



A chain rule involving vector functions of bounded variation

Jean Jacques Moreau, Michel Valadier

► To cite this version:

Jean Jacques Moreau, Michel Valadier. A chain rule involving vector functions of bounded variation. Journal of Functional Analysis, 1987, 74 (2), pp.333 - 345. 10.1016/0022-1236(87)90029-2 . hal-01788917

HAL Id: hal-01788917

<https://hal.science/hal-01788917>

Submitted on 9 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Chain Rule Involving Vector Functions of Bounded Variation

J. J. MOREAU AND M. VALADIER

*Institut de Mathématiques, Université des Sciences et Techniques du Languedoc,
34060 Montpellier-Cédex, France*

Communicated by H. Brezis

By $f \in \text{lbv}(I, X)$, we mean that f is a function of a real interval I to a Banach space X , with bounded variation on every compact subinterval of I ; to such f , an X -valued measure df , called its differential measure, classically corresponds. Let Ω be an open convex subset of X and $\gamma: \Omega \rightarrow \mathbb{R}$. Two situations are investigated where the function $\gamma \circ f: t \rightarrow \gamma(f(t))$ belongs to $\text{lbv}(I, \mathbb{R})$ and some properties of the real measure $d(\gamma \circ f)$ are established. In the first case, γ is supposed convex and continuous in Ω . The subdifferential $\partial\gamma$ is invoked in the sense of Convex Analysis; under the ordering of real measures, $d(\gamma \circ f)$ is shown to satisfy some inequalities. This generalizes previous results of one of the authors, aimed at deriving energy-like inequalities in nonsmooth mechanical evolution problems. In the second case, γ is supposed Lipschitz on every bounded subset of Ω and Clarke's generalized gradient of γ is used. In both situations, if γ happens to be Gâteaux-differentiable, and $f \in \text{lbv}(I, X)$ continuous, a chain rule of the familiar form is found to hold. Finally, for γ Fréchet-differentiable, an expression of $d(\gamma \circ f)$ is obtained. © 1987 Academic

Press, Inc

1. INTRODUCTION

Let I denote a real interval and X a real Banach space. By $f \in \text{lbv}(I, X)$ we shall mean that f is a function of I to X with *locally bounded variation*, i.e., it has a bounded variation on every compact subinterval of I . Classically, with every such function f is associated an X -valued measure on I , denoted in the sequel by df and commonly called the *differential measure* (or Stieltjes measure) of f . Precisely, df is a σ -additive X -valued set function, defined on the class of the relatively compact Borel subsets of I , with finite variation. Conversely, every X -valued set function meeting these requirements equals the differential measure of some "cumulative distribution function." (See e.g. [20], [21].)

For the sake of consistency, all the measures considered in this article will similarly be denoted with a d , however archaic this may look.

In usual situations, there exists (non uniquely) a nonnegative real

measure $d\mu$ on I and an element f'_μ of $L^1_{\text{loc}}(I, d\mu; X)$ such that $df = f'_\mu d\mu$. This certainly holds if the Banach space X is reflexive (more generally, X may be specified as possessing the "Radon-Nikodym property"; cf. [1, Chap. III, Sect. 1]; in particular, one may take $d\mu = |df|$, the "absolute value" or "variation measure" of the vector measure df .

The primary object of this article is to extend, in two different ways, the following results previously established by one of the authors [2], with a view to derive energy-like inequalities in some mechanical evolution problems [3]: these concerned the special case where X equals a real Hilbert space H , with norm denoted by $\|\cdot\|$. For every $f \in \text{lbv}(I, H)$, the real function $t \rightarrow \|f(t)\|^2$ belongs to $\text{lbv}(I, \mathbb{R})$; it has been found that its differential measure is expressed by

$$d(\|f\|^2) = (f^+ + f^-) \cdot df. \quad (1.1)$$

The dot in the right-hand side refers to the scalar product in H ; this right-hand side is naturally understood as a real measure, since the functions $t \rightarrow f^+(t)$ and $t \rightarrow f^-(t)$, the right-limit and left-limit of f at every point, as elements of $\text{lbv}(I, H)$, are universally locally integrable; more concretely, a representation $df = f'_\mu d\mu$ as above may be used in order to calculate this right-hand side. Furthermore, the following inequalities hold in the sense of the ordering of real measures

$$2f^- \cdot df \leq d(\|f\|^2) \leq 2f^+ \cdot df. \quad (1.2)$$

In the special case $H = \mathbb{R}$, equality (1.1) was stated by Daniell [4] as early as 1918.

For the generalizations in view, the squared norm in H is replaced by some function γ from an open subset of X to \mathbb{R} , which shall be supposed either *convex and continuous* or *Lipschitz*.

Concerning the notations f^+ and f^- for the respective right and left limits in the case of a possible endpoint of I , let us agree on the following: if I includes its possible left end, say t_l , then by convention $f^-(t_l) = f(t_l)$; symmetrically, if I includes its possible right end, say t_r , then $f^+(t_r) = f(t_r)$. These conventions secure that, generally, for every $[a, b] \subset I$, $df([a, b]) = f^+(b) - f^-(a)$.

2. STATEMENT OF RESULTS

In Convex Analysis $\partial\gamma(x)$ denotes the *subdifferential* of γ at the point x , i.e., the subset of the dual space X' of X consisting of the elements x' such that the affine function $u \rightarrow \langle u - x, x' \rangle + \gamma(x)$ minorizes γ . In the case where γ is convex and Gâteaux-differentiable at the point x , this trivially reduces to the single element $\nabla\gamma(x)$, the gradient of γ at this point.

THEOREM 1. Let $f \in \text{lbv}(I, X)$ with differential measure $df = f'_\mu d\mu$; here $d\mu$ is a nonnegative real measure on I and $f'_\mu \in L^1_{\text{loc}}(I, d\mu; X)$. Let Ω be an open convex subset of X and $\gamma: \Omega \rightarrow \mathbb{R}$ be continuous and convex. Assume that, for every compact subinterval $[a, b]$ of I , the closure of $f([a, b])$ is contained in Ω .

Then $\gamma \circ f: t \rightarrow \gamma(f(t))$ belongs to $\text{lbv}(I, \mathbb{R})$ and its differential measure possesses, relative to $d\mu$, a density $(\gamma \circ f)'_\mu \in L^1_{\text{loc}}(I, d\mu; \mathbb{R})$.

The real functions

$$\varphi_*: t \rightarrow \sup\{\langle g, f'_\mu(t) \rangle: g \in \partial\gamma(f^-(t))\}$$

$$\varphi^*: t \rightarrow \inf\{\langle g, f'_\mu(t) \rangle: g \in \partial\gamma(f^+(t))\}$$

belong to $L^1_{\text{loc}}(I, d\mu; \mathbb{R})$ and the inequalities

$$\varphi_* \leq (\gamma \circ f)'_\mu \leq \varphi^* \quad (2.1)$$

hold $d\mu$ -a.e. in I .

In the set $I_c = \{t \in I: df(\{t\}) = 0\}$, equality $\varphi_* = \varphi^*$ holds $d\mu$ -a.e. (in other words, $d\mu$ -a.e. in this set the subdifferential $\partial\gamma(f^+(t))$, equivalently $\partial\gamma(f^-(t))$, is contained in an affine hyperplane of X' , orthogonal to $f'_\mu(t)$).

Comments on the Gâteaux-Differentiability of Convex Functions. If γ is Gâteaux-differentiable at every point of Ω , then convexity ensures the continuity of this function (in fact γ equals, in that case, the supremum of a collection of continuous affine functions; hence it is l.s.c. and therefore continuous throughout the open convex set Ω , because X is a barrelled space [5]).

The following implications are also useful; if $\gamma: \Omega \rightarrow \mathbb{R}$ is convex and continuous, then at every point $x \in \Omega$, the subdifferential $\partial\gamma(x)$ is nonempty; γ is Gâteaux-differentiable at x if and only if $\partial\gamma(x)$ reduces to a singleton; a sufficient condition for that is the strict convexity of the conjugate function γ^* of X' to $]-\infty, +\infty]$.

When Theorem 1 is specialized to the case of a convex Gâteaux-differentiable function γ , with gradient at a point x denoted by $\nabla\gamma(x)$, it yields that a chain rule of the familiar form

$$(\gamma \circ f)'_\mu = \langle \nabla\gamma(f), f'_\mu \rangle \quad (2.2)$$

holds $d\mu$ -a.e. in I_c (and, a fortiori, in the set of the continuity points of f); generally, $d\mu$ -a.e. in I , one has

$$\langle \nabla\gamma(f^-), f'_\mu \rangle \leq (\gamma \circ f)'_\mu \leq \langle \nabla\gamma(f^+), f'_\mu \rangle$$

inequalities which extend (1.2).

Brezis [6, lemme 3.3 p. 73] has obtained an equality of the same sort as (2.2) by assuming $f \in W_{\text{loc}}^{1,2}(I, H)$ (H is a Hilbert space), but requiring only of γ to be convex and l.s.c. from H to $]-\infty, +\infty]$: if there exists $g \in L_{\text{loc}}^2(I, H)$ such that $g(t) \in \partial\gamma(f(t))$ holds Lebesgue-a.e. then, with $d\mu$ equal to the Lebesgue measure of I , he proves that $\gamma \circ f$ is locally absolutely continuous and that, for almost every t ,

$$\forall h \in \partial\gamma(f(t)): (\gamma \circ f)'_{\mu}(t) = \langle h, f'_{\mu}(t) \rangle.$$

For the second type of results of this paper Ω is still supposed open and convex, but the convexity assumption for $\gamma: \Omega \rightarrow \mathbb{R}$ is dropped; instead, this real function is assumed Lipschitz on every bounded subset of Ω . Again, we shall need the closure of every $f([a, b])$ to be contained in Ω . By $\partial\gamma(x)$ will be denoted *Clarke's generalized gradient* [7] at the point x , a convex weakly* compact subset of X' .

THEOREM 2. *With $X, \Omega, I, f, df, d\mu, f'_{\mu}$ as before, suppose that $\gamma: \Omega \rightarrow \mathbb{R}$ is Lipschitz on every bounded subset of Ω . Then $\gamma \circ f$ belongs to $\text{lbv}(I, \mathbb{R})$ and $d(\gamma \circ f)$ admits, relatively to $d\mu$, a density $(\gamma \circ f)'_{\mu} \in L_{\text{loc}}^1(I, d\mu; \mathbb{R})$.*

For $d\mu$ -almost every t in I there exists x in $[f^{-}(t), f^{+}(t)]$ (the closed line segment in X with endpoints the left limit and the right limit of f at t) and $x' \in \partial\gamma(x)$ (Clarke's generalized gradient of γ at x) such that

$$(\gamma \circ f)'_{\mu}(t) = \langle x', f'_{\mu}(t) \rangle. \quad (2.3)$$

Comments on the Gateaux Differentiability of Lipschitz Functions. Suppose γ is Lipschitz on every bounded subset of the open subset Ω of X . Then Clarke's generalized gradient $\partial\gamma(x)$ is a singleton for every x in Ω if and only if γ is Gâteaux-differentiable in Ω with gradient mapping $x \rightarrow \nabla\gamma(x)$ continuous from Ω to X'_s [18]. This simplifies the application of Theorem 2. If in addition f is continuous, a chain rule of the form (2.2) $d\mu$ -a.e. holds and the convexity of Ω is not required.

The third theorem of this paper is aimed at constructing an expression of $d(\gamma \circ f)$, in order to generalize equality (1.1). A hint is found in the monograph of Vol'pert [8, Sect. 13.2, p. 248]; this author considers a function f , with values in \mathbb{R}^n , defined in an open subset U of \mathbb{R}^n . Such a function is said to be of locally bounded variation if its partial derivatives in the sense of distributions in U are measures; this generalizes $\text{lbv}(I)$ in a specific way involving the Lebesgue measure of U essentially. In contrast, for $n = 1$ (and for an interval I nonnecessarily open in \mathbb{R}) the concept of a locally bounded variation relies only on the ordering of \mathbb{R} , without any reference to Lebesgue measure. Restricting ourselves to $n = 1$ will enable us to consider as before a function f with values in an arbitrary Banach space.

THEOREM 3. *Let $X, \Omega, I, f, df, d\mu, f'_\mu$ be as in the preceding theorems; suppose that $\gamma: \Omega \rightarrow \mathbb{R}$ is continuously Fréchet-differentiable, with gradient at point x denoted by $\nabla\gamma(x)$. Then $\gamma \circ f \in \text{lbv}(I, \mathbb{R})$ and the differential measure $d(\gamma \circ f)$ admits as density relative to $d\mu$ the function $t \rightarrow \langle \theta(t), f'_\mu(t) \rangle$, where $\theta: I \rightarrow X'$ is defined as*

$$\theta(t) = \int_0^1 \nabla\gamma[(1-r)f^-(t) + rf^+(t)] dr.$$

3. A DERIVATION PROPERTY

The following extends to vector measures a result of Jeffery [9].

PROPOSITION 1. *Let I denote a real interval, X a real Banach space, $d\mu$ a nonnegative Radon measure on I , dv an X -valued measure on I admitting a density $dv/d\mu \in L^1_{\text{loc}}(I, d\mu; X)$. Then, for $d\mu$ -almost every t ,*

$$\frac{dv}{d\mu}(t) = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{dv([t, t + \varepsilon])}{d\mu([t, t + \varepsilon])} = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{dv([t - \varepsilon, t])}{d\mu([t - \varepsilon, t])}. \quad (3.1)$$

By $d\mu([t, t + \varepsilon])$, we mean the $d\mu$ -measure of $I \cap [t, t + \varepsilon]$ and similar abridged notations apply to dv and to the interval $I \cap [t - \varepsilon, t]$.

Proof. The writing in (3.1) makes sense only if t does not belong to the respective subsets of I ,

$$I_r = \{t \in I: \exists \alpha > 0, d\mu(I \cap [t, t + \alpha]) = 0\}$$

$$I_l = \{t \in I: \exists \beta > 0, d\mu(I \cap [t - \beta, t]) = 0\}.$$

Let us check that these two subsets are $d\mu$ -negligible. In fact, both I_r and I_l contain I_0 , the greatest open subset of I throughout which the measure $d\mu$ vanishes; I_0 equals the union of an at most countable collections (J_n) of disjoint subintervals of I and $d\mu(I_0) = 0$. Any element t_e of $I_r \setminus I_0$ (except trivially the possible right end of I) is the left end of a nonzero interval closed on the left with zero $d\mu$ -measure; hence $d\mu(\{t_e\}) = 0$ and t_e equals the left end of some of the intervals (J_n) . This shows that $I_r \setminus I_0$ is at most countable; therefore $d\mu(I_r) = 0$ and similar reasoning yields $d\mu(I_l) = 0$.

Let h denote a representative of the element $dv/d\mu$ of $L^1_{\text{loc}}(I, d\mu; X)$; since this vector function is locally Bochner-integrable, it is strongly measurable. Hence there exists a separable subspace Z of X such that $h(t) \in Z$ for $d\mu$ -almost every t ; let (z_n) denote a dense sequence in Z .

Let us apply Jeffery's theorem to the nonnegative real measure dv_n which

admits as density relative to $d\mu$ the function $t \rightarrow \|h(t) - z_n\|$: there exists N_n , a $d\mu$ -negligible subset of I such that, for $t \notin N_n$,

$$\|h(t) - z_n\| = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{dv_n([t, t + \varepsilon])}{d\mu([t, t + \varepsilon])}.$$

Let N denote the union of $\{t: h(t) \notin Z\}$ with the totality of the sets N_n . Let $t \notin N$ and $\eta > 0$; there exists an integer n such that

$$\|h(t) - z_n\| \leq \eta. \quad (3.2)$$

Since $t \notin N_n$, there exists $\varepsilon_0 > 0$ such that, for every ε in $]0, \varepsilon_0]$,

$$\left| \frac{dv_n([t, t + \varepsilon])}{d\mu([t, t + \varepsilon])} - \|h(t) - z_n\| \right| \leq \eta. \quad (3.3)$$

In view of inequalities

$$\begin{aligned} \int_{[t, t + \varepsilon]} \|h(s) - h(t)\| d\mu(s) &\leq \int_{[t, t + \varepsilon]} (\|h(s) - z_n\| + \|h(t) - z_n\|) d\mu(s) \\ &\leq dv_n([t, t + \varepsilon]) + \eta d\mu([t, t + \varepsilon]) \end{aligned}$$

one obtains

$$\frac{1}{d\mu([t, t + \varepsilon])} \int_{[t, t + \varepsilon]} \|h(s) - h(t)\| d\mu(s) \leq \frac{dv_n([t, t + \varepsilon])}{d\mu([t, t + \varepsilon])} + \eta.$$

Through the use of (3.2) and (3.3), this yields that, for every ε in $]0, \varepsilon_0]$ the left-hand side is less than or equal to 3η . This left-hand side constitutes in turn an upper bound of

$$\left\| \frac{dv([t, t + \varepsilon])}{d\mu([t, t + \varepsilon])} - h(t) \right\| = \frac{1}{d\mu([t, t + \varepsilon])} \left\| \int_{[t, t + \varepsilon]} (h(s) - h(t)) d\mu(s) \right\|.$$

So the first equality in (3.1) is established and similar reasoning, with intervals of the form $[t - \varepsilon, t]$, applies to the second one.

Remark 1. In view of the convention made, that $d\mu([t, t + \varepsilon])$ should be understood as the $d\mu$ -measure of $I \cap [t, t + \varepsilon]$, the above does not exclude that t equals t_r , the possible right end of I . In such a case $d\mu(\{t_r\}) > 0$ (otherwise, t_r would be comprised in the excluded $d\mu$ -negligible set) and, for every $\varepsilon > 0$,

$$\frac{dv([t_r, t_r + \varepsilon])}{d\mu([t_r, t_r + \varepsilon])} = \frac{dv(\{t_r\})}{d\mu(\{t_r\})}$$

so that the first equality in (3.1) holds trivially. A similar remark applies to the second equality and the possible left end of I .

Remark 2. For an alternative proof, valid when X has the Radon–Nikodym property (thus applicable to Jeffery’s original case $X = \mathbb{R}$), see [17]: there is used a technique of “jump unfolding” which reduces lbv functions to Lipschitz functions.

4. THE CONVEX CASE

Throughout this section, f and γ are assumed to satisfy the assumptions of Theorem 1.

Observe that from $f \in \text{lbv}(I, X)$ it is easily deduced that the image $f([a, b])$ of every compact subinterval of I has compact closure in X ; Theorem 1 supposes this closure contained in Ω . Assumption $f(I) \subset \Omega$ would not be sufficient in order that $\gamma \circ f \in \text{lbv}(I, \mathbb{R})$. Counterexample: take $X = \mathbb{R}$, $\Omega =] -\pi/2, \pi/2[$, $\gamma(x) = tg^2 x$, $I = [0, \pi/2]$ and

$$f(t) = \begin{cases} t & \text{if } t \in [0, \pi/2[\\ 0 & \text{if } t = \pi/2. \end{cases}$$

LEMMA 1. *Denote by X'_s the dual space of X , equipped with the weak* topology. Under the assumptions made, the multifunction $x \rightarrow \partial\gamma(x)$ is upper semicontinuous from Ω to X'_s , with nonempty convex compact values. For every compact subset K of Ω , the real function γ is Lipschitz on K .*

Proof. By the continuity of γ , every point in Ω possesses a neighborhood, say V , throughout which γ is bounded from above. Therefore (cf. [10] or, for more details [11, Sect. 11.e]; see also a proof in [12, Theorem 10]) $\partial\gamma$ is an upper semicontinuous multifunction of V to X'_s , with nonempty convex compact values and the image of V under $\partial\gamma$ is an equicontinuous subset of X' .

Since K may be covered by a finite collection of such neighborhoods, the image of K under $\partial\gamma$ is an equicontinuous subset of X' , thus contained in a ball centered at the origin; denote its radius by C . Let x and y be elements of K ; for every $x' \in \partial\gamma(x)$ one has

$$\gamma(y) \leq \gamma(x) + \langle x', y - x \rangle \leq \gamma(x) + C \|y - x\|$$

and the similar inequality after exchanging x and y . So the Lipschitz property is proved (for finite dimensional X , this is Theorem 24.7 in [13]).

LEMMA 2. *The function $\gamma \circ f$ belongs to $\text{lbv}(I, \mathbb{R})$; its differential measure is continuous relatively to $|df|$, the absolute value of the vector measure df .*

Proof. Let $[a, b]$ denote a compact subinterval of I . If $C \geq 0$ is a Lipschitz constant of γ on the compact set $\text{cl } f([a, b])$, one readily obtains

$$\text{var}(\gamma \circ f; a, b) \leq C \text{var}(f; a, b),$$

which proves that $\gamma \circ f$ has a locally bounded variation.

Let the compact interval $[a', b']$ be a neighborhood of $[a, b]$ in I . This implies $b' > b$, except in the case where b happens to be the right end of I ; symmetrically, $a' < a$ except if a is the left end of I . Recall that, in Section 1, the conventions $f^+(b) = f(b)$ and $f^-(a) = f(a)$ has been made in these respective events. Let C' denote a Lipschitz constant of γ on the compact set $\text{cl } f([a', b'])$. For every $[s, t] \subset [a, b]$ one has

$$\begin{aligned} |d(\gamma \circ f)([s, t])| &= |\gamma(f^+(t)) - \gamma(f^-(s))| \\ &\leq C' \|f^+(t) - f^-(s)\| = C' \|df([s, t])\| \\ &\leq C' |df|([s, t]). \end{aligned}$$

This proves the $|df|$ -continuity of $d(\gamma \circ f)$.

LEMMA 3. *Let $d\mu$ be a nonnegative real Radon measure on I such that df is $d\mu$ -continuous. Suppose (this is automatic if X has the Radon–Nikodym property) that df admits, relatively to $d\mu$ a density $f'_\mu \in L^1_{\text{loc}}(I, d\mu; X)$. Then the real functions defined as*

$$\begin{aligned} \varphi_*: \quad t &\rightarrow \sup \{ \langle g, f'_\mu(t) \rangle : g \in \partial\gamma(f^-(t)) \} \\ \varphi^*: \quad t &\rightarrow \inf \{ \langle g, f'_\mu(t) \rangle : g \in \partial\gamma(f^+(t)) \} \end{aligned}$$

are $d\mu$ -measurable (or even Borelian, if the chosen representative of the element f'_μ of L^1_{loc} is constructed as the limit of a sequence of Borelian step-functions) and locally $d\mu$ -integrable.

Proof. Take f'_μ as the limit of a sequence of Borelian step-functions. We are to show that φ_* is Borelian on every $[a, b] \subset I$. Let $K = \text{cl } f([a, b])$, a compact subset of Ω to which corresponds a constant C as in the proof of Lemma 1. For u in X and x in K , denote by $\Phi(u, x)$ the value at point u of the support function of $\partial\gamma(x)$; then

$$\varphi_*(t) = \Phi(f'_\mu(t), f^-(t)). \quad (4.1)$$

Since f'_μ and f^- equal the limits of sequences of Borelian step-functions, the function $t \rightarrow (f'_\mu(t), f^-(t))$ is Borelian from $]a, b]$ to $X \times K$. We just have to show that Φ is Borelian.

In view of Lemma 1, one has

- (i) $\forall u \in X, \forall x \in K, \Phi(u, x) \leq C\|u\|,$
- (ii) $\forall x \in K, u \rightarrow \Phi(u, x)$ is a continuous real function in X ,
- (iii) $\forall u \in X, x \rightarrow \Phi(u, x)$ is an u.s.c. real function in K .

For every positive integer k , define

$$\Phi_k(u, x) = \sup_{v \in K} (\Phi(u, v) - k\|v - x\|) \quad (4.2)$$

which is finite throughout $X \times K$, in view of (i). Due to (ii), $u \rightarrow \Phi_k(u, x)$ is l.s.c., hence Borelian; $x \rightarrow \Phi_k(u, x)$ is k -Lipschitz, hence continuous. Since K is separable and metrizable, it follows (cf. [16, Lemma III.14]) that Φ_k is Borelian on $X \times K$.

That Φ is Borelian results from

$$\Phi(u, x) = \inf_k \Phi_k(u, x).$$

In fact $\Phi \leq \Phi_k$; one has to check that for fixed u, x and $\varepsilon > 0$, there exists k such that $\Phi_k(u, x) < \Phi(u, x) + \varepsilon$. Now, due to (iii) there exists $\beta > 0$ such that for $\|v - x\| < \beta$ in (4.2), $\Phi(u, x) + \varepsilon$ majorizes $\Phi(u, v)$ thus majorizes also $\Phi(u, v) - k\|v - x\|$; on the contrary, for $\|v - x\| \geq \beta$ one has

$$\Phi(u, v) - k\|v - x\| \leq \max_K \Phi(u, \cdot) - k\beta$$

which may be made less than $\Phi(u, x) + \varepsilon$ by taking k large enough.

The above technique may be traced back to Castaing [14]; a similar argument was also used by Hausdorff [15, Sect. 42, pp. 282–283].

Finally observe that, due to (i), one has

$$|\varphi_*(t)| \leq C\|f'(t)\|,$$

an inequality which proves that φ_* is locally $d\mu$ -integrable.

Similar reasoning applies to φ^* .

Remark. The preceding does not rely on any separability assumption. If X is separable, the measurability of φ_* and φ^* may also be derived from Castaing's representation of the multifunction $\partial\gamma$ (see, e.g., [16]).

Proof of Theorem 1. It just remains to complete the preceding lemmata by the calculation that follows. In view of Proposition 1, there exists a $d\mu$ -negligible subset N of I such that, for $t \notin N$, one has, with nonzero denominators,

$$(\gamma \circ f)'_{\mu}(t) = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{d(\gamma \circ f)([t, t + \varepsilon])}{d\mu([t, t + \varepsilon])} = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{d(\gamma \circ f)([t - \varepsilon, t])}{d\mu([t - \varepsilon, t])}$$

and

$$f'_\mu(t) = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{df([t, t + \varepsilon])}{d\mu([t, t + \varepsilon])} = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{df([t - \varepsilon, t])}{d\mu([t - \varepsilon, t])}.$$

Now, if $g \in \partial\gamma(f^-(t))$,

$$d(\gamma \circ f)([t, t + \varepsilon]) = \gamma(f^+(t + \varepsilon)) - \gamma(f^-(t)) \geq \langle g, f^+(t + \varepsilon) - f^-(t) \rangle.$$

After dividing by $d\mu([t, t + \varepsilon])$ one makes ε tend to zero, hence

$$(\gamma \circ f)'_\mu(t) \geq \langle g, f'_\mu(t) \rangle.$$

By taking the supremum for g ranging through $\partial\gamma(f^-(t))$ one obtains the first inequality (2.1).

Strictly speaking, the above supposes that t is not the right end of I ; otherwise the convention made in the Introduction, that $f^+(t) = f(t)$, has to be invoked.

The second inequality in (2.1) is proved symmetrically.

As for the last statement in Theorem 1, observe that $df(\{t\})$ vanishes if and only if $f^-(t) = f^+(t)$. In that case $\varphi_* \geq \varphi^*$; in view of (2.1) this yields $\varphi_* = \varphi^*$ $d\mu$ -a.e. in I_c .

5. PROOF OF THEOREM 2

The fact that $\gamma \circ f$ belongs to $\text{lbv}(I, \mathbb{R})$ and the existence of $(\gamma \circ f)'_\mu \in L^1_{\text{loc}}(I, d\mu; \mathbb{R})$ are proved through similar inequalities as before.

First, suppose $d\mu(\{t\}) > 0$; then

$$\begin{aligned} (\gamma \circ f)'_\mu(t) &= \frac{\gamma(f^+(t)) - \gamma(f^-(t))}{d\mu(\{t\})} \\ f'_\mu(t) &= \frac{1}{d\mu(\{t\})} (f^+(t) - f^-(t)). \end{aligned}$$

If $f^+(t) = f^-(t) = x$, any element x' of $\partial\gamma(x)$ satisfies (2.3) since $f'_\mu(t) = 0$ in that case. Otherwise, use the mean value theorem of Lebourg [18]: there exists x in $]f^-(t), f^+(t)[$ and x' in $\partial\gamma(x)$ such that $\gamma(f^+(t)) - \gamma(f^-(t)) = \langle x', f^+(t) - f^-(t) \rangle$.

Now, suppose $d\mu(\{t\}) = 0$, hence $f^-(t) = f^+(t)$. With the exception of a $d\mu$ -negligible subset of I (in particular, let us exclude the possible right end of I), Proposition 1 ensures that

$$\begin{aligned}
(\gamma \circ f)'_{\mu}(t) &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{\gamma(f^+(t+\varepsilon)) - \gamma(f^-(t))}{d\mu([t, t+\varepsilon])} \\
f'_{\mu}(t) &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{1}{d\mu([t, t+\varepsilon])} (f^+(t+\varepsilon) - f^-(t)).
\end{aligned} \tag{5.1}$$

We use again Lebourg's mean value theorem: there exist x_{ε} in $]f^-(t), f^+(t+\varepsilon)[$ and x'_{ε} in $\partial\gamma(x_{\varepsilon})$ such that

$$\frac{\gamma(f^+(t+\varepsilon)) - \gamma(f^-(t))}{d\mu([t, t+\varepsilon])} = \left\langle x'_{\varepsilon}, \frac{1}{d\mu([t, t+\varepsilon])} (f^+(t+\varepsilon) - f^-(t)) \right\rangle.$$

Let the compact interval $[a, b]$ be a neighborhood of t in I . In view of the assumptions made, $\text{cl } f([a, b])$ is a compact subset of Ω ; it contains $f^-(t)$ and also $f^+(t+\varepsilon)$ if ε is small enough. The multifunction $\partial\gamma$ is known [7, 19] to be upper semicontinuous from Ω to X' , (the dual space of X equipped with the weak* topology) with compact values. When $\varepsilon > 0$ tends to zero, $f^+(t+\varepsilon)$ tends strongly to $f^+(t) = f^-(t)$; so does x_{ε} . Therefore the net x'_{ε} possesses a cluster point in X' , say x' , belonging to $\partial\gamma(f^-(t))$. Since the convergence in (5.1) is strong, this establishes (2.3).

6. PROOF OF THEOREM 3

By the same arguments as before, $\gamma \circ f$ belongs to $\text{lbv}(I, \mathbb{R})$ and there exists $(\gamma \circ f)'_{\mu} \in L^1_{\text{loc}}(I, d\mu; \mathbb{R})$.

Put $I_c = \{t \in I: d\mu(\{t\}) = 0\}$; let us first establish the asserted formula for $t \in I_c$. Then we are allowed to neglect the possibility of t being an endpoint of I . In view of Proposition 1 one has, except for t is some $d\mu$ -negligible subset,

$$\begin{aligned}
(\gamma \circ f)'_{\mu}(t) &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{d(\gamma \circ f)([t, t+\varepsilon])}{d\mu([t, t+\varepsilon])} \\
&= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{\gamma(f^+(t+\varepsilon)) - \gamma(f^-(t))}{d\mu([t, t+\varepsilon])}
\end{aligned} \tag{6.1}$$

and

$$f'_{\mu}(t) = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \frac{1}{d\mu([t, t+\varepsilon])} (f^+(t+\varepsilon) - f^-(t)).$$

The latter equivalently writes down as

$$f^+(t+\varepsilon) - f^-(t) = d\mu([t, t+\varepsilon])(f'_{\mu}(t) + \eta(\varepsilon)), \tag{6.2}$$

where the function $\varepsilon \rightarrow \eta(\varepsilon) \in X$ tends normwise to zero when ε tends to zero; also $d\mu([t, t + \varepsilon])$ tends to zero, since $t \in I_c$.

The Fréchet-differentiability of γ at some point x means that

$$\gamma(x + h) = \gamma(x) + \langle \nabla \gamma(x), h \rangle + \|h\| \zeta(h),$$

where the real function ζ tends to zero with $\|h\|$. By making $x = f^-(t)$ and $h = f^+(t + \varepsilon) - f^-(t)$, as expressed by (6.2), this yields

$$\begin{aligned} \gamma(f^+(t + \varepsilon)) - \gamma(f^-(t)) &= \langle \nabla \gamma(f^-(t)), d\mu([t, t + \varepsilon])(f'_\mu(t) + \eta(\varepsilon)) \rangle \\ &\quad + d\mu([t, t + \varepsilon])\|f'_\mu(t) + \eta(\varepsilon)\| \zeta(h). \end{aligned}$$

When ε tends to zero, $f'_\mu(t) + \eta(\varepsilon)$ remains bounded in X , while $\zeta(h)$ tends to zero; thus the limit appearing in the last member of (6.1) equals $\langle \nabla \gamma(f^-(t), f'_\mu(t)) \rangle$. Since in the present case $d\mu(\{t\}) = 0$, one has $df(\{t\}) = 0$, i.e., $f^+(t) = f^-(t)$. Therefore the asserted formula is true, as the expression of θ reduces to

$$\theta(t) = \nabla \gamma(f^-(t)) = \nabla \gamma(f^+(t)).$$

Let us consider now the case where $d\mu(\{t\}) > 0$, i.e., t is an atom of $d\mu$. Then

$$(\gamma \circ f)'_\mu(t) = \frac{d(\gamma \circ f)(\{t\})}{d\mu(\{t\})} = \frac{\gamma(f^+(t)) - \gamma(f^-(t))}{d\mu(\{t\})} \quad (6.3)$$

and

$$f'_\mu(t) = \frac{1}{d\mu(\{t\})} (f^+(t) - f^-(t)) \quad (6.4)$$

(recall the conventions made in the Introduction if t is an endpoint of I). Since it has been assumed that $\text{cl } f([a, b]) \subset \Omega$ for every $[a, b] \subset I$, the points $f^+(t)$ and $f^-(t)$ belong to Ω ; the continuous differentiability of γ throughout this open convex set yields

$$\begin{aligned} \gamma(f^+(t)) - \gamma(f^-(t)) &= \int_0^1 \langle \nabla \gamma[(1-r)f^-(t) + rf^+(t)], f^+(t) - f^-(t) \rangle dr \\ &= \langle \theta(t), f^+(t) - f^-(t) \rangle. \end{aligned}$$

In view of (6.3) and (6.4), this establishes the formula.

REFERENCES

1. J. DIESTEL AND J. J. UHL, Vector measures, *Amer. Math. Soc. Math. Surveys* **15** (1977).
2. J. J. MOREAU, Sur les mesures différentielles de fonctions vectorielles et certains problèmes d'évolution, *C. R. Acad. Sci. Paris* **282** (1976), 837-840.

3. J. J. MOREAU, Evolution problem associated with a moving convex set in a Hilbert space, *J. Differential Equations* **26** (1977), 347–374.
4. P. J. DANIELL, Differentiation with respect to a function of limited variation, *Trans. Amer. Math. Soc.* **19** (1918), 353–362.
5. R. T. ROCKAFELLAR, Level sets and continuity of conjugate convex functions, *Trans. Amer. Math. Soc.* **123** (1966), 46–63.
6. H. BREZIS, “Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert,” North-Holland, Amsterdam, 1973.
7. F. H. CLARKE, Generalized gradients of Lipschitz functionals, *Advan. in Math.* **40** (1981), 52–67.
8. A. I. VOL’PERT, The space BV and quasilinear equations, *Math. USSR-Sb.* **2**, No. 2 (1965), 225–267; translated from Russian.
9. R. L. JEFFERY, Non-absolutely convergent integrals with respect to functions of bounded variation, *Trans. Amer. Math. Soc.* **34** (1932), 645–675.
10. J. J. MOREAU, Semi-continuité du sous-gradient d’une fonctionnelle, *C. R. Acad. Sci. Paris* **260** (1965), 1067–1070.
11. J. J. MOREAU, “Fonctionnelles convexes,” Séminaire sur les Equations aux Dérivées Partielles, Collège de France, Paris 1967, multigraphié, 108 pp.
12. M. VALADIER, Sous-différentiabilité de fonctions convexes à valeurs dans un espace vectoriel ordonné, *Math. Scand.* **30** (1972), 65–74.
13. R. T. ROCKAFELLAR, “Convex Analysis,” Princeton Univ. Press, Princeton, N.J., 1970.
14. C. CASTAING, A propos de l’existence des sections séparément mesurables et séparément continues, *Sém. Anal. Convexe, Montpellier* **6** (1976), 6.
15. F. HAUSDORFF, “Set Theory,” Chelsea, 1957; Engl. transl. from “Mengenlehre,” 3rd ed., 1937.
16. C. CASTAING AND M. VALADIER, Convex analysis and measurable multifunctions, “Lectures Notes in Math.,” No. 580, Springer-Verlag, Berlin, 1977.
17. J. J. MOREAU AND M. VALADIER, Dérivation d’une mesure vectorielle sur un intervalle, *Sém. Anal. Convexe, Montpellier* **14** (1984), 1.
18. G. LEBOURG, Valeur moyenne pour gradient généralisé, *C. R. Acad. Sci. Paris, Ser. A* **281** (1975), 795–797.
19. F. H. CLARKE, A new approach to Lagrange multipliers, *Math. Oper. Res.* **1** (1976), 167–174.
20. N. DINCULEANU, “Vector Measures,” Pergamon, Oxford, 1967.
21. J. J. MOREAU, Bounded Variation in time, in (J. J. Moreau, P. D. Panagiotopoulos, and G. Strang, Eds.), “Topics in Nonsmooth Mechanics,” Birkhäuser, Basel, 1987, to appear.