Bollettino Unione Matematica Italiana

François Murat, Cristina Trombetti

A chain rule formula for the composition of a vector-valued function by a piecewise smooth function

Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. **6-B** (2003), n.3, p. 581–595.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2003_8_6B_3_581_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2003.

Bollettino U. M. I.
(8) **6**-B (2003), 581-595

A Chain Rule Formula for the Composition of a Vector-Valued Function by a Piecewise Smooth Function.

FRANÇOIS MURAT - CRISTINA TROMBETTI

Sunto. – *Si enuncia e si dimostra una formula di derivazione per funzioni $T(u)$ ottenute componendo una funzione a valori vettoriali $u \in W^{1,r}(\Omega; \mathbb{R}^M)$ con una funzione T globalmente lipschitziana e C^1 a tratti. Si dimostra inoltre che l'applicazione $u \to T(u)$ è continua da $W^{1,r}(\Omega; \mathbb{R}^M)$ in $W^{1,r}(\Omega)$ rispetto alle topologie forti di questi spazi.*

Summary. – *We state and prove a chain rule formula for the composition $T(u)$ of a vector-valued function $u \in W^{1,r}(\Omega; \mathbb{R}^M)$ by a globally Lipschitz-continuous, piecewise C^1 function T. We also prove that the map $u \to T(u)$ is continuous from $W^{1,r}(\Omega; \mathbb{R}^M)$ into $W^{1,r}(\Omega)$ for the strong topologies of these spaces.*

1. – Introduction.

The prototype of the result that we will prove in the present paper is the following:

PROPOSITION 1.1. – *Let Ω be an open set of $\mathbb{R}^N$, $N \geq 1$, and let r be a real number with $1 \leq r < +\infty$. Let u_1 and u_2 be two functions of $W^{1,r}(\Omega)$, and let $S : \mathbb{R}^2 \to \mathbb{R}$ be the supremum defined by*

$$
(1.1) \quad \begin{cases} S(y_1, y_2) = \sup(y_1, y_2) \\ \qquad = \chi_{\{y_1 < y_2\}} y_2 + \chi_{\{y_1 > y_2\}} y_1 + \chi_{\{y_1 = y_2\}} y_1 \\ \qquad = \chi_{\{y_1 < y_2\}} y_2 + \chi_{\{y_1 > y_2\}} y_1 + \chi_{\{y_1 = y_2\}} y_2 \quad \forall (y_1, y_2) \in \mathbb{R}^2. \end{cases}
$$

Then the function $v = S(u_1, u_2) = \sup(u_1, u_2)$ belongs to $W^{1,r}(\Omega)$ and its gradient is given by

$$(1.2)\qquad \begin{cases} Dv = \chi_{\{u_1 < u_2\}} Du_2 + \chi_{\{u_1 > u_2\}} Du_1 + \chi_{\{u_1 = u_2\}} Du_1 \\ \quad = \chi_{\{u_1 < u_2\}} Du_2 + \chi_{\{u_1 > u_2\}} Du_1 + \chi_{\{u_1 = u_2\}} Du_2 \quad \text{a.e. in } \Omega\,. \end{cases}$$

Moreover the map $u \to S(u)$ is sequentially continuous from $W^{1,r}(\Omega; \mathbb{R}^2)$ into $W^{1,r}(\Omega)$ for the weak topologies of these spaces, and continuous for the strong topologies of these spaces.

In this statement, as well as in the whole of the present paper, Df and f' denote the derivative of the function f, while χ_E denotes the characteristic function of the set E and $\{f < g\}$ the set $\{z : f(z) < g(z)\}$.

Observe that formula (1.2) is very simple, and that there is a strong analogy between this formula and the formula (1.1) which defines the function S.

Observe also that formula (1.2) implies that

$$(1.3)\qquad Du_1(x) = Du_2(x) \quad \text{a.e.} \quad x \in \{x \in \Omega : u_1(x) = u_2(x)\}.$$

We will actually prove a general result (see Theorem 2.1 below) where the pair (u_1, u_2) is replaced by a vector-valued function $u = (u_m)_{1 \leq m \leq M}$ of $W^{1,r}(\Omega; \mathbb{R}^M)$, and where the supremum $S(y_1, y_2) = \sup(y_1, y_2) : \mathbb{R}^2 \to \mathbb{R}$ is replaced by a globally Lipschitz-continuous, piecewise affine (or even piecewise C^1) function $T : \mathbb{R}^M \to \mathbb{R}$.

The proof of Proposition 1.1 easily follows from the classical result (see e.g. [KS] or [S]) which asserts that

$$(1.4)\qquad \text{if } w \in W^{1,r}(\Omega), \quad \text{then } Dw(x) = 0 \quad \text{a.e.} \quad x \in \{x \in \Omega : w(x) = 0\}\,.$$

Indeed from the first line of formula (1.1), one deduces that

$$v = \chi_{\{u_1 < u_2\}} u_2 + \chi_{\{u_1 \geq u_2\}} u_1 \quad \text{a.e. in } \Omega\,,$$

which implies that

$$(1.5)\qquad \begin{cases} \{x \in \Omega : u_1(x) < u_2(x)\} \subset \{x \in \Omega : v(x) = u_2(x)\}\,, \\ \{x \in \Omega : u_1(x) \geq u_2(x)\} \subset \{x \in \Omega : v(x) = u_1(x)\}\,. \end{cases}$$

Once one knows that v belongs to $W^{1,r}(\Omega)$, property (1.4) implies that

$$D(v - u_2)(x) = 0 \quad \text{a.e.} \quad x \in \{x \in \Omega : v(x) - u_2(x) = 0\}\,,$$

$$D(v - u_1)(x) = 0 \quad \text{a.e.} \quad x \in \{x \in \Omega : v(x) - u_1(x) = 0\}\,,$$

which together with (1.5) implies that

$$Dv(x) = Du_2(x) \quad \text{a.e.} \quad x \in \{x \in \Omega : u_1(x) < u_2(x)\},$$
$$Dv(x) = Du_1(x) \quad \text{a.e.} \quad x \in \{x \in \Omega : u_1(x) \geqslant u_2(x)\}.$$

This proves the first line of formula (1.2). The proof of the second one is similar.

Therefore the result of the present paper could be considered as «well known». We nevertheless decided to write it for four reasons.

The first one is that we never saw in the literature the chain rule formula for the composition $T(u)$ of a vector-valued function u by a globally Lipschitz-continuous, piecewise C^1 function T written as in (1.2), or more in general as in (2.4) below, except in the special case of truncations in [L p. 29], even if generalizations have been proposed e.g. in [AD p. 701], [B p. 79] and [MM1 p. 298 and 315] for general Lipschitz-continuous functions T.

The second one is that in the scalar-valued case ($M = 1$), one very often sees the following formulation (see e.g. [BM p. 554], [KS p. 54], [MM2 p. 353], [S p. 15]):

PROPOSITION 1.2. – *Let Ω be an open set of $\mathbb{R}^N$, and let $T : \mathbb{R} \to \mathbb{R}$ be a Lipschitz-continuous, piecewise C^1 function with $T(0) = 0$. By piecewise C^1, we mean that there exist $c_\alpha \in \mathbb{R}$, $1 \leqslant \alpha \leqslant P$, with $-\infty = c_0 < c_1 < c_2 < \ldots < c_{P-1} < < c_P < c_{P+1} = +\infty$ such that on the interval $\{c_\alpha \leqslant y \leqslant c_{\alpha+1}\}$, $0 \leqslant \alpha \leqslant P$, the function T coincides with a function $T_\alpha : \mathbb{R} \to \mathbb{R}$ which is Lipschitz-continuous and C^1 on the whole of $\mathbb{R}$. Then for every $u \in W^{1,r}(\Omega)$, with $1 \leqslant r < +\infty$, the function $T(u)$ belongs to $W^{1,r}(\Omega)$ and one has*

$$(DT(u))(x) = T'(u(x))\, Du(x) \quad \text{a.e.} \quad x \in \Omega, \tag{1.6}$$

where the following abuse of notation is made: the function $T'(u)\,Du$ is defined to be equal to 0 on the set $E = \bigcup_{1 \leqslant \alpha \leqslant P} \{x \in \Omega : u(x) = c_\alpha\}$ (which is the set where the function $T'(u)$ is not defined).

The above abuse of notation is «justified» by the fact that $Du = 0$ almost everywhere on the set E, a fact which follows from (1.4). This abuse of notation can be avoided by writing, in place of formula (1.6),

$$\begin{cases} (DT(u))(x) = \sum\limits_{0 \leqslant \alpha \leqslant P} \chi_{\{c_\alpha < y < c_{\alpha+1}\}}(u(x))\, T'_\alpha(u(x))\, Du(x) + \\ \qquad + \sum\limits_{1 \leqslant \alpha \leqslant P} \chi_{\{y = c_\alpha\}}(u(x))\, 0 \quad \text{a.e.} \quad x \in \Omega, \end{cases} \tag{1.7}$$

in which every function $\chi_{\{c_\alpha < y' < c_{\alpha+1}\}}(y)\, T'_\alpha(y)$ and $\chi_{\{y' = c_\alpha\}}(y)$ is now defined for every $y \in \mathbb{R}$. But while formula (1.7) does not suffer any ambiguity, and can be generalized in the vector-valued case in formula (1.2) or in formula (2.4) below, formula (1.6) does not have a clear generalization even in the simple case where $T(y_1, y_2) = S(y_1, y_2) = \sup(y_1, y_2)$: indeed in this case the analogue of formula (1.6) would read as

$$
\text{(1.8)} \quad \begin{cases} (DS(u_1, u_2))(x) = \\ \qquad = \left(\dfrac{\partial S}{\partial y_1}\right)(u_1(x), u_2(x))\, Du_1(x) + \left(\dfrac{\partial S}{\partial y_2}\right)(u_1(x), u_2(x))\, Du_2(x), \end{cases}
$$

but when, for example, $u_1 = u_2$ on the whole of Ω, the derivatives $\left(\frac{\partial S}{\partial y_1}\right)(u_1(x), u_2(x))$ and $\left(\frac{\partial S}{\partial y_2}\right)(u_1(x), u_2(x))$ are not defined in any point $x \in \Omega$, and since $Du_1(x) = Du_2(x)$ is not zero in general, it is not clear how formula (1.8) should be understood.

The third reason is that the result of Proposition 1.1 is not trivial, since it is not so straightforward even when u_1 and u_2 belong to $C^1(\Omega)$: indeed in such a case, the set $\{x \in \Omega : u_1(x) < u_2(x)\}$ is open, and it is clear that $Dv(x) = Du_2(x)$ on this set; but the set $\{x \in \Omega : u_1(x) = u_2(x)\}$ can be a very nasty closed set, and to prove that $Du_1(x) = Du_2(x)$ almost everywhere on this set requires some effort (actually the same effort as in the $W^{1,r}(\Omega)$ case). Observe moreover that when u_1 and u_2 belong to $C^1(\Omega)$, the right hand side of (1.2) is defined for every point $x \in \Omega$, while the left hand side is the gradient of a Lipschitz-continuous function, which is therefore defined only almost everywhere, and that (1.2) only holds for almost every point $x \in \Omega$.

The fourth reason is that formula (2.4) below, which generalizes formula (1.2), allows us to prove that the map $u \to T(u)$ is continuous from $W^{1,r}(\Omega; \mathbb{R}^M)$ into $W^{1,r}(\Omega)$ for the strong topologies of these spaces, when T is a globally Lipschitz-continuous, piecewise C^1 function. This result is new, as far as we know.

2. – Statement of the result.

In the present paper, Ω is an arbitrary open set of $\mathbb{R}^N$, with $N \geq 1$ (no smoothness is assumed on $\partial\Omega$, and Ω is not assumed to be bounded), and r is a real number with $1 \leq r < +\infty$.

We consider a vector-valued function $u = (u_m)_{1 \leq m \leq M} \in W^{1,r}(\Omega; \mathbb{R}^M)$, and a function $T : \mathbb{R}^M \to \mathbb{R}$ which is globally Lipschitz-continuous on $\mathbb{R}^M$, and which is piecewise C^1 in the sense that we describe now.

The model example.

The prototype example consists in the case where $M = 2$ (but where N is arbitrary) and where T is a globally Lipschitz-continuous, piecewise affine function defined as follows.

First the space $\mathbb{R}^M = \mathbb{R}^2$ is decomposed into

$$\mathbb{R}^M = \mathbb{R}^2 = \bigcup_{\alpha \in I} P^\alpha, \quad I \text{ finite}, \quad \text{with } P^\alpha \cap P^\beta = \emptyset \text{ if } \alpha \neq \beta,$$

i.e. as a union of a finite number of disjoints sets P^α (the pieces) which are either polygons, possibly unbounded, or edges, which are one-dimensional segments, possibly unbounded, or finally vertices, which are points. The pieces can be either open, or closed, or neither open nor closed subsets of $\mathbb{R}^N$ (see an example below).

The function T is then defined on $\mathbb{R}^M = \mathbb{R}^2$ as a globally Lipschitz-continuous function which coincides on each piece P^α with some affine function T^α defined on the whole of $\mathbb{R}^M = \mathbb{R}^2$, which means that for every superscript $\alpha \in I$, there exists an affine function T^α defined on the whole of $\mathbb{R}^M = \mathbb{R}^2$ (and not only on the corresponding piece P^α) such that

$$T(y) = T^\alpha(y) \qquad \forall y \in P^\alpha.$$

For example in the case considered in Proposition 1.1 one can take

$$\mathbb{R}^2 = P^2 \cup P^1 \cup P^0,$$

with

$$P^2 = \{y \in \mathbb{R}^2 \colon y_1 < y_2\}, \quad P^1 = \{y \in \mathbb{R}^2 \colon y_1 > y_2\}, \quad P^0 = \{y \in \mathbb{R}^2 \colon y_1 = y_2\},$$

$$T^2(y) = y_2, \quad T^1(y) = y_1, \quad T^0(y) = y_1.$$

Note that different choices can be made as far as T^0 is concerned, namely $T^0(y) = ty_1 + (1-t)\, y_2$ with $t \in \mathbb{R}$, but also as far as the decomposition of $\mathbb{R}^2$ is concerned: indeed one can take $\mathbb{R}^2 = P^2 \cup P^3$, with P^3 defined by $P^3 = P^1 \cup P^0$, or make some stranger choices, like $\mathbb{R}^2 = P^2 \cup P^1 \cup P^4 \cup P^5$, with

$$P^4 = \{y \in \mathbb{R}^2 \colon y_1 = y_2 > 0\}, \qquad P^5 = \{y \in \mathbb{R}^2 \colon y_1 = y_2 \leqslant 0\},$$

or even $\mathbb{R}^2 = P^6 \cup P^7$, with P^6 and P^7 defined by $P^6 = P^2 \cup P^4$ and $P^7 = P^1 \cup P^5$ (the latest pieces are neither open nor closed). This example clearly shows that the definition of the pieces P^α and of the functions T^α is not unique when a globally Lipschitz-continuous, piecewise C^1 function T is given.

The general setting.

We now describe the general setting which is a natural generalization of the model example.

The space $\mathbb{R}^M$ is decomposed into a finite union of disjoint Borel sets P^α (the pieces), i.e.,

$$\mathbb{R}^M = \bigcup_{\alpha \in I} P^\alpha, \quad I \text{ finite}, \quad P^\alpha \text{ Borel subset of } \mathbb{R}^M, \quad P^\alpha \cap P^\beta = \emptyset \;\; \forall \alpha \neq \beta. \tag{2.1}$$

Let us emphasize that the pieces P^α are only assumed (*) to be Borel subsets of $\mathbb{R}^M$; however in the applications, the pieces will usually be smooth open or closed subsets of smooth manifolds of dimension $M, M-1, \ldots, 2, 1$, or 0.

The function $T : \mathbb{R}^M \to \mathbb{R}$ is then defined as a Lipschitz-continuous function defined on the whole of $\mathbb{R}^M$ (which is therefore defined in every point, and not only almost everywhere), which coincides on each piece P^α with some given function $T^\alpha : \mathbb{R}^M \to \mathbb{R}$ (which is defined on the whole of $\mathbb{R}^M$ and not only on the piece P^α) which is globally Lipschitz-continuous and C^1, i.e.

$$\begin{cases} T : \mathbb{R}^M \to \mathbb{R} & \text{Lipschitz-continuous}, \\ T^\alpha : \mathbb{R}^M \to \mathbb{R} & \text{Lipschitz-continuous and } C^1(\mathbb{R}^M) \;\; \forall \alpha \in I, \\ T(y) = T^\alpha(y) & \forall y \in P^\alpha \;\; \forall \alpha \in I. \end{cases} \tag{2.2}$$

In other terms

$$T(y) = \sum_{\alpha \in I} \chi_{P^\alpha}(y)\, T^\alpha(y). \tag{2.3}$$

As already said in the case of the model example, for a given function T, the pieces P^α and the functions T^α are not defined in a unique way. This implies in particular that «compatibility relations» hold (see Remark 2.3 below).

THEOREM 2.1. – *Let $T : \mathbb{R}^M \to \mathbb{R}$ be a globally Lipschitz-continuous, piecewise C^1 function in the sense defined by* (2.1) *and* (2.2), *which satisfies $T(0) = 0$. Then for every $u \in W^{1,r}(\Omega; \mathbb{R}^M)$, the function $T(u)$ belongs to $W^{1,r}(\Omega)$ and one has*

$$(DT(u))(x) = \sum_{\alpha \in I} \chi_{P^\alpha}(u(x)) \;(DT^\alpha)(u(x))\;\, Du(x) \quad \text{a.e. } \; x \in \Omega. \tag{2.4}$$

Moreover $T(u)$ belongs to $W_0^{1,r}(\Omega)$ when u belongs to $W_0^{1,r}(\Omega; \mathbb{R}^M)$.

Finally the map $u \to T(u)$ is sequentially continuous from $W^{1,r}(\Omega; \mathbb{R}^M)$ into $W^{1,r}(\Omega)$ for the weak topologies of these spaces, and continuous for the strong topologies of these spaces.

(*) The fact that Theorem 2.1 and Proposition 2.2 below hold true under this very weak assumption on the smoothness of the pieces was pointed out to us by Gianni Dal Maso and Maria Giovanna Mora.

As far as we know, the chain rule formula (2.4) and the continuity in the strong topologies are new results.

Observe that there is a strong analogy between formulas (2.3) and (2.4) (like it was the case between formulas (1.1) and (1.2)).

REMARK 2.1. – In formula (2.4), $DT(u)$ is a (row) vector of size N, DT^α is a (row) vector of size M, and Du is an $M \times N$ matrix, with respective entries

$$(DT(u))_n = \frac{\partial T(u)}{\partial x_n}, \qquad (DT^\alpha)_m = \frac{\partial T^\alpha}{\partial y_m}, \qquad (Du)_{m,n} = \frac{\partial u_m}{\partial x_n},$$

$$1 \leqslant n \leqslant N, \ 1 \leqslant m \leqslant M,$$

and formula (2.4) written in components thus reads as

$$\frac{\partial T(u)}{\partial x_n}(x) = \sum_{\alpha \in I} \chi_{P^\alpha}(u(x)) \sum_{1 \leqslant m \leqslant M} \frac{\partial T^\alpha}{\partial y_m}(u(x)) \frac{\partial u_m}{\partial x_n}(x) \quad \text{a.e.} \quad x \in \Omega,$$

for every n, $1 \leqslant n \leqslant N$. ■

REMARK 2.2. – In Theorem 2.1, we have assumed that

$$T(0) = 0.$$

If we do not make this assumption, all the results of Theorem 2.1 continue to hold true, except the three following ones: when the measure of Ω is not finite, the assertion $T(u) \in W^{1,r}(\Omega)$ should be replaced by $T(u) \in L^r_{\text{loc}}(\Omega)$ with $DT(u) \in L^r(\Omega; \mathbb{R}^N)$; similar changes have to be made as far as the continuity of the map $u \to T(u)$ is concerned; finally, $T(u)$ does not in general belong to $W_0^{1,r}(\Omega)$ when u belongs to $W_0^{1,r}(\Omega; \mathbb{R}^M)$ even if the measure of Ω is finite. ■

PROPOSITION 2.2. – *Consider two functions T^1, T^2: $\mathbb{R}^M \to \mathbb{R}$ which are globally Lipschitz-continuous and C^1. Then for every $u \in W^{1,r}(\Omega; \mathbb{R}^M)$ one has*

$$(2.5) \qquad (DT^i(u))(x) = (DT^i)(u(x))\, Du(x) \quad \textit{a.e.} \quad x \in \Omega,$$

for $i = 1, 2$, and

$$(2.6) \qquad \begin{cases} (DT^1)(u(x))\, Du(x) = (DT^2)(u(x))\, Du(x) \\ \qquad\qquad\qquad \textit{a.e.} \quad x \in \{x \in \Omega : T^1(u(x)) = T^2(u(x))\}. \end{cases}$$

REMARK 2.3. – If the functions T^1 and T^2 coincide on some piece P^α, then $T^1(u(x)) = T^2(u(x))$ for almost every $x \in \Omega$ such that $u(x) \in P^\alpha$, and therefore, by Proposition 2.2 one has

$$\tag{2.7} \begin{cases} (DT^1)(u(x))\, Du(x) = (DT^2)(u(x))\, Du(x) \\ \qquad\qquad\qquad\qquad \text{a.e.} \quad x \in \Omega \ \text{ such that } \ u(x) \in P^\alpha. \end{cases}$$

In particular, if $T^1 = T^2 = T$ on P^α, then, because of (2.5), the common value in (2.7) is nothing but $(DT(u))(x)$ for almost every $x \in \Omega$ such that $u(x)$ belongs to P^α.

This observation has strong consequences since there are many «compatibility relations» between the various functions T^α. For example in the piecewise affine case described in the model example, when two polygons P^α and $P^{\alpha'}$, corresponding to two functions T^α and $T^{\alpha'}$, share a common edge P^β, corresponding to a function T^β, one necessarily has

$$T^\alpha(y) = T^{\alpha'}(y) = T^\beta(y) = T(y) \qquad \forall y \in P^\beta;$$

this implies that

$$DT^\alpha(u(x))\, Du(x) = DT^{\alpha'}(u(x))\, Du(x) = DT^\beta(u(x))\, Du(x)$$

$$\text{a.e.} \quad x \in \Omega \ \text{ such that } \ u(x) \in P^\beta;$$

when the two polygons P^α and $P^{\alpha'}$ and their common edge P^β share a common vertex P^γ, corresponding to a function T^γ, one necessarily has

$$T^\alpha(P^\gamma) = T^{\alpha'}(P^\gamma) = T^\beta(P^\gamma) = T^\gamma(P^\gamma) = T(P^\gamma);$$

this implies that

$$DT^\alpha(u(x))\, Du(x) = DT^{\alpha'}(u(x))\, Du(x) = DT^\beta(u(x))\, Du(x) = DT^\gamma(u(x))\, Du(x)$$

$$\text{a.e.} \ x \in \Omega \ \text{ such that } \ u(x) \in P^\gamma;$$

moreover this common value is 0, as it easily seen by considering the constant function T^δ defined by $T^\delta(y) = T(P^\gamma)$.

Similar «compatibility relations» hold in the general case. ■

3. – Proofs.

First step. In this first step we assume only that the function $T : \mathbb{R}^M \to \mathbb{R}$ satisfies $T(0) = 0$ and is globally Lipschitz-continuous, i.e. satisfies for some constant C_0

$$\tag{3.1} |T(y) - T(y')| \le C_0 |y - y'| \quad \forall y,\ y' \in \mathbb{R}^M.$$

Under these hypotheses, it is well known that for every $u \in W^{1,r}(\Omega;\mathbb{R}^M)$ one has $T(u) \in W^{1,r}(\Omega)$, but let us sketch a proof.

For $v \in C^1_{\mathrm{loc}}(\Omega;\mathbb{R}^M)$ and $S \in C^1(\mathbb{R}^M)$, one has $S(v) \in C^1_{\mathrm{loc}}(\Omega)$ with the classical chain rule

$$(DS(v))(x) = (DS)(v(x))\, Dv(x) \qquad \forall x \in \Omega\,.$$

We approximate u by a sequence of functions $v_n \in C^1_{\mathrm{loc}}(\Omega;\mathbb{R}^M)$ which converge to u in the strong topology of $W^{1,r}_{\mathrm{loc}}(\Omega;\mathbb{R}^M)$, and T by a sequence of functions $S_n \in C^1(\mathbb{R}^M)$ which converge to T uniformly on $\mathbb{R}^M$ and satisfy

$$S_n(0)=0, \qquad |DS_n(y)| \leqslant C_0 \qquad \forall y \in \mathbb{R}^M;$$

convolutions provide such approximate sequences. Since

$$\begin{aligned}|S_n(v_n)(x) - T(u)(x)| &\leqslant |S_n(v_n)(x) - S_n(u)(x)| + |S_n(u)(x) - T(u)(x)| \leqslant \\ &\leqslant C_0\,|v_n(x) - u(x)| + |S_n(u)(x) - T(u)(x)|\,,\end{aligned}$$

and since the last term converges uniformly to 0 on Ω, we obtain that $S_n(v_n)$ tends to $T(u)$ strongly in $L^r_{\mathrm{loc}}(\Omega)$.

Therefore $DS_n(v_n)$ converges to $DT(u)$ in the sense of distributions. This convergence, the estimate

$$|DS_n(v_n)(x)| \leqslant |(DS_n)(v_n(x))|\,|Dv_n(x)| \leqslant C_0\,|Dv_n(x)| \quad \forall x \in \Omega\,,$$

and the strong convergence of Dv_n to Du in $L^r_{\mathrm{loc}}(\Omega;\mathbb{R}^N)$ allows us to pass to the limit in

$$|\langle DS_n(v_n)),\varphi\rangle| \leqslant C_0 \int_\Omega |Dv_n(x)|\,|\varphi(x)|\,dx \quad \forall \varphi \in C^\infty_c(\Omega;\mathbb{R}^N)\,,$$

obtaining

$$|\langle DT(u),\varphi\rangle| \leqslant C_0 \int_\Omega |Du(x)|\,|\varphi(x)|\,dx \quad \forall \varphi \in C^\infty_c(\Omega;\mathbb{R}^N).$$

This implies that $DT(u) \in L^r(\Omega;\mathbb{R}^N)$ and that

$$|DT(u)(x)| \leqslant C_0\,|Du(x)| \quad \text{a.e. } x \in \Omega\,. \tag{3.2}$$

On the other hand, since $T(0)=0$, we have

$$|T(u(x))| \leqslant C_0\,|u(x)| \quad \text{a.e. } x \in \Omega\,. \tag{3.3}$$

Therefore $T(u)$ belongs to $L^r(\Omega)$ and (3.2) proves that $T(u)$ belongs to $W^{1,r}(\Omega)$.

Moreover if $u \in W^{1,r}_0(\Omega;\mathbb{R}^M)$, we can approximate u in $W^{1,r}_0(\Omega;\mathbb{R}^M)$ by a sequence of functions $v_n \in C^\infty_c(\Omega;\mathbb{R}^M)$ and, since $S_n(v_n)$ then belongs to $C^1_c(\Omega)$, the previous proof implies that $T(u)$ belongs to $W^{1,r}_0(\Omega)$.

Second step. Under the assumptions of the first step, we deduce from (3.1) that for every u and v in $L^r_{\text{loc}}(\Omega;\mathbb{R}^M)$, one has

$$|T(u)(x) - T(v)(x)| \leqslant C_0\,|u(x) - v(x)| \quad \text{a.e.} \quad x \in \Omega\,.$$

Therefore the map $u \to T(u)$ is continuous (and even Lipschitz-continuous) from the strong topology of $L^r_{\text{loc}}(\Omega;\mathbb{R}^M)$ into the strong topology of $L^r_{\text{loc}}(\Omega)$.

On the other hand it is also well known that, under the hypotheses made on T, the map $u \to T(u)$ is sequentially continuous from $W^{1,r}(\Omega;\mathbb{R}^M)$ into $W^{1,r}(\Omega)$ for the weak topologies of these spaces, but let us sketch a proof.

Consider a sequence u_n which converges to u in the weak topology of $W^{1,r}(\Omega;\mathbb{R}^M)$. This is equivalent to assume that u_n tends to u weakly in $L^r(\Omega;\mathbb{R}^M)$ and that Du_n tends to Du weakly in $L^r(\Omega;\mathbb{R}^{M\times N})$, and in view of Rellich's compactness theorem, that u_n tends to u strongly in $L^r_{\text{loc}}(\Omega;\mathbb{R}^M)$. By the strong L^r continuity we have just proved, we deduce that $T(u_n)$ tends to $T(u)$ strongly in $L^r_{\text{loc}}(\Omega)$. On the other hand, from (3.3) and (3.2) we deduce that

$$(3.4)\qquad |T(u_n)(x)| \leqslant C_0\,|u_n(x)|\,, \qquad |DT(u_n)(x)| \leqslant C_0\,|Du_n(x)| \qquad \text{a.e.} \quad x \in \Omega\,.$$

Therefore $T(u_n)$ and $DT(u_n)$ are bounded in $L^r(\Omega;\mathbb{R}^M)$ and in $L^r(\Omega;\mathbb{R}^{M\times N})$, and the strong convergence of $T(u_n)$ to $T(u)$ in $L^r_{\text{loc}}(\Omega)$ implies the desired weak continuity when $1 < r < +\infty$. When $r = 1$, one further deduces from (3.4) that $|T(u_n)|$ and $|DT(u_n)|$ are equi-integrable in $L^1(\Omega)$, since this is the case for $|u_n|$ and $|Du_n|$; combined with the strong convergence of $T(u_n)$ to $T(u)$ in $L^1_{\text{loc}}(\Omega)$, this implies the desired weak continuity.

Third step. If further to the assumptions made in the first step, we assume that T belongs to $C^1(\mathbb{R}^M)$, we can take $S_n = T$ in the proof of the first step, and passing to the limit in the classical chain rule for functions of C^1, namely

$$(DT(v_n))(x) = (DT)(v_n(x))\,Dv_n(x) \quad \forall x \in \Omega\,,$$

we deduce the chain rule

$$(3.5)\qquad (DT(u))(x) = (DT)(u(x))\,Du(x) \quad \text{a.e.} \quad x \in \Omega\,,$$

for every $u \in W^{1,r}(\Omega;\mathbb{R}^M)$ and for every function $T : \mathbb{R}^M \to \mathbb{R}$ which is globally Lipschitz-continuous and C^1. This proves (2.5).

Fourth step. From now on, we make all the assumptions of Theorem 2.1.

For every $\alpha \in I$ we deduce from the first and third steps that $T^\alpha(u)$ belongs to $W^{1,r}(\Omega)$ and that

$$(3.6)\qquad (DT^\alpha(u))(x) = (DT^\alpha)(u(x))\,Du(x) \quad \text{a.e.} \quad x \in \Omega\,.$$

Let U^α be the measurable set defined by

$$U^\alpha = \{x \in \Omega : u(x) \in P^\alpha\}\,; \tag{3.7}$$

here we used the fact that P^α is a Borel set. From

$$T(u)(x) = T^\alpha(u)(x) \quad \text{a.e.} \quad x \in U^\alpha, \tag{3.8}$$

we deduce that

$$(DT(u))(x) = (DT^\alpha(u))(x) \quad \text{a.e.} \quad x \in U^\alpha: \tag{3.9}$$

indeed we have proved that the function $w = T(u) - T^\alpha(u)$ belongs to $W^{1,r}(\Omega)$; from property (1.4) one deduces that $Dw = 0$ almost everywhere on the set $\{x \in \Omega : w(x) = 0\}$, which implies (3.9) since by (3.8) one has $U^\alpha \subset \{x \in \Omega : w(x) = 0\}$.

From (3.9) and (3.6) we deduce that

$$(DT(u))(x) = (DT^\alpha)(u(x))\, Du(x) \quad \text{a.e.} \quad x \in U^\alpha, \tag{3.10}$$

which immediately implies (2.4) since

$$\chi_{U^\alpha}(x) = \chi_{P^\alpha}(u(x)) \quad \text{a.e.} \quad x \in \Omega\,.$$

Observe that formula (2.5) has already been proved in the third step and that formula (2.6) follows from (1.4) applied to $w = T^1(u) - T^2(u)$ and from (2.5). Therefore proposition 2.2 is proved.

Fifth step. It remains to prove that under the assumptions of Theorem 2.1, the map $u \to T(u)$ is continuous from $W^{1,r}(\Omega; \mathbb{R}^M)$ into $W^{1,r}(\Omega)$ for the strong topologies of these spaces. This is the goal of the present and of the next steps.

Since it is sufficient to prove the sequential continuity in order to prove the continuity in strong topologies, we consider a sequence of functions u_n of $W^{1,r}(\Omega; \mathbb{R}^M)$ which converges to u strongly in $W^{1,r}(\Omega; \mathbb{R}^M)$. We will prove that from any given subsequence of the original sequence $\{n\}$ we can extract a new subsequence along which $T(u_n)$ converges to $T(u)$ strongly in $W^{1,r}(\Omega)$. This implies that $T(u_n)$ converges to $T(u)$ strongly in $W^{1,r}(\Omega)$ along the whole sequence $\{n\}$, hence the desired continuity

Extracting from the given subsequence of the original sequence $\{n\}$ a new subsequence, that we still denote by $\{n\}$, we can assume that for this subsequence

$$u_n(x) \to u(x) \quad \text{and} \quad Du_n(x) \to Du(x) \quad \text{a.e.} \quad x \in \Omega\,, \quad n \in \{n\}\,. \tag{3.11}$$

From the first continuity result proved in the second step, we know that $T(u_n)$ converges to $T(u)$ in the strong topology of $L^r(\Omega)$. We will prove in the seventh step below that if (3.11) is satisfied, then one has

$$(DT(u_n))(x) \to (DT(u))(x) \quad \text{a.e.} \quad x \in \Omega, \quad n \in \{n\}. \tag{3.12}$$

Since by (3.2) we have

$$|DT(u_n)(x)| \leq C_0 \, |Du_n(x)| \quad \text{a.e.} \quad x \in \Omega,$$

this result and Vitali's theorem will imply the strong convergence of $DT(u_n)$ to $DT(u)$ in $L^r(\Omega; \mathbb{R}^{M \times N})$ and will complete the proof of Theorem 2.1.

From now on, for every function of $L^1(\Omega)$, we consider its representative defined at its Lebesgue's points (we could as well have done that starting from the beginning of the proof of Theorem 2.1). The set of the Lebesgue's points differs from Ω by a set of measure zero. Since only a countable number of functions are involved in the proof below, we are at liberty to consider a unique set $G \subset \Omega$ (the «good set») with $\operatorname{meas}(\Omega \backslash G) = 0$ such that for every $n \in \{n\}$, the functions $u_n(x)$, $Du_n(x)$, $(DT(u_n))(x)$, as well as $u(x)$, $Du(x)$ and $(DT(u))(x)$ are defined for every $x \in G$ (and not only for almost every x), such that (see (3.11))

$$u_n(x) \to u(x) \quad \text{and} \quad Du_n(x) \to Du(x) \ \ \forall x \in G, \quad n \in \{n\}, \tag{3.13}$$

and such that, according to formula (2.4),

$$(DT(u_n))(x) = \sum_{\gamma \in I} \chi_{P^\gamma}(u_n(x)) \, (DT^\gamma)(u_n(x)) \, Du_n(x) \quad \forall x \in G \ \ \forall n \in \{n\}, \tag{3.14}$$

$$(DT(u))(x) = \sum_{\alpha \in I} \chi_{P^\alpha}(u(x)) \, (DT^\alpha)(u(x)) \, Du(x) \quad \forall x \in G. \tag{3.15}$$

Moreover, since we know from (2.6) that for every given α and β in I we have

$$\begin{cases} (DT^\alpha)(u(x)) \, Du(x) = (DT^\beta)(u(x)) \, Du(x) \\ \qquad\qquad\qquad \text{a.e.} \quad x \in \{x \in \Omega : \ T^\alpha(u(x)) = T^\beta(u(x))\}, \end{cases}$$

we can assume, by removing from G a finite number (remember that I is finite) of sets of measure zero, that for a new set, that we still denote by G, still with $\operatorname{meas}(\Omega \backslash G) = 0$, we have

$$\begin{cases} (DT^\alpha)(u(x)) \, Du(x) = (DT^\beta)(u(x)) \, Du(x) \\ \qquad\qquad\qquad \forall x \in \{x \in G : T^\alpha(u(x)) = T^\beta(u(x))\}. \end{cases} \tag{3.16}$$

Sixth step. At this point, the sequence $\{n\}$ and the set G are fixed, with the properties (3.13), (3.14), (3.15) and (3.16). We will prove that

$$(DT(u_n))(x) \to (DT(u))(x) \quad \forall x \in G, \quad n \in \{n\}, \tag{3.17}$$

which is nothing but (3.12).

From (3.14), (3.13) and the fact that every function T^γ belongs to $C^1(\mathbb{R}^M)$ we deduce that

$$\text{(3.18)}\qquad \begin{cases} (DT(u_n))(x) = \sum_{\gamma\in I}\chi_{P^\gamma}(u_n(x))\,(DT^\gamma)(u(x))\,Du(x) + \omega_n(x) \\ \qquad\qquad\qquad\qquad \text{with } \omega_n(x)\to 0 \quad \forall x\in G,\quad n\in\{n\}; \end{cases}$$

we define

$$\text{(3.19)}\qquad Y_n(x) = \sum_{\gamma\in I}\chi_{P^\gamma}(u_n(x))\,(DT^\gamma)(u(x))\,Du(x)\ \ \forall\, x\in G\ \ \forall\, n\in\{n\}.$$

For every $x\in G$ and every $\beta\in I$, we define the set $\{m(x,\beta)\}\subset\{n\}$ as the set of those indices $n\in\{n\}$ which are such that $u_n(x)\in P^\beta$. Therefore for every fixed $x\in G$ the sequence $\{n\}$ is the union of the disjoint sets of indices $\{m(x,\beta)\}$ when β runs in I. We define $J(x)\subset I$ as the set of those β such that the set of indices $\{m(x,\beta)\}$ is infinite. Since $\{n\}$ is infinite, the set $J(x)$ is not empty. Moreover there exists some $n_0(x)$ such that whenever $n>n_0(x)$, every index $n\in\{n\}$ belongs to a (infinite) subsequence $\{m(x,\beta)\}$ for some $\beta\in J(x)$.
Since for every $x\in G$ and for every $\beta\in I$ we have

$$\chi_{P^\beta}(u_n(x)) = 1,\quad \chi_{P^\gamma}(u_n(x)) = 0\ \ \forall\,\gamma\neq\beta,\quad \forall\, n\in\{m(x,\beta)\},$$

formula (3.19) implies that

$$\text{(3.20)}\qquad Y_n(x) = (DT^\beta)(u(x))\,Du(x)\ \ \forall\, x\in G\ \ \forall\,\beta\ \ \forall\, n\in\{m(x,\beta)\}.$$

Similarly, since (2.3) implies that

$$T(u_n(x)) = \sum_{\gamma\in I}\chi_{P^\gamma}(u_n(x))\,T^\gamma(u_n(x))\ \ \forall\, x\in G\ \ \forall\, n\in\{n\},$$

we have

$$\text{(3.21)}\qquad T(u_n(x)) = T^\beta(u_n(x))\ \ \forall\, x\in G\ \ \forall\,\beta\ \ \forall\, n\in\{m(x,\beta)\}.$$

When $\beta\in J(x)$, by passing to the limit in (3.21) along the (infinite) subsequence $\{m(x,\beta)\}$, we deduce from (3.13) and from the continuity of T and T^β that

$$\text{(3.22)}\qquad T(u(x)) = T^\beta(u(x))\quad \forall\, x\in G\quad \forall\,\beta\in J(x).$$

For every $\alpha \in I$ we define, as in (3.7), the set

$$U^{\alpha} = \{x \in G : u(x) \in P^{\alpha}\},$$

and for every $\alpha, \beta \in I$, the set

$$U^{\alpha,\beta} = \{x \in G : u(x) \in P^{\alpha},\ \beta \in J(x)\}.$$

Since $T(u(x)) = T^{\alpha}(u(x))$ for every $x \in U^{\alpha}$, (3.22) implies that

$$T^{\alpha}(u(x)) = T^{\beta}(u(x)) \quad \forall\, x \in U^{\alpha,\beta}.$$

This implies, in view of (3.16), that for every $\alpha, \beta \in I$

$$(DT^{\alpha})(u(x))\, Du(x) = (DT^{\beta})(u(x))\, Du(x) \quad \forall\, x \in U^{\alpha,\beta}.$$

Turning back to (3.20), we have

$$\text{(3.23)} \qquad Y_n(x) = DT^{\alpha}(u(x))\, Du(x) \quad \forall\, x \in U^{\alpha,\beta} \quad \forall\, n \in \{m(x,\beta)\}\,.$$

Passing to the limit in (3.18) along the (infinite) subsequence $\{m(x,\beta)\}$, and using (3.23) and the definition (3.19) of Y_n, we have proved that for every $\alpha, \beta \in I$

$$\text{(3.24)} \quad (DT(u_n))(x) \to DT^{\alpha}(u(x))\, Du(x) \quad \forall\, x \in U^{\alpha,\beta}, \quad n \in \{m(x,\beta)\}\,.$$

But since $U^{\alpha,\beta} \subset U^{\alpha}$, the right hand side of (3.24) coincides with $(DT(u))(x)$ because of (3.15) and of the definition of U^{α}, and (3.24) is nothing but

$$(DT(u_n))(x) \to (DT(u))(x) \quad \forall\, x \in U^{\alpha,\beta}, \quad n \in \{m(x,\beta)\}.$$

Since this holds true for every $\alpha \in I$, we have proved that

$$\text{(3.25)} \qquad (DT(u_n))(x) \to (DT(u))(x) \quad \forall\, x \in G \ \ \forall \beta \in J(x), \quad n \in \{m(x,\beta)\}.$$

Fix now $x \in G$, and recall the definition of $n_0(x)$ given above. Since every $n > n_0(x)$ belongs to a subsequence $\{m(x,\beta)\}$ for some $\beta \in J(x)$, (3.25) is nothing but

$$(DT(u_n))(x) \to (DT(u))(x) \quad \forall\, x \in G, \quad n \in \{n\},$$

i.e. (3.17). This completes the proof of Theorem 2.1. ∎

Acknowledgments. The authors would like to thank Gianni Dal Maso and Maria Giovanna Mora for their contribution concerning the non smoothness of the pieces.

This work was done during the visits of the first author to the Dipartimento di Matematica «R. Caccioppoli» dell'Università di Napoli «Federico II». This author wants to thank all the members of the Dipartimento for their warm hospitality.

REFERENCES

[AD] L. AMBROSIO - G. DAL MASO, *A general chain rule for distributional derivatives*, Proc. Amer. Math. Soc., **108** (1990), 691-702.

[BM] L. BOCCARDO - F. MURAT, *Remarques sur l'homogénéisation de certains problèmes quasi-linéaires*, Portugaliæ Math., **41** (1982), 535-562.

[B] F. BOUCHUT, *Renormalized solutions to the Vlasov equation with coefficients of bounded variation*, Arch. Rat. Mech. Anal., **157** (2001), 75-90.

[KS] D. KINDERLHERER - G. STAMPACCHIA, *An introduction to variational inequalities and their applications*, Academic Press, New York (1980).

[L] R. LANDES, *On the existence of weak solutions of perturbated systems with critical growth*, J. reine angew. Math., **393** (1989), 21-38.

[MM1] M. MARCUS - V. J. MIZEL, *Absolute continuity on tracks and mappings of Sobolev spaces*, Arch. Rat. Mech. Anal., **45** (1972), 294-320.

[MM2] M. MARCUS - V. J. MIZEL, *Nemitsky operators on Sobolev spaces*, Arch. Rat. Mech. Anal., **51** (1973), 347-370.

[S] G. STAMPACCHIA, *Equations elliptiques du second ordre à coefficients discontinus*, Séminaire de Mathématiques Supérieures, **16**, Les Presses de l'Université de Montréal, Montréal (1966).

François Murat: Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie (Paris 6), Boîte courrier 187
75252 Paris cedex 05, France

Cristina Trombetti: Dipartimento di Matematica e Applicazioni «R. Caccioppoli»
Università di Napoli «Federico II», Via Cintia, 80126 Napoli, Italy

Pervenuta in Redazione
il 26 gennaio 2002 e in forma rivista l'8 luglio 2002