Indiana University Mathematics Journal ©
Article electronically published on February 18th, 2008

Young Measures, Superposition and Transport

PATRICK BERNARD

ABSTRACT. We discuss a space of Young measures in connection with some variational problems. We use it to present a proof of the Theorem of Tonelli on the existence of minimizing curves. We generalize a recent result of Ambrosio, Gigli and Savaré on the decomposition of the weak solutions of the transport equation. We also prove, in the context of Mather theory, the equality between Closed measures and Holonomic measures.

1. INTRODUCTION

It is by now a well understood fact that Young measures are a very useful tool in variational problems. In his book, Young exposes how the relaxation to appropriate spaces of Young measure allow to treat with great elegance the problem of length-minimizing curves. In the present paper, we present an extension of Young's approach to the non-parametric situation, and describe some applications. This provides a new proof of the theorem of Tonelli on the existence of curves minimizing a fiberwise convex action. The objects which appear in this program are related to some dynamical optimal transportation problems and to a variational approach of the Euler equation due to Arnold and Brenier, see [5, 7–11]. Our initial motivation has been to clarify our understanding of these objects.

We expose in Sections 2 and 3 the definition and main properties of the measures we will work with: Young measures, transport measures and generalized curves.

It should come as a reward and as an indication of the usefulness of this theory that we can provide in Section 4 a short and, we believe, elegant proof of the famous theorem of Tonelli on the existence of action-minimizing curves. We also underline the formal similarity between the problem of action minimizing curves and some dynamic optimal transportation problem as discussed in [7] and other papers. We obtain general existence results for these questions. In order to study the minimizing measures in a general framework, we need to pursue the study of transport measures.

This is what we do in Section 5, where we state, discuss and prove Theorem 5.2, which is certainly the most important result of the present paper. We call it Young's superposition principle for it is directly inspired by a result which appears in the appendix of Young's book. We propose some applications to the continuity equation and to the decomposition of optimal transport measures, that is to the full understanding of the relation between dynamic optimal transportations and action-minimizing curves. It should be noted that although Young's superposition principle is more general than another superposition principle recently obtained by Ambrosio, Gigli and Savaré in [3], many of its application to the study of transport measures minimizing the action defined by a fiberwise convex integrand could in fact be obtained from this especially important particular case.

In Section 6, we apply and adapt the ideas of Theorem 5.2 to study the closed measures which appear in Mather's theory of minimizing measures. In [15] Mather introduced and studied invariant measures of a Lagrangian system which minimize the action. These measures turn out to have a remarkable property. Later, Mañé introduced a class of probability measures, Holonomic measures, which contain the invariant measures of all Lagrangian flows, and which have the property that minimizing closed measures are invariant. Then Bangert introduced the larger class of closed measure and proved, for some specific Lagrangians, that minimizing closed measures are invariant. This was generalized by Fathi and Siconolfi to a much larger class of C^2 Lagrangians. Young's superposition principle allows to generalize these results to non-regular integrands (with the appropriate definition of invariance). We also prove that the holonomic measure of Mañé and the closed measures of Bangert are the same objects. We finish with some generalities of measure theory in the appendix.

I thank Alessio Figalli and Boris Buffoni for their help at different stages of the elaboration of the present work.

I finished to write this paper in De Giorgi center, Pisa. This was an occasion to visit the beautiful Camposanto. There, in a corner, under a scaffolding, is the sober grave of Leonida Tonelli, 1885-1946, Accademia dei Lincei.

2. YOUNG MEASURES

We define the space of Young measure we will use, and recall some general results on the topology of this space. Let (X, d) be a complete and separable metric space. We denote by $(\mathcal{P}_1(X), d)$ the Kantorovich-Rubinstein space of Borel probability measures on X with finite first moment, see the appendix. Recall that $(\mathcal{P}_1(X), d)$ is a complete and separable metric space. Let $I = [a, b]$ be a compact interval and let λ be the normalized Lebesgue measure on I. We denote by $\tilde{\mathcal{Y}}_1(I, X)$ the set of measurable maps

$$I \ni t \longmapsto \mu_t \in \mathcal{P}_1(X).$$

There is a natural map

$$\eta_t \longmapsto \lambda \otimes \eta_t$$

from $\tilde{\mathcal{Y}}_1(I,X)$ to $\mathcal{P}_1(I\times X)$, where we denote by $\lambda\otimes\eta_t$ the only measure which satisfies

$$\int_{I\times X} f(t,x)\,\mathrm{d}(\lambda\otimes\eta_t)(t,x)=\int_I\int_X f(t,x)\,\mathrm{d}\eta_t(x)\,\mathrm{d}\lambda(t)$$

for each bounded Borel function $f: I\times X\to\mathbb{R}$. The disintegration theorem states that the image of this map is the set $\mathcal{Y}_1(I,X)$ of probability measures $\eta\in\mathcal{P}_1(I\times X)$ whose marginal on the component I is the measure λ. We call these measures Young measures. Moreover, two elements of $\tilde{\mathcal{Y}}_1(I,X)$ have the same image if and only if they are almost everywhere equal. Note that $\mathcal{Y}_1(I,X)$ is a closed subset of the Kantorovich-Rubinstein space $\mathcal{P}_1(I\times X)$. We endow it from now on with the induced distance. The map

$$\eta\longmapsto\int_{I\times X} f(t,x)\,\mathrm{d}\eta \tag{2.1}$$

is continuous on $\mathcal{Y}_1(I,X)$ for all continuous function $f(t,x): I\times X\to\mathbb{R}$ such that $|f(t,x)|/(1+d(x_0,x))$ is bounded for some $x_0\in X$. This continuity holds for many more functions f.

Definition 2.1. A Caratheodory integrand is a Borel function $f(t,x): I\times X\to\mathbb{R}$ which is continuous in the second variable. A normal integrand is a Borel function $f(t,x): I\times X\to(-\infty,\infty]$ which is lower semi-continuous in the second variable.

Proposition 2.2. *The map (2.1) is continuous on $\mathcal{Y}_1(I,X)$ if f is a Caratheodory integrand such that $|f(t,x)|/(1+d(x_0,x))$ is bounded for some $x_0\in X$. It is lower semi-continuous if f is a normal integrand such that $f(t,x)/(1+d(x_0,x))$ is bounded from below.*

Proof. We follow [6, Lemma II.1.1, p. 142] for the first part. By the Scorza-Dragoni Theorem, (see [6, Theorem I.1.1, p. 132].) there exists a sequence J_n of compact subsets on I such that f is continuous on $J_n\times X$ and such that $\lambda(J_n)\to 1$ as $n\to\infty$. Then, there exists a sequence of continuous functions f_n such that $|f_n(t,x)|/(1+d(x_0,x))$ is bounded, independently of n, and such that $f_n=f$ on $J_n\times X$. It follows that the map (2.1) is the uniform limit of the continuous maps $\eta\longmapsto\int f_n\,\mathrm{d}\eta$, and therefore it is continuous.

In order to prove the second part of the statement, we first write the integrand $f(t,x)=(1+d(x_0,x))g(t,x)$ with a normal integrand g which is bounded from below. Then g is the increasing pointwise limit of a sequence g_n of bounded Caratheodory integrands, see [6, Theorem I.1.2, p. 138]. Finally, the map (2.1) is the increasing limit of the continuous maps

$$\eta\longmapsto\int(1+d(x_0,x))g_n(t,x)\,\mathrm{d}\eta(t,x),$$

and therefore it is lower semi-continuous. □

Theorem 2.3. *Let $f(t,x)$ be a normal integrand. Assume that there exists a proper function $\ell : X \to [0,\infty)$ and an integrable function $g : I \to \mathbb{R}$ such that $f(t,x) \geqslant \ell(x)(1+d(x,x_0)) + g(t)$. Then for each $C \in \mathbb{R}$, the set of Young measures $\eta \in \mathcal{Y}_1(I,X)$ which satisfy $\int f \,\mathrm{d}\eta \leqslant C$ is compact.*

Proof. Since the map $\eta \longmapsto \int f\,\mathrm{d}\eta$ is lower semi-continuous, it is enough to prove that the set of Young measures η which satisfy $\int \ell(x)\,\mathrm{d}(x,x_0)\,\mathrm{d}\eta \leqslant C$ is compact. This set is obviously 1-tight, see the Appendix. □

3. TRANSPORT MEASURES AND GENERALIZED CURVES

In the present section, we set $X = TM$, where M is a complete Riemannian manifold without boundary. We endow this tangent space TM with a complete distance d such that the quotient

$$\frac{1+d((x_0,0),(x,v))}{1+\|v\|_x}$$

and its inverse are bounded on TM for one (and then any) point $x_0 \in M$. The discussions below do not depend on the choice of this distance d. In order to prove that such a distance exists, we can isometrically embed M into a Euclidean space $\mathbb{R}^d$ and restrict the distance

$$D((x,v),(x',v')) = \min(1, |x'-x|) + |v'-v|,$$

where $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$. We fix a compact interval $I=[a,b]$. and denote by $C_1(I\times TM)$ the set of continuous functions $f : I\times TM \to \mathbb{R}$ such that

$$\|f\|_1 := \sup_{(t,x,v)\in I\times TM} \frac{\|f(t,x,v)\|}{1+\|v\|_x} < \infty.$$

Definition 3.1. A *transport measure* is a measure $\eta \in \mathcal{Y}_1(I,TM)$ which satisfies the relation

$$\int_{I\times TM} \partial_t g + \partial_x g\cdot v \,\mathrm{d}\eta(t,x,v) = 0 \tag{3.1}$$

for all smooth compactly supported functions $g :]a,b[\times M \to \mathbb{R}$. We denote by $\mathcal{T}(I,M) \subset \mathcal{Y}_1(I,TM)$ the set of all transport measures. Given two probability measures μ_i and μ_f on M, we say that the transport measure η is a transport measure between μ_i and μ_f if, in addition, we have

$$\int_{I\times TM} \partial_t g + \partial_x g\cdot v\,\mathrm{d}\eta(t,x,v) = \int_M g_b(x)\,\mathrm{d}\mu_f(x) - \int_M g_a(x)\,\mathrm{d}\mu_i(x)$$

for each smooth compactly supported function $g : [a,b] \times M \to \mathbb{R}$. We denote by $\mathcal{T}_{\mu_i}^{\mu_f}(I,M)$ the set of transport measures between μ_i and μ_f.

Note that $\mathcal{T}(I,M)$ and $\mathcal{T}_{\mu_i}^{\mu_f}(I,M)$ are closed subsets of $\mathcal{Y}_1(I,M)$. Recalling that we denote by $\mathcal{P}(M)$ the set of Borel probability measures endowed with the narrow topology, we have the following result:

Lemma 3.2. *Let $\eta \in \mathcal{T}(I,M)$ be a transport measure. There exists a continuous family $\mu_t : I \to \mathcal{P}(M)$ of probability measures on M and a disintegration $\eta_t \in \tilde{\mathcal{Y}}_1(I,TM)$ of η such that, for each t, μ_t is the marginal of η_t on the base M. We then have $\eta \in \mathcal{T}_{\mu_a}^{\mu_b}(I,M)$.*

Proof. Let us choose a disintegration η_t of η, and let $\tilde{\mu}_t$ be the marginal of η_t on M. We want to prove that there is a narrowly continuous map $\mu_t : I \to \mathcal{P}(M)$ which is equal to $\tilde{\mu}_t$ for almost each t. In view of general remarks recalled in the Appendix, it is enough to prove that, for each smooth and compactly supported function $f : M \to \mathbb{R}$, the function $t \longmapsto F(t) := \int f \, \mathrm{d}\mu_t$ is equal almost everywhere to a continuous function. By applying the equation (3.1) to functions $g(t,x) = \varphi(t)f(x)$, we get that $F'(t) = \int df_x \cdot v \, \mathrm{d}\eta_t(x,v)$ in the sense of distributions. It implies that the function F is equal almost everywhere to an absolutely continuous function.

Lemma 3.3. *Let $g(t,x) : I \times M \longmapsto \mathbb{R}$ be a C^1 bounded and Lipschitz function. Then for each interval $[\alpha,\beta] \subset [a,b]$, we have*

$$\int_{[\alpha,\beta]\times TM} \partial_t g + \partial_x g \cdot v \, \mathrm{d}\eta = \int_M g_\beta \, \mathrm{d}\mu_\beta - \int_M g_\alpha \, \mathrm{d}\mu_\alpha. \tag{3.2}$$

Proof. Let us first assume that g is a smooth compactly supported function. Let us set $F(t) = \int g_t \, \mathrm{d}\mu_t$. It is easy to prove using (3.1) that

$$F'(t) = \int_{TM} \partial_t g + \partial_x g \cdot v \, \mathrm{d}\eta_t$$

in the sense of distribution. The desired equality follows by integration. If g is C^1 and compactly supported, then we prove (3.2) by approximating g by smooth compactly supported functions. Let us expose a bit more carefully how the equality can be extended to bounded and Lipschitz functions which are not necessarily compactly supported. We consider an increasing sequence $\xi_n : M \to [0,1]$ of smooth equi-Lipschitz compactly supported functions such that, for each relatively compact open set U, we have $\xi_n = 1$ on U after a certain rank. Then (3.2) holds for the function $g\xi_n$:

$$\int_{[\alpha,\beta]\times TM} \xi_n \partial_t g + \xi_n \partial_x g \cdot v + g \partial_x \xi_n \cdot v \, \mathrm{d}\eta = \int_M g_\beta \xi_n \, \mathrm{d}\mu_\beta - \int_M g_\alpha \xi_n \, \mathrm{d}\mu_\alpha.$$

Thanks to the dominated convergence theorem, we get (3.2) at the limit. □

Definition 3.4. The transport measure η is called a generalized curve if μ_t is a dirac measure for each $t \in I$. Then, there exists a continuous curve $\gamma(t) : I \to M$ such that $\mu_t = \delta_{\gamma(t)}$ for each t. We say that η is a generalized curve above γ. We denote by $\mathcal{G}(I,M)$ the set of generalized curves.

A continuous curve $\gamma : I \to M$ is absolutely continuous if and only if the function $\varphi \circ \gamma : I \to \mathbb{R}$ is absolutely continuous for each smooth and compactly supported function $\varphi : M \to \mathbb{R}$. We denote by $W^{1,1}(I,M)$ the set of absolutely continuous curves. We say that a sequence γ_n is converging to γ in $W^{1,1}(I,M)$ if the function sequence

$$d\big((\gamma_n(t), \dot\gamma_n(t)), (\gamma(t), \dot\gamma(t))\big)$$

is converging to zero in L^1, or equivalently if the following three conditions are satisfied:

- The sequence γ_n is converging uniformly to γ.
- The sequence $(\gamma_n(t), \dot\gamma_n(t)) : I \to TM$ is converging in measure to $(\gamma(t), \dot\gamma(t))$.
- The sequence $\|\dot\gamma_n(t)\|_{\gamma_n(t)}$ is equi-integrable, or equivalently it is relatively weakly compact in $L^1(I,\mathbb{R})$.

It is well known that smooth curves are dense in $W^{1,1}(I,M)$.

Lemma 3.5. *Let $\Gamma \in \mathcal{T}(I,M)$ be a generalized curve. Then there exists an absolutely continuous curve $\gamma(t)$ such that Γ is a generalized curve above γ and there exists a measurable family Γ_t of probability measures on $T_{\gamma(t)}M$ such that $\Gamma = dt \otimes \delta_{\gamma(t)} \otimes \Gamma_t$, which means that*

$$\int_{I\times TM} f(t,x,v)\, \mathrm{d}\Gamma(t,x,v) = \int_I \int_{T_{\gamma(t)}M} f(t,\gamma(t),v)\, \mathrm{d}\Gamma_t(v)\, \mathrm{d}t$$

for each $f \in L^1(\Gamma)$. In order that this formula defines a generalized curve above the absolutely continuous curve γ, it is necessary and sufficient that the function $t \longmapsto \int_{T_{\gamma(t)}M} \|v\|_{\gamma(t)}\, \mathrm{d}\Gamma_t(v)$ is λ-integrable on I, and that $\int_{T_{\gamma(t)}M} v\, \mathrm{d}\Gamma_t(v) = \dot\gamma(t)$ for almost all t.

Proof. Let Γ be a generalized curve over γ. By the disintegration theorem, the measure Γ can be written in the form $\Gamma = dt \otimes \delta_{\gamma(t)} \otimes \Gamma_t$ with some measurable family Γ_t of probability measures on $T_{\gamma(t)}M$. We want to prove that the curve $\gamma(t)$ is absolutely continuous and that

$$\dot\gamma(t) = \int_{\mathbb{R}^d} v\, \mathrm{d}\Gamma_t(v)$$

for almost all t. It is enough to prove that, for each smooth compactly supported function $\varphi : M \to \mathbb{R}$, we have

$$(\varphi \circ \gamma)'(t) = d\varphi_{\gamma(t)} \cdot \int_{T_{\gamma(t)}M} v \, \mathrm{d}\Gamma_t(v)$$

in the sense of distributions. For each smooth compactly supported function $f(t) :]a, b[\to \mathbb{R}$ we can apply the equation (3.1) to the function $g(t, x) = f(t)\varphi(x)$, and get

$$\begin{aligned} 0 &= \int_{I\times TM} f'(t)\varphi(x) + f(t)\,\mathrm{d}\varphi_x \cdot v \,\mathrm{d}\Gamma(t,x,v) \\ &= \int_0^1 f'(t)\varphi(\gamma(t))\,\mathrm{d}t + \int_0^1 f(t) \int_{T_{\gamma(t)}M} \mathrm{d}\varphi_{\gamma(t)} \cdot v \,\mathrm{d}\Gamma_t(v)\,\mathrm{d}t. \end{aligned}$$

This implies that $\varphi \circ \gamma$ is absolutely continuous and that

$$(\varphi \circ \gamma)'(t) = \int_{T_{\gamma(t)}M} \mathrm{d}\varphi_{\gamma(t)} \cdot v \,\mathrm{d}\Gamma_t(v) = d\varphi_{\gamma(t)} \cdot \int_{T_{\gamma(t)}M} v \,\mathrm{d}\Gamma_t(v)$$

which is the desired result. □

Theorem 3.6. *The set $\mathcal{G}(I, M)$ of generalized curves is closed in $\mathcal{Y}_1(I, TM)$. In addition, the map $\mathcal{G} \to C^0(I, \mathbb{R}^d)$ which, to a generalized curve Γ above γ, associates the curve γ, is continuous.*

Proof. Let Γ_n be a sequence of generalized curves converging in $\mathcal{P}_1(I \times TM)$ to a limit η. We have to prove that η is a generalized curve. The family η, Γ_1, $\Gamma_2, \ldots, \Gamma_n, \ldots$ is compact in $\mathcal{P}_1(I \times TM)$, hence it has uniformly integrable first moment. This implies that the sequence γ_n of associated curves is equi-absolutely continuous. Taking a subsequence, we can assume that the sequence γ_n has a limit γ in $C^0(I, M)$. It is not hard to check, then, that η is a generalized curve above γ. □

If $\gamma : I \to M$ is absolutely continuous, then we will denote by $\bar{\gamma}$ the generalized curve above γ given by

$$\int_X f(t, x, v)\,\mathrm{d}\bar{\gamma}(t, x, v) = \int_I f(t, \gamma(t), \dot{\gamma}(t))\,\mathrm{d}t$$

for each bounded Borel function f. In other words, we have

$$\bar{\gamma} = dt \otimes \delta_{\gamma(t)} \otimes \delta_{\dot{\gamma}(t)}.$$

We denote by $C(I, M) \subset \mathcal{T}(I, M)$ the set of transport measures which are of that form.

Lemma 3.7. *The map*

$$W^{1,1}(I,M) \to \mathcal{G}(I,M)\,, \qquad \gamma \longmapsto \bar{\gamma}$$

is continuous.

Proof. Let $\gamma_n \in W^{1,1}(I,M)$ be a sequence which converges to γ. We have to prove that

$$\int_I f(t,\gamma_n(t),\dot\gamma_n(t))\,\mathrm{d}\lambda \to \int_I f(t,\gamma(t),\dot\gamma(t))\,\mathrm{d}\lambda$$

for each $f \in C_1(I \times TM)$. Since the sequence $(\gamma_n(t),\dot\gamma_n(t))$ is converging in measure to $(\gamma(t),\dot\gamma(t))$, we can suppose by extracting a subsequence that it is converging almost everywhere. The desired convergence follows from the observation that the sequence of real functions

$$t \longmapsto f(t,\gamma_n(t),\dot\gamma_n(t))$$

is converging almost everywhere to $f(t,\gamma(t),\dot\gamma(t))$ and is equi-integrable because

$$|f(t,\gamma_n(t),\dot\gamma_n(t))| \leqslant \|f\|_1(1+\|\dot\gamma_n(t)\|_{\gamma_n(t)})$$

and, by definition of the convergence in $W^{1,1}$, the sequence $\|\dot\gamma_n(t)\|_{\gamma_n(t)}$ is equi-integrable. □

Let us mention, for completeness, the following result:

Theorem 3.8. *The set $\mathcal{G}(I,M)$ of generalized curves is the closure, in $\mathcal{Y}_1(I,TM)$, of the set $C(I,M)$ of curves.*

4. TONELLI THEOREM AND OPTIMAL TRANSPORTATION

In the present section, we use transport measures and generalized curves to expose some results on the existence of certain minimizers. The results are well known, but the presentation is somewhat original. We consider a normal integrand $L : [a,b] \times TM \to \mathbb{R} \cup \{+\infty\}$. We say that L is fiberwise convex if, for each fixed (t,x), the function $v \longmapsto L(t,x,v)$ is convex on T_xM. The role of convexity in minimization problems is enlightened by the following standard observation:

Lemma 4.1. *Let L be a fiberwise convex normal integrand. If Γ is a generalized curve above γ, then*

$$\int L\,\mathrm{d}\Gamma \geqslant \int_a^b L(t,\gamma(t),\dot\gamma(t))\,\mathrm{d}t = \int L\,\mathrm{d}\bar\gamma.$$

Proof. For each t, we have

$$\int_{T_{\gamma(t)}M} L(t, \gamma(t), v)\, \mathrm{d}\Gamma_t(v) \geqslant L(t, \gamma(t), \dot{\gamma}(t))$$

by Jensen's inequality. We obtain

$$\int L\, \mathrm{d}\Gamma = \int_0^1 \int_{T_{\gamma(t)}M} L(t, \gamma(t), v)\, \mathrm{d}\Gamma_t(v)\, \mathrm{d}t$$
$$\geqslant \int_0^1 L(t, \gamma(t), \dot{\gamma}(t))\, \mathrm{d}t = \int L\, \mathrm{d}\bar{\gamma}$$ □

We now discuss the classical problem of the existence of minimizing curves. We fix two points x_i and x_f in M, and consider the set $AC_{x_i}^{x_f}$ of absolutely continuous curves $\gamma : I \to M$ such that $\gamma(a) = x_i$ and $\gamma(b) = x_f$. We also consider the set

$$\mathcal{G}_{x_i}^{x_f} = \mathcal{G}(I, M) \cap \mathcal{T}_{\delta_{x_i}}^{\delta_{x_f}}(I, M)$$

of generalized curves above elements of $AC_{x_i}^{x_f}$. Note that $\mathcal{G}_{x_i}^{x_f}$ is closed in $\mathcal{T}(I, M)$. The action of an absolutely continuous curve γ is the integral

$$\int_a^b L(t, \gamma(t), \dot{\gamma(t)})\, \mathrm{d}t,$$

the action of a transport measure η is the integral $\int_{I\times TM} L\, \mathrm{d}\eta$. The following result is well known:

Theorem 4.2. *Let $L(t, x, v) : [a, b] \times TM \to \mathbb{R} \cup +\infty$ be a normal integrand. We assume that the integrand L satisfies:*

(L1) *the quotient*

$$\frac{L(t, x, v)}{1 + \|v\|_x}$$

is bounded from below and proper.

For each $C \in \mathbb{R}$ the set

$$\mathcal{A}_C^g := \left\{ \Gamma \in \mathcal{G}_{x_i}^{x_f} \mid \int L\, \mathrm{d}\Gamma \leqslant C \right\} \subset \mathcal{G}_{x_i}^{x_f}$$

is compact, and if L is fiberwise convex, the set

$$\mathcal{A}_C := \left\{ \gamma \in AC_{x_i}^{x_f} \mid \int_a^b L(t, \gamma(t), \dot{\gamma}(t)) \leqslant C \right\} \subset C(I, M)$$

is compact for the uniform topology.

As a major consequence, we obtain that the action reaches its minimum on $\mathcal{G}_{x_i}^{x_f}$ if there exists a generalized curve of finite action in $\mathcal{G}_{x_i}^{x_f}$. If in addition the integrand is fiberwise convex, then the action also reaches its minimum on $AC_{x_i}^{x_f}$, and we have

$$\min_{\Gamma\in\mathcal{G}_{x_i}^{x_f}} \int L\,\mathrm{d}\gamma = \min_{\gamma\in AC_{x_i}^{x_f}} \int_a^b L(t,\gamma(t),\dot\gamma(t))\,\mathrm{d}t.$$

Proof. The compactness of $\mathcal{A}_C^g$ follows from Theorem 2.3. If L is fiberwise convex, then, by Lemma 4.1, the set $\mathcal{A}_C$ is the image of the compact set $\mathcal{A}_C^g$ by the continuous map $\Gamma \longmapsto \gamma$ (the map which, to a generalized curves Γ above γ, associates the curve γ). ◻

In applications, it is useful to have the following stronger and still standard result:

Theorem 4.3 (Tonelli). *The same conclusions* (*as Theorem 4.2*) *hold if the hypothesis* (L1) *on the integrand is replaced by the two following ones:*

(L2) *The integrand L is uniformly superlinear over each compact subset of M. It means that, for each compact $K \subset M$, there exists a function $\ell : \mathbb{R}^+ \to \mathbb{R}$ such that* $\lim_{r\to\infty} \ell(r)/r = \infty$ *and such that* $L(t,x,v) \geqslant \ell(\|v\|_x)$ *for each* $(t,x,v) \in [a,b]\times T_K M$.

(L3) *There exists a positive constant c such that* $L(t,x,v) \geqslant c(\|v\|_x - 1)$.

Proof. We have to prove that the set of generalized curves $\Gamma \in \mathcal{G}_{x_i}^{x_f}$ which satisfy $\int L\,\mathrm{d}\Gamma \leqslant C$ is compact. Using (L3), we see that, if Γ is a generalized curve over γ, then

$$\int_a^b \|\dot\gamma(t)\|_{\gamma(t)}\,\mathrm{d}t \leqslant \frac{C+(b-a)}{c}.$$

Let K be the closed ball (for the Riemaniann distance on M) of center x_i and radius $(C+b-a)/c$. This ball is compact because M is complete. Let us define a modified integrand L_K by $L_K(t,x,v) = L(t,x,v)$ if $x\in K$ and $L_K(t,x,v) = +\infty$ if $x \notin K$. A generalized curve $\Gamma \in \mathcal{G}_{x_i}^{x_f}$ satisfies $\int L\,\mathrm{d}\Gamma \leqslant C$ if and only if it satisfies $\int L_K\,\mathrm{d}\Gamma \leqslant C$. Since L_K satisfies (L1), we conclude by Theorem 4.2. ◻

We can extend these considerations to more general boundary conditions. Our presentation allows to see the following dynamic optimal transportation problem as a natural generalisation of Tonelli theorem.

Theorem 4.4. *Let L be a normal integrand which satisfies:*

(L4) *The integrand L is uniformly superlinear: there exists a function $\ell : \mathbb{R}^+ \to \mathbb{R}$ such that* $\lim_{r\to\infty} \ell(r)/r = \infty$ *and such that* $L(t,x,v) \geqslant \ell(\|v\|_x)$ *for each* $(t,x,v) \in [a,b]\times TM$.

Let μ_i and μ_f be two Borel probability measures on M. Then for each $C \in \mathbb{R}$, the set $\mathcal{B}_C$ of transport measures $\eta \in \mathcal{T}_{\mu_i}^{\mu_f}$ which satisfy $\int L \,\mathrm{d}\eta \leqslant C$ is compact.

Note that (L4) implies (L2) and (L3).

Proof. The conclusion would be obvious if L satisfied (L1), but (L4) is weaker. For each $\varepsilon > 0$, there exists a constant R such that

$$\int_{\|v\|_x \geqslant R} (1 + \|v\|_x) \,\mathrm{d}\eta(t, x, v) \leqslant \varepsilon$$

for each $\eta \in \mathcal{B}_C$. This is a direct consequence of (L4). We claim that there exists a compact ball $B \subset M$ such that $\eta(I \times T_B M) \geqslant 1 - \varepsilon/(1 + R)$ for each $\eta \in \mathcal{B}_C$. Assuming the claim, we have

$$\int_{\{(t,x,v) \in I \times TM \mid x \notin B \text{ or } \|v\|_x \geqslant R\}} (1 + \|v\|_x) \,\mathrm{d}\eta(t, x, v) \leqslant 2\varepsilon$$

for each $\eta \in \mathcal{B}_C$. Therefore, $\mathcal{B}_C$ is 1-tight and thus compact. Let us now prove the claim. For each $\Delta > 0$, there exists a C^1, bounded and 1-Lipschitz function $g : M \to [0, \Delta]$ such that $g = \Delta$ outside of a compact ball B and such that $\int g \,\mathrm{d}\mu_i \leqslant 1$. For $\eta \in \mathcal{B}_C$, we have

$$\int_M g \,\mathrm{d}\mu_t = \int_M g \,\mathrm{d}\mu_i + \int_{[a,t] \times TM} \mathrm{d}g_x \cdot v \,\mathrm{d}\eta \leqslant 1 + \frac{C + b - a}{c}$$

where c is the constant of (L3). We conclude that

$$\int g \,\mathrm{d}\eta = \int_a^b \int_M g \,\mathrm{d}\mu_t \,\mathrm{d}t \leqslant (b - a)\left(1 + \frac{C + b - a}{c}\right)$$

for each $\eta \in \mathcal{B}_C$. It follows that

$$\eta(I \times TM - I \times T_B M) \leqslant \frac{(b - a)(1 + (C + b - a)/c)}{\Delta}.$$

Since Δ can be chosen arbitrarily, the claim is proved. □

Some general comments are needed before we can describe the additional conclusions satisfied for fiberwise convex Lagrangians. If $\eta \in \mathcal{Y}_1(I, TM)$ is a Young measure, then we call μ the image of η by the projection $I \times TM \to I \times M$. We can desintegrate η with respect to this projection and obtain a measurable family $\eta_{t,x}$ of probability measures on $T_x M$ such that $\eta = \mu \otimes \eta_{t,x}$. We define the vector-field $V(t, x) : I \times M \to TM$ by the expression

$$V(t, x) = \int_{T_x M} v \,\mathrm{d}\eta_{t,x}(v).$$

Note that $V(t,x)$ is a Borel time-dependant vector-field, and that the integrability condition

$$\int \|V(t,x)\|_x \,\mathrm{d}\mu(t,x) < \infty$$

is satisfied.

Lemma 4.5. *The Young measure $\eta \in \mathcal{Y}_1(I, TM)$ is a transport measure if and only if the continuity equation*

$$\text{(PDE)} \qquad \partial_t \mu + \operatorname{div}(V\mu) = 0.$$

holds in the sense of distributions.

The couple (V,μ) is what we called in [7] the transport current associated to the transport measure η. Such objects were previously introduced by Benamou and Brenier, see [5], [10] and [11].

Proof. A test function is a smooth and compactly supported function on $]a,b[\times M$. The measure η is a transport measure if and only if

$$\int_{I\times M}\int_{T_xM} \partial_t g(t,x) + \partial_x g(t,x)\cdot v \,\mathrm{d}\eta_{t,x}(v)\,\mathrm{d}\mu(t,x) = 0$$

for each test function. The equation (PDE) holds in the sense of distributions if and only if

$$\int_{I\times M} \partial_t g(t,x) + \partial_x g(t,x)\cdot V(t,x)\,\mathrm{d}\mu(t,x) = 0$$

for each test function g. The equivalence follows from the observation that

$$\int_{T_xM} \partial_x g(t,x)\cdot v \,\mathrm{d}\eta_{t,x} = \partial_x g(t,x)\cdot \int_{T_xM} v\,\mathrm{d}\eta_{t,x} = \partial_x g(t,x)\cdot V(t,x)$$

by definition of V. ∎

Conversely, consider a Borel vector-field $V(t,x): I\times M \to TM$ and a probability measure μ on $I\times M$ whose marginal on I is λ. Assume that (PDE) holds and that the integrability condition $\int \|V\|\,\mathrm{d}\mu < \infty$ is satisfied. Then, the measure $\tilde{V}_\sharp\mu$ is a transport measure, where $\tilde{V}(t,x) = (t, V(t,x)) \in I\times TM$. The following generalization of Lemma 4.1 is now obvious:

Lemma 4.6. *Let L be a fiberwise convex normal integrand. If η is a transport measure, and μ and V are associated to it as above, then $\tilde{V}_\sharp\mu$ is a transport measure, and*

$$\int L\,\mathrm{d}\eta \geqslant \int L\,\mathrm{d}(\tilde{V}_\sharp\mu).$$

As a consequence, if there exists a transport measure minimizing the action in $\mathcal{T}_{\mu_i}^{\mu_f}(I, M)$*, then there exists a minimizing transport measure in* $\mathcal{T}_{\mu_i}^{\mu_f}(I, M)$ *which is concentrated on the graph of a Borel vector-field.*

5. The Superposition Principle

The main stream of this section consists of writing transport measures as superpositions of generalized curves. This is the adaptation to the non-parametric setting of a theory sketch in the appendix of Young's book.

5.1. Young's superposition principle. We first adapt an important result of Young:

Theorem 5.1 (Young)**.** *The set* $\mathcal{T}(I, \mathbb{R}^d)$ *of transport measures is the closed convex envelop in* $\mathcal{Y}_1(I, \mathbb{R}^{2d})$ *of the set* $C(I, \mathbb{R}^d)$ *of curves (and hence also of the set* $\mathcal{G}(I, \mathbb{R}^d)$ *of generalized curves).*

Let us immediately mention the restatement of Young's result, which we will use:

Theorem 5.2. *If* η *is a transport measure on a complete manifold* M*, then there exists a Borel measure* ν *on* $\mathcal{G}(I, M)$ *such that* $\eta = \int_{\mathcal{G}} \Gamma \, \mathrm{d}\nu(\Gamma)$*, which means that*

$$\int_{I \times TM} f \, \mathrm{d}\eta = \int_{\mathcal{G}} \int_{I \times TM} f \, \mathrm{d}\Gamma \, \mathrm{d}\nu(\Gamma) \tag{5.1}$$

for each function $f \in L^1(\eta)$*. We then say that* ν *is a decomposition of* η*.*

Let us make a few simple remarks before proving these results.

Proposition 5.3. *If* η *is concentrated on the Borel subset* $Y \subset I \times TM$*, and if* ν *is a decomposition of* η*, then* ν*-almost every generalized curve* Γ *is concentrated on* Y*.*

Proof. Apply (5.1) with $f = 0$ on Y and $f = 1$ outside of Y. We get

$$\int_{I \times TM} f \, \mathrm{d}\Gamma = 0$$

for ν-almost all Γ, which means that Γ is concentrated on Y. □

For each $t \in I$, let $ev_t : \mathcal{G}(I, M) \to M$ be the continuous map obtained by composing the natural projection $\mathcal{G} \to C^0(I, M)$ and the evaluation map $\gamma \longmapsto \gamma(t)$.

Proposition 5.4. *If* ν *is a decomposition of* η*, and if* μ_t *is the continuous family of probability measures on* M *associated to* η*, then* $\mu_t = (ev_t)_\sharp \nu$*,* *Proof.* We denote by γ_Γ the continuous curve associated to the generalized curve Γ. It is enough to prove that

$$\int f(t,x)\,\mathrm{d}\eta = \int_{\mathcal{G}} f(t,\gamma_\Gamma(t))\,\mathrm{d}t\,\mathrm{d}\nu(\Gamma)$$

for each continuous and bounded function $f : I \times M \to \mathbb{R}$. This follows from the fact that

$$\int_X f(t,x)\,\mathrm{d}\Gamma(t,x,v) = \int_0^1 f(t,\gamma_\Gamma(t))\,\mathrm{d}t$$

for each generalized curve Γ. □

Finally, let us explain how Theorem 5.2 follows from Theorem 5.1. We isometrically embed the manifold M as a closed subset of some Euclidean space $\mathbb{R}^d$. Then the transport measures and generalized curves on M are just the transport measures and generalized curves on $\mathbb{R}^d$ which are supported on $I \times TM \subset I \times \mathbb{R}^{2d}$. Let η be a transport measure. In view of Young's theorem and of the appendix, η admits a decomposition ν by generalized curves on $\mathbb{R}^d$. By Proposition 5.3 above, ν almost every generalized curve Γ is supported on $I \times TM$, hence ν can be seen as a probability measure on $\mathcal{G}(I,M)$.

5.2. Proof of Young's superposition principle. We prove the superposition principle by duality, following the sketch of proof proposed by Young in his book. By Proposition B.5 of the appendix, it is enough to prove that, for each function $f \in C_1(I \times \mathbb{R}^{2d})$ such that

$$\int_0^1 f(t,\gamma(t),\dot\gamma(t))\,\mathrm{d}t \geqslant 0 \quad \forall\, \gamma \in W^{1,1}(I,\mathbb{R}^d) \tag{5.2}$$

we have $\int f\,\mathrm{d}\eta \geqslant 0$ for all transport measures $\eta \in \mathcal{T}(I,\mathbb{R}^d)$. It is sufficient to obtain the conclusion for functions $f \in C_1(I \times TM)$ which satisfy

$$\int_0^1 f(t,\gamma(t),\dot\gamma(t))\,\mathrm{d}t \geqslant 1 \quad \forall\, \gamma \in W^{1,1}(I,\mathbb{R}^d). \tag{5.3}$$

Indeed, if this is proved, and if f satisfies (5.2), then for each $\varepsilon > 0$, the function $(f+\varepsilon)/\varepsilon$ satisfies (5.3), hence $\int f\,\mathrm{d}\eta \geqslant -\varepsilon$ for each transport measure η, and finally $\int f\,\mathrm{d}\eta \geqslant 0$.

Let us fix a function $f \in C_1(I \times \mathbb{R}^{2d})$, assume (5.3), and define the value function $u : I \times \mathbb{R}^d \to \mathbb{R}$ by

$$u(t,x) := \inf_{\gamma \in W^{1,1}(\mathbb{R},\mathbb{R}^d),\, \gamma(t)=x} \int_0^t f(s,\gamma(s),\dot\gamma(s))\,\mathrm{d}s.$$

We have the equality

$$u(t,x) = \inf_{\gamma\in W^{1,1}(\mathbb{R},\mathbb{R}^d),\, \gamma(t)=x} u(s,\gamma(s)) + \int_s^t f(\sigma,\gamma(\sigma),\dot\gamma(\sigma))\,\mathrm{d}\sigma$$

for each $s \leqslant t$ and each x. This equality is called the dynamic programming principle.

Lemma 5.5. *We have*

$$u(t,y) \leqslant u(s,x) + \|f\|_1((t-s) + |y-x|)$$

for each $s \leqslant t$ in I and each x, y in $\mathbb{R}^d$.

Proof. Just observe that

$$u(t,y) \leqslant u(s,x) + \int_s^t f\left(\sigma,\ x+\sigma\,\frac{y-x}{t-s},\ \frac{y-x}{t-s}\right)\mathrm{d}\sigma$$
$$\leqslant u(s,x) + (t-s)\|f\|_1\left(1+\frac{|y-x|}{t-s}\right). \qquad \square$$

Lemma 5.6. *The value function u is bounded and upper semi-continuous. In addition, we have $u(0,x) = 0$ and $u(1,x) \geqslant 1$ for all x.*

Proof. The inequality $u(1,x) \geqslant 1$ follows from (5.3). For each $\gamma \in W^{1,1}(\mathbb{R},\mathbb{R}^d)$, let us consider the function

$$u_\gamma(t,x) := \int_0^t f(s,\gamma(s)+x-\gamma(t),\dot\gamma(s))\,\mathrm{d}s$$

which is continuous and bounded. Observing that $u = \inf_{\gamma\in W^{1,1}(\mathbb{R},\mathbb{R}^d)} u_\gamma$, we conclude that the function u is upper semi-continuous and bounded from above. It follows from Lemma 5.5 that $u(t,x) \geqslant u(1,x)+\|f\|_1(t-1) \geqslant 1+\|f\|_1(t-1)$ is bounded from below. $\square$

Lemma 5.7. *There exist sequences $u_n : I\times\mathbb{R}^d \to \mathbb{R}$ and $f_n : I\times\mathbb{R}^{2d}\to\mathbb{R}$ of functions such that:*

- *The sequence f_n is bounded in $C_1(I\times\mathbb{R}^{2d})$ and $f_n \to f$ pointwise.*
- *The functions u_n are smooth, bounded and Lipschitz. They satisfy*

$$u_n(0,x) = u_n(1,x) = 0 \quad \text{for all } n \text{ and all } x.$$

- *The inequality*

$$u_n(t,\gamma(t)) - u_n(s,\gamma(s)) \leqslant \int_s^t f_n(\sigma,\gamma(\sigma),\dot\gamma(\sigma))\,\mathrm{d}\sigma$$

holds for each $s \leqslant t$ in $\mathbb{R}$ and each absolutely continuous curve $\gamma:\mathbb{R}\to\mathbb{R}^d$.

Proof. There exists $\delta > 0$ such that $u(t,x) < \frac{1}{2}$ when $t \leqslant a + \delta$ and $u(t,x) > \frac{1}{2}$ when $t \geqslant b - \delta$. It is convenient to consider the function $\tilde{f}_n : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ which is equal to f on $[a+2/n,\ b-2/n] \times \mathbb{R}^d \times \mathbb{R}^d$ and to 0 outside of this set, and the function $\tilde{u}_n : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ which is equal to $u - \frac{1}{2}$ on $[a+2/n,\ b-2/n] \times \mathbb{R}^d$, to 0 outside of this set. Note that

$$\tilde{u}_n(t,\gamma(t)) - \tilde{u}_n(s,\gamma(s)) \leqslant \int_s^t \tilde{f}_n(\sigma,\gamma(\sigma),\dot{\gamma}(\sigma))\,\mathrm{d}\sigma$$

for each n, each $s \leqslant t$ in $\mathbb{R}$ and each absolutely continuous curve $\gamma : \mathbb{R} \to \mathbb{R}^d$.

Let $\rho_n(t,x) : \mathbb{R} \times \mathbb{R}^d \to [0,\infty)$ be a sequence of convolution kernels, that is of smooth non-negative functions such that $\int_{\mathbb{R}\times\mathbb{R}^d} \rho_n(t,x)\,\mathrm{d}x\,\mathrm{d}t = 1$ and such that ρ_n is supported on the ball of center 0 and radius $1/n$. Let us define the functions $u_n = \rho_n * \tilde{u}_n : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$:

$$u_n(t,x) = \int_{\mathbb{R}\times\mathbb{R}^d} \tilde{u}_n(t-\sigma,\ x-y)\rho_n(\sigma,y)\,\mathrm{d}\sigma\,\mathrm{d}y,$$

and $f_n : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ by

$$f_n(t,x,v) = \int_{\mathbb{R}\times\mathbb{R}^d} \tilde{f}_n(t-\sigma,\ x-y,\ v)\rho_n(\sigma,y)\,\mathrm{d}\sigma\,\mathrm{d}y.$$

For each fixed curve γ and each n, the inequality

$$\tilde{u}_n(t-\sigma,\ \gamma(t-\sigma)-y) - \tilde{u}_n(s-\sigma,\ \gamma(s-\sigma)-y) \leqslant \int_s^t \tilde{f}_n(\zeta-\sigma,\ \gamma(\zeta-\sigma)-y,\ \dot{\gamma}(\zeta-\sigma))\,\mathrm{d}\sigma$$

holds for each (σ,y), and then the third point of the lemma is obtained by integration. □

Let η be a transport measure. We want to prove that $\int f\,\mathrm{d}\eta \geqslant 0$. Let us set

$$h_n(t,x,v) := f_n(t,x,v) - \partial_t u_n(t,x) - \partial_x u_n(t,x)\cdot v$$

in such a way that

$$\int_s^t h_n(\sigma,\gamma(\sigma),\dot{\gamma}(\sigma))\,\mathrm{d}\sigma \geqslant 0$$

for all absolutely continuous curves γ and all $s \leqslant t$ in $\mathbb{R}$. We deduce that h_n is a non-negative function, and then $\int h_n\,\mathrm{d}\eta \geqslant 0$. We have the equality (3.2) for u_n:

$$\int_{I\times\mathbb{R}^{2d}} \partial_t u_n(t,x) + \partial_x u_n(t,x)\cdot v\,\mathrm{d}\eta(t,x,v) = 0$$

which implies that $\int f_n \, d\eta = \int h_n \, d\eta \geqslant 0$. At the limit $n \to \infty$, we conclude that $\int f \, d\eta \geqslant 0$. This ends the proof of Theorem 5.1. ◻

5.3. Application to the continuity equation. Young's superposition principle implies elegant results of Ambrosio, Gigli and Savaré concerning the continuity equation

(PDE) $$\partial_t \mu + \operatorname{div}(V\mu) = 0.$$

It is well known that close relations exist between (PDE) and the following

(ODE) $$\dot{\gamma}(t) = V(t, \gamma(t)).$$

More precisely, if $\gamma(t)$ is an absolutely continuous solution of (ODE), then $\mu := dt \otimes \delta_{\gamma(t)}$ is a weak solution of (PDE). We call these solutions elementary. The relations between (PDE) and (ODE) are enlightened by the following result, which was obtained by Ambrosio, Gigli and Savaré [2, 3], in the line of anterior works of Smirnov [18] and Bangert [4]:

Theorem 5.8 (Ambrosio, Gigli, Savaré). *Let $V : I \times M \to TM$ be a Borel time-dependant vector field. Every probability measure μ on $I \times M$ which solves (PDE) in the sense of distributions and satisfies the integrability condition*

$$\int \|V(t,x)\|_x \, d\mu < \infty$$

is a superposition of elementary solutions. More precisely, there exists a Borel probability measure ν on $\mathcal{G}(I, M)$ such that $\mu = dt \otimes (ev_t)_\sharp \nu$, and ν-almost every generalized curve is a curve and is a solution of (ODE).

Proof. In order to see the relation between this result and Young's superposition principle, observe that weak solutions of (PDE) are in bijection with transport measures which are concentrated on the graph of V (which is a Borel subset of X). More precisely, if η is such a transport measure, then its marginal μ on $[0, 1] \times \mathbb{R}^d$ is a weak solution of (PDE). Conversely, if μ is a solution of (PDE), then its lifting to the graph of V is a transport measure. Now the transport measure η associated to the solution μ can be written as a superposition of generalized curves which are concentrated on the graph of V. But it is obvious that a generalized curve which is concentrated on the graph of V is nothing but an absolutely continuous solution of (ODE). ◻

Note that the result can be applied in $\mathbb{R}^d$ endowed with the complete metric

$$g_x(v, w) = \frac{\langle v, w \rangle}{(1 + |x|)^2}.$$

The integrability condition then reads

$$\int_{I\times\mathbb{R}^d} \frac{|V(t,x)|}{1+|x|}\,\mathrm{d}\mu(t,x) < \infty$$

as in [2].

5.4. Application to optimal transport. Let L be a normal integrand. A generalized curve Γ is called minimizing if it is minimizing the action with fixed boundary points. If η is minimizing the action in $\mathcal{T}_{\mu_i}^{\mu_f}(I,M)$, then η can be decomposed into minimizing generalized curves. The decompositions ν of η are minimizing the action

$$\int_{\mathcal{G}(I,M)} \int_{I\times TM} L\,\mathrm{d}\Gamma\,\mathrm{d}\nu(\gamma)$$

on the set of probability measures ν on $\mathcal{P}_1(M)$ such that $(ev_a)_\sharp\nu = \mu_i$ and $(ev_b)_\sharp\nu = \mu_f$.

If in addition the integrand L is fiberwise convex, and if there exists a minimizing transport measure η in $\mathcal{T}_{\mu_i}^{\mu_f}(I,M)$, then there exists a minimizing transport measure in $\mathcal{T}_{\mu_i}^{\mu_f}(I,M)$ which is concentrated on the graph of a Borel vector-field $V(t,x)$. This minimizing measure can be decomposed into minimizing curves which are solutions of (ODE).

6. HOLONOMIC AND CLOSED MEASURES

In the theory of Mather minimizing measures, several spaces of measures were introduced on $\mathbb{T}\times TM$. In order to be coherent with the exposition of the rest of the present paper, we shall view them, in an equivalent way, as transport measures in $\mathcal{T}([0,1],M)$.

6.1. Closed measures. They have been used in the context of Lagrangian dynamics by Bangert in [4].

Definition 6.1. A measure $\eta \in \mathcal{T}([0,1],M)$ is called closed if there exists a probability measure μ on M such that $\eta \in \mathcal{T}_\mu^\mu([0,1],M)$. We denote by $\mathcal{F}(M)$ the set of closed measures, so that

$$\mathcal{F}(M) = \bigcup_{\mu\in\mathcal{P}(M)} \mathcal{T}_\mu^\mu([0,1],M) \subset \mathcal{T}([0,1],M).$$

We now expose a superposition principle for closed measures in the spirit of Smirnov [18], Bangert [4] and De Pascal, Gelli and Granieri [12]. Let us first define the set $\mathcal{G}(\mathbb{R},M)$ of measures Γ on $\mathbb{R}\times TM$ such that, for each $[a,b]\subset\mathbb{R}$, the rescaled restriction

$$\Gamma_{[a,b]} := \Gamma\big|_{[a,b]\times TM}/(b-a)$$

is a generalized curve in $\mathcal{G}([a,b],M)$. Denoting by d^k the distance on $\mathcal{G}([-k,k],M)$, we have a distance

$$d(\Gamma,\Gamma') = \sum_{k=1}^{\infty} \frac{d^k(\Gamma_{[-k,k]},\Gamma'_{[-k,k]})}{2^k}$$

on $\mathcal{G}(\mathbb{R},M)$. Clearly, a sequence Γ^n of elements of $\mathcal{G}(\mathbb{R},M)$ is converging to Γ if and only if we have $\Gamma^n_{[a,b]} \to \Gamma_{[a,b]}$ for each $[a,b] \subset \mathbb{R}$. It is not hard to check that $\mathcal{G}(\mathbb{R},M)$ is a complete and separable metric space. Let $\tau : \mathbb{R} \times TM \to \mathbb{R} \times TM$ be the translation $(t,x,v) \longmapsto (t+1,x,v)$. The map $\tau_\sharp : \mathcal{G}(\mathbb{R},M) \to \mathcal{G}(\mathbb{R},M)$ is continuous. Consequently, the map

$$\tau_{\sharp\sharp} : \mathcal{P}(\mathcal{G}(\mathbb{R},M)) \to \mathcal{P}(\mathcal{G}(\mathbb{R},M))$$

is continuous. A probability measure ν on $\mathcal{G}(\mathbb{R},M)$ is called translation invariant if $\tau_{\sharp\sharp}\nu = \nu$.

For each compact time interval I, we denote by

$$P_I : \mathcal{G}(\mathbb{R},M) \to \mathcal{G}(I,M)$$

the map $\Gamma \longmapsto \Gamma_I$. The Borel σ-algebra of $\mathcal{G}(\mathbb{R},M)$ is also the σ-algebra induced by the projections P_I, $I \subset \mathbb{R}$.

Theorem 6.2. *If η is a closed measure on M, then there exists a translation-invariant probability measure ϑ on $\mathcal{G}(\mathbb{R},M)$ such that*

$$\int_{[0,1]\times TM} f(t,x,v)\,\mathrm{d}\eta(t,x,v) = \int_{\mathcal{G}(\mathbb{R},M)} \int_{[0,1]\times TM} f(t,x,v)\,\mathrm{d}\Gamma(t,x,v)\,\mathrm{d}\vartheta(\Gamma)$$

for each function $f \in L^1(\eta)$. We call ϑ a solenoidal decomposition of η.

Proof. The proof is based on Young's superposition principle and on general constructions of measure theory. We have $\eta \in \mathcal{T}^\mu_\mu$ for some probability measure μ on M. Let ν be a decomposition of η in the sense of Theorem 5.2. We claim that, for each $k \in \mathbb{N}$, there exists a Borel probability measure ν^k on $\mathcal{G}([0,k],M)$ such that $P_{[\ell,\ell+1]\sharp}\nu^k = \tau^\ell_\sharp \nu$ for each $\ell \in \{0,1,\ldots,k-1\}$. Then, by standard extension theorems, (for example Theorem V.4.1 of [17]) there exists a unique probability ϑ on $\mathcal{G}(\mathbb{R},M)$ such that $P_{[0,1]\sharp}\vartheta = \nu$, and it is translation invariant.

We have to prove the existence of the measures ν^k. Let ν_x be the disintegration of ν with respect to the map ev_0. In other words, $M \ni x \longmapsto \nu_x$ is a measurable family of Borel probability measures on $\mathcal{G}([0,1],M)$ such that ν_x is concentrated on the set of generalized curves Γ which satisfy $ev_0(\Gamma) = x$ and such that

$$\int_{\mathcal{G}([0,1],M)} f(\Gamma)\,\mathrm{d}\nu = \int_M \int_{\mathcal{G}([0,1],M)} f(\Gamma)\,\mathrm{d}\nu_x(\Gamma)\,\mathrm{d}\mu(x)$$

for each bounded Borel function on $\mathcal{G}([0,1],M)$. Let us denote

$$\mathcal{Y}_1([0,k-1],M)\times \mathcal{Y}_1([k-1,k],M)\to \mathcal{Y}_1([0,k],M)$$
$$(Y,Z)\longmapsto Y\star Z$$

the natural gluing. Note that this map is continuous. We can now define the sequence ν^k by recurrence setting $\nu^1:=\nu$ and

$$\int_{\mathcal{G}([0,k],M)} f(\Gamma)\,\mathrm{d}\nu^k(\Gamma)$$
$$=\int_{\mathcal{G}([0,k-1],M)}\int_{\mathcal{G}([0,1],M)} f(Y\star\tau_\sharp^{k-1}Z)\,\mathrm{d}\nu_{ev_{k-1}(Y)}(Z)\,\mathrm{d}\nu^{k-1}(Y)$$

for each bounded continuous function f on $\mathcal{G}([0,k],M)$. Note in this expression that $Y\star\tau_\sharp^{k-1}Z$ is indeed a generalized curve for $\nu_{ev_{k-1}(Y)}$-almost all Z because the measure $\nu_{ev_{k-1}(Y)}$ is supported on the set of generalized curves Γ which satisfy $ev_0(\Gamma)=ev_{k-1}(Y)$. □

Let $L:[0,1[\times TM\to\mathbb{R}\cup\{\infty\}$ be a normal integrand. We extend L by periodicity to a function on $\mathbb{R}\times TM$. We say that the generalized curve $\Gamma\in\mathcal{G}(\mathbb{R},M)$ is globally minimizing the action if Γ_I is minimizing in $\mathcal{G}(I,M)$ (with fixed endpoints) for each compact interval I. Similarly, an absolutely continuous curve $\gamma:\mathbb{R}\to M$ is called globally minimizing if it is minimizing the action with fixed endpoints on each compact interval of time. If η is a closed measure which minimizes the action in $\mathcal{F}$, and if ϑ is a solenoidal decomposition of η, then ϑ-almost every generalized curve is globally minimizing. If, in addition, the integrand L is fiberwise strictly convex, then each minimizing closed measure η is concentrated on the graph of a Borel vector field $V(t,x)$, this was observed in [12] and can be proved as the similar statements in Section 4. In addition, if ϑ is a solenoidal decomposition of η, then ϑ-almost every generalized curve $\Gamma\in\mathcal{G}(\mathbb{R},M)$ is a curve, is a solution of (ODE) (with the vector field V extended to $\mathbb{R}\times M$ by periodicity), and is globally minimizing. This property is the generalization in our setting of the theorems of Mañé [16], Bangert [4], Fathi and Siconolfi [14] stating, under additional assumptions on L, that minimizing closed measures are invariant.

6.2. Holonomic measures. Our analysis of Closed measures makes it well suited to minimization problems. However, Mañé first introduced in [16] the *a priori* smaller set of holonomic measures. For historical reasons, we believe it is worth proving here the equality between holonomic measures and closed measures. Let $T\in\mathbb{N}$ and $\gamma:\mathbb{R}\to M$ be a T-periodic absolutely continuous curve. We denote by $\tilde{\gamma}$ the closed measure defined by

$$\int_{[0,1]\times TM} f\,\mathrm{d}\tilde{\gamma}=\frac{1}{T}\int_0^T f(t-[t],\gamma(t),\dot{\gamma}(t))\,\mathrm{d}t$$

where $[t]$ is the integral part of t.

Definition 6.3. The set $\mathcal{H}(M)$ of holonomic measures is the closure, in $\mathcal{T}([0,1],M)$, of the set of all measures of the form $\tilde{\gamma}$, for smooth T-periodic curves γ, $T \in \mathbb{N}$.

Lemma 6.4. *The set $\mathcal{H}(M)$ of holonomic measures can equivalently be defined as the closure in $\mathcal{T}([0,1],M)$, of the set of all measures of the form $\tilde{\gamma}$, for all absolutely continuous T-periodic curves γ, $T \in \mathbb{N}$.*

Proof. It is sufficient to prove that each measure $\tilde{\gamma}$, where γ is a T-periodic absolutely continuous curve, belongs to $\mathcal{H}(M)$. Let γ be such a curve. Let γ_n be a sequence of smooth T-periodic curves which converge to γ in $W^{1,1}([0,T],M)$. Then we prove as in Lemma 3.7 that $\tilde{\gamma}_n \to \tilde{\gamma}$. □

We recall a first remark of Ricardo Mañé:

Lemma 6.5. *The set $\mathcal{H}(M)$ is convex if M is connected.*

Proof. Let η_1 and η_2 be holonomic measures, and let λ_1 and λ_2 in $[0,1]$ be such that $\lambda_1 + \lambda_2 = 1$. We want to prove that $\lambda_1\eta_1 + \lambda_2\eta_2$ is holonomic. Since η_i is holonomic, there exist sequences of integers T_n^i and sequences of smooth curves $\gamma_n^i(t) : \mathbb{R} \to M$ of period T_n^i such that $\tilde{\gamma}_n^i \to \eta_i$. By possibly replacing the periods T_n^i by multiples, we can suppose without loss of generality that

$$T_n^1/T_n^2 \to \lambda_1/\lambda_2.$$

Let $(0, x^i, v^i)$ be a point in the support of η^i. Since $\tilde{\gamma}_n^i \to \eta^i$, there exists a sequence t_n^i of times such that

$$(t_n^i - [t_n^i], \gamma(t_n^i), \dot{\gamma}(t_n^i)) \to (0, x^i, v^i).$$

We can suppose that $t_n^i \to 0$ by replacing the curves $\gamma_n^i(t)$ by $\gamma_n^i(t - [t_n^i])$. We consider the sequence γ_n of absolutely continuous curves of period $T_n^1 + 2 + T_n^2$ such that

$$\gamma_n = \begin{cases} \gamma_n^1 & \text{on } [t_n^1,\ T_n^1 + t_n^1], \\ \gamma_n^2 & \text{on } [1 + T_n^1 + t_n^2,\ 1 + T_n^1 + T_n^2 + t_n^2], \end{cases}$$

and γ_n is a minimizing geodesic on the remaining intervals. It is not hard to see that

$$\tilde{\gamma}_n \to \lambda_1\eta_1 + \lambda_2\eta_2$$

as $n \to \infty$, so that this measure is holonomic. □

The following result is a piece of unproved folklore:

Theorem 6.6. *If M is a compact connected manifold, then each closed measure is holonomic:*

$$\mathcal{H}(M) = \mathcal{F}(M).$$

Proof. By Lemma 6.5 and Proposition B.5, it is enough to prove that if $f \in C_1([0,1] \times TM)$ satisfies

$$\int_0^T f(t-[t], \gamma(t), \dot\gamma(t))\, \mathrm{d}t \geqslant 0$$

for each $T \in \mathbb{N}$ and each absolutely continuous T-periodic curve $\gamma : \mathbb{R} \to M$, then $\int f \,\mathrm{d}\eta \geqslant 0$ for each closed measure η.

We fix an integrand $f \in C_1([0,1] \times TM)$, and extend $f|_{[0,1[}$ to a 1-periodic function on $\mathbb{R} \times TM$ without changing the name. We have

$$\int_0^T f(t, \gamma(t), \dot\gamma(t))\, \mathrm{d}t \geqslant 0$$

for each $T \in \mathbb{N}$ and each absolutely continuous T-periodic curve γ. As in the proof of Young's principle, we consider the value function $u : [0,\infty) \times M \to \mathbb{R}$ defined by

$$u(t,x) := \inf_{\gamma(t)=x} \int_0^t f(s, \gamma(s), \dot\gamma(s))\, \mathrm{d}s$$

where the infimum is taken on the set of absolutely continuous curves $\gamma : [0,t] \to M$ which satisfy $\gamma(t) = x$.

Lemma 6.7. *The function u is upper semi-continuous and locally bounded on $[0,\infty) \times N$. In addition, it is bounded from below.*

Proof. The proof of the first part is similar to the proof of Lemma 5.6. We prove that the value function is bounded from below. There exists a constant C such that

$$\int_0^s f(t, \gamma(t), \dot\gamma(t))\, \mathrm{d}t \geqslant -C$$

for each $s \geqslant 0$ and each absolutely continuous curve $\gamma : [0,s] \to M$. Indeed, we can consider the periodic curve $x(t) : [0,[s]+2] \to M$ such that $x = \gamma$ on $[0,s]$ and $x(t)$ is a minimizing geodesic on $[s,[s]+2]$ between $\gamma(s)$ and $\gamma(0)$. We have

$$\begin{aligned} 0 &\leqslant \int_0^{[s]+2} f(t, x(t), \dot x(t))\, \mathrm{d}t \\ &\leqslant \int_0^s f(t, \gamma(t), \dot\gamma(t))\, \mathrm{d}t + 2\|f\|_1 (1 + D(\gamma(0), \gamma(s))), \end{aligned}$$

where D is the Riemannian distance on M, which is bounded. From the definition of u, it follows that $u(s,x) \geqslant -C$ for each (s,x) in $[0,\infty) \times M$. □

As in the proof of Young's principle, we have the following result:

Lemma 6.8. *There exist sequences $u_n : [1,\infty)\times M \to \mathbb{R}$ and $f_n : \mathbb{R}\times TM \to \mathbb{R}$ of functions such that:*

- *The functions f_n are 1-periodic in t. They are continuous and satisfy a uniform estimate $|f_n(t,x,v)| \leqslant C(1+\|v\|_x)$. Finally, we have $f_n \to f$ almost everywhere.*
- *The functions u_n are smooth, locally bounded and bounded from below.*
- *The inequality*

$$u_n(t,\gamma(t)) - u_n(s,\gamma(s)) \leqslant \int_s^t f_n(\sigma,\gamma(\sigma),\dot\gamma(\sigma))\,\mathrm{d}\sigma$$

holds for each $1 \leqslant s \leqslant t$ in $\mathbb{R}$ and each absolutely continuous curve $\gamma : \mathbb{R} \to M$.

Proof. We regularize as in the proof of Young's principle. There is, however, a small difficulty related to the fact that we now work on a manifold. In order to solve this difficulty, we embed M as a Riemannian submanifold of some Euclidean space $\mathbb{R}^d$—one could also regularize in a more intrinsic way in the spirit of De Rham [13]. Then, we consider a tubular neighborhood U of M in $\mathbb{R}^d$ and the associated smooth projection $\pi : U \to M$. We set, for $t \geqslant 1$ and $(x,v) \in TM$,

$$u_n(t,x) = \int_{\mathbb{R}\times\mathbb{R}^d} u(t-\sigma,\ \pi(x-y))\rho_n(\sigma,y)\,\mathrm{d}\sigma\,\mathrm{d}y$$

$$f_n(t,x,v) = \int_{\mathbb{R}\times\mathbb{R}^d} f(t-\sigma,\ \pi(x-y),\ d\pi_{(x-y)}\cdot v)\rho_n(\sigma,y)\,\mathrm{d}\sigma\,\mathrm{d}y.$$

Let $\gamma : \mathbb{R} \to M$ be an absolutely continuous curve. For each fixed small $y \in \mathbb{R}^d$, the curve $\pi(\gamma(t) - y)$ is absolutely continuous, and we have

$$\begin{aligned} &u(t-\sigma,\ \pi(\gamma(t-\sigma)-y)) - u(s-\sigma,\ \pi(\gamma(s-\sigma)-y)) \\ &\qquad \leqslant \int_s^t f\Big(\zeta-\sigma,\ \pi(\gamma(\zeta-\sigma)-y),\ d\pi_{(\gamma(\zeta-\sigma)-y)}\cdot\dot\gamma(\zeta-\sigma)\Big)\,\mathrm{d}\zeta \end{aligned}$$

for all small s. The third point of the lemma follows by integration. The other points are standard. ◻

Let $\eta \in \mathcal{T}_\mu^\mu([0,1],M)$ be a closed measure. We want to prove that $\int f\,\mathrm{d}\eta \geqslant 0$. We see, as in the proof of Young's principle, that

$$\int_M u_n(i+1,\ x) - u_n(i,x)\,\mathrm{d}\mu(x) \leqslant \int_{[0,1]\times TM} f_n\,\mathrm{d}\eta$$

for each integer $i \geqslant 1$. By summation, we obtain, for each $T \in \mathbb{N}$,

$$\int_M \frac{u_n(T+1,\ x) - u_n(1,x)}{T}\,\mathrm{d}\mu(x) \leqslant \int_{[0,1]\times TM} f_n\,\mathrm{d}\eta.$$

At the limit $T \to \infty$, we obtain that $\int f_n \, \mathrm{d}\eta \geqslant 0$, and then at the limit $n \to \infty$, we get $\int f \, \mathrm{d}\eta \geqslant 0$, as desired. This ends the proof of Theorem 3.6. □

APPENDIX A. KANTOROVICH-RUBINSTEIN SPACE

Good references for the material exposed here are [3] and [20]. Let (X, d) be a complete and separable metric space. Let $\mathcal{P}_1(X)$ be the set of Borel probability measures on X with finite first moment, that is the set of Borel probability measures μ on X such that the integral

$$\int_X \mathrm{d}(x_0, x) \, \mathrm{d}\mu(x)$$

is finite for one (and then each) point $x_0 \in X$.

A coupling between two probability measures μ and η is a probability measure λ on X^2 whose marginals are μ and η, or in other words such that

$$\int_{X^2} f(x) + g(y) \, \mathrm{d}\lambda(x, y) = \int_X f(x) \, \mathrm{d}\mu(x) + \int_X g(y) \, \mathrm{d}\eta(y)$$

for all continuous functions f and g on X.

We recall the definition of the Kantorovich-Rubinstein distance d on $\mathcal{P}_1(X)$:

$$d_1(\mu, \eta) = \min_{\lambda} \int_{X \times X} d(x, y) \, \mathrm{d}\lambda(x, y)$$

where the minimum is taken on the set of couplings λ between μ and η.

Let us denote by $C_1(X)$ the set of continuous functions f on X such that

$$\sup_{x \in X} \frac{|f(x)|}{1 + d(x_0, x)} < \infty$$

for one (and then any) point $x_0 \in X$. The topology on $\mathcal{P}(X)$ defined by the distance d is precisely the weak topology associated to the linear forms $\mu \longmapsto \int f \, \mathrm{d}\mu$, $f \in C_1(X)$. In other words, we have $d(\mu_n, \mu) \to 0$ if and only if

$$\int f \, \mathrm{d}\mu_n \to \int f \, \mathrm{d}\mu$$

for all $f \in C_1(X)$. There is an interesting duality formula for the distance:

$$d_1(\mu, \eta) = \sup_f \int_X f(x) \, \mathrm{d}(\mu - \eta)(x)$$

where the supremum is taken on the set of 1-Lipschitz functions $f : X \longmapsto \mathbb{R}$. An important remark is that, if the distance d on X is bounded, then the associated Kantorovich-Rubinstein space is just the space $\mathcal{P}(X)$ of all Borel probability measures on X endowed with the narrow topology. Since it is always possible to replace a given distance d by another distance which is bounded and generates the same topology, our discussion includes the study of the narrow topology on $\mathcal{P}(X)$.

The metric space $(\mathcal{P}_1(X), d)$ is complete and separable, see [3]. The relatively compact subsets of $\mathcal{P}_1(X)$ are those which are 1-tight:

Definition A.1. The subset $Y \subset \mathcal{P}(X)$ is called 1-tight if one of the following equivalent properties holds:

- For each $\varepsilon > 0$, there exists a compact set $K \subset X$ and a point x_0 such that $\int_{X-K} (1 + d(x_0, x))\, \mathrm{d}\mu \leqslant \varepsilon$ for each $\mu \in Y$.
- There exists a function $f : X \to [0, \infty]$ whose sublevels are compact, a constant C and a point x_0 such that $\int_X (1 + d(x_0, x)) f(x)\, \mathrm{d}\mu \leqslant C$ for each $\mu \in Y$.
- The family Y is tight with uniformly integrable first moment. The first means that, for each $\varepsilon > 0$, there exists a compact set $K \subset X$ such that $\mu(X - K) \leqslant \varepsilon$ for each $\mu \in Y$. The second means that for each $\varepsilon > 0$, there exists a ball B in X such that $\int_{X-B} \mathrm{d}(x_0, x)\, \mathrm{d}\mu \leqslant \varepsilon$ for each $\mu \in Y$.

Note that 1-tightness is just tightness if the distance d is bounded.

Lemma A.2. *A sequence μ_n converges to μ in $\mathcal{P}_1(X)$ if and only if the family $\{\mu_n, n \in \mathbb{N}\}$ is 1-tight and if μ_n narrowly converges to μ, which means that $\int f\, \mathrm{d}\mu_n \to \int f\, \mathrm{d}\mu$ for each bounded continuous function f. It is enough that the family μ_n is converging narrowly to μ and has uniformly integrable first moment.*

Let us now assume that X is a finite dimensional manifold.

Lemma A.3. *A sequence μ_n converges to μ in $\mathcal{P}_1(X)$ if and only if the sequence has uniformly integrable first moment and converges to μ in the sense of distributions.*

Still assuming that X is a manifold, we finish with the following result:

Lemma A.4. *Let μ_t, $t \in I$ be a measurable family of probability measures on X, where I is an interval of $\mathbb{R}$. In order that μ_t is equal almost everywhere to a narrowly continuous map, it is enough that, for each compactly supported smooth function $f : X \longmapsto \mathbb{R}$, the function $t \longmapsto \int f\, \mathrm{d}\mu_t$ is equal almost everywhere to a continuous function.*

Appendix B. Superpositions

We continue with the notation of the first appendix. Let ν be a Borel probability measure on the complete metric space $\mathcal{P}_1(X)$. We say that ν represents the measure $\eta \in \mathcal{P}_1(X)$ if the equality

$$\int_X f \,\mathrm{d}\eta = \int_{\mathcal{P}_1(X)} \int_X f \,\mathrm{d}\mu \,\mathrm{d}\nu(\mu) \tag{B.1}$$

holds for each function $f \in L^1(\eta)$. Let us first check that the right hand side is meaningful:

Lemma B.1. *The linear map* $\mu \longmapsto \int_X f \,\mathrm{d}\mu$ *is Borel measurable on* $\mathcal{P}_1(X)$ *when* f *is a non-negative Borel function on* X. *Each probability measure* ν *on* $\mathcal{P}_1(X)$ *represents one (and only one) element* $\eta \in \mathcal{P}(X)$. *We have* $\eta \in \mathcal{P}_1(X)$ *if and only if*

$$\int_{\mathcal{P}_1(X)} \int_X \mathrm{d}(x_0, x) \,\mathrm{d}\mu(x) \,\mathrm{d}\nu(\mu) < \infty$$

for one point x_0.

Proof. It is clear that (B.1) defines a (unique) Borel measure η if it is meaningful for each non-negative Borel function. So we have to prove the first statement. Since the conclusion holds when f is continuous and bounded (for then the map $\mu \longmapsto \int f \,\mathrm{d}\mu$ is continuous), it is a consequence of the following standard statement. □

Lemma B.2. *Let* E *be a vector space of real-valued functions on* X. *Assume that* E *contains all bounded continuous functions and is closed under monotone convergence. Then* E *contains all non-negative Borel functions.*

Proof. Let $\mathcal{B}$ be the set of subsets of X whose caracteristic function belongs to E. It is not hard to see that $\mathcal{B}$ contains closed sets, that it is closed under increasing union, and that if $A \subset B$ are two elements of $\mathcal{B}$, then $B \setminus A$ is an element of $\mathcal{B}$. The classical Dynkin class theorem then implies that $\mathcal{B}$ contains all the Borel sets. But then E contains all Borel non-negative functions. □

This statement also implies the following result:

Lemma B.3. *In order that* (B.1) *holds for each function* $f \in L^1(\eta)$, *it is sufficient that it holds for all bounded continuous functions.*

Proposition B.4. *Let* $\mathcal{G}$ *be a closed subset of* $\mathcal{P}_1(X)$, *and let* $\mathcal{T}$ *be the closed convex envelop of* $\mathcal{G}$ *in* $\mathcal{P}_1(X)$. *Each measure* $\eta \in \mathcal{T}$ *is represented by a measure* ν *which is supported on* $\mathcal{G}$ *(we say that* μ *is a superposition of elements of* $\mathcal{G}$*).*

Proof. Let us consider the set S of elements of $\mathcal{P}(X)$ which are superpositions of elements of $\mathcal{G}$. It is obvious that the set S is convex, and contains $\mathcal{G}$. So we have to prove that this set is closed. Let us consider a sequence η_n in S, which has a limit η in $\mathcal{P}_1(X)$. There exists a sequence ν_n of Borel probability measures on $\mathcal{P}_1(X)$ which represents η_n. Since the family $\{\eta, \eta_1, \ldots, \eta_n, \ldots\}$ is compact in

$\mathcal{P}_1(X)$, it is 1-tight, hence there exists a function $f : X \to [0, \infty]$ whose sublevels $f^{-1}([0, c])$ are compact and such that the integral

$$\int_X (1 + d(x_0, x))f(x)\,\mathrm{d}\eta_n(x) = \int_{\mathcal{P}_1(X)} \int_X (1 + d(x_0, x))f(x)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu_n(\mu)$$

is a bounded sequence. The map

$$\mu \longmapsto \int_X (1 + d(x_0, x))f(x)\,\mathrm{d}\mu(x)$$

has compact sublevels on $\mathcal{P}_1(X)$, hence the boundedness of the sequence above implies that the sequence ν_n is a tight sequence of probability measures on $\mathcal{P}_1(X)$. By the standard Prohorov theorem, we can assume that ν_n has a limit ν for the narrow topology, which means that

$$\int_{\mathcal{P}(X)} F(\eta)\,\mathrm{d}\nu_n(\eta) \to \int_{\mathcal{P}(X)} F(\eta)\,\mathrm{d}\nu(\eta)$$

for each bounded and continuous function F on $\mathcal{P}_1(X)$. For each continuous and bounded function f on X, the affine function $\mu \longmapsto \int_X f\,\mathrm{d}\mu$ is continuous and bounded on $\mathcal{P}(X)$, hence

$$\int_{\mathcal{P}(X)} \int_X f\,\mathrm{d}\mu\,\mathrm{d}\nu_n(\mu) \to \int_{\mathcal{P}(X)} \int_X f\,\mathrm{d}\mu\,\mathrm{d}\nu(\mu).$$

Recalling that

$$\int_{\mathcal{P}(X)} \int_X f\,\mathrm{d}\mu\,\mathrm{d}\nu_n(\mu) = \int_X f\,\mathrm{d}\eta_n \to \int_X f\,\mathrm{d}\eta,$$

we conclude that

$$\int_X f\,\mathrm{d}\eta = \int_{\mathcal{P}(X)} \int_X f\,\mathrm{d}\mu\,\mathrm{d}\nu(\mu)$$

for each bounded and continuous function f on X. This implies that ν represents μ. Since the measures ν_n are supported on the closed set $\mathcal{G}$, the limit ν is supported on $\mathcal{G}$. We have proved that $\mu \in S$. □

We finish with an obvious remark on closed convex subsets of $\mathcal{P}_1(X)$.

Proposition B.5. *Let C be a closed convex subset of $\mathcal{P}_1(X)$, and let C^+ be the set of functions $f \in C_1(X)$ such that $\int_X f\,\mathrm{d}\mu \geqslant 0$ for each $\mu \in C$. Then C is the set of measures $\mu \in \mathcal{P}_1(X)$ such that $\int_X f\,\mathrm{d}\mu \geqslant 0$ for each $f \in C^+$.*

REFERENCES

[1] L. AMBROSIO, *Lecture notes on transport equation and Cauchy problem for BV vector fields and applications*.

[2] ———, *Transport equation and Cauchy problem for non-smooth vector fields*.

[3] L. AMBROSIO, N. GIGLI, and G. SAVARÉ, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005, ISBN 978-3-7643-2428-5, 3-7643-2428-7. MR 2129498 (2006k:49001)

[4] V. BANGERT, *Minimal measures and minimizing closed normal one-currents*, Geom. Funct. Anal. **9** (1999), 413–427, http://dx.doi.org/10.1007/s000390050093. MR 1708452 (2000m:49058)

[5] J.-D. BENAMOU and Y. BRENIER, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), 375–393, http://dx.doi.org/10.1007/s002110050002. MR 1738163 (2000m:65111)

[6] H. BERLIOCCHI and J.-M. LASRY, *Intégrandes normales et mesures paramétrées en calcul des variations*, Bull. Soc. Math. France **101** (1973), 129–184. MR 0344980 (49 #9719) (French)

[7] P. BERNARD and B. BUFFONI, *Optimal mass transportation and Mather theory*, J. Eur. Math. Soc. (JEMS) **9** (2007), 85–121. MR 2283105 (2007j:49050)

[8] Y. BRENIER, *The least action principle and the related concept of generalized flows for incompressible perfect fluids*, J. Amer. Math. Soc. **2** (1989), 225–255, http://dx.doi.org/10.2307/1990977. MR 969419 (90g:58012)

[9] ———, *The dual least action problem for an ideal, incompressible fluid*, Arch. Rational Mech. Anal. **122** (1993), 323–351, http://dx.doi.org/10.1007/BF00375139. MR 1217592 (94f:58039)

[10] ———, *Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations*, Comm. Pure Appl. Math. **52** (1999), 411–452, http://dx.doi.org/10.1002/(SICI)1097-0312(199904)52:4¡411::AID-CPA1¿3.0.CO;2-3. MR 1658919 (2000c:58020)

[11] ———, *Extended Monge-Kantorovich theory*, Proc. Optimal Transportation and Applications, Martina Franca, 2001, Lecture Notes in Math., vol. 1813, Springer, Berlin, 2003, pp. 91–121. MR 2006306

[12] L. DE PASCALE, M. S. GELLI, and L. GRANIERI, *Minimal measures, one-dimensional currents and the Monge-Kantorovich problem*, Calc. Var. Partial Differential Equations **27** (2006), 1–23. MR 2241304 (2007h:37095)

[13] G. DE RHAM, *Variétés différentiables*, Hermann, 1960.

[14] A. FATHI and A. SICONOLFI, *Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, Invent. Math. **155** (2004), 363–388, http://dx.doi.org/10.1007/s00222-003-0323-6. MR 2031431 (2004m:37114)

[15] J. N. MATHER, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Z. **207** (1991), 169–207, http://dx.doi.org/10.1007/BF02571383. MR 1109661 (92m:58048)

[16] R. MAÑÉ, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (1996), 273–310. MR 1384478 (97d:58118)

[17] K. R. PARTHASARATHY, *Probability Measures on Metric Spaces*, Probability and Mathematical Statistics, No. 3, Academic Press Inc., New York, 1967. MR 0226684 (37 #2271)

[18] S. K. SMIRNOV, *Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows*, Algebra i Analiz **5** (1993), 206–238, English translation: *St. Petersburg Math. J.*, **5**, (1994), 841–867, issn=1061-0022. MR 1246427 (94i:49054) (Russian, with Russian summary)

[19] M. VALADIER, *A course on Young measures*, Rend. Istit. Mat. Univ. Trieste **26** (1994), 349–394, Workshop on Measure Theory and Real Analysis (Italian), Grado, 1993. MR 1408956 (97k:28009)

[20] C. VILLANI, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003, ISBN 0-8218-3312-X. MR 1964483 (2004e:90003)

[21] L. C. YOUNG, *Lectures on the Calculus of Variations and Optimal Control Theory*, second edition, Chelsea, 1980.

PATRICK BERNARD:
CEREMADE, UMR CNRS 7534
Université Paris-Dauphine
Pl. du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16, France.
E-MAIL: `Patrick.Bernard@ceremade.dauphine.fr`

KEY WORDS AND PHRASES:
] Young measures; minimization; tonelli theorem; Mather measures; closed measures; transport.

2000 MATHEMATICS SUBJECT CLASSIFICATION: 28A33; 49J45; 37J50.

Received: *November 28th, 2006; revised: March 8th, 2007.*