APPLICATIONES MATHEMATICAE
26,3 (1999), pp. 299–314

N. BELILI and H. HEINICH (Rouen)

MASS TRANSPORT PROBLEM AND DERIVATION

Abstract. A characterization of the transport property is given. New properties for strongly nonatomic probabilities are established. We study the relationship between the nondifferentiability of a real function f and the fact that the probability measure $\lambda_{f^*} := \lambda \circ (f^*)^{-1}$, where $f^*(x) := (x, f(x))$ and λ is the Lebesgue measure, has the transport property.

1. Introduction and notations. The mass transport problem studied in this paper finds its origin at the end of the 18th century with Monge's paper on "clearings and fillings" (cf. [22]). A modern formulation could be as follows: let P and Q be two probabilities on E; can we find a measurable function $\varphi : E \to E$ transporting P into Q and minimizing a given cost c?

The mass transport problem has attracted a lot of attention in recent years and has found applications in mathematical sciences such as statistics, economics and fluid mechanics (see [5], [23] and [25]).

Let us introduce a few notations. Let (E, d) be a complete separable metric space, and $\mathcal{M}$ the set of all Borel probability measures defined on E. For $P, Q \in \mathcal{M}$, we denote by $\mathcal{M}(P, Q)$ the set of all probability measures defined on $E \times E$ whose marginal distributions are, respectively, P and Q. Another form of mass transport problem, formulated by Kantorovich (cf. [16] and [17]), is to evaluate the functional

$$\mathcal{K}_c(P,Q) := \inf\Big\{\int c(x,y)\, d\mu : \mu \in \mathcal{M}(P,Q)\Big\},$$

where the cost $c(x, y)$ is a measurable function ≥ 0 on $E \times E$. The functional $\mathcal{K}_c(P, Q)$ is called *the Kantorovich functional* (or *Kantorovich metric* if c is

1991 *Mathematics Subject Classification*: Primary 62H20; Secondary 26A27.

Key words and phrases: Monge–Kantorovich transportation problem, optimal coupling, cyclic monotonicity, Lévy–Wasserstein distance, (c-c)-surface, strongly nonatomic probability.

the distance d). An important case is when c is the square of the Euclidean distance on $E=\mathbb{R}^d$; $\mathcal{K}_c(\mu,\nu)^{1/2}$ is then the Lévy–Wasserstein metric.

Given two random variables X and Y defined on the probability space $(\Omega,\mathcal{A},\mathbb{P})$ with distribution laws P and Q, we say that (X,Y) is a *c-optimal coupling* (c-o.c.) between P and Q with respect to $\mathbb{P}$ if

$$\int c(X,Y)\,d\mathbb{P}=\mathcal{K}_c(P,Q).$$

In recent years, the characterization of the solution, for particular cost functions $c(x,y)$, of the Monge–Kantorovich problem has been given in terms of c-subgradients of generalized convex functions (cf. [18], [19], [4], [6], [24], [28], [29], [9] and [13]).

It is natural to find a condition on P so that, for any other probability Q, there should be a measurable function $\varphi: E\to E$ such that $(X,\varphi(X))$ is a c-optimal coupling between P and Q.

When E is a Hilbert separable space, Cuesta-Albertos and Matrán introduced in [8] the notion of *strongly nonatomic* probabilities, realizing the above condition for a quadratic cost. Abdellaoui and Heinich have shown in [1] that for such a probability P, we have $P(\nabla F \text{ exists})=1$ for any convex function F with $P(F\in\mathbb{R})=1$.

More recently, McCann [20], Gangbo and McCann [13] have given another class of probability satisfying the above condition, for more general costs. This class is the set of probabilities on $\mathbb{R}^d$ which vanish on subsets having Hausdorff dimension $\le d-1$. A note from [13] says that the class of probability which vanish on (*c-c*)-*surfaces* goes as well. McCann [21] conjectures and starts proving that this class is the only one that solves the transport problem for a quadratic cost.

In the present paper, in order to simplify matters, we consider $E=\mathbb{R}^d$ with the inner product $\langle\cdot,\cdot\rangle$ and the quadratic cost. The article is organized as follows: in the next section, we give a characterization of the probabilities which have the transport property, preceded by the necessary definitions; in Section 3, we introduce the notion of strongly nonatomic probabilities and we investigate their properties. Finally, in Section 4, we examine the relationship between the nondifferentiability of a real function f and the fact that the probability measure $\lambda_{f^*}:=\lambda\circ(f^*)^{-1}$, where $f^*(x):=(x,f(x))$ and λ is the Lebesgue measure, has the transport property. For example, we show that the probability λ_{B^*}, where B is the Brownian motion, is strongly nonatomic a.e.; on the other hand, if we take $f(x)=\int_0^x B(t)\,dt$, the probability λ_{f^*} has the transport property a.e. but it is not strongly nonatomic a.e.

2. Characterization of transport property. The purpose of this section is to give a characterization of the probabilities which have the transport property. It is useful to review some facts of life concerning convex functions $F : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$; the case of F identically ∞ is excluded by convention. Rockafellar [26] provides the standard reference. Whenever F is finite at x, its graph admits a supporting hyperplane: there is some $\varphi(x) \in \mathbb{R}^d$ such that

$$\langle y - x, \varphi(x)\rangle + F(x) \le F(y) \qquad \forall y \in \mathbb{R}^d. \tag{1}$$

In this case $\varphi(x)$ is called a *subgradient* of F at x, motivating the following definitions:

DEFINITION 1. The *subdifferential* of a convex function F on $\mathbb{R}^d$ is the subset $\partial F \subset \mathbb{R}^d \times \mathbb{R}^d$ of pairs $(x, \varphi(x))$ satisfying (1) for all $y \in \mathbb{R}^d$.

DEFINITION 2. A subset $S \subset \mathbb{R}^d \times \mathbb{R}^d$ is called *cyclically monotone* (c.m.) if it satisfies

$$\sum_{i=1}^{k} \langle x_{i+1} - x_i, y_i\rangle \le 0$$

for any finite number of points $(x_1, y_1), \dots, (x_k, y_k) \in S$, with $x_{k+1} := x_1$.

Rockafellar's main result in [26] exhibits the connection between gradients of convex functions and cyclically monotone sets: any cyclically monotone set $S \subset \mathbb{R}^d \times \mathbb{R}^d$ is contained in the subdifferential of some convex function on $\mathbb{R}^d$.

A function $\varphi : D \subset \mathbb{R}^d \to \mathbb{R}^d$ is *cyclically monotone* if it has cyclically monotone graph. As we have recalled, if P is a strongly nonatomic probability (see below) and Q another probability on $\mathbb{R}^d$, having second moments, there is a cyclically monotone function φ such that $(X, \varphi(X))$ is optimal for (P, Q). The function φ is P-a.e. in a gradient of a convex function (cf. [1] and [8]). The suppression of the moment condition leads to the transport property and to the differentiation of convex functions.

The study of differentiation of convex functions leads Zajíček (cf. [32]) to introduce the following definition.

DEFINITION 3. We define $M \subset \mathbb{R}^d$ to be a *(c-c)-surface* of dimension k, $k < d$, if there is a permutation σ of $\{1, \dots, n\}$ and $2n - 2k$ convex functions f_i, g_i, $k < i \le d$, defined throughout $\mathbb{R}^k$, such that $M = (x_1, \dots, x_d)$, where $x_{\sigma(j)} = y_j$, $y_i = f_i(y_1, \dots, y_k) - g_i(y_1, \dots, y_k)$ for $k < i \le d$. In other words, M is the graph of a difference of convex functions.

Zajíček's main theorem in [32] gives a characterization of the set of all points at which a continuous function is not differentiable.

THEOREM 1 (cf. [32]). *A subset $M \subset \mathbb{R}^d$ is a subset of the set of all points at which a continuous convex function f is not differentiable in $\mathbb{R}^d$ if and only if M can be covered by countably many (c-c)-surfaces.*

We can now introduce the transport property.

Definition 4. A Borel probability measure P on $\mathbb{R}^d$ has the *transport property* (t.p.) if it can be mapped to any other Borel probability measure by a cyclically monotone map.

For example, a probability which vanishes on Borel subsets having Hausdorff dimension $\le d-1$ has the transport property (cf. [20] and [13]). More generally, our main result in this section is the following.

Theorem 2. *The following assertions are equivalent for a Borel probability measure P on $\mathbb{R}^d$*:

(i) *The probability P assigns zero to every* (c-c)*-surface.*

(ii) *For any convex function F, $P(F\in\mathbb{R})=1$ implies $P(\nabla F$ exists$)=1$.*

(iii) *The probability P has the transport property.*

(iv) *The probability P is nonatomic and*, *if (φ_n) is a sequence of cyclically monotone functions which converges in law*, *then φ_n converges almost everywhere.*

Proof. (i)⇔(ii) is obvious by the previous Theorem 1 of Zajíček.

(iv)⇒(iii). Let Q be a probability and (Q_n) a sequence of nonatomic probabilities which converges in law to Q. If P is a nonatomic probability, there is, for each n, a cyclically monotone function φ_n such that $P\circ(\varphi_n)^{-1} = Q_n$ (cf. [1]). Now, suppose P satisfies (iv). Then φ_n converges a.e. to a cyclically monotone function φ, and $Q = P\circ(\varphi)^{-1}$.

(iii)⇒(i). Assume that there exists a (c-c)-surface M such that $P(M) > 0$. Theorem 1 provides a convex function F on $\mathbb{R}^d$ whose differentiability fails throughout M. It is easy to ensure the differentiability of the Legendre transform F^* of F.

Now, by Proposition 7 of the appendix, there exist two cyclically monotone measurable functions $\tau^+(x)$ and $\tau^-(x)$ in $\partial F(x)$ such that $\langle\tau^+(x), e_1\rangle \ge \langle\tau^-(x), e_1\rangle$ P-a.e., where $e_1 := (1,0,\ldots,0)\in\mathbb{R}^d$. If Q^+ and Q^- are the images of P under τ^+ and τ^- respectively, then $Q := (Q^+ + Q^-)/2$ cannot be attained as the image of P under any cyclically monotone map.

For a contradiction, we suppose that Q is the image of P under a cyclically monotone map τ. Since $s := \nabla F^*$ is a continuous map such that $s(\tau^+(x)) = s(\tau^-(x)) = x$, it is easy to prove, using the duality techniques of Brenier [4] or Gangbo and McCann [13], that $Q\circ(s\times \mathrm{id})^{-1}$ is the *only* measure with cyclically monotone support having P and Q as its marginals.

This uniqueness implies that $P\circ(\mathrm{id}\times\tau)^{-1}$ is the same measure as $Q\circ(s\times\mathrm{id})^{-1}$, from which one deduces that $s(\tau(x)) = x$ holds P-almost everywhere. In other words, after modifying τ on a set of measure zero, it will be true that τ lies in the subgradient of F.

Now, define $\tau' := (\tau^+ + \tau^-)/2$, and divide M into two disjoint measurable subsets $M := M^+ \cup M^-$ by defining

$$M^+ := \{x \in M : \langle \tau(x) - \tau'(x), e_1 \rangle \geq 0\}.$$

Also, divide $S := s^{-1}(M)$ into disjoint measurable subsets S^+ and S^- given by

$$S^+ := \{y \in S : \langle y - \tau'(s(y)), e_1 \rangle \geq 0\}.$$

Finally, let $Y := (s^{-1}(M^+) \cap S^+) \cup (s^{-1}(M^-) \cap S^-)$. Then $\tau^{-1}(Y) = M^+ \cup M^- = M$, so $P(\tau^{-1}(Y)) = P(M)$. On the other hand, since $Q^+(S^-) = 0$ while $Q^-(S^+) = 0$, one sees that

$$\begin{aligned} 2Q(Y) &= Q^+(s^{-1}(M^+) \cap S^+) + Q^-(s^{-1}(M^-) \cap S^-) \\ &= P(M^+) + P(M^-) = P(M). \end{aligned}$$

Since $P(\tau^{-1}(Y)) = 2Q(Y) = P(M) \neq 0$, the probability Q cannot be the image of P under τ, a contradiction.

(ii)⇒(iv). Let $\mathcal{M}_0$ be the set of probabilities satisfying (ii). For any $a \in \mathbb{R}^d$, we have

$$P \in \mathcal{M}_0 \quad \text{if and only if} \quad P_a \in \mathcal{M}_0,$$

where $P_a(\cdot) = P(\cdot - a)$. Indeed, if $F_a(x) = F(x + a)$, we have $P \circ (F_a)^{-1} = P_a \circ (F)^{-1}$ and $P \circ (\nabla F_a)^{-1} = P_a \circ (\nabla F)^{-1}$.

(a) Assume that the sequence (φ_n) of cyclically monotone functions converges in law to a function φ. Then $P(\underline{\lim} \|\varphi_n\| < \infty) = 1$. By translation (see above) we can suppose that $\underline{\lim} \|\varphi_n(0)\| < \infty$. Choose convex functions F_n so that $F_n(0) = 0$ and $\varphi_n(x) \in \partial F_n(x)$. The inequalities $\langle x, \varphi_n(0) \rangle \leq F_n(x) \leq \langle x, \varphi_n(x) \rangle$ show that

$$P(\mathcal{C} := \{\underline{\lim}\, F_n(x) \in \mathbb{R}\}) = 1.$$

For $x \in \mathcal{C}$, there is a subsequence $(n_k^x)_k$ such that $F_{n_k^x}(x)$ converges. Take a set $D := \{x_p : p \in \mathbb{N},\ x_0 = 0\}$ which is dense in $\mathcal{C}$ and contained in the support of P. There is a subsequence (n_k^*) so that $F_{n_k^*}(x_p) \to \underline{F}(x_p) := \underline{\lim} F_n(x_p)$ for all $p \in \mathbb{N}$. We deduce by Theorem 10.8 of Rockafellar (cf. [26]) that $F_{n_k^*}(x) \to \underline{F}(x)$ for all $x \in \mathcal{C}$. As $P(\underline{F} \in \mathbb{R}) = 1$ and $P \in \mathcal{M}_0$, we have $\nabla F_{n_k^*} \to \nabla \underline{F}$ a.e. (cf. [26], Theorem 25.7) and $\nabla \underline{F} = \varphi$ in law.

(b) The previous argument remains valid if, from the beginning, we substitute the initial sequence in (a) by a subsequence $(\widetilde{n}_k)$. We extract a new subsequence $(\widetilde{n}_k^*)$ so that $F_{\widetilde{n}_k^*} \to F$ and we deduce that $\nabla \underline{F} = \nabla F$ in law. Lemma 2 of [15] gives the equality $\nabla \underline{F} = \nabla F$ a.e. Therefore, we have proved that φ_n converges in probability. This property is sufficient to prove all equivalences of Theorem 1. Nevertheless, let us prove the a.e. convergence.

For any subsequence $(\widetilde{n}_k)$, let $\widetilde{\mathcal{C}} := \{\underline{\lim} F_{\widetilde{n}_k}(x) \in \mathbb{R}\}$; we have $\mathcal{C} \subset \widetilde{\mathcal{C}}$. Then the sequence (x_p) satisfies: there is a subsequence $(\widetilde{n}_k^*) \subset (\widetilde{n}_k)$ such that $F_{\widetilde{n}_k^*}(x_p)$ converges for all p. For a point a, suppose $F_{\widetilde{n}_k}(a) \to \infty$ and $a \in \operatorname{convex}(x_0, \ldots, x_j)$. There is a subsequence $(\widetilde{n}_k^*)$ such that $F_{\widetilde{n}_k^*}(x_p)$ converges and we obtain a contradiction, so $P(\sup F_n \in \mathbb{R}) = 1$. Write $\overline{F} = \overline{\lim} F_n$ a.e.; we have seen that $\nabla \underline{F} = \nabla \overline{F}$ a.e. The adaptation of Theorems 24.8 and 24.9 of Rockafellar [26] shows that all limit convex functions are equal a.e. Thus $\underline{F} = \overline{F}$ a.e. and, for all p, $F_n(x_p)$ converges. Indeed, let F and G be two limits, so that, for one p, we have $F(x_p) < G(x_p)$. There exist $\varepsilon > 0$ and a ball $B(x_p, \eta)$ such that $F(x) \le \varepsilon + G(x)$ for $x \in B(x_p, \eta) \cap \mathcal{C}$. We have a contradiction with the equality a.e., because x_p is in the support of P. We finally conclude that F_n converges a.e., and this completes the proof of Theorem 2. ∎

Remarks. (1) The implication (ii)⇒(iv) gives a new proof of the main theorems in [31], [10], [15] and [2].

(2) Let $\mathcal{M}_t$ be the set of all probabilities with the transport property. The following assertions are now easy to prove:

- For any probabilities $P \in \mathcal{M}_t$ and Q, we have the uniqueness (a.e.) of the cyclically monotone function φ such that $Q = P \circ \varphi^{-1}$.
- If $P \in \mathcal{M}_t$ and Q is absolutely continuous with respect to P, then $Q \in \mathcal{M}_t$.
- $\mathcal{M}_t$ is an extremal convex set: let P and Q be two probabilities and $\alpha \in]0,1[$. Then $\alpha P + (1-\alpha)Q \in \mathcal{M}_t$ if and only if P and Q are in $\mathcal{M}_t$.
- For a probability P, there exists a unique a.e. Borel set A such that $P|_A \in \mathcal{M}_t$ and $P|_{A^c} \notin \mathcal{M}_t$ where $P|_A(\cdot) = P(A \cap \cdot)/P(A)$. This property results from the fact that if, for a sequence (A_n), we have $P|_{A_n} \in \mathcal{M}_t$, then $P|_{\bigcup A_n} \in \mathcal{M}_t$.

3. On strongly nonatomic probability. The notion of strongly nonatomic probability, introduced by Cuesta-Albertos and Matrán (cf. [8]), is bound to the choice of an orthonormal basis. Only such bases will be considered in this study. In this section, we will provide some properties of this class of probabilities.

A set $G \subset \mathbb{R}^d$ is a *graph* for a basis $\mathbf{b} = (e_1, \ldots, e_d)$ if there exists a Borel set $A \subset e_d^\perp$ (the orthogonal space to $\mathbb{R}e_d$) and a measurable function $f : A \to \mathbb{R}e_d$ such that $G = \{(x, f(x)) : x \in A\}$. We also write $G = G^{\mathbf{b}}$. Let $\mathbf{b} = (e_1, \ldots, e_d)$ be a basis of $\mathbb{R}^d$. For a probability measure P on $\mathbb{R}^d$, $P_{e_i^\perp}$ is the marginal law of P on $e_i^\perp$; and for $P_{e_i^\perp}$-almost every $y \in e_i^\perp$, π_y is the conditional law $P(\cdot \mid y)$.

Definition 5. A probability measure P on $\mathbb{R}^d$ is *strongly nonatomic* (s.n.) if there exists a basis $\mathbf{b} = (e_1, \ldots, e_d)$ such that $\forall i \in \{1, \ldots, d\}$, $\pi_y(\cdot)$ is nonatomic $P_{e_i^\perp}$-a.e.

By negation, P is not strongly nonatomic if, for any $\mathbf{b} = (e_1, \ldots, e_d)$, there exists i and a Borel set $A \subset e_i^\perp$, $P_{e_i^\perp}(A) > 0$, such that π_y has a nonatomic component for all $y \in A$.

In [27] we can find a more general version of the next equivalence: Let P be a probability measure on $\mathbb{R}^p \times \mathbb{R}^q$ and P_1 its marginal law on $\mathbb{R}^p$. The following assertions are equivalent:

(i) $P_1(\{x : \pi_x \text{ has a nonatomic component}\}) > 0$.

(ii) There exists a measurable function $f : \mathbb{R}^p \to \mathbb{R}^q$ such that

$$P(\{(x, f(x)) : x \in \mathbb{R}^p\}) > 0.$$

This yields the following result.

Proposition 1. *A probability measure P on $\mathbb{R}^d$ is strongly nonatomic if and only if there is a basis $\mathbf{b}$ such that for any graph $G^{\mathbf{b}}$, we have $P(G^{\mathbf{b}}) = 0$.*

The next assertion is now obvious: if P is strongly nonatomic, then every probability Q which is absolutely continuous with respect to P is also strongly nonatomic.

The following property completes Lemma 1 of [15].

Proposition 2. *A probability measure P on $\mathbb{R}^d$ is strongly nonatomic for a basis $\mathbf{b} = (e_1, \ldots, e_d)$ if and only if for every Borel set B, the set $\mathcal{E}_B := \{x : x \in B,\ (\forall \varepsilon = \pm 1)\ (\forall i)\ (\exists$ a sequence $(u_n = u_n(x, \varepsilon, i)) > 0$ converging to 0 and $(x + \varepsilon u_n e_i)_n \subset B\}$ is P-a.e. equal to B.*

Proof. The direct part is established in [15]. For the converse, if the probability P is not strongly nonatomic, there exists a graph G in a basis $\mathbf{b}$ such that $P(G) > 0$. This graph fails the relevant property. ■

Now, we are going to introduce a wider class of probabilities with similar properties (cf. Theorem 4).

Definition 6. A probability measure P is *quasi-strongly nonatomic* (q.s.n.) (respectively, *purely not strongly nonatomic* (p.n.s.n.)) if for every A with $P(A) > 0$, the probability $P|_A$ is s.n. (resp. not s.n.).

We write $\mathcal{M}_q$ (resp. $\mathcal{M}_q^{\mathrm{c}}$) the set of quasi-strongly nonatomic (resp. purely not strongly nonatomic) probabilities. We obtain the following dichotomy.

Proposition 3. *For any probability measure P on $\mathbb{R}^d$, there exists a unique a.e. Borel set A such that $P|_A \in \mathcal{M}_q$ and $P|_{A^{\mathrm{c}}} \in \mathcal{M}_q^{\mathrm{c}}$.*

The proof is a consequence of the stability under countable unions of the set $\{B : P|_B \in \mathcal{M}_q^{c}\}$.

It is easy to deduce the following assertion: A probability measure P is in $\mathcal{M}_q$ if and only if there is a Borel set A with $P|_A, P|_{A^c} \in \mathcal{M}_q$.

The next property expresses the fact that $\mathcal{M}_q$ is an extremal convex set.

PROPOSITION 4. *Let P and Q be two probability measures on $\mathbb{R}^d$ and $\alpha \in]0,1[$. Then $\alpha P + (1-\alpha)Q \in \mathcal{M}_q$ if and only if $P, Q \in \mathcal{M}_q$.*

Proof. Let $R = \alpha P + (1-\alpha)Q$, and assume that $R \in \mathcal{M}_q$. Let A be a Borel set with $R(A) > 0$. As $P|_A$ and $Q|_A$ are absolutely continuous with respect to $R|_A$, we deduce that P and Q are in $\mathcal{M}_q$. Conversely, write $Q = Q_1 + Q_2$, where Q_1 is absolutely continuous with respect to P and $Q_2 \perp P$ (i.e., P and Q_2 are mutually singular). Let D be a Borel set such that $P(D) = 0 = Q_2(D^c)$. For $A \subset D^c$, $R|_A = \alpha P|_A + (1-\alpha)Q_1|_A$ is absolutely continuous with respect to P, thus $R|_A \in \mathcal{M}_q$. For $A \subset D$, $R|_A = Q_2|_A = Q|_A$, thus $R|_A \in \mathcal{M}_q$. We conclude with the use of Proposition 3. ∎

Finally, one can also show that $\mathcal{M}_q \subset \mathcal{M}_t$.

4. Transport and derivation. We limit this part to the probabilities on $\mathbb{R}^2$ which have the form $P = \lambda \circ (f^*)^{-1} := \lambda_{f^*}$, where $f^*(x) := (x, f(x))$ and f is a real measurable function defined on a Borel set A such that $\lambda(A) = 1$. We examine the relationship between the nondifferentiability of the real function f and the fact that the probability measure λ_{f^*} has the transport property.

Proposition 2 shows that, if the probability P is strongly nonatomic, then $\lambda(\{x : f'(x) \text{ exists}\}) = 0$. This fact motivates the study.

Let us introduce the following vocabulary.

- Two bases $\mathbf{b}^i = (e_1^i, e_2^i)$, $i = 1, 2$, are *different* if $|\langle e_1^1, e_1^2\rangle| \in]0,1[$.
- For a real function f defined on A with $\lambda(A) > 0$, we denote by f_B the restriction to $B \subset A$.
- The function f is *purely nonderivable* (p.n.d.) on A if for every Borel set $B \subset A$ with $\lambda(B) > 0$, f_B is not almost everywhere derivable (i.e., $\lambda(\{x : f_B' \text{ exists}\}) = 0$).
- For f defined on a Borel set A and $x \in A$, the *Dini derivate numbers* are

$$D^+ f(x) = \overline{\lim_{u\to x_+}} \frac{f(u)-f(x)}{u-x}, \qquad D_+ f(x) = \underline{\lim_{u\to x_+}} \frac{f(u)-f(x)}{u-x},$$
$$D^- f(x) = \overline{\lim_{u\to x_-}} \frac{f(u)-f(x)}{u-x}, \qquad D_- f(x) = \underline{\lim_{u\to x_-}} \frac{f(u)-f(x)}{u-x}.$$

We can now recall the Denjoy–Young–Saks Theorem (cf. [30]).

THEOREM 3 (cf. [30]). *Let f be a real function defined on an interval I. Then, with the possible exception of a null set, I can be decomposed into four sets*:

(1) $\{D^+f = D_+f = D^-f = D_-f \in \mathbb{R}\}$ (*on which f has a finite derivative*);
(2) $\{D^+f = \infty,\ D_-f = -\infty,\ D_+f = D^-f \in \mathbb{R}\}$;
(3) $\{D^+f = D_-f \in \mathbb{R},\ D_+f = -\infty,\ D^-f = \infty\}$;
(4) $\{D^+f = D^-f = \infty,\ D_+f = D_-f = -\infty\}$.

More generalizations and applications can be found in [3], [33], [14] and [7].

The following result characterizes the strongly nonatomic probabilities and shows these are independent of the basis choice.

THEOREM 4. *Let $f : A \subset \mathbb{R} \to \mathbb{R}$, $\lambda(A) = 1$, be a measurable function. Then the following assertions are equivalent*:

(1) *For any Borel set $B \subset A$ with $\lambda(B) > 0$, all $\delta \in \mathbb{R}$, and a.e. $x \in B$, the set $\delta^B(x) := \{u \in B : f(u) - f(x) = \delta(u - x)\}$ is uncountable.*

(2) *For any Borel set $B \subset A$ with $\lambda(B) > 0$, all $\delta \in \mathbb{R}$, and a.e. $x \in B$, the set $\delta^B(x)$ is not finite.*

(3) *For all bases* **b** *which are different from the initial basis, λ_{f^*} is strongly nonatomic.*

(4) *The probability λ_{f^*} is strongly nonatomic.*

(5) *The probability λ_{f^*} is quasi-strongly nonatomic.*

(6) *The function f is purely nonderivable.*

Proof. The assertion (2) shows that a Borel set B with $\lambda_{f^*}(B) > 0$ fails to be a graph in another basis. By Proposition 1, we have (2)⇒(3).

It is obvious that (3)⇒(1)⇒(2) and (3)⇒(4)⇒(5)⇒(6). The last implication is a consequence of Proposition 2.

(6)⇒(3). We suppose that, for a basis **b** (different from the initial basis), there exists $B \subset A$ such that $\lambda(B) > 0$ and the graph of $(f_B)^*$ is also a graph in **b**. Then

$$\lambda(\{D_-(f_B) > -\infty\} \cup \{D^+(f_B) < \infty\}) > 0.$$

The lemma of Saks [30] provides a subset C of B such that $\lambda(C) > 0$ and f_C is derivable, thus f is not purely nonderivable. ∎

This theorem can be made precise when we consider the intermediate value property, which is also called the Darboux Property.

For any interval I, we denote by $\mathcal{D}B_1(I)$ the set of measurable functions defined on I of the first class of Baire and which have the Darboux Property (see [7]).

THEOREM 5. *Let $f \in \mathcal{D}B_1(I)$. Then the following assertions are equivalent*:

(1) $\lambda(\{x : f'(x) \text{ exists}\}) = 0$.

(2) *For any interval $J \subset I$ and any $\delta \in \mathbb{R}$, J is the set of a.e. accumulation points of $\delta^J(x)$.*

(3) *We have $D^+f = D^-f = \infty$ and $D_+f = D_-f = -\infty$ a.e.*

Furthermore, any one of these assertions implies that $\lambda_{f^*} \in \mathcal{M}_t$.

Proof. In order to simplify the notations, we only consider the case $\delta = 0$. The function $g(x) = f(x) - \delta x$ gives the general case. We write $\#(A)$ for the cardinality of a set A.

Let $f \in \mathcal{D}B_1(I)$. Then

$$(2) \qquad \{x : f'(x) \text{ exists}\} = \operatorname*{ess\,sup}_{\delta,J}\{x : x \text{ is not an accumulation point of } \delta^J(x) - \{x\}\},$$

where $\delta \in \mathbb{R}$ and J is a subinterval of I. First, we prove that

$$\{x : \#(J \cap f^{-1}f(x)) < \infty\} \subset \{x : f'(x) \text{ exists}\}.$$

Let $B := \{x : J \cap f^{-1}f(x) = x\}$ and let x_1, x_2 be two points of B such that $x_1 < x_2$ and, for example, $f(x_1) < f(x_2)$. If, for $u \in [x_1, x_2]$, we have $f(u) > f(x_2)$, the Darboux Property shows that the line $f(x_2)$ cuts the graph of f at a point with abscissa in $]x_1, u[$, which is contradictory. So, for all $u \in]x_1, x_2[$, we have $f(x_1) < f(u) < f(x_2)$. This fact also shows that f is strictly nondecreasing on $E \cap [x_1, x_2]$. There exists a negligible set N such that $B \cap [x_1, x_2] \subset \mathcal{D} := \{x : f'(x) \text{ exists}\}$ a.e.

We will give an elementary proof. For $u \in [x_1, x_2]$ let $g(u) := \sup_{x \in B \cap [x_1,u]} f(x)$ and $h(u) := \inf_{x \in E \cap [u,x_2]} f(x)$. The functions g and h are nondecreasing on $[x_1, x_2]$ and $g(u) \le f(u) \le h(u)$. Furthermore, g and h are equal to f on $E \cap [x_1, x_2]$. For $x \in B \cap [x_1, x_2]$ and $u \in [x_1, x_2]$ we have

$$\frac{g(u) - g(x)}{u - x} \le \frac{f(u) - f(x)}{u - x} \le \frac{h(u) - h(x)}{u - x}$$

if $x < u$, and the reverse inequalities if $u < x$. As g and h are derivable a.e. on $[x_1, x_2]$, we have

$$\begin{aligned} g'(x) &\le \varliminf_{u \to x_+} \frac{f(u) - f(x)}{u - x} = D_+f(x) \\ &\le \varlimsup_{u \to x_+} \frac{f(u) - f(x)}{u - x} = D^+f(x) \le h'(x). \end{aligned}$$

The reverse inequalities give

$$h'(x) \le D_-f(x) = \varlimsup_{u \to x_-} \frac{f(u) - f(x)}{u - x} \le D^-f(x) \le g'(x).$$

Thus we have the result.

In fact, we have shown that for all $\delta \in \mathbb{R}$, and for any subinterval J, the set $\delta_1^J := \{x : \delta^J(x) = x\}$ is a.e. included in $\mathcal{D}$. Thus $\operatorname{ess\,sup}_{J,\delta} \delta_1^J \subset \mathcal{D}$.

Now, we assume that x is not an accumulation point of $\delta^I(x)$, i.e. $x \notin \overline{\delta^I(x)}$. There exists an interval J such that $x \in \delta_1^J$. Then, a.e.,

$$\operatorname*{ess\,sup}_{J,\delta}\{x : x \text{ is not an accumulation point of } \delta^J(x)\} \subset \mathcal{D}.$$

Notice that we can choose J as an open interval with rational ends. Conversely, if $x \in \mathcal{D}$, there exist $\delta \in \mathbb{Q}$ and J such that $x \in \delta_1^J$. Finally, we have a.e.

$$\{x : f'(x) \text{ exists}\} = \operatorname*{ess\,sup}_{J,\delta}\{x : x \text{ is not an accumulation point of } \delta^J(x)\}.$$

The proof of Theorem 5 is now clear. For (1)⇒(2), if $\lambda(\{x : f'(x) \text{ exists}\}) = 0$, then almost every x is an accumulation point of $\delta(x)$ for each δ. (3)⇒(1) is obvious and (2)⇒(3) results from the Denjoy–Young–Saks Theorem.

The transport property can be established by showing that, for any rectifiable curve S, we have $\lambda_{f^*}(S) = 0$. This can also be established as a consequence of the following proposition.

Proposition 5. *Let f be a function such that $f \in \mathcal{D}B_1(I)$, $\lambda(I) = 1$ and $\lambda(f'$ exists$) = 0$. The following assertions are equivalent and both hold*:

(a) *The probability λ_{f^*} is nonatomic*, *i.e.* $(\forall y)\ \lambda(\{x : f(x) = y\}) = 0$.

(b) *For all continuous functions g on I if $\lambda(\{x : g'(x)$ exists$\}) > 0$, then* $\lambda(f = g) = 0$.

Before proving this proposition, we finish the proof of Theorem 5.

It is sufficient to take two convex functions for f and g; then the function $h := f - g$ is in $\mathcal{D}B_1(I)$ and satisfies $\lambda(\{x : h'(x) \text{ exists}\}) = 0$. Then, from Proposition 5, $\lambda(h = 0 = f - g) = 0$, which proves Theorem 5. ■

Proof of Proposition 5. First, we show that if $f \in \mathcal{D}B_1(I)$, then $|f| \in \mathcal{D}B_1(I)$; furthermore, if $\lambda(f' \text{ exists}) = 0$, then $\lambda(|f|' \text{ exists}) = 0$.

Indeed, let $x < y$, let $a \in [|f|(x), |f|(y)]$, and assume that $f(x) = |f|(x)$ and $-f(y) = |f|(y)$. There exists $u \in [x, y]$ such that $f(u) = -f(x)$. Thus there exists $v \in [u, y]$ such that $f(v) = -a$ and $|f|(v) = a$. We have proved that $|f|$ satisfies the Darboux Property.

Assume now that $\lambda(\{x : f'(x) \text{ exists}\}) = 0$, and that Theorem 5(3) is satisfied. For $x \notin N$ with $\lambda(N) = 0$, there is a sequence $(x_n) \to x_+$ such that

$$\frac{f(x_n) - f(x)}{x_n - x} \to \infty;$$

if $f(x) > 0$, then

$$\frac{f(x_n) - f(x)}{x_n - x} \le \frac{|f|(x_n) - |f|(x)}{x_n - x} \to \infty,$$

which shows that $|f|$ is not derivable at the point x. The other cases are similar.

To show that (b)⇒(a), it suffices to take $g \equiv y$.

(a)⇒(b), one sees that the function $h = f - g$ is in $\mathcal{D}B_1(I)$, and from (a) we have

$$\lambda(h = 0 = f - g) = 0.$$

To complete the proof, we show that (a) holds. Let

$$N := \{x : f' \text{ and } |f|' \text{ exist}\}.$$

The set N is negligible. The first part of Theorem 5 shows that for a.e. x, $|f|^+(x) = \infty$, and $|f|_+(x) = -\infty$. If furthermore $f(x) = 0$, we have a contradiction, and $\{x : f(x) = 0\}$ is a negligible set. By translation we have $\lambda(\{x : f = y\}) = 0$. ∎

Examples. Let $B(t)$, $t \in [0,1]$, be the standard Brownian motion.

(1) The probability λ_{B^*} is strongly nonatomic on $\mathbb{R}^2$ for a.e. ω. Indeed, the function $t \mapsto B(t)$ is continuous and not derivable for a.e. ω. And, for all δ, the set of accumulation points of $\delta^{[0,1]}(x)$ has no isolated point. Thus it is uncountable and we have Theorem 4(1).

(2) Let $f(x) = \int_0^x B(t)\,dt$. The probability λ_{f^*} is in $\mathcal{M}_t$, but it is not strongly nonatomic. The function f is derivable, so from Theorem 4, λ_{f^*} is not strongly nonatomic. Furthermore, if on a set A with $\lambda(A) > 0$, we have $f = g - h$, where g and h are two convex functions, then $B(t)$ is of bounded variation on A, which is contradictory. Thus λ_{f^*} has the transport property.

A contrario, we have the following proposition.

Proposition 6. *Let f be a continuously differentiable function on* $[0,1]$. *Then there is a nonatomic probability on the graph of f which does not have the transport property.*

Proof. We can assume, without loss of generality, $f([0,1]) \subset [0,1]$. Let G be the graph of f, and let $E = [0,1] \times [0,1]$. Let p_G be the projection on the graph: $p_G(z) := \{g : g \in G,\ d(z,g) = d(z,G)\}$. It is easy to verify the following assertion: If F is a compact $\subset G^{\mathrm{c}} \cap E$ such that $\lambda \otimes \lambda(F) > 0$, then $p_G(F)$ is uncountable.

This last point ensures the existence of a probability measure ν which is nonatomic on $[0,1]$ and such that $P = \nu_{f^*}$ and $P(p_G(G^{\mathrm{c}})) = 1$. This probability is not in $\mathcal{M}_t$.

For $g \in p_G(G^c)$, let $z_g \in E$ be the farthest point on the normal $\mathcal{N}_g$ (oriented from the positive axis) and such that $g \in p_G(z)$. Let $I_g := [g, M_g[$. We define on $p_G(G^c) \otimes \mathcal{B}(G^c)$ the kernel

$$N(g, B) := \frac{1}{\lambda(I_g)} \lambda(B \cap I_g).$$

Let Q be the probability defined by

$$Q(B) := \int_{p_G(G^c)} N(g, B)\, dP(g).$$

We have $P = Q \circ p_G^{-1}$ and

$$\inf\Big\{\int d^2(x, y)\, d\mu : \mu \in \mathcal{M}(P, Q)\Big\} = \int d^2(z, p_G(z))\, dQ(z).$$

By negation, we suppose $P \in \mathcal{M}_t$. Then there exists a cyclically monotone function φ such that $Q = P \circ \varphi^{-1}$ and

$$\inf\Big\{\int d^2(x, y)\, d\mu : \mu \in \mathcal{M}(P, Q)\Big\} = \int d^2(z, \varphi(g))\, dP(g).$$

So, P-a.e. we have $\varphi(g) \in [z_g, g]$. As $\{\varphi(p_G(G^c)) \cap I_g\} = \{\varphi(g)\}$, we obtain

$$Q(\varphi \circ p_G(G^c)) = \int_{p_G(G^c)} \frac{1}{\lambda(I_g)} \lambda(G^c \cap I_g)\, dP(g) = 0,$$

which is absurd. ∎

Remarks. (1) From Theorem 2, there is a function h which is the difference of two convex functions such that $\nu(f = h) > 0$.

(2) The results of Section 4 can be generalized by considering a nonatomic probability η on $\mathbb{R}$ instead of the Lebesgue measure λ and by replacing the ordinary derivation, related to λ, by that related to η (see [12]).

5. Appendix. Given a convex function F on $\mathbb{R}^d$, in this appendix we prove the existence of a measurable function in $\partial F(x)$ with certain properties. We note that the existence of a Borel function, with different properties, in a more general setting can be found in [11].

Lemma 1. *Let $D = (x_n)$ be a dense sequence in a metric space (E, d), and let $\varphi : D \to \mathbb{R}^d$. Then φ can be extended to a measurable function $\underline{\varphi} : E \to \overline{\mathbb{R}}^d$ and there exists a map from E into $\mathbb{N}^{\mathbb{N}}$, denoted by $y \mapsto n_k^y$, such that $\underline{\varphi}(y) = \lim_k \varphi(x_{n_k^y})$.*

Proof. With each $y \in E$, we associate the sequence $(\widetilde{n}_k^y)_k$ where $\widetilde{n}_k^y := \inf\{n : d(x_n, y) < 1/k\}$. Let $\varphi_k(y) := \varphi(x_{\widetilde{n}_k^y})$ and $N_p := \{q : \varphi(x_q) = \varphi(x_p)\}$. Then

$$\bigcup_{q \in N_p} \{y : \widetilde{n}_k^y = q\} = \{y : \varphi_k(y) = \varphi(x_p)\}.$$

Indeed, if $\widetilde{n}_k^y = q$ for $q \in N_p$, then $\varphi_k(y) = \varphi(x_p)$. Conversely, if $\varphi_k(y) = \varphi(x_p)$, the integer $q = \widetilde{n}_k^y$ is in N_p, which shows the equality. Since every $\{y : \widetilde{n}_k^y = p\}$ is a Borel set, the function $\underline{\varphi} := \underline{\lim} f_k$ is Borel on $\overline{\mathbb{R}}^d$. For each i, one has $\widetilde{n}_k^{x_i} = i$ when k is large enough, thus $\underline{\varphi}(x_i) = \varphi(x_i)$.

To prove the second statement, let

$$n_k^y := \inf\{i : \sup(d(x_i, y), |\underline{\varphi}(y) - \varphi(x_i)|) < 1/k\}.$$

The number n_k^y is well defined, because $d(x_{n_k^y}, y) \to 0$ and $\underline{\varphi}(y) = \underline{\lim}\, \varphi(x_{n_k^y})$. For each $y \in E$, we have $\underline{\varphi}(y) = \lim \varphi(x_{n_k^y})$, which proves Lemma 1 in the one-dimensional case.

For higher dimensions, it is enough to treat the $\mathbb{R}^2$ case. Let g be a second function defined on D such that the map $y \mapsto g_k(y) = g(x_{n_k^y})$ is Borel. Finally, the coupling (f, g) can be extended to $(\underline{f}, \underline{g})$ satisfying the conditions lemma. ■

Remark. Let φ be a cyclically monotone function defined on a countable set $D \subset \mathbb{R}^d$. Then the above extension is cyclically monotone on $\overline{D}$. Indeed, take m points from $\overline{D}$, and, with the previous notations, set $x_k^i := x_{n_k^{y^i}}$. For every $i \in \{1, \ldots, m\}$, we have $x_k^i \to y^i$ and $\varphi(x_k^i) \to \underline{\varphi}(y^i)$ as $k \to \infty$. By the hypotheses, we have

$$\sum_{k=0}^{m} \langle x_k^{i+1} - x_k^i, \varphi(x_k^i)\rangle \geq 0.$$

Letting $k \to \infty$, we get the expected result.

Proposition 7. *If $F < \infty$ is a convex function on $\mathbb{R}^d$, then there exists a measurable function φ defined on* $\operatorname{dom} F$ *with $\varphi(x) \in \partial F(x)$. Furthermore, if P is a Borel probability on $\mathbb{R}^d$ such that $P(F$ is not differentiable$) > 0$, then there exist two measurable functions $\tau^+(x)$ and $\tau^-(x)$ in $\partial F(x)$ such that $P(\langle \tau^+(x), e_1\rangle \geq \langle \tau^-(x), e_1\rangle) = 1$, where $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^d$.*

Proof. Let E be the set where ∇F exists and let $D \subset E$ be a dense countable subset of $\operatorname{dom} F$. If the interior of $\operatorname{dom} F$ is empty, we can restrict ourselves to a subspace generated by $\operatorname{dom} F$ and apply Theorem 25.5 of Rockafellar [26].

From Lemma 1, there exists a measurable map $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ such that $\varphi(x) \in \partial F(x)$ for $x \in \operatorname{dom} F$. Indeed, for every $x_n \in D$ we have $\langle y - x_n, \nabla F(x_n)\rangle + F(x_n) \leq F(y)$. For $x \in \operatorname{dom} F$ we take a suitable subsequence denoted by $(x_n^*) \subset D$ such that $x_n^* \to x$ and $\nabla F(x_n^*)$ converges to $F(x)$. It follows that $\langle y - x, \varphi(x)\rangle + F(x) \leq F(y)$.

By using Lemma 1 and Zajíček's theorem, we can prove the second statement of the proposition. ■

Acknowledgments. We are grateful to R. McCann for many helpful discussions on the content of this work, and to C. Dellacherie and P. Raynaud de Fitte for comments on the paper. We also thank the referee for valuable suggestions.

References

[1] T. Abdellaoui et H. Heinich, *Sur la distance de deux lois dans le cas vectoriel*, C. R. Acad. Sci. Paris 319 (1994), 397–400.

[2] N. Belili, *A theorem of stochastic representation related to the Monge–Kantorovich problem*, Statist. Probab. Lett. 41 (1999), 139–143.

[3] C. L. Belna, M. J. Cargo, M. J. Evans and P. D. Humke, *Analogues of the Denjoy–Young–Saks theorem*, Trans. Amer. Math. Soc. 271 (1982), 253–260.

[4] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris 305 (1987), 805–808.

[5] —, *The dual least action problem for an ideal incompressible fluid*, Arch. Rational Mech. Anal. 122 (1991), 323–351.

[6] —, *Polar factorization and monotone rearragement of vector-valued functions*, Comm. Pure Appl. Math. 44 (1991), 375–417.

[7] A. Bruckner, *Differentiation of Real Functions*, CRM Monogr. Ser. 5, Amer. Math. Soc., 1994.

[8] J. A. Cuesta-Albertos and C. Matrán-Bea, *Notes on the Wasserstein metric in Hilbert spaces*, Ann. Probab. 17 (1989), 1264–1276.

[9] J. A. Cuesta-Albertos, C. Matrán-Bea and A. Tuero-Diaz, *On lower bounds for the L^2-Wasserstein metric in a Hilbert space*, J. Theoret. Probab. 9 (1996), 263–283.

[10] —, —, —, *Propreties of the optimal maps for the l^2-Monge–Kantorovich transportation problem*, preprint, 1996.

[11] C. Dellacherie et P. A. Meyer, *Probabilités et potentiel*, Hermann, Paris, 1983.

[12] A. Faure, *Sur le théorème de Denjoy–Young–Saks*, C. R. Acad. Sci. Paris 320 (1995), 415–418.

[13] W. Gangbo and R. J. McCann, *The geometry of optimal transportation*, Acta Math. 177 (1996), 113–161.

[14] K. M. Garg, *Applications of Denjoy analogue. II: Local structure of level sets and Dini derivates. III: Distribution of various typical level sets*, Acta Math. Acad. Sci. Hungar. 14 (1963), 183–195.

[15] H. Heinich et J. C. Lootgieter, *Convergence des fonctions monotones*, C. R. Acad. Sci. Paris 322 (1996), 869–874.

[16] L. V. Kantorovich, *On the translocation of masses*, C. R. (Doklady) Acad. Sci. URSS (N.S.) 37 (1942), 199–201.

[17] —, *On a problem of Monge*, Uspekhi Mat. Nauk 3 (1948), 225–226 (in Russian).

[18] M. Knott and C. S. Smith, *On the optimal transportation of distributions*, J. Optim. Theory Appl. 52 (1987), 323–329.

[19] —, —, *On Hoeffding–Fréchet bounds and cyclic monotone relations*, J. Multivariate Anal. 40 (1992), 328–334.

[20] R. J. McCann, *Existence and uniqueness of monotone measure-preserving maps*, Duke Math. J. 80 (1995), 309–323.

[21] —, personal communication, 1996.

[22] G. Monge, *Mémoire sur la théorie des déblais et des remblais*, dans : Histoires de l'Académie Royale des Sciences de Paris, avec les mémoires de Mathématiques et de Physique pour la même année, 1781, 257–263.
[23] S. T. Rachev, *The Monge–Kantorovich mass transference problem and its stochastic applications*, Theory Probab. Appl. 29 (1984), 647–676.
[24] S. T. Rachev and L. Rüschendorf, *A characterization of random variables with minimum* l^2*-distance*, J. Multivariate Anal. 32 (1990), 48–54.
[25] —, —, *Mass Transportation Problems*, Springer, New York, 1998.
[26] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.
[27] V. A. Rohlin, *On the fundamental ideas of measure theory*, Amer. Math. Soc. Transl. 71 (1952), 1–54.
[28] L. Rüschendorf, *Fréchet bounds and their applications*, in: S. Kotz, G. Dall'Aglio and G. Salinetti (eds.), *Advances in Probability Distributions with Given Marginals*: *Beyond the Copulas*, Kluwer, Dordrecht, 1991, 141–176.
[29] —, *Optimal solutions of multivariate coupling problems*, Appl. Math. (Warsaw) 23 (1995), 325–338.
[30] S. Saks, *Sur les nombres dérivés des fonctions*, Fund. Math. 5 (1924), 98–104.
[31] A. Tuero-Diaz, *On the stochastic convergence of representations based on Wasserstein metrics*, Ann. Probab. 21 (1993), 72–85.
[32] L. Zajiček, *On the differentiation of convex functions in finite and infinite dimensional spaces*, Czechoslovak Math. J. 29 (1979), 340–348.
[33] —, *On the symmetry of Dini derivates of arbitrary functions*, Comment. Math. Univ. Carolin. 22 (1981), 195–209.

Nacereddine Belili
UPRES-A. CNRS 6085
Analyse et Modèles Stochastiques
Université de Rouen
76821 Mont Saint Aignan Cedex, France
E-mail: Nacereddine.Belili@univ-rouen.fr

Henri Heinich
UPRES-A. CNRS 6085, INSA de Rouen
Département de Génie Mathématiques
Place E. Blondel
76131 Mont Saint Aignan Cedex, France
E-mail: heinich@insa-rouen.fr

Received on 7.12.1998;
revised version on 13.4.1999