximation

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

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

\begin{equation}
(1.11) \quad \int_{\Omega_1} \sigma_x \, d\mu(x) = \int_{\Omega_1} \sigma'_x \, d\mu(x) \Leftrightarrow \sigma_x = \sigma'_x \quad \mu\text{-almost everywhere}
\end{equation}

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

\begin{equation}
(1.12) \quad \left| \int_{\Omega_1} \sigma_x \, d\mu(x) \right| = \int_{\Omega_1} |\sigma_x| \, d\mu(x).
\end{equation}

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

\[
\pi_\nu = \{ y \in \mathbb{R}^n : (y, \nu) = 0 \},
\]

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

\[
u_y(t) = u(y + t\nu) \quad \forall t \in \mathbb{R}.
\]

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

\[
\langle D u, \nu \rangle = \int_{\pi_\nu} D_{u_y} \, d\mathcal{H}_{n-1}(y),
\]

\[
\langle J u, \nu \rangle = \int_{\pi_\nu} J_{u_y} \, d\mathcal{H}_{n-1}(y).
\]

In addition,

\begin{equation}
(1.14) \quad S_{u_y} = \{ t \in \mathbb{R} : y + t\nu \in S_u \}
\end{equation}

and

\begin{equation}
(1.15) \quad \ddot{u}_y(t) = \ddot{u}(y + t\nu) \quad \forall t \in \mathbb{R}\backslash S_{u_y},
\end{equation}

for $\mathcal{H}_{n-1}$-almost every $y \in \pi_\nu$.
Theorem 1.2. Let \( \mu, \nu \) be Radon measures in \( \mathbb{R} \), and assume that \( \mu \geq 0 \). Then the set
\[
S_{\mu} = \{ t \in \mathbb{R} : \mu([t, s[) = 0 \text{ for some } s > t \}
\]
belongs to \( \mathcal{B}(\mathbb{R}) \) and \( \mu(S_{\mu}) = 0 \). Moreover, the Borel functions
\[
\Delta_-(t) = \liminf_{s \to t^+} \frac{\nu([t, s[)}{\mu([t, s[)}, \quad \Delta_+(t) = \limsup_{s \to t^+} \frac{\nu([t, s[)}{\mu([t, s[)}
\]
are equal \( \mu \)-almost everywhere on \( \mathbb{R} \setminus S_{\mu} \) and they are both versions of \( \nu/\mu \).

2. STATEMENT AND PROOF OF THE MAIN RESULT

We can now prove the main theorem of this paper.

Theorem 2.1. Let \( \Omega \subset \mathbb{R}^n \) be an open set, let \( u \in BV(\Omega; \mathbb{R}^m) \), and let
\[
T^u_x = \left\{ y \in \mathbb{R}^m : y = \tilde{u}(x) + \frac{Du}{|Du|}(x)z \text{ for some } z \in \mathbb{R}^n \right\}
\]
for every \( x \in \Omega \setminus S_u \). Let \( f : \mathbb{R}^m \to \mathbb{R}^k \) be a Lipschitz continuous function such that \( f(0) = 0 \), and let \( v = f(u) : \Omega \to \mathbb{R}^k \). Then \( v \in BV(\Omega; \mathbb{R}^k) \) and
\[
Jv = (f(u^+) - f(u^-)) \otimes \nu_u \cdot \mathcal{H}^{n-1}|_{S_u}.
\]
In addition, for \( |Du| \)-almost every \( x \in \Omega \) the restriction of the function \( f \) to \( T^u_x \)
is differentiable at \( \tilde{u}(x) \) and
\[
\tilde{D}v = \nabla (f|_{T^u_x})(\tilde{u}) \frac{\tilde{D}u}{|\tilde{D}u|} \cdot |\tilde{D}u|.
\]

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
\[
\lim_{h \to 0^+} g(\omega(h)) = L \iff \lim_{h \to 0^+} g(h) = L
\]
for any function \( g : ]0, +\infty[ \to \mathbb{R} \).

Remark 2.2. Let \( g : \mathbb{R}^n \to \mathbb{R} \) be a Lipschitz continuous function and assume that
\[
L(z) = \lim_{h \to 0^+} \frac{g(hz) - g(0)}{h}
\]
exists for every \( z \in \mathbb{Q}^n \) and that \( L \) is a linear function of \( z \). Then \( g \) is differentiable at 0.

Remark 2.3. Let \( A : \mathbb{R}^n \to \mathbb{R}^m \) be a linear function, and let \( f : \mathbb{R}^m \to \mathbb{R} \) be a function. Then the restriction of \( f \) to the range of \( A \) is differentiable at 0 if and only if \( f(A) : \mathbb{R}^n \to \mathbb{R} \) is differentiable at 0 and
\[
\nabla(f|_{\text{Im}(A)})(0)A = \nabla(f(A))(0).
\]
Proof of Theorem 2.1. We begin by showing that \( v \in BV(\Omega; \mathbb{R}^k) \) and
\[
|Dv|(B) \leq K|Du|(B) \quad \forall B \in \mathcal{B}(\Omega),
\]
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 \( (u_h) \subset C^1(\Omega; \mathbb{R}^m) \) converging to \( u \) in \( L^1(\Omega; \mathbb{R}^m) \) 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 \( \Omega \), and the definition of differential implies that \( |\nabla v_h| \leq K|\nabla u_h| \) almost everywhere in \( \Omega \). The lower semicontinuity of the total variation and (1.2) yield
\[
|Dv|(\Omega) \leq \liminf_{h \to +\infty} |Dv_h|(\Omega) = \liminf_{h \to +\infty} \int_{\Omega} |\nabla v_h| \, dx 
\leq K \liminf_{h \to +\infty} \int_{\Omega} |\nabla u_h| \, dx = K|Du|(\Omega).
\]
Since \( f(0) = 0 \), we have also
\[
\int_{\Omega} |v| \, dx \leq K \int_{\Omega} |u| \, dx;
\]
therefore \( u \in BV(\Omega; \mathbb{R}^k) \). Repeating the same argument for every open set \( A \subset \Omega \), we get (2.4) for every \( B \in \mathcal{B}(\Omega) \), because \( |Dv| \), \( |Du| \) are Radon measures. To prove (2.2), first we observe that
\[
S_v \subset S_u, \quad \tilde{v}(x) = f(\tilde{u}(x)) \quad \forall x \in \Omega \setminus S_u.
\]
In fact, for every \( \varepsilon > 0 \) we have
\[
\{ y \in B_{\rho}(x) : |v(y) - f(\tilde{u}(x))| > \varepsilon \} \subset \{ y \in B_{\rho}(x) : |u(y) - \tilde{u}(x)| > \varepsilon/K \},
\]
hence
\[
\lim_{\rho \to 0^+} \frac{|\{ y \in B_{\rho}(x) : |v(y) - f(\tilde{u}(x))| > \varepsilon \}|}{\rho^n} = 0
\]
whenever \( x \in \Omega \setminus S_u \). By a similar argument, if \( x \in S_u \) is a point such that there exists a triplet \( (u^+, u^-, \nu_v) \) satisfying (1.6), (1.7), then
\[
(v^+(x) - v^-(x)) \otimes \nu_v = (f(u^+(x)) - f(u^-(x))) \otimes \nu_v \quad \text{if } x \in S_v
\]
and \( f(u^-)(x) = f(u^+(x)) \) if \( x \in S_u \setminus S_v \). Hence, by (1.8) we get
\[
Jv(B) = \int_{B \cap S_v} (v^+ - v^-) \otimes \nu_v, d\mathcal{H}^{n-1}_{\nu_v} = \int_{B \cap S_v} (f(u^+) - f(u^-)) \otimes \nu_v, d\mathcal{H}^{n-1}_{\nu_v}
= \int_{B \cap S_u} (f(u^+) - f(u^-)) \otimes \nu_v, d\mathcal{H}^{n-1}_{\nu_v}
\]
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 \( \Omega = \mathbb{R}^n \). 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 \(|\tilde{D}v|(S_u \setminus S_v) = 0\), so that (2.4) and (2.5) imply that \(Dv = \tilde{D}v + Jv\) is the Radon-Nikodym decomposition of \(Dv\) in absolutely continuous and singular part with respect to \(|\tilde{D}u|\). By Theorem 1.2, we have

\[
(2.6) \quad \frac{\tilde{D}v}{|\tilde{D}u|}(t) = \lim_{s \to t^+} \frac{Dv([t, s])}{|\tilde{D}u|([t, s])}, \quad \frac{\tilde{D}u}{|\tilde{D}u|}(t) = \lim_{s \to t^+} \frac{Du([t, s])}{|\tilde{D}u|([t, s])}
\]

\(|\tilde{D}u|\)-almost everywhere in \(\mathbb{R}\). 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 \(\tilde{w}\) such that \(\tilde{w} = w\) almost everywhere and \(\lim_{s \to t^-} \tilde{w}(s) = \tilde{w}(t)\) for every \(t \in \mathbb{R}\). These conditions imply

\[
(2.7) \quad \check{u}(t) = Du([-\infty, t]), \quad \check{v}(t) = Dv([-\infty, t]) \quad \forall t \in \mathbb{R}
\]

and

\[
(2.8) \quad \check{v}(t) = f(\check{u}(t)) \quad \forall t \in \mathbb{R}.
\]

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

\[
\frac{\check{v}(s) - \check{v}(t)}{|\tilde{D}u|([t, s])} = \frac{f(\check{u}(s)) - f(\check{u}(t))}{|\tilde{D}u|([t, s])} = \frac{f(\check{u}(s)) - f(\check{u}(t)) + \frac{\tilde{D}u}{|\tilde{D}u|}(t)|\tilde{D}u|([t, s])}{|\tilde{D}u|([t, s])}
\]

\[+ \frac{f(\check{u}(t) + \frac{\tilde{D}u}{|\tilde{D}u|}(t)|\tilde{D}u|([t, s])) - f(\check{u}(t))}{|\tilde{D}u|([t, s])}
\]

for every \(s > t\). Using the Lipschitz condition on \(f\) we find

\[
\left| \frac{\check{v}(s) - \check{v}(t)}{|\tilde{D}u|([t, s])} - \frac{f(\check{u}(t) + \frac{\tilde{D}u}{|\tilde{D}u|}(t)|\tilde{D}u|([t, s])) - f(\check{u}(t))}{|\tilde{D}u|([t, s])} \right| \leq K \frac{\check{u}(s) - \check{u}(t)}{|\tilde{D}u|([t, s]) - |\tilde{D}u|([t, s])}
\]

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

\[
\frac{\tilde{D}v}{|\tilde{D}u|}(t) = \lim_{h \to 0^+} \frac{f(\check{u}(t) + h \frac{\tilde{D}u}{|\tilde{D}u|}(t)) - f(\check{u}(t))}{h} \quad |\tilde{D}u|\text{-a.e. in } \mathbb{R}.
\]
By (2.7), \( \hat{u}(x) = \check{u}(x) \) for every \( x \in \mathbb{R} \setminus S \), moreover, applying the same argument to the functions \( u'(t) = u(-t) \), \( v'(t) = f(u'(t)) = v(-t) \), we get

\[
\frac{\tilde{D}v}{|\tilde{D}u|}(t) = \lim_{h \to 0} \frac{f(\hat{u}(t) + h \frac{\tilde{D}u}{|\tilde{D}u|}(t)) - f(\hat{u}(t))}{h}, \quad |\tilde{D}u|-\text{a.e. in } \mathbb{R}
\]

and our statement is proved.

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

\[
\lim_{h \to 0} \frac{f(\hat{u}(y + tv) + h \frac{\tilde{D}u}{|\tilde{D}u|}(t)) - f(\hat{u}(y + tv))}{h} = \frac{\tilde{D}v}{|\tilde{D}u|}(t) \quad |\tilde{D}u|-\text{a.e. in } \mathbb{R}
\]

for \( \mathcal{H}_{n-1} \)-almost every \( y \in \pi_\nu \). We claim that

\[
(2.9) \quad \frac{\langle \tilde{D}u, \nu \rangle}{|\langle \tilde{D}u, \nu \rangle|} \frac{(y + tv)}{t} = \frac{\tilde{D}v}{|\tilde{D}u|}(t) \quad |\tilde{D}u|-\text{a.e. in } \mathbb{R}
\]

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

\[
\int_{\pi_\nu} \frac{\tilde{D}u}{|\tilde{D}u|} \cdot \tilde{D}u y \ d\mathcal{H}_{n-1}(y) = \int_{\pi_\nu} \tilde{D}u y \ d\mathcal{H}_{n-1}(y)
\]

\[
= \langle \tilde{D}u, \nu \rangle \cdot \langle \tilde{D}u, \nu \rangle = \int_{\pi_\nu} \frac{\langle \tilde{D}u, \nu \rangle}{|\langle \tilde{D}u, \nu \rangle|} (y + t \nu) \cdot |\tilde{D}u y | d\mathcal{H}_{n-1}(y)
\]

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

\[
(2.10) \quad \frac{\langle \tilde{D}v, \nu \rangle}{|\langle \tilde{D}u, \nu \rangle|} \frac{(y + tv)}{t} = \frac{\tilde{D}v}{|\tilde{D}u|}(t) \quad |\tilde{D}u|-\text{a.e. in } \mathbb{R}
\]

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

\[
\lim_{h \to 0} \frac{f(\hat{u}(y + tv) + h \frac{\tilde{D}u}{|\tilde{D}u|}(y + tv)) - f(\hat{u}(y + tv))}{h} = \frac{\langle \tilde{D}v, \nu \rangle}{|\langle \tilde{D}u, \nu \rangle|} \frac{(y + tv)}{t} \quad |\tilde{D}u|\text{-a.e. in } \mathbb{R}
\]

for \( \mathcal{H}_{n-1} \)-almost every \( y \in \pi_\nu \), and using again (1.12), (1.13) we get

\[
\lim_{h \to 0} \frac{f(\hat{u}(x) + h \frac{\tilde{D}u}{|\tilde{D}u|}(x)) - f(\hat{u}(x))}{h} = \frac{\langle \tilde{D}v, \nu \rangle}{|\langle \tilde{D}u, \nu \rangle|}(x) \quad |\tilde{D}u|\text{-a.e. in } \mathbb{R}
\]
\[ f(\tilde{u}(x) + h \frac{\langle \tilde{D}u, \nu \rangle}{|\tilde{D}u|}(x)) \frac{\langle \tilde{D}u, \nu \rangle}{|\tilde{D}u|}(x) - f(\tilde{u}(x)) \]

\[ = \frac{|\langle \tilde{D}u, \nu \rangle|}{|\tilde{D}u|} (x) \frac{\langle \tilde{D}v, \nu \rangle}{|\tilde{D}u|}(x) \]

\[\text{for } |\tilde{D}u|\text{-a.e. in } \mathbb{R}^n.\] Finally, since

\[ \frac{|\langle \tilde{D}u, \nu \rangle|}{|\tilde{D}u|} \frac{\langle \tilde{D}u, \nu \rangle}{|\tilde{D}u|} = \langle \tilde{D}u, \nu \rangle \]

\[\text{and since both sides of (2.11) are zero } |\tilde{D}u|\text{-almost everywhere on } |\tilde{D}u|\text{-negligible sets, we conclude that}\]

\[ \lim_{h \to 0} \frac{f(\tilde{u}(x) + h \frac{\langle \tilde{D}u, \nu \rangle}{|\tilde{D}u|}(x)) - f(\tilde{u}(x))}{h} = \langle \tilde{D}v, \nu \rangle \]

\[\text{for } |\tilde{D}u|\text{-a.e. in } \mathbb{R}^n.\] Since \( \nu \) is arbitrary, by Remarks 2.2 and 2.3 the restriction of \( f \) to the affine space \( T_xu \) is differentiable at \( \tilde{u}(x) \) for \( |\tilde{D}u| \)-almost every \( x \in \mathbb{R}^n \) 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 \( u \in W^{1,p}(\Omega; \mathbb{R}^m) \) for some \( p \), \( 1 \leq p \leq +\infty \).

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

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

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

\[ \tilde{D}v = \nabla f(\tilde{u}) \cdot \tilde{D}u. \]

**Proof.** Since \( \tilde{D}u/|\tilde{D}u| = \pm 1 \) \( |\tilde{D}u|\)-almost everywhere in \( \Omega \), 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:

\[ \tilde{D}v = g(\tilde{u}) \cdot \tilde{D}u \]
where \( g: \mathbb{R} \to \mathbb{R} \) 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 deduce that (see, for instance, [1])

\[
|\tilde{u}(\tilde{u}^{-1}(E))| = 0 \quad \forall E \in \mathcal{B}(\mathbb{R}) \text{ with } |E| = 0.
\]

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

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

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 \( u \in W^{1,1}_{\text{loc}}(\Omega; \mathbb{R}^m) \) the set \( S_u \) is \( \mathcal{H}^{n-1} \)-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**

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


Dipartimento di Matematica, II Università di Roma, Tor Vergata, 00173 Roma, Italy
Sissa, Strada Costiera 11, 34014 Trieste, Italy

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use