

On distribution functions of the derivatives of weakly differentiable mappings

by

Victor J. MIZEL

Department of Mathematics,
Carnegie Mellon University,
Pittsburgh, Pa. 15213-3890, U.S.A.

ABSTRACT. — In this paper I describe and resolve affirmatively the following question: For a function on the unit interval possessing an absolutely continuous first derivative, does the joint distribution of the function and its first derivative fully determine the distribution of the second derivative? I also describe some consequences and extensions of the result.

Key words : Joint distribution, Hausdorff 1-measure.

RÉSUMÉ. — Dans ce travail nous abordons et résolvons affirmativement la question suivante : Soit une fonction réelle définie sur l'intervalle $[0, 1]$ et dotée d'une dérivée première absolument continue. Est-ce que la distribution conjointe de la fonction et de sa dérivée première détermine complètement la distribution de la dérivée seconde? Nous décrivons quelques conséquences et généralisations de ce résultat.

A recent investigation into the behavior under homogenization of second order materials with negative capillarity [CMM] led to the formulation of

Classification A.M.S. : 28 A 75, 26 A 46.

Annales de l'Institut Henri Poincaré - Analyse non linéaire - 0294-1449
Vol. 11/94/02/\$4.00/

© 1994 L'Association Publications de l'Institut Henri Poincaré. Published by Elsevier B.V. All rights reserved

certain questions concerning distribution functions of absolutely continuous mappings on a real interval, none of which appears to have been examined previously. The present article describes and resolves one of these questions: For a function on the unit interval possessing an absolutely continuous first derivative, does the joint distribution of the function and its first derivative fully determine the distribution of the second derivative?

Given any measurable mapping $w: \Omega \rightarrow \mathbb{R}^m$, with Ω a bounded measurable subset of \mathbb{R}^N , the (mass) distribution $\mu = \mu^w$ denotes that Borel measure on \mathbb{R}^m defined by

$$\mu(B) = L_N \{ x \in \Omega : w(x) \in B \}, \quad B \in \mathbb{B}(\mathbb{R}^m), \quad (0)$$

where L_N denotes N -dimensional Lebesgue measure and $\mathbb{B}(\mathbb{R}^m)$ denotes the Borel σ -algebra on \mathbb{R}^m [$N=1$ for most of this paper]. For brevity, we will also write μ^w for w a measurable L_N equivalence class.

Given a function $u \in W^{2,1}(0,1)$, where $W^{2,1}(0,1)$ denotes the Sobolev space of real functions on $(0,1)$ possessing two summable generalized derivatives, there are several associated (mass) distributions to be considered: one for each of the functions u, u', u'' as well as for the mappings $f=(u, u'), g=(u, u', u'')$, etc. We are here interested in the linkage between certain of these distributions, in particular between the distribution $\mu := \mu^{u''}$ and the distribution $\pi := \mu^f$; π is hereafter called a *joint* distribution for emphasis. Our main result shows that μ is characterized by π ; in particular, if u, v are such that $\mu^{(u, u')} = \mu^{(v, v')}$ then it necessarily follows that $\mu^{u''} = \mu^{v''}$.

Our arguments will utilize two results stated below (cf. [S] Ch. IX; [F] Theorem 2.10.10). In stating the first result we utilize terminology stemming from the following definition of Bouligand ([S], p. 263; [F], p. 233; [MM1], p. 304), see also [AE].

DEFINITION. — An element $v \in \mathbb{R}^m$ is a *bilateral unit tangent* to the closed set S of \mathbb{R}^m at the point $y \in S$ provided that there exist sequences $\{y_i\}, \{y'_i\}$ in S converging to y and satisfying

$$(y_i - y)/|y_i - y| \rightarrow v, \quad (y'_i - y)/|y'_i - y| \rightarrow -v$$

The collection of all such v is denoted by $U_S(y)$, and the cone consisting of all lines $\mathbb{R}v$ with $v \in U_S(y)$ is denoted by $T_S(y)$ and called the contingent cone to S at y .

The result of interest for us is the fact that since $(u, u') = f: [0, 1] \rightarrow \mathbb{R}^2$ is an absolutely continuous curve, the following conclusions apply [MM1], §3 (in what follows H_1 denotes one-dimensional Hausdorff measure on \mathbb{R}^m):

LEMMA. — If $S \subset \mathbb{R}^m$ is the track of an absolutely continuous curve f then for H_1 -a.e. $y \in S$

- (a) $U_S(y)$ consists of a single pair of opposing unit vectors,
- (b) $f^{-1}(y)$ is a finite set,

- (c) $f'(t) \neq 0$ for all $t \in f^{-1}(y)$,
- (d) $U_S(y) = \{ \pm f'(t) / |f'(t)| : t \in f^{-1}(y) \}$.

Furthermore, f has the following property

$$(N)L_1(A) = 0 \Rightarrow H_1(f(A)) = 0, \quad A \subset \text{Dom}(f).$$

The second result is a measure theoretic fact.

PROPOSITION. — Suppose X and Y are separable metric spaces, ν is a metric outer measure on Y , f maps X into Y , and $f(A)$ is ν -measurable whenever A is a Borel subset of X . If

$$\zeta(S) = \nu[f(S)] \quad \text{for } S \subset X$$

and ψ is the metric outer measure over X resulting from ζ by Caratheodory's construction on the family of all Borel subsets of X , then

$$\psi(A) = \int N(f|A, y) d\nu(y) \quad \text{for every Borel set } A \subset X,$$

where for any mapping $g: Z \rightarrow Y$

$$N(g, y) = \# \{ x \in Z \mid g(x) = y \} \in \{ 0, 1, \dots, n, \dots, \infty \}.$$

It will be seen below that our proof that μ is determined by π hinges on utilizing the preceding results to express both measures in terms of integration over the set $S = f([0, 1])$. We proceed to state and present the proof of Theorem A.

THEOREM A. — For each u in $W^{2,1}(0, 1)$, the joint distribution $\pi = \mu^{(u, u')} =: \mu^f$ determines the distribution $\mu = \mu^{u''}$. More precisely, for every A in $\mathbb{B}(\mathbb{R})$ one has the formula [with $y = (y_0, y_1) \in \mathbb{R}^2$]:

$$\mu(A) = \int_{E_A} [N(f, y) / y_1] dy_0 + L_1(Z) \delta_{\{0\}}(A), \tag{1}$$

where $Z = \{ x \in [0, 1] : u'(x) = 0 \}$, $\delta_{\{0\}}$ is the unit mass at $\{0\}$ and $E_A = \{ y \in f([0, 1] \setminus Z) : y_1 dy_1 / dy_0 \in A \}$, with dy_1 / dy_0 denoting the (H_1 -a.e. unique) slope of the lines in the cone $T_{f[0, 1]}(y)$.

Likewise, π satisfies the relation

$$\pi(B) = \int_{F_B} [N(f, y) / y_1] dy_0 + \pi^s(B \cap (\mathbb{R} \times \{0\})), \quad \forall B \in \mathbb{B}(\mathbb{R}^2), \tag{2}$$

where $F_B = f([0, 1]) \cap B$, and π^s is supported on an L_1 null set of $\mathbb{R} \times \{0\}$ with $\pi^s(\mathbb{R} \times \{0\}) = L_1(Z)$. Thus π determines the integrand in (1) through (2), and consequently determines μ .

Remarks. — The line integrals in (1) and (2) are to be interpreted as integrals with respect to arclength, i.e. with respect to H_1 (dy_0 is the horizontal projection of dH_1).

Since the symbol dy_1/dy_0 appearing in the definition of E_A refers to the slope of the (bilateral) tangent line to the set $\mathbf{f}([0, 1])$ at the point $y \in \mathbf{f}([0, 1])$ one has, formally, $y_1 dy_1/dy_0 = u''$, provided that y is a point where the contingent cone $T_S(y)$ to $S = \mathbf{f}([0, 1])$ consists of a single line.

Proof. — Note that whenever Y, Z are disjoint measurable subsets of $[0, 1]$ it follows from (0) that the distribution function $\mu_3 = \mu^{\mathbf{f}|Y \cup Z}$ is simply the sum of $\mu_1 = \mu^{\mathbf{f}|Y}$ and $\mu_2 = \mu^{\mathbf{f}|Z}$. In particular this holds for $Y = [0, 1] \setminus Z$, with Z as in the statement of the theorem. Now set for each $\varepsilon > 0$, $Y_\varepsilon = \{x \in [0, 1] : |u'(x)| > \varepsilon\}$, $\pi_\varepsilon = \mu^{\mathbf{f}|Y_\varepsilon}$, and $\mu_\varepsilon = \mu^{u''|Y_\varepsilon}$. Then it follows from (0) and the Vitali-Hahn-Saks theorem that

$$\lim_{\varepsilon \rightarrow 0} \pi_\varepsilon = \mu^{\mathbf{f}|Y} =: \pi_0,$$

in the sense that the total variation of the difference measure approaches zero, $\text{var}(\pi_0 - \pi_\varepsilon) \rightarrow 0$, and similarly,

$$\lim_{\varepsilon \rightarrow 0} \mu_\varepsilon = \mu^{u''|Y} =: \mu_0.$$

Now the behavior of μ restricted to $\mathbb{B}(\mathbb{R} \setminus [-\varepsilon, \varepsilon])$ for all positive ε determines μ since $\mu(\mathbb{R}) = 1$. Thus it will suffice for our purposes to prove that for each $\varepsilon > 0$ μ_ε is given by the line integral formula

$$\mu_\varepsilon(A) = \int_{E_{\varepsilon, A}} [N(\mathbf{f}, y)/y_1] dy_0, \quad \text{all } A \in \mathbb{B}(\mathbb{R}), \tag{3}$$

with $E_{\varepsilon, A} = \{y \in \mathbf{f}(Y_\varepsilon) : y_1 dy_1/dy_0 \in A\}$, while π_ε determines the integrand in (3) via

$$\pi_\varepsilon(B) = \int_{F_{\varepsilon, B}} [N(\mathbf{f}, y)/y_1] dy_0, \quad \text{all } B \in \mathbb{B}(\mathbb{R}^2), \tag{4}$$

with $F_{\varepsilon, B} = \mathbf{f}(Y_\varepsilon) \cap B$.

In (3) and (4) we have utilized the fact that, by the definition of Y_ε ,

$$\mathbf{f}(Y_\varepsilon) = \mathbf{f}([0, 1]) \cap \{y \in \mathbb{R}^2 : |y_1| > \varepsilon\},$$

so that

$$N(\mathbf{f}|Y_\varepsilon, y) = N(\mathbf{f}, y), \quad \text{for each } y \in \mathbf{f}(Y_\varepsilon). \tag{5}$$

We have also denoted by $y_1 dy_1/dy_0$ the product of y_1 with the slope of the tangent line to the set $\mathbf{f}([0, 1])$ at the point $y = (y_0, y_1)$ [by the lemma there is a unique tangent line at H_1 -a.e. $y \in \mathbf{f}([0, 1])$]. The existence of the integrals in (3) and (4) as integrals with respect to H_1 [taking $dy_0 = \cos \vartheta(y) dH_1$ where $\vartheta(y)$ denotes the angle with the horizontal made by the tangent line to $\mathbf{f}([0, 1])$ at y] follows by the Proposition with $X = [0, 1]$ and $v = H_1$ on $Y = \mathbb{R}^2$. Namely, by the Lemma the absolute continuity of \mathbf{f} ensures that this mapping has the (N) property of Lusin [S]

from which it easily follows that $\mathbf{f}(E)$ is H_1 -measurable for each L_1 -measurable subset E of $[0, 1]$. Thus on defining the set function Π by

$$\Pi(A) = H_1(\mathbf{f}(A)) \text{ on subsets } A \subset [0, 1],$$

and letting β denote the metric (outer) measure generated by $\Pi|_{\mathbb{B}([0, 1])}$, it follows that

$$\beta(A) = \int_{\mathbf{f}(A)} N(\mathbf{f}|A, y) dH_1(y) \text{ for all } A \in \mathbb{B}([0, 1]). \tag{6}$$

It follows from (6) that the line integral in (4) is also well defined, for each $\varepsilon > 0$. The validity of (4) is now an easy consequence of the formal relation $dx = du/u'$, since the equality is evident for $B = \mathbf{f}(J)$ on each component J of the open sets

$$C_+(\varepsilon) = \{x | u'(x) > \varepsilon\}, \quad C_-(\varepsilon) = \{x | u'(x) < -\varepsilon\}.$$

The validity of formula (3) utilizes the fact that the function u'' can be factored as $h \circ \mathbf{f}$ where $h: \mathbb{R} \rightarrow \mathbb{R}$ is the measurable mapping defined H_1 -a.e. by $h(y) = y_1 dy_1/dy_0$. More precisely, if C, D, E denote the subsets of $[0, 1]$ consisting, respectively, of points x where $u''(x)$ is not defined as a real number, of points x in Z at which Z doesn't have unit density, and of points x such that $\mathbf{f}([0, 1])$ fails to have a unique (nonvertical) tangent line, then

$$L_1(C) = L_1(D) = L_1(E) = 0$$

(the validity of the last relation follows from the proof of the Lemma, cf. e.g. [S], Ch. IX, Lemma 3.1). Hence by setting $X = [0, 1] \setminus (C \cup D \cup E)$ one finds that $\mu = \mu^{u''|X}$. Moreover $u'' : X \rightarrow \mathbb{R}$ can be factored as $u'' = h \circ \mathbf{f}$, where $h: \mathbb{R} \rightarrow \mathbb{R}$ is given by $h(y) = y_1 dy_1/dy_0$. It follows from this that (3) is valid for each set $A \in \mathbb{B}(\mathbb{R})$ such that $A \subset \mathbb{R} \setminus \{0\}$. The validity of (3) for all A follows immediately.

Remarks. — We observe that the arguments used in the proof of Theorem A demonstrate that π also characterizes μ^g , where $g = (u, u', u'')$. So π characterizes $\mu^{(u', u'')}$, and hence if $u \in W^{3,1}(0, 1)$ it also characterizes $\mu^{u'''}$, and so on.

Although Theorem A leads fairly straightforwardly (using [MM2]) to the conclusion that for $\Omega \in \mathbb{R}^N$ and $u \in W^{2,p}(\Omega)$, $p > N$, the joint distribution $\pi = \mu^f$ where $f = (u, \text{grad } u)$, characterizes each $\mu_i = \mu^{u, x^i x^i}$, $i = 1, \dots, N$, as well as $\mu^* = \mu^{\Delta u}$, one has to utilize the additional fact that π characterizes the distributions $\mu_a = \mu^{(\partial_a)^2 u}$ for all second directional derivatives $(\partial_a)^2$ to obtain a multidimensional analogue of Theorem A involving nonsymmetric differential operators. Of course the use of rectifiable currents would be more suitable for this multidimensional context ([F], ch. 4 or [M], ch. 4). In any event Theorem A supplies a key result for the relaxation

analysis of the second order model in [CMM], as will be shown in a subsequent publication.

ACKNOWLEDGEMENTS

The work during this project was partially supported by the National Science Foundation under Grants DMS 90 02562 and DMS 92 01221.

The author wishes to express thanks to Bernard Coleman, Dick James and Moshe Marcus for very enlightening conversations on the subject of this manuscript, as well as to the referee for some useful suggestions.

REFERENCES

- [AE] J.-P. AUBIN and I. EKELAND, *Applied Nonlinear Analysis*, Wiley-Interscience, New York, 1984.
- [CMM] B. D. COLEMAN, M. MARCUS and V. J. MIZEL, On the Thermodynamics of Periodic Phases, *Arch. Ration. Mech. and Analysis*, Vol. **117**, 1992, pp. 321-347.
- [F] H. FEDERER, *Geometric Measure Theory*, Springer Verlag, New York, 1969.
- [M] F. MORGAN, *Geometric Measure Theory a beginner's guide*, Academic Press, Boston, 1988.
- [MM1] M. MARCUS and V. J. MIZEL, Absolute Continuity on Tracks and Mappings of Sobolev Spaces, *Arch. Ration. Mech. and Analysis*, Vol. **45**, 1972, pp. 294-320.
- [MM2] M. MARCUS and V. J. MIZEL, Transformation by Functions in Sobolev Spaces and Lower Semicontinuity for Parametric Variational Problems, *Bull. Amer. Math. Sci.*, Vol. **79**, 1973, pp. 790-795.
- [S] S. SAKS, *Theory of the Integral*, Dover, reprint, 1964.

(Manuscript received March 17, 1993;

revised July 5, 1993.)