A PROOF OF SUDAKOV THEOREM WITH STRICTLY CONVEX NORMS

LAURA CARAVENNA

ABSTRACT. We establish a solution to the Monge problem in $\mathbb{R}^N$, with an asymmetric, strictly convex norm cost function, when the initial measure is absolutely continuous. We focus on the strategy, based on disintegration of measures, initially proposed by Sudakov. As known, there is a gap to fill. The missing step is completed when the unit ball is strictly convex, but not necessarily differentiable nor uniformly convex. The key disintegration is achieved following a similar proof for a variational problem.

CONTENTS

1. INTRODUCTION

The present paper concerns the existence of deterministic transport plans for the Monge-Kantorovich problem, with a strictly convex norm cost function in $\mathbb{R}^N$. Let μ be an initial probability measure, absolutely continuous, and ν a final Borel probability measure. A cost function is defined by a generally non symmetric norm $|\cdot|_{D^*}$: let

$$c(x,y) = |y-x|_{D^*}.$$

We are interested in establishing the existence of maps T, such that $T_\sharp\mu = \nu$, minimizing the functional

$$T \mapsto \int_{\mathbb{R}^N} |T(x)-x|_{D^*}\, d\mu(x).$$

This is almost the original problem proposed by Monge in 1781 ([14]), but the norm is not Euclidean.

The modern approach passes through the Kantorovich formulation ([11], [12]). Rather than a map $T : X \mapsto Y$, a transport is defined as a coupling of μ, ν: a probability measure π on the product space $X \times Y$ having marginals μ, ν. The cost of the new transports is defined as

$$\pi \mapsto \int c(x,y)\, d\pi(x,y).$$

In the Monge problem, one just moves the mass present at x to some point $T(x)$. This weaker formulation gives instead the amount $\pi_x(S)$ of the mass at x which is spread in a region S. Actually, π_x is obtained disintegrating the transport plan π w.r.t. the projection on the first variable. A coupling reduces to a map when the measure π_x, for μ-a.e. x, is concentrated at one point. Conversely, a transport map T induces the transport plan $\pi = (\mathrm{Id}, T)_\sharp\mu$.

One can regard the new formulation as the relaxation of the Monge problem. Assuming, more generally, c lower semicontinuous, one can now deduce the existence of solutions by the direct method of calculus of variations. In order to recover solutions in the sense of Monge, then, one proves that some optimal transport plan is deterministic.

Preprint SISSA 64/2008/M (September 2008)

The issue has already been studied extensively. We focus only on the Monge problem with norm cost functions, presenting thus a limited literature. For a broad overview one can consult [18], [3]. A solution was initially given in 1976 by V. N. Sudakov ([16]). However, in 2000 it was pointed out that the proof refers to a lemma, about disintegration of measures, which in general does not hold ([1]); therefore there is a gap. Before this was known, another approach to the Monge problem in $\mathbb{R}^N$, based on partial differential equations, was given in [10]. Despite some additional regularity on μ, ν, they introduced new interesting ideas. Strategies at least partially in the spirit of Sudakov one were instead pursued independently in [8], [17], and then [4], improving the result. They achieved the solution to the Monge problem with an absolute continuity hypothesis only on the first marginal μ, and cost functions satisfying some kind of uniform convexity property. In [3] the thesis is instead gained for a particular norm, crystalline, which is neither strictly convex, nor symmetric. The problem with merely strictly convex norms has recently been addressed also in [9], with a different technique, in convex domains.

One year ago, a variational problem, with an analogous structure, was solved with a lemma similar to the one needed here, proving a suitable disintegration theorem ([5], [7]). There, the regularity of the norm was bypassed. The aim of this paper is to adapt the construction given in [7] to this setting. This shows that the strategy started from Sudakov actually works, when one assumes D^* strictly convex.

Before presenting the scheme of the article, we give two general definitions.

Definition 1.1 (Potential). A *potential* is a 1-Lipschitz map, $\phi : \mathbb{R}^N \mapsto \mathbb{R}$ such that

$$\phi(x) - \phi(y) \leq |y - x|_{D^*} \qquad \forall x, y \in \mathbb{R}^N. \tag{1.1}$$

We are interested in the set where the decrease of ϕ is the maximal allowed, as in the literature. In fact, the mass present at x is constrained to move only towards those points y where $\phi(x) - \phi(y) = |y - x|_{D^*}$, for a special class of potentials related to the transport problem — the Kantorovich potentials. Such a potential does exist assuming that the optimal cost is finite — otherwise every transport is optimal.

Definition 1.2 (Transport set). Given a potential ϕ, we define as *transport set* the set $\mathcal{T}$ made of segments $(\!|x, y|\!)$, without the endpoints, for every couple (x, y) such that in (1.1) equality holds:

$$\mathcal{T} = \bigcup_{(x,y)\in\partial_c\phi} (\!|x, y|\!) \qquad \text{where } \partial_c\phi = \Big\{(x, y) : \ \phi(x) - \phi(y) = |y - x|_{D^*}\Big\} \subset \mathbb{R}^{2N}.$$

Similarly, we will also consider the transport set with all the endpoints: $\mathcal{T}_{\mathrm{e}} = \bigcup_{(x,y)\in\partial_c\phi\setminus\{y=x\}} [\![x, y]\!]$.

This introduction will end collecting the notations. After that, the plan is the follwing.

- In Sect. 2 we present the main construction, adapted from [7]. The decomposition of $\mathcal{T}$ in transport rays is studied. Introductory notations are given in Subs. 2.1. After that, Subs. 2.2 contains the partition of $\mathcal{T}$ into model sets, which take into account the structure of the vector field of rays' directions. Subs. 2.3, then, establishes the main point, completing Sudakov strategy. The Lebesgue measure on the transport set is explicitly disintegrated w.r.t. the partition in transport rays. In particular, it is proved that the disintegrated measures are absolutely continuous w.r.t. the measure $\mathcal{H}^1$ on the segments. Finally, in Sect. 2.4 we study the divergence of the vector field defined on $\mathcal{T}$ as the direction of the rays, and null elsewhere. It turns out that, generally, it is a series of measure, and that a kind of divergence formula holds on model sets.
- In Sect. 3 the proof starting from Sudakov is completed, in the case D^* strictly convex. The transport problem in $\mathbb{R}^N$, with $\mu \ll \mathcal{L}^N$, is reduced to the one dimensional case, disintegrating the measures w.r.t. the equivalence relation given by the membership in a transport ray. The one dimensional case, well known since old, is solved with the selection in [3] — having, by Sect. 3, disintegrated measures still absolutely continuous w.r.t. $\mathcal{H}^1$ on each the ray. Putting side by side the one dimensional solutions, a global map is constructed.
- In Sect. 4 there is some counterexample. Firstly we recall the one in [3], showing that disintegrating a compact, positive $\mathcal{L}^N$-measure set w.r.t. the membership to disjoint segments, with Borel direction, the conditional probabilities can be atomic. Then, it is proved that the transport set $\mathcal{T}$ in general does not fill the space. In the same examples, one can see how the divergence of the vector field of rays' directions, defined as zero out of $\mathcal{T}$, can fail to be a measure.
- In Appendix A we recall the disintegration of measures, in the form presented in [6].

1.1. **Notation.** The following table lists some notations of this article.

w.r.t.	with respect to
s.t.	such that
a.a.	almost all w.r.t. a measure, which means out of a set of that measure zero
a.e.	almost everywhere, as above w.r.t. a measure
D	A convex, bounded subset of $\mathbb{R}^N$ with nonempty interior
D^*	The dual convex set of D: $D^* = \Big\{d : \ d\cdot\ell \le 1 \ \forall \ell \in D\Big\}$ (it holds $(D^*)^* = D$)
∂D	The border of D
δD	The support cone of D at $\ell \in \partial D$: $\delta D(\ell) = \Big\{d \in \partial D^* : \ d\cdot\ell = 1 = \sup_{\hat\ell \in \partial D} d\cdot\hat\ell\Big\}$
e_k	The k-element of a fixed orthonormal basis of $\mathbb{R}^N$
$\lvert\cdot\rvert$	The Euclidean norm of a vector
$\lvert\cdot\rvert_\infty$	The maximum of the component of a vector
$\lvert\cdot\rvert_{D^*}$	The asymmetric norm given by the Minkowski functional $$\lvert x\rvert_{D^*} = \inf\Big\{k : \ x \in kD^*\Big\} = \sup\Big\{x\cdot\ell, \text{with } \ell \in D\Big\}$$
$\Gamma \subset \mathbb{R}^{2N}$	A subset Γ of $\mathbb{R}^N$ is $\lvert\cdot\rvert_{D^*}$-*monotone* if for every collection of points $\{(x_i,y_i)\}_{i=1}^n \in \Gamma$ $$\sum_{i=1}^n \lvert y_i - x_i\rvert_{D^*} \le \sum_{i=1}^n \lvert y_i - x_{i+1}\rvert_{D^*}, \qquad \text{where we set } x_{n+1} := x_1$$
π	A probability measure on $\mathbb{R}^{2N}$ is $\lvert\cdot\rvert_{D^*}$-*monotone* if its support is $\lvert\cdot\rvert_{D^*}$-monotone
$\partial_c\phi$	The c-sub-differential of ϕ, with $c(x,y) = \lvert y-x\rvert_{D^*}$: $$\partial_c\phi = \Big\{(x,y) : \ \phi(x) - \phi(y) = \lvert y - x\rvert_{D^*}\Big\}$$
$\partial^- f(x)$	The *sub-differential* of a function $\phi : \mathbb{R}^N \mapsto \mathbb{R}$ at some point $x \in \mathbb{R}^N$ is the set of vectors $v^* \in \mathbb{R}^N$ such that $f(y) \ge f(x) + \langle x^*, y - x\rangle$ for all $y \in \mathbb{R}^N$.
Id	The identity function, $\mathrm{Id}(x) = x$
$\mathbb{1}_S$	The function vanishing out of S, equal to one on S (where $S \subset \mathbb{R}^N$)
$(\!\lvert a,b\rvert\!)$	The segment in $\mathbb{R}^N$ from a to b, without the endpoints
$[\![a,b]\!]$	The segment in $\mathbb{R}^N$ from a to b, including the endpoints
$\triangle$	The symmetric difference between two sets
$\mathcal{H}^\alpha$	The α-dimensional Hausdorff measure in $\mathbb{R}^N$
$\mathcal{L}^N$	The Lebesgue measure on $\mathbb{R}^N$
$\ll$	Denotes that a measure is absolutely continuous w.r.t. another one
$T_\sharp$	The push forward with a measurable map T (see Appendix A)
$\Pi(\mu,\nu)$	The set of transport plans between two probability measures μ and ν

We recall now from [2] the definition of rectifiable set and a rectifiability criterion.

Definition 1.3 (Rectifiable set). Let $E \subset \mathbb{R}^N$ be an $\mathcal{H}^k$-measurable set. We say that E is k-*countably rectifiable* if there exist countable many Lipschitz functions $f_i : \mathbb{R}^k \mapsto \mathbb{R}^N$ such that $E \subset \cup f_i(\mathbb{R}^k)$.

Definition 1.4 (k-cone). Let $\pi \subset \mathbb{R}^N$ be a k-plane and $M > 0$. Denote with π also the projection onto π, with $\pi^\perp$ the one onto the orthogonal to π. The cone $K_M(\pi)$ with axis π and opening M is defined by

$$K_M(\pi) = \big\{x \in \mathbb{R}^N : \lvert\pi^\perp x\rvert \le M\lvert\pi x\rvert\big\}.$$

Theorem 1.5 (Th.2.61, [2])**.** *Let $S \subset \mathbb{R}^N$ and assume that for any $x \in S$ there exists $\rho(x) > 0$, $M(x) > 0$ and a k-plane $\pi(x) \subset \mathbb{R}^N$ such that*

$$S \cap B_{\rho(x)}(x) \subset x + K_{M(x)}(\pi(x)).$$

Then S is contained in the union of countably many Lipschitz k-graphs whose Lipschitz constants do not exceed $2\sup_x M(x)$.

Given a multivaled fucntion $F : X \to Y$, the counterimage of a set $S \subset Y$ is defined as the set of $x \in X$ such that $F(x) \cap S \neq \emptyset$. We say that F is Borel if the counterimage of an open set is Borel.

2. Disintegration in Transport Rays

In the present section, we suppose to have a potential ϕ (Def. 1.1) which defines a particular subset of $\mathbb{R}^N$, the transport set $\mathcal{T}$ (Def. 1.2).

Firstly, in Subs. 2.1 the structure of the transport set is analyzed: by the strict convexity of D^*, this set is made of disjoint oriented segments, the transport rays. Secondly, in Subs. 2.2 the transport set is partitioned into model sets, which are basically sheaf of rays.

The membership in a ray defines then an equivalence relation on $\mathcal{T}$. In Subs. 2.3, the goal is to study the disintegration of the Lebesgue measure, on the transport set, w.r.t. this equivalence relation. The point is to show that the disintegrated measures, which will be concentrated on the rays, are absolutely continuous w.r.t. the one dimensional Hausdorff measure on the rays. This is done first on model sets, then in the whole $\mathcal{T}$.

Finally, is Subs. 2.4 the density of the disintegrated measures w.r.t. the Hausdorff one dimensional measure is related to the divergence of the vector field of the directions of the rays. We cannot say that this distribution is a Radon measure, since in general it is not true. Nevertheless, it turns out to be a series of measures, converging in the topology of distributions. The absolutely continuous part of those measures, which defines a measurable function on $\mathcal{T}$, is the coefficient for an ODE for the above density.

2.1. Elementary structure of the Transport Set.

In the present subsection, we demonstrate the basic (and well-known) fact that the transport set is made of oriented segments — the transport rays — which can intersect only at endpoints. Moreover, we show that the set of those points of intersections is countably rectifiable.

We first define the multivalued functions associating to a point the transport rays through that point, the relative directions and endpoints — and we prove them to be Borel (Lem. 2.2). We easily see that, in $\mathcal{T}$, there is exactly one ray through each point: due to the strict convexity of D^*, two rays can intersect only at a common starting point or final point. This set, which is where there are more outgoing or incoming rays, is $\mathcal{H}^{N-1}$-rectifiable (Lem. 2.4). We get thus a decomposition of the whole transport set, including the endpoints, in segments, and a $\mathcal{H}^{N-1}$-rectifiable set of points —where more rays intersect. In the following subsections, we will study some regularity of this partition, w.r.t. the disintegration of the Lebesgue measure. The fact that the set of endpoints of the rays is $\mathcal{L}^N$-negligible will be proved later on (Lem. 2.16).

Definition 2.1. The *outgoing rays* from $x \in \mathbb{R}^N$ are defined as

$$\mathcal{P}(x) := \{y : \ \phi(y) = \phi(x) - |y - x|_{D^*}\}.$$

The *incoming rays* at x are then given by

$$\mathcal{P}^{-1}(x) = \{y : \ \phi(x) = \phi(y) - |x - y|_{D^*}\}.$$

The *rays* at x are then defined as $\mathcal{R}(x) = \mathcal{P}(x) \cup \mathcal{P}^{-1}(x)$.

The transport set with the endpoints, which we denote with $\mathcal{T}_{\mathrm{e}}$, therefore, is the subset of $\mathbb{R}^N$ where there is some non degenerate transport ray: those x such that $\mathcal{R}(x) \neq \{x\}$. Similarly, $\mathcal{T}$ is the set where both $\mathcal{P}(x) \neq \{x\}$ and $\mathcal{P}^{-1}(x) \neq \{x\}$.

In the following is shown that, at $\mathcal{L}^N$-a.e. point $x \in \mathcal{T}_{\mathrm{e}}$, it is possible to define a vector field giving the direction of the ray through x:

$$d(x) := \frac{y - x}{|y - x|}\mathbb{1}_{\mathcal{P}(x)}(y) + \frac{x - y}{|y - x|}\mathbb{1}_{\mathcal{P}^{-1}(x)}(y) \qquad \text{for some } y \neq x \text{ on the ray through } x.$$

Notice immediately that the set $\mathcal{P}(x)$ is really a union of closed segments with endpoint x, which we call rays. In fact, ϕ's Lipschitz condition (1.1) implies that, for every $y \in \mathcal{P}(x)$, ϕ must decrease linearly from x to y at the maximal rate allowed:

$$\phi(x + t(y - x)) = \phi(x) - t|y - x|_{D^*} \qquad \text{for all } y \in \mathcal{P}(x),\ t \in [0, 1]. \tag{2.1}$$

Moreover, notice also that, due to strict convexity, two rays can intersect only at some point which is a beginning point for both, or a common final point. In fact, suppose two rays intersect in y, take $x \in \mathcal{P}^{-1}(y)$, $z \in \mathcal{P}(y)$. Therefore,

$$\phi(z) \overset{z\in\mathcal{P}(y)}{=} \phi(y) - |z - y|_{D^*} \overset{y\in\mathcal{P}(x)}{=} \phi(x) - |y - x|_{D^*} - |z - y|_{D^*} \leq \phi(x) - |z - x|_{D^*}.$$

Again by Lipschitz condition (1.1), equality must hold: then $|z - y|_{D^*} = |y - x|_{D^*} + |z - y|_{D^*}$. Since D^* is strictly convex, this implies that x, y, z must be aligned.

As a consequence, in order to show that in $\mathcal{T}_{\mathrm{e}}$ there exists a vector field of directions, one has to show that there is at most one transport ray even at $\mathcal{L}^N$-a.e. endpoint. This is not trivial because, up to now, we can't say that the set of endpoints is $\mathcal{L}^N$-negligible. This does not follow from the fact that the set e.g. of initial points is Borel and from each point starts at least a segment which does not intersect the others, and the field of directions of those segments is Borel. One can see the counterexample reported in Sect. 4 (Ex. 4.1, see [13], [3]). There, on the set of endpoints, one should then study carefully the multivalued map giving the directions of those rays:

$$\mathcal{D}(x) := \left\{ \frac{y - x}{|y - x|} \mathbb{1}_{\mathcal{P}(x)}(y) + \frac{x - y}{|y - x|} \mathbb{1}_{\mathcal{P}^{-1}(x)}(y) \right\}_{y \in \mathcal{R}(x)} \qquad \text{for all } x \in \mathcal{T}_{\mathrm{e}}. \tag{2.2}$$

We show firstly that the above maps $\mathcal{P}$, $\mathcal{D}$ are Borel maps.

Lemma 2.2. *The multivalued functions $\mathcal{P}$, $\mathcal{P}^{-1}$, $\mathcal{R}$, $\mathcal{D}$ have a σ-compact graph. In particular, the inverse image — in the sense of multivalued functions — of a compact set is σ-compact. Therefore, the transport sets $\mathcal{T}$ and $\mathcal{T}_{\mathrm{e}}$ are σ-compact.*

Proof. Firstly, consider the graph of $\mathcal{P}$: it is closed. In fact, take a sequence (x_k, z_k), with $z_k \in \mathcal{P}(x_k)$, converging to a point (x, z). Then, since $\phi(z_k) = \phi(x_k) - |z_k - x_k|_{\mathcal{D}^*}$, by continuity we have that $\phi(z) = \phi(x) - |z - x|_{\mathcal{D}^*}$. Therefore the limit point (x, z) belongs to $\mathrm{Graph}(\mathcal{P}(x))$. Since the graph is closed, then both the image and the counterimage of a closed set are σ-compact. In particular, this means that $\mathcal{P}$, $\mathcal{P}^{-1}$ and $\mathcal{R}$ are Borel. Secondly, since the graph of $\mathcal{P}$ is closed, both the graphs of $\mathcal{P} \setminus \mathrm{Id}$ and $\mathcal{P}^{-1} \setminus \mathrm{Id}$ are still σ-compact. In particular, the intersection and the union of their images must be σ-compact. These are, respectively, the transport sets $\mathcal{T}$, $\mathcal{T}_{\mathrm{e}}$. Finally, the map $\mathcal{D}$ is exactly the composite map $x \in \mathcal{T} \to \mathrm{dir}(x, \mathcal{R}^{-1}(x) \setminus \{x\})$, where $\mathrm{dir}(x, \cdot) = (\cdot - x)/|\cdot - x|$. In particular, by the continuity of the map of directions on $\mathbb{R}^N \times \mathbb{R}^N \setminus \{x = y\}$, its graph is again σ-compact. □

Remark 2.3. The fact that the inverse image of a multivalued function is compact implies that the inverse image of an open set is Borel, since it is σ-compact. In the case it is single-valued, this means, in turn, that the map is Borel.

The next point is to show that the transport rays define a partition of $\mathcal{T}$ into segments, up to a $\mathcal{L}^N$-negligible set. This is a consequence of the strict convexity of the norm. On the one hand, the strict convexity implies the differentiability of ∂D: then, at any $\ell \in \partial D$, the support set $\delta D(\ell)$ consists of a single vector d. At $\mathcal{L}^N$-a.e. point x of $\mathcal{T}_{\mathrm{e}}$, moreover, $-\nabla\phi(x) \in \partial D$ and the direction of each ray through x must belong to $\delta D(-\nabla\phi)$, thus there is just one possible choice (see Section 4). On the other hand, one can get a stronger result studying d more carefully.

Lemma 2.4. *On $\mathcal{T}_{\mathrm{e}}$, $\mathcal{D}$ is single valued out of a $\mathcal{H}^{N-1}$-countably rectifiable set.*

Proof. We prove the claim on $\mathcal{P}^{-1}(\mathbb{R}^N)$. On $\mathcal{P}(\mathbb{R}^N)$ the proof is analogous.

`Step 0.` As already seen, due to strict convexity of D^*, $\mathcal{D}$ is single valued on a transport ray.

`Step 1.` Due to D^* strict convexity, given two different $d_1, d_2 \in \partial D^*$, the respective support cones $\delta D^*(d_1)$, $\delta D^*(d_2)$ are separate.

`Step 2.` Consider now a point x where there are two different directions d^1, $d^2 \in \mathcal{D}(x)$. Suppose to have a sequence $(x_k, d_k) \in \mathrm{Graph}(\mathcal{D})$ converging to (x, d^1); take points y_k, y_1, y_2 on the relative rays, with the same $|\cdot|_{D^*}$-distance from x, and with y_k converging to y_1. Then we prove that x_k does not approach

x from the negative cone centered in x and spanned by the vectors $\delta D^*(d^2/|d^2|_{D^*}) - \delta D^*(d^1/|d^1|_{D^*})$. Precisely, we mean that from every subsequence one can extract a subsequence for which there exist $v^i \in \delta D^*(d^1/|d^1|_{D^*})$, $i = 1, 2$, satisfying

$$\left\langle \lim_i \frac{x_k - x}{|x_k - x|}, v^2 - v^1 \right\rangle \geq 0.$$

By compactness, every sequence x_k, has a subsequence such that the directions $\frac{x_k - x}{|x_k - x|_{D^*}}$ converge. Now, for every subsequence x_k for which the directions $\frac{x_k - x}{|x_k - x|_{D^*}}$ converge to some direction $\ell \in S^{N-1}$, we prove that

$$\langle v^2(\ell, y^2) - v^1(\ell, y^1), \ell \rangle \geq 0,$$

where v^2, v^1 are suitable elements in the sub-differentials of $|\cdot|_{D^*}$ at, respectively, $y^2 - x$ and $y^1 - x$. This is based on the existence of a vector $\partial|b|_{D^*}$, belonging to the subdifferential of $|\cdot|_{D^*}$ at b, and depending also on a direction $\ell \in S^{N-1}$, such that the equality

$$|a|_{D^*} = |b|_{D^*} + \langle \partial|b|_{D^*}, a - b\rangle + o(|a - b|), \tag{2.3}$$

holds for every $a \in \mathbb{R}^N$ converging to $b \in \mathbb{R}^N$ with $\frac{a-b}{|a-b|_{D^*}}$ converging to ℓ. In fact, as a consequence, one can choose vectors $v^2 \in \partial^-|y^2 - x|_{D^*}$, $v^1_k \in \partial^-|y^1_k - x|_{D^*}$ in order to have

$$\begin{aligned}
\phi(x) + \langle v^2, x - x_k\rangle + o(|x - x_k|) &\overset{(2.3)}{=} \phi(x) - |y^2 - x|_{D^*} + |y^2 - x_k|_{D^*} \\
&= \phi(y^2) + |y^2 - x_k|_{D^*} \geq \phi(x_k) = \phi(y^1_k) + |y^1_k - x_k|_{D^*} \\
&\geq \phi(x) - |y^1_k - x|_{D^*} + |y^1_k - x_k|_{D^*} \overset{(2.3)}{=} \phi(x) + \langle v^1_k, x - x_k\rangle + o(|x - x_k|).
\end{aligned}$$

This yields, considering every subsequence for which v^1_k converge to some v^1, necessarily in $\partial^-|y^1 - x|_{D^*}$

$$\langle v^2 - v^1, \ell\rangle = \lim_k \left\langle v^2 - v^1_k, \frac{x - x_k}{|x - x_k|} \right\rangle \geq 0.$$

showing that every limit direction of $x_k - x$ lies in the positive cone spanned by the vectors $\delta D^*(d^2/|d^2|_{D^*}) - \delta D^*(d^1/|d^1|_{D^*})$.

`Step 3.` Choose a dense sequence $\{y_i\}$ on S^{N-1} and call B^n_i the closed ball centered in y_i with radius $\frac{1}{2n}$. Define $J_{i,j,n,m}$ to be the set where $\mathcal{D}(x)$ contains at least two directions, belonging respectively to the closed balls $B^n_{i,j}$, with relative distance more than $\frac{1}{n}$ and relative rays of length more than $\frac{1}{m}$. The thesis will follow collecting those countably many sets $J_{i,j,n,m}$, since we are showing that each of them is $\mathcal{H}^{N-1}$ countably rectifiable. To this purpose, it suffices to prove that the cone condition in Th. 1.5 holds: for every $x \in J_{i,j,n,m}$ there is no sequence in $J_{i,j,n,m}$ converging to x in a suitable cone centered in x. In fact, consider a sequence $x_k \in J_{i,j,n,m}$ converging to x. By construction, it turns out that every subsequence has a sub-subsequence such that there exist $d^k_1 \in \mathcal{D}(x_k) \cap B^n_i$ and $d^k_2 \in \mathcal{D}(x_k) \cap B^n_j$ converging to some $d_1 \in \mathcal{D}(x) \cap B^n_i$ and $d_2 \in \mathcal{D}(x) \cap B^n_j$; moreover, $y^i_k = x_k + d^k_i/m$ converges to $y^i = x + d_i/m$. Applying twice the second step, it follows, then, that each of those subsequences must approach x out of the whole cone, centered in x, through $\delta D^*(\widehat{B}^n_i) - \delta D^*(\widehat{B}^n_j)$, where $\widehat{B}^n_\cdot$ are the radial projections, with center x, of $B^n_\cdot$ onto $\{x : |x|_{D^*=1}\}$. Therefore, the same holds for the whole sequence. □

The above lemma ensures that we have a vector field giving at a.e. any point $x \in \mathcal{T}_{\mathrm{e}}$ the direction of the ray passing there:

$$d(x) \quad \text{s.t.} \quad \mathcal{D}(x) := \{d(x)\}.$$

Notice that it is single valued on a Borel set with the same measure as $\mathcal{T}$. On this domain, the function d is Borel, by Lem. 2.2, being just a restriction of the Borel multivalued map $\mathcal{D}$. Since, by the strong triangle inequality, rays cannot bifurcate, we are allowed to consider their endpoints, possibly at infinity. After compactifying $\mathbb{R}^N$, define on $\mathcal{T}_{\mathrm{e}}$

$$a(x) = \big\{x + td, \text{ where } t \text{ is the minimal value for which } \phi(x) = \phi(x + td) + td,\ d \in \mathcal{D}(x)\big\},$$

$$b(x) = \big\{x + td, \text{ where } t \text{ is the maximal value for which } \phi(x) = \phi(x + td) + td,\ d \in \mathcal{D}(x)\big\}.$$

Both these functions are Borel, and $\mathcal{L}^N$-a.e. single valued. Moreover, their image is $\mathcal{H}^N$-negligible; in particular, $a(x) \neq x$ for $\mathcal{H}^N$ a.e. $x \in \mathcal{T}$. We postpone the proof of these two facts to the following (see resp. Lem. 2.9 and Lem. 2.16), since it will be easier after the construction of the following subsection.

2.2. **Partition of $\mathcal{T}$ into model sets.** Here we decompose the transport set $\mathcal{T}$ into particular sets, which take account of the structure of the vector field. They will be called sheaf sets and d-cylinders. We show that sets of this kind approximate $\mathcal{T}$, in a sense that we will specify. This will be fundamental in the following, since the estimates will be proved first in a model set like those, then extended on the whole $\mathcal{T}$, by approximation.

Definition 2.5 (Sheaf set). The sheaf sets $\mathcal{Z}$, $\mathcal{Z}_{\mathfrak{e}}$ are defined to be σ-compact subsets of $\mathcal{T}$ of the form

$$\mathcal{Z} = \mathcal{Z}(Z) = \cup_{y\in Z} (\!| a(y), b(y) |\!) \qquad \mathcal{Z}_{\mathfrak{e}} = \mathcal{Z}_{\mathfrak{e}}(Z) = \cup_{y\in Z} [\![a(y), b(y)]\!]$$

for some σ-compact Z contained in a hyperplane of $\mathbb{R}^N$, intersecting each $(\!| a(y), b(y) |\!)$ in one point. We define Z to be a basis, while the relative axis is a unit vector, in the direction of the rays, orthogonal to the above hyperplane.

The first point is to prove that one can cover $\mathcal{T}$ (resp. $\mathcal{T}_{\mathfrak{e}}$) with countably many possibly disjoint sets $\mathcal{Z}_i$ (resp. $\mathcal{Z}_{\mathfrak{e}i}$). Fix some $1 > \varepsilon > 0$. Consider a finite number of points $\mathfrak{e}_j \in S^{N-1}$ such that $S^{N-1} \subset \cup_{j=1}^{J} B_\varepsilon(\mathfrak{e}_j)$; define, then, the following finite, disjoint covering $\{S_j^{N-1}\}$ of S^{N-1}:

$$S_j^{N-1} = \left\{ d \in S^{N-1} : \ d\cdot \mathfrak{e}_j \geq 1-\varepsilon \right\} \setminus \bigcup_{i=1}^{j-1} S_i^{N-1}.$$

Lemma 2.6. *The following sets are sheaf sets covering $\mathcal{T}$ (resp. $\mathcal{T}_{\mathfrak{e}}$):*

$$\text{for } j = 1,\dots,J,\ k\in\mathbb{N},\ \ell, -m \in \mathbb{Z}\cup\{-\infty\},\ \ell < m$$

$$\mathcal{Z}_{jk\ell m} = \left\{ x \in \mathcal{T} : \ d(x) \in S_j^{N-1},\ \ell, m \text{ extremal values s.t. } 2^{-k}[\ell-1, m+1] \subset \mathcal{R}(x)\cdot \mathfrak{e}_{\mathrm{j}} \right\}$$

$$\mathcal{Z}^{\mathfrak{e}}_{jk\ell m} = \left\{ x \in \mathcal{T}_{\mathfrak{e}} : \ \exists d \in S_j^{N-1} \cap \mathcal{D}(x),\ \ell, m \text{ extr. val. s.t. } 2^{-k}[\ell-1, m+1] \subset \big(\mathcal{R}(x) \cap \{x+\mathbb{R}d\}\big)\cdot \mathfrak{e}_{\mathrm{j}} \right\}.$$

$\{\mathcal{Z}_{ik\ell m}\}_{i,\ell,m}$ is a partition of $\mathcal{T}$, it refines when k increases. Two different $\mathcal{Z}^{\mathfrak{e}}_{ik\ell m}$ can instead intersect each other, but only in points where $\mathcal{D}$ is multivalued. We denote with $Z_{jk\ell m}$ a basis of $\mathcal{Z}_{jk\ell m}$.

Proof. Consider a point on a ray. Then $d(x) \in S_j^{N-1}$ for exactly one j. Moreover, since $\mathcal{R}(x)\cdot\mathfrak{e}_{\mathrm{j}}$ is a nonempty interval, for k sufficiently large we can define maximal values of ℓ, m such that $2^{-k}[\ell-1, m+1] \subset \mathcal{R}(x)\cdot\mathfrak{e}_{\mathrm{j}}$. Therefore $x \in \mathcal{Z}_{jk\ell m}$, or $\mathcal{Z}^{\mathfrak{e}}_{jk\ell m}$, in the case x is an endpoint. This proves that we have a covering of $\mathcal{T}$ (resp. $\mathcal{T}_{\mathfrak{e}}$). It remains to show that the above sets are σ-compact: then, intersecting $\mathcal{Z}_{jk\ell m}$ with an hyperplane with projection on $\mathbb{R}\mathfrak{e}_j$ belonging to $2^{-k}(\ell, m)$, we will have a σ-compact basis $Z_{jk\ell m}$. It is clear that the covering, then, can be refined to a partition into sheaf sets with bounded basis.

To see that the above sets are σ-compact, one first observes that the following one are closed: since S_j^{N-1} is σ-compact, consider a covering of it with compact sets $\mathfrak{S}_j^n$, for $n \in \mathbb{N}$; define then

$$C_{j\alpha\beta n} = \left\{ x : \ d(x) \in \mathfrak{S}_j^n \quad \mathcal{R}(x)\cdot \mathfrak{e}_j \leq \alpha \supset [\alpha, \beta] \right\}.$$

In particular, both $C_{j\alpha\beta n}$ and its complementary are σ-compact. Then one has the thesis by

$$\begin{aligned} \mathcal{Z}_{jk\ell m} &= \cup_n C_{j,2^{-k}(\ell-1),2^{-k}(m+1),n} \setminus \big(C_{j,2^{-k}(\ell-2),2^{-k}(m+1),n} \cup C_{j,2^{-k}(\ell-1),2^{-k}(m+2),n} \big) \\ &= \cup_n C_{j,2^{-k}(\ell-1),2^{-k}(m+1),n} \cap \cup_h K_h^n = \cup_{n,h} C_{j,2^{-k}(\ell-1),2^{-k}(m+1),n} \cap K_h^n. \end{aligned}$$

where we have replaced the complementary of $C_{j,2^{-k}(\ell-2),2^{-k}(m+1),n} \cup C_{j,2^{-k}(\ell-1),2^{-k}(m+2),n}$ by the union of suitable compacts K_h^n, clearly depending also on j, k, ℓ, m.

□

The next point is to extract a disjoint covering made of cylinders subordinated to d.

Definition 2.7 (d-cylinder). A cylinder subordinated to the vector field d, is a σ-compact set of the form

$$\mathcal{K} = \left\{ \sigma^t(Z) : \ t \in [h^-, h^+] \right\} \subset \mathcal{Z}(Z) \qquad \text{where } \sigma^t(y) = y + \frac{t d(y)}{d(y)\cdot \mathfrak{e}},$$

for some σ-compact Z contained in a hyperplane of $\mathbb{R}^N$, a direction $\mathfrak{e} \in S^{N-1}$, real values $h^- < h^+$. We call $\mathfrak{e}$ the axis, $\sigma^{h^\pm}(Z)$ the bases.

Lemma 2.8. *With the notations of Lemma 2.6, $\mathcal{T}$ is covered by the d-cylinders*

$$\mathcal{K}_{jk\ell m} = \left\{\sigma^t y = y + \frac{td(y)}{d(y)\cdot e_j} \quad \text{with } y \in Z_{jk\ell m},\ t \in 2^{-k}[\ell, m]\right\}.$$

Therefore, a partition is given by the d-cylinders $\{\mathcal{K}^{\pm}_{jk\ell m} = \mathcal{K}_{jk\ell m} \setminus \cup_{h<k}\mathcal{K}_{jh\ell m}\}$.

Proof. The proof is similar to the one of Lemma 2.6: just cut the sets $\mathcal{Z}_{jk\ell m}$ with strips orthogonal to $\mathfrak{e}_j$. Moreover, the partition given in the statement is still made by d-cylinders because, when k increases of a unity, the sheaf $\mathcal{Z}_{jk\ell m}$ generally splits into slightly longer four pieces: we are removing the central d-cylinder, already present in a d-cylinder corresponding to a lower k, and taking the 'boundary' ones. □

Lemma 2.9. *The (multivalued) functions a, b are Borel on the transport set with endpoints $\mathcal{T}_{\mathfrak{e}}$.*

Proof. A first way could be to show that their graph is σ-compact (as for Lem. 2.2). Define instead the following intermediate sets between a d-cylinder and a sheaf set:

$$\mathcal{V}^-_{jk\ell m} = \mathcal{Z}^{\mathfrak{e}}_{jk\ell m} \cap \{x :\ x\cdot\mathfrak{e}_j \le 2^{-k}m\} \qquad \mathcal{V}^+_{jk\ell m} = \mathcal{Z}^{\mathfrak{e}}_{jk\ell m} \cap \{x :\ x\cdot\mathfrak{e}_j \ge 2^{-k}\ell\},$$

where $\{\mathcal{Z}^{\mathfrak{e}}_{jk\ell m}\}$ is the partition defined in Lemma 2.6. Define the Borel function pushing, along rays, each point in $\mathcal{V}^-_{jk\ell m}$ to the upper basis:

$$\sigma^+\mathbb{1}_{\mathcal{V}^-_{jk\ell m}}(x) = \begin{cases} \sigma^{2^{-k}m - x\cdot\mathfrak{e}_j}x & \text{if } x \in \mathcal{Z}_{jk\ell m} \\ \sigma^{2^{-k}m - y\cdot\mathfrak{e}_j}y \text{ for a } y \in \mathcal{R}(x)\cap Z_{jk\ell m} & \text{if } x \text{ is a beginning point }. \\ \emptyset & \text{if } x \notin \mathcal{V}^-_{jk\ell m} \end{cases}$$

Then, the Borel functions $\cup_{j\ell m}\sigma^+\mathbb{1}_{\mathcal{V}^-_{jk\ell m}}(x)$, multivalued on a $\mathcal{H}^{N-1}$-countably rectifiable set, converge pointwise to b when k increases. The same happens for a, considering an analogous sequence $\cup_{j\ell m}\sigma^-\mathbb{1}_{\mathcal{V}^+_{jk\ell m}}(x)$. □

Remark 2.10. Fix the attention on a sheaf set with axis e_1 and basis $Z \subset \{x\cdot e_1 = 0\}$. The composite map

$$\begin{array}{ccccc} \mathcal{Z}(z) \subset \mathbb{R}^N & \to & \mathbb{R}\times\mathbb{R}^N & \to & Z + (-1,1)e_1 \subset \mathbb{R}^N \\ z & \to & (z\cdot e_1, \sigma^{-z\cdot e_1}z) = (t,x) & \to & \left(x + \dfrac{t\arctan t}{\arctan(b(x)\cdot e_1)}e_1\mathbb{1}_{t\ge 0} + \dfrac{t\arctan t}{\arctan(a(x)\cdot e_1)}e_1\mathbb{1}_{t\le 0}\right) \end{array}$$

is a Borel and invertible change of variable from $\mathcal{Z}(Z)$ to the cylinder $Z + (-1,1)e_1$, with Borel inverse. This will turn out to carry negligible sets into negligible sets (see Cor. 2.18).

Remark 2.11. Consider a d-cylinder of the above partition

$$\mathcal{K} = \left\{\sigma^t(Z) :\ t \in [h^-, h^+]\right\}.$$

Then, partitioning it into countably many new d-cylinders and a negligible set, we will see that one can assume Z to be compact, and a, d, b to be continuous on it. In fact, applying repeatedly Lusin theorem one can find a sequence of compacts covering $\mathcal{H}^{N-1}$-almost all Z. Moreover, the local disintegration formula (2.15) will ensure that, when replacing Z with a subset of equal $\mathcal{H}^{N-1}$ measure, the Lebesgue measure of the new d-cylinder does not vary.

2.3. **Explicit disintegration of $\mathcal{L}^N$.** In this subsection we arrive to the explicit disintegration of the Lebesgue measure on $\mathcal{T}$, w.r.t. the partition in rays. Initially, the ambient space is restricted to a model set, which can be a sheaf set or a d-cylinder. The main advantage is that there is a sequence of vector fields — piecewise radial in connected, open sets with Lipschitz boundary — converging pointwise to d. They are the direction of the rays relative to potentials approximating ϕ. Taking advantage of that approximation, we first show a basic estimate on the push forward, by d, of the Hausdorff $N-1$-dimensional measure on hyperplanes orthogonal to the axis of the cylinder. This is the main result in Sub-subs. 2.3.1. It will lead to the disintegration of the Lebesgue measure on the d-cylinder, w.r.t. the partition defined by transport rays — topic of Sub-subs. 2.3.2. In particular, it is proved that the disintegrated measures are absolutely continuous w.r.t. the Hausdorff one dimensional measure on the rays. We recall that this is nontrivial, since some regularity of the field of directions is needed (see Ex.4.1).

2.3.1. *Fundamental estimate: the sheaf set $\mathcal{Z}$.* We study here the problem on a d-cylinder and on the relative sheaf set. Firstly, we show with an example how the vector field d can be approximated with a piecewise radial vector field d_I. Secondly, with that approximation, we prove that the vector field d does not shrink to zero $\mathcal{H}^{N-1}$-positive orthogonal sections. Thirdly, we conclude that the disintegration of the Lebesgue measure on such cylinders enjoys the desired property: the disintegrated measures on the rays are absolutely continuous w.r.t. $\mathcal{H}^1$.

Fix the attention on a sheaf set $\mathcal{Z}_{\mathtt{e}}$ with axis e_1 and a bounded basis $Z \subset \{x : \ e_1 \cdot x = 0\}$: assume that, for suitable $h^\pm$,

$$\mathcal{Z}_{\mathtt{e}} = \cup_{y\in Z} [\![a(y), b(y)]\!], \qquad e_1 \cdot a \restriction_Z < h^- \leq 0, \qquad e_1 \cdot b \restriction_Z > h^+ \geq 0.$$

Example 2.12 (Local approximation of the vector field d). Suppose $h^- < 0$. Consider the Borel functions moving points along rays, parametrized with the projection on the e_1 axis,

$$x \longrightarrow \sigma^t(x) := x + \frac{t}{d(x)\cdot e_1} d(x).$$

In order to avoid to work with infinite values, consider the auxiliary function $\tilde{a}(x) := \sigma^{h^-}(x)$. Choose now a dense sequence $\{\mathtt{a}_i\}$ in $\sigma^{h^-}Z$. Approximate the potential ϕ with the sequence of potentials

$$\phi_I(x) = \max\big\{\phi(\mathtt{a}_i) - |x - \mathtt{a}_i|_{D^*} : \ i = 1, \dots, I\big\}. \tag{2.4}$$

Since ϕ is uniformly continuous on $\sigma^{h^-}Z$, as a consequence of the representation formula for ϕ, we see easily that ϕ_I decreases to ϕ on the closure of $\mathcal{Z}_{\mathtt{e}} \cap \{x \cdot e_1 \geq h^-\}$. There, consider now the vector fields of ray's directions

$$d_I(x) = \sum_{i=1}^{I} d^i(x) \mathbb{1}_{\overline{\Omega_i}}(x) \qquad \text{with} \qquad d^i(x) = \frac{x - \mathtt{a}_i}{|x - \mathtt{a}_i|}, \tag{2.5}$$

where the open sets Ω_i are

$$\begin{aligned}\Omega_i &= \Big\{x : \ \phi(x) - |x - \mathtt{a}_i|_{D^*} > \phi(x) - |x - \mathtt{a}_j|_{D^*}, \ j \in \{1 \dots I\} \setminus i\Big\} \\ &= \text{interior of } \Big\{x : \ \phi(\mathtt{a}_i) = \phi_I(x) + |x - \mathtt{a}_i|_{D^*}\Big\}.\end{aligned}$$

They partition $\mathbb{R}^N$, together with their border. Notice that this border is $\mathcal{H}^{N-1}$-countably rectifiable: for example apply Lemma 2.4, since it is where the field of rays' direction associated to ϕ_I is multivalued. We show that the sequence d_I converges $\mathcal{H}^N$-a.e. to d on $\mathcal{Z}_{\mathtt{e}} \cap \{x \cdot e_1 > h^-\}$. More precisely, every selection of the d_I converges pointwise to d on $\mathcal{Z}_{\mathtt{e}} \cap \{x \cdot e_1 > h^-\}$. Consider any sequence $\{d_{I_j}(x)\}_j$ convergent to some $\bar{d}$. The corresponding points $\mathtt{a}_{i_j}$ satisfy

$$\phi_{I_j}(\mathtt{a}_{i_j}) = \phi_{I_j}(x) + |x - \mathtt{a}_{i_j}|_{D^*};$$

therefore, they will converge to some point $\mathtt{a}$ s.t. $\bar{d} = (x - \mathtt{a})/|x - \mathtt{a}|$ and $\mathtt{a} \cdot e_1 = h^-$; in particular, $\mathtt{a} \neq x$. Then, taking the limit in the last equation, one gets that $\phi(\mathtt{a}) = \phi(x) + |x - \mathtt{a}|_{D^*}$. In particular, where d is single valued, $d = (x - z)/|x - z| = \bar{d}$ follows.

Moreover, by the explicit formula of the d_I, just considering separately on each piece in Ω_i as in [7], one computes that

$$\frac{d}{dt}\big(\mathcal{H}^{n-1}(\sigma^t_{d_I}S)\big) \leq \frac{n-1}{t - h^-}\mathcal{H}^{n-1}(\sigma^t_{d_I}S) \qquad \text{for all measurable set } S \subset \sigma_I^{-\bar{t}}\sigma^{\bar{t}}Z,\, t \leq \bar{t}, \tag{2.6}$$

where $\sigma^t_{d_I}$, similarly to σ^t, moves points along the rays relative to ϕ_I.

We study now the push forward, with the vector field d, of the measure $\mathcal{H}^{N-1}$ on the orthogonal sections of the d-cylinder

$$\mathcal{K} = \mathcal{Z} \cap \{h^- \leq e_1 \cdot x \leq h^+\} = \cup_{t\in[h^-,h^+]}\sigma^t Z, \qquad \text{and } a \restriction_{\mathcal{K}} \cdot e_1 \leq h^-,\, b \restriction_{\mathcal{K}} \cdot e_1 \geq h^+.$$

Lemma 2.13 (Absolutely continuous push forward)**.** *For $h^- < s \le t < h^+$ the following estimate holds:*

$$\left(\frac{h^+ - t}{h^+ - s}\right)^{N-1} \mathcal{H}^{N-1}(\sigma^s S) \le \mathcal{H}^{N-1}(\sigma^t S) \le \left(\frac{t - h^-}{s - h^-}\right)^{N-1} \mathcal{H}^{N-1}(\sigma^s S) \qquad \forall S \subset Z.$$

Moreover, for $h^- \le s \le t < h^+$ the left inequality still holds, and for $h^- < s \le t \le h^+$ the right one.

Proof. Fix $h^- < s \le t \le h^+$. Consider $S \subset Z$ and assume firstly that $\mathcal{H}^{N-1}(\sigma^t S) > 0$. Approximate the vector field d as in Ex. 2.12. There, we proved pointwise convergence on $\mathcal{Z}_{\mathtt{e}} \cap \{x \cdot e_1 > h^-\}$. Choose any $\eta > 0$. By Egoroff theorem, the convergence of d_I to d is uniform on a compact subset $A_\eta \subset \sigma^t S$ such that

$$\mathcal{H}^{n-1}(A_\eta) \ge \mathcal{H}^{n-1}(\sigma^t S) - \eta. \tag{2.7}$$

Eventually restricting it, we can also assume that d, $\{d_I\}$ are continuous on A_η, by Lusin theorem. Let A_η evolve with d_I and d. By d_I's uniform convergence, it follows than that $\sigma_{d_I}^{s-t}(A_\eta)$ converges in Hausdorff metric to $\sigma_d^{s-t}(A_\eta)$. Moreover, by the explicit formula (2.6) for the regular d_I,

$$\mathcal{H}^{n-1}(A_\eta) \equiv \mathcal{H}^{n-1}(\sigma_{d_I}^0 A_\eta) \le \left(\frac{t - h^-}{s - h^-}\right)^{N-1} \mathcal{H}^{n-1}(\sigma_{d_I}^{s-t} A_\eta). \tag{2.8}$$

By the semicontinuity of $\mathcal{H}^{N-1}$ w.r.t. Hausdorff convergence then

$$\limsup_{I \to \infty} \mathcal{H}^{n-1}(\sigma_{d_I}^{s-t} A_\eta) \le \mathcal{H}^{n-1}(\sigma_d^{s-t} A_\eta) \le \mathcal{H}^{n-1}(\sigma^s S). \tag{2.9}$$

Collecting (2.7), (2.8) and (2.9) we get the right estimate, by the arbitrariness of η. In particular, $\mathcal{H}^{N-1}(\sigma^t S) > 0$ implies $\mathcal{H}^{N-1}(\sigma^s S) > 0$.

Secondly, assume $\mathcal{H}^{N-1}(\sigma^s S) > 0$ and $h^- \le s \le t < h^+$. One can now prove the opposite inequality in a similar way, truncating and approximating $b(Z)$ instead of $a(Z)$. In particular, this left estimate implies $\mathcal{H}^{N-1}(\sigma^t S) > 0$.

As a consequence, $\mathcal{H}^{N-1}(\sigma^s S) = 0$ if and only if $\mathcal{H}^{N-1}(\sigma^t S) = 0$ for all $s, t \in (h^-, h^+)$ — therefore the statement still holds in a trivial way when the $\mathcal{H}^N$-measure vanishes. □

Remark 2.14. The consequences of this fundamental formula are given in Sub-subs. 2.3.2. We just anticipate immediately that it states exactly that the push forward of the $\mathcal{H}^{N-1}$-measure on 'orthogonal' hyperplanes remains absolutely continuous w.r.t. the Lebesgue measure. Suppose $\mathcal{H}^{N-1}(Z(h^-)) > 0$. The inequality

$$\left(\frac{h^+ - t}{h^+ - h^-}\right) \mathcal{H}^{N-1}(Z(h^-)) \le \mathcal{H}^{N-1}\big(Z(t)\big) \tag{2.10}$$

shows that the $\mathcal{H}^{N-1}$ measure will not shrink to 0 if the distance of $b(Z)$ from $\sigma_d^s Z$ is not zero. Then the set of initial and end points, $\cup_x a(x) \cup b(x)$, is $\mathcal{H}^N$-negligible (Lem. 2.16). As a consequence, we can cover $\mathcal{H}^N$-almost all $\mathcal{T}_{\mathtt{e}}$ with countably many d-cylinders — of positive $\mathcal{H}^N$-measure if $\mathcal{T}_{\mathtt{e}}$ has positive $\mathcal{H}^N$-measure.

2.3.2. *Disintegration of the Lebesgue measure on $\mathcal{Z}$.* This section collects consequences of the fundamental estimates of Lemma 2.13. Firstly, the set of endpoints of transport rays is $\mathcal{H}^N$-negligible (Lem. 2.16). Then, we fix the attention on model d-cylinders. We explicit the fact that the push forward, w.r.t. the map σ^t, of the $\mathcal{H}^{N-1}$-measure on orthogonal hyperplanes remains absolutely continuous w.r.t. $\mathcal{H}^{N-1}$. This also allows to change variables, in order to pass from $\mathcal{L}^N$-measurable functions on d-cylinders to $\mathcal{L}^N$-measurable functions on usual cylinders. Some regularity properties of the above density are presented. Finally, the estimate leads to the explicit disintegration of the Lebesgue measure, first on the model sets and then, since the endpoints are negligible, on the whole transport set (Th. 2.21).

Remark 2.15. We underline that the results of this subsection are, more generally, based on the following ingredients: we are considering the image set of a piecewise Lipschitz semigroup, which satisfies the absolutely continuous push forward estimate of Lemma 2.13.

Lemma 2.16. *The set of endpoints of transport rays is negligible: $\mathcal{L}^N(\mathcal{T}_{\mathtt{e}} \setminus \mathcal{T}) = 0$.*

Proof. We analyze just $\mathcal{A} = \cup_x a(x)$, the other case is symmetric. Suppose $\mathcal{H}^N(\mathcal{A}) > 0$. Consider a Lebesgue point of both the sets $\mathcal{A}$ and $\mathcal{L}$, say the origin. Since we have the decomposition of Subs. 2.2, it is enough to prove the negligibility e.g. of the initial points of the set $\mathcal{L}$ where $d \in B_\eta(e_1)$, for some small $\eta > 0$, and $\mathcal{H}^1(\mathcal{P}(x) \cdot e_1) > 1$. For every $\varepsilon > 0$, then, and every r sufficiently small, there exists $T \subset [0, r]$ with $\mathcal{H}^1(T) > (1-\varepsilon)r$ such that for all $\lambda \in T$

$$\mathcal{H}^{N-1}(H_\lambda) \geq (1-\varepsilon)r^{N-1} \qquad \text{where } H_\lambda = \mathcal{L} \cap \mathcal{A} \cap \{x \cdot e_1 = \lambda,\ |x - \lambda e_1|_\infty \leq r\}. \tag{2.11}$$

Choose, now, $s < t$, both in T, with $|t - s| < \varepsilon r$. By Lemma 2.13, then

$$\mathcal{H}^{N-1}(\sigma^{t-s} H_s) \geq \left(\frac{1-t}{1-s}\right)^{N-1} \mathcal{H}^{N-1}(H_s) \overset{(2.11)}{\geq} (1-2\varepsilon) r^{N-1}.$$

Moreover, since $d \in B_\eta(e_1)$, we have that $\mathcal{H}^{N-1}\big(\sigma^{t-s} H_s \setminus \{x \cdot e_1 = t,\ |x - te_1|_\infty \leq r\}\big) \leq 2\eta r^{N-1}$. Since points in $\sigma^{t-s} H_s$ do not stay in $\mathcal{A}$, then we reach a contradiction with the estimate (2.11) for $\lambda = t$: we would have

$$r^{N-1} = \big|\{x \cdot e_1 = t,\ |x - te_1|_\infty \leq r\}\big| \geq (1-\varepsilon) r^{N-1} + (1 - 2\varepsilon - 2\eta) r^{N-1} = (2 - 3\varepsilon - 2\eta) r^{N-1}. \qquad \square$$

Lemma 2.17. *With the notations of Lemma 2.13, the push forward of the measure $\mathcal{H}^{N-1} \restriction_Z$ by the map σ^t can be written as*

$$\sigma^t_\sharp \mathcal{H}^{N-1} \restriction_Z (y) = \alpha^t(y) \mathcal{H}^{N-1} \restriction_{\sigma^t Z} (y), \qquad (\sigma^{-t})_\sharp \mathcal{H}^{N-1} \restriction_{\sigma^t Z} (y) = \frac{1}{\alpha^t(\sigma^t y)} \mathcal{H}^{N-1} \restriction_Z (y).$$

Moreover, when $h^- < 0 < h^+$, then one has uniform bounds on the $\mathcal{H}^{N-1}$-measurable function α^t:

$$\begin{aligned} &\left(\frac{h^+ - t}{h^+}\right)^{N-1} \leq \frac{1}{\alpha^t} \leq \left(\frac{t - h^-}{-h^-}\right)^{N-1} && \textit{for } t \geq 0, \\ &\left(\frac{t - h^-}{-h^-}\right)^{N-1} \leq \frac{1}{\alpha^t} \leq \left(\frac{h^+ - t}{h^+}\right)^{N-1} && \textit{for } t < 0. \end{aligned}$$

Proof. Lemma 2.13 ensures that the measures $\sigma^t_\sharp \mathcal{H}^{N-1} \restriction_Z$ and $\mathcal{H}^{N-1} \restriction_{\sigma^t Z}$ are absolutely continuous one with respect to the other. Radon-Nikodym theorem provides the the existence of the above function α^t, which is the Radon-Nikodym derivative of $\sigma^t_\sharp \mathcal{H}^{N-1} \restriction_Z$ w.r.t. $\mathcal{H}^{N-1} \restriction_{\sigma^t Z}$. For the inverse mapping σ^{-t}, the Radon-Nikodim derivative is instead $\alpha^t(\sigma^t(y))^{-1}$. The last estimate, then, is straightforward from Lem. 2.13, with $s = 0$. $\square$

Corollary 2.18. *The map $\sigma^t(x) : [h^-, h^+] \times Z \mapsto \mathcal{Z}$ is invertible, linear in t and Borel in x (thus Borel in (t, x)). It induces also an isomorphism between the $\mathcal{L}^N$-measurable functions on $[h^-, h^+] \times Z$ and on $\mathcal{Z}$, since images and inverse images of $\mathcal{L}^N$-zero measure sets are $\mathcal{L}^N$-negligible.*

Proof. What has to be proved is that the maps σ^t, $(\sigma^t)^{-1}$ bring null measure sets into null measure sets. We show just one verse, the other one is similar. By direct computation, if $N \subset \mathcal{Z}$ is $\mathcal{H}^N$-negligible, then

$$0 = \int_{\mathcal{Z}} \mathbb{1}_N(y) d\mathcal{H}^N(y) = \int_{h^-}^{h^+} \left\{ \int_{\sigma^t(Z)} \mathbb{1}_N(y) d\mathcal{H}^{N-1}(y) \right\} dt = \int_{h^-}^{h^+} \left\{ \int_Z \frac{\mathbb{1}_N}{\alpha^t}(\sigma^t y) d\mathcal{H}^{N-1}(y) \right\} dt.$$

Consequently, being α^t positive, for $\mathcal{H}^1$-a.e. t we have that $\mathcal{H}^{N-1}\big(\{y \in Z :\ \sigma^t(y) \in N\}\big) = 0$. Therefore

$$\mathcal{H}^N\big((\sigma^t)^{-1} N\big) = \int_{[h^-, h^+] \times Z} \mathbb{1}_{(\sigma^t)^{-1} N} d\mathcal{H}^N = \int_{h^-}^{h^+} \left\{ \int_{\{y \in Z:\ \sigma^t(y) \in N\}} d\mathcal{H}^{N-1}(y) \right\} dt = 0. \qquad \square$$

In particular, define $\tilde{\alpha}(t, y) := \frac{1}{\alpha^t(\sigma^t y)}$. In the following, $\tilde{\alpha}$ will enter in the main theorem, the explicit disintegration of the Lebesgue measure. Before proving it, we remark some regularity and estimates for this density — again consequence of the fundamental estimate.

Corollary 2.19. *The function $\tilde{\alpha}(t, y) = \frac{(\sigma^{-t})_\sharp \mathcal{H}^{N-1} \restriction_{\sigma^t Z}}{\mathcal{H}^{N-1} \restriction_Z}$ is measurable in y, locally Lipschitz in t (thus measurable in (t, y)). Moreover, consider any $\mathtt{a}, \mathtt{b}$ drawing a sub-ray through y, possibly converging to*

$a(y)$, $b(y)$. Then, the following estimates hold for $\mathcal{H}^{N-1}$-a.e. $y \in Z$:

$$-\left(\frac{N-1}{\mathsf{b}(y)\cdot e_1 - t}\right)\tilde{\alpha}(t,y) \le \frac{d}{dt}\tilde{\alpha}(t,y) \le \left(\frac{N-1}{t-\mathsf{a}(y)\cdot e_1}\right)\tilde{\alpha}(t,y), \tag{2.12}$$

$$\left(\frac{|\mathsf{b}(y)-\sigma^t y|}{|\mathsf{b}(y)-y|}\right)^{N-1}(-1)^{\mathbb{1}_{t<0}} \le \tilde{\alpha}(t,y)(-1)^{\mathbb{1}_{t<0}} \le \left(\frac{|\sigma^t y-\mathsf{a}(y)|}{|y-\mathsf{a}(y)|}\right)^{N-1}(-1)^{\mathbb{1}_{t<0}}. \tag{2.13}$$

Moreover,

$$\int_{\mathsf{a}\cdot e_1}^{\mathsf{b}\cdot e_1}\left|\frac{d}{dt}\tilde{\alpha}(t,y)\right| \le 2\left(\frac{|\mathsf{b}-\mathsf{a}|^{N-1}}{|\mathsf{b}|^{N-1}}+\frac{|\mathsf{b}-\mathsf{a}|^{N-1}}{|\mathsf{a}|^{N-1}}-1\right). \tag{2.14}$$

Proof. Applying Lem. 2.13 and Cor. 2.18, for $h^- < s < t < h^+$ and every measurable $S \subset Z$, we have

$$\left(\frac{h^+-t}{h^+-s}\right)^{N-1}\int_S \tilde{\alpha}(s,y)d\mathcal{H}^{N-1}(y) \le \int_S \tilde{\alpha}(t,y)d\mathcal{H}^{N-1}(y) \le \left(\frac{t-h^-}{s-h^-}\right)^{N-1}\int_S \tilde{\alpha}(s,y)d\mathcal{H}^{N-1}(y).$$

As a consequence, there is a dense sequence $\{t_i\}_{i\in N}$ in (h^-,h^+), such that, for $\mathcal{H}^{N-1}$-a.e. $y \in Z$, the following Lipschitz estimate holds ($t_j \ge t_i$):

$$\left[\left(\frac{h^+-t_j}{h^+-t_i}\right)^{N-1}-1\right]\tilde{\alpha}(t_i,y) \le \tilde{\alpha}(t_j,y)-\tilde{\alpha}(t_i,y) \le \left[\left(\frac{t_j-h^-}{t_i-h^-}\right)^{N-1}-1\right]\tilde{\alpha}(t_i,y).$$

Therefore, there is a locally Lipschitz extension of $\tilde{\alpha}(t_i,y)$ from $\{t_i\}$ to (h^-,h^+). By the above integral estimate this limit function, at any t, must be a representative of the $\mathcal{L}^1(\mathcal{H}^{N-1})$ function $\tilde{\alpha}(t,y)$ — just take $t \to s^+$. By the above pointwise estimate, taking the derivative, we get (2.12). Eq. 2.12, moreover, implies the following monotonicity:

$$\frac{d}{dt}\left(\frac{\tilde{\alpha}(t,y)}{(e_1\cdot\mathsf{b}-t)^{n-1}}\right) \ge 0 \qquad \text{and} \qquad \frac{d}{dt}\left(\frac{\tilde{\alpha}(t,y)}{(t-e_1\cdot\mathsf{a})^{n-1}}\right) \le 0.$$

Then, since $\frac{e_1\cdot\mathsf{b}-t}{e_1\cdot\mathsf{b}} = \frac{|\mathsf{b}-\sigma^t y|}{|\mathsf{b}-y|}$, $\frac{t-e_1\cdot\mathsf{a}}{-e_1\cdot\mathsf{a}} = \frac{|\mathsf{a}-\sigma^t y|}{|\mathsf{a}-y|}$ and $\tilde{\alpha}(0,\cdot)\equiv 1$, we obtain exactly (2.13). Furthermore

$$\begin{aligned}
\int_{\mathsf{a}\cdot e_1}^0\left|\frac{d}{dt}\tilde{\alpha}(t,y)\right|dt &\overset{(2.12)}{\le} \int_{\left\{\frac{d\tilde{\alpha}(t,y)}{dt}>0\right\}\cap\{t<0\}}\frac{d}{dt}\tilde{\alpha}(t,y)\,dt + \int_{\mathsf{a}\cdot e_1}^0\frac{(N-1)\tilde{\alpha}(t,y)}{\mathsf{b}\cdot e_1-t}\,dt\\
&\overset{(2.12)}{\le} \int_{\mathsf{a}\cdot e_1}^0\frac{d}{dt}\tilde{\alpha}(t,y)\,dt + 2\int_{\mathsf{a}\cdot e_1}^0\frac{(N-1)\tilde{\alpha}(t,y)}{\mathsf{b}\cdot e_1-t}\,dt\\
&\overset{(2.13)}{\le} 1+2\int_{\mathsf{a}\cdot e_1}^0\frac{(N-1)(e_1\cdot\mathsf{b}-t)^{N-2}}{(e_1\cdot\mathsf{b})^{N-1}}\,dt = 1+2\left(\frac{|\mathsf{b}-\mathsf{a}|^{N-1}}{|\mathsf{b}|^{N-1}}-1\right).
\end{aligned}$$

Summing the symmetric estimate on $(0,\mathsf{b}\cdot e_1)$, we get

$$\int_{\mathsf{a}\cdot e_1}^{\mathsf{b}\cdot e_1}\left|\frac{d}{dt}\tilde{\alpha}(t,y)\right| \le 2\left(\frac{|\mathsf{b}-\mathsf{a}|^{N-1}}{|\mathsf{b}|^{N-1}}+\frac{|\mathsf{b}-\mathsf{a}|^{N-1}}{|\mathsf{a}|^{N-1}}-1\right). \qquad \square$$

We present now the disintegration of the Lebesgue measure, first on a model set, then on the whole transport set.

Lemma 2.20. *On $\mathcal{K} = \{\sigma^t Z\}_{t\in(h^-,h^+)}$, we have the following disintegration of the Lebesgue measure*

$$\int_{\mathcal{K}}\varphi(x)\,d\mathcal{L}^N(x) = \int_{\mathsf{y}\in Z}\left\{\int_{h^-}^{h^+}\varphi(\sigma^t y)\tilde{\alpha}(t,y)\,d\mathcal{H}^1(t)\right\}d\mathcal{H}^{N-1}(y), \tag{2.15}$$

where $\tilde{\alpha}(t,\cdot)$ is the Radon-Nykodim derivative of $(\sigma^{-t})_\sharp\mathcal{H}^{N-1}\lceil_{\sigma^t Z}$ w.r.t. $\mathcal{H}^{N-1}\lceil_Z$.

Proof. Consider any integrable function φ. Then, since $(\sigma^{-t})_\sharp\mathcal{H}^{N-1}\lceil_{\sigma^t S} = \tilde{\alpha}(t,\cdot)\mathcal{H}^{N-1}\lceil_Z$ and since $\phi\circ\sigma^t\lceil_Z$ is still $\mathcal{L}^N$-measurable (Cor. 2.18), we have

$$\int_Z\varphi(\sigma^t y)\tilde{\alpha}(t,y)\,d\mathcal{H}^{N-1}(y) = \int_{\sigma^t Z}\varphi(y)\,d\mathcal{H}^{N-1}(y) = \int_{\mathcal{K}\cap\{x\cdot e_1=t\}}\varphi(y)\,d\mathcal{H}^{N-1}(y).$$

Integrating this equality, for $t \in (h^-, h^+)$

$$\int_{\mathcal{K}} \varphi(y)\, d\mathcal{L}^N(y) = \int_{h^-}^{h^+} \int_{\mathcal{K}\cap\{x\cdot e_1 = t\}} \varphi(y)\, d\mathcal{H}^{N-1}(y)\, dt = \int_{h^-}^{h^+} \int_Z \varphi(\sigma^t y)\tilde{\alpha}(t,y)\, d\mathcal{H}^{N-1}(y)\, dt.$$

Finally, since $\tilde{\alpha}$ is measurable (Cor. 2.19) and locally integrable, by the above estimate and Tonelli theorem applied to the negative and positive part, Fubini theorem provides the thesis. ☐

We present now the main theorem. The notation is the following.

Partition the transport set $\mathcal{T}_{\mathfrak{e}}$ into sheaf sets $\mathcal{Z}_i$ as in Subs. 2.2, let $\{Z_i, \mathfrak{d}_i\}$ be a set of bases and relative axes. Denote with $\mathcal{S}$ the quotient set of $\mathcal{T}_{\mathfrak{e}}$ w.r.t. the membership to transport rays, identified with $\cup_i Z_i$.

Consider the map moving points along rays, $\sigma^t(x) = x + \frac{t}{d(x)\cdot\mathfrak{d}_i} d(x)$ on $\mathcal{Z}_i$. Let $\tilde{\alpha}(t,\cdot)$, be the Radon-Nykodim derivative of $(\sigma^{-t})_\sharp \mathcal{H}^{N-1} \llcorner_{\sigma^t Z_i}$ w.r.t. $\mathcal{H}^{N-1} \llcorner_{Z_i}$ and set $c(t,y) = \tilde{\alpha}\big(d(y)\cdot(t\mathfrak{d}_i - y), y\big) d(y)\cdot\mathfrak{d}_i$.

Then, we have the following result.

Theorem 2.21. *One has then the following disintegration of the Lebesgue measure on $\mathcal{T}_{\mathfrak{e}}$*

$$\int_{\mathcal{T}_{\mathfrak{e}}} \varphi(x)\, d\mathcal{L}^N(x) = \int_{y\in\mathcal{S}} \left\{ \int_{a(y)\cdot d(y)}^{b(y)\cdot d(y)} \varphi(y + (t - y\cdot d(y))d(y)) c(t,y) d\mathcal{H}^1(t) \right\} d\mathcal{H}^{N-1}(y), \tag{2.16}$$

where $\mathcal{S}$, defined above, is a countable union of σ-compact subsets of hyperplanes.

Remark 2.22. As a consequence of Cor. 2.19, c is measurable in y and locally Lipschitz in t.

Remark 2.23 (Dependence on the partition). Suppose to partition the transport set in a different family of sheaf sets $\mathcal{Z}'_i$, with the quotient space identified with the union $\mathcal{S}'$ of the new basis. Then, one can refine the partitions $\{\mathcal{Z}_i\}_i$ and $\{\mathcal{Z}'_i\}_i$ into a family of sheaf sets $\{\widehat{\mathcal{Z}}_i\}_i$. Consider the change of variables in a single sheaf set $\widehat{\mathcal{Z}}_i$. If we consider $Z \subset \{x\cdot v + c = 0\}$ and $Z' \subset \{x\cdot v' + c'\} = 0$, then

$$y + \big(t - y\cdot d(y)\big)d(y) = y' + \big(t' - y'\cdot d(y)\big)d(y) \quad \text{with} \quad y' = y - \frac{c' + y\cdot v'}{d(y)\cdot v'} d(y), \qquad t' = t.$$

Moreover, we have the disintegration formulas

$$\begin{aligned}\int_{\mathcal{Z}} \varphi(x)\, d\mathcal{L}^N(x) &= \int_{y\in Z} \left\{ \int_{a(y)\cdot d(y)}^{b(y)\cdot d(y)} \varphi(y + (t - y\cdot d(y))d(y)) c(t,y) d\mathcal{H}^1(t) \right\} d\mathcal{H}^{N-1}(y) \\ &= \int_{y\in Z'} \left\{ \int_{a(y)\cdot d(y)}^{b(y)\cdot d(y)} \varphi(y + (t - y\cdot d(y))d(y)) c'(t,y) d\mathcal{H}^1(t) \right\} d\mathcal{H}^{N-1}(y),\end{aligned}$$

where c is the density relative to Z, c' to Z'. The relation between the two densities c, c' is the following:

$$c'(t,x) = c(t, T^{-1}x)\beta(x),$$

where we denote with T the map from Z to Z' and with β the following Radon-Nicodym derivative

$$T(t) := y - \frac{c' + y\cdot v'}{d(y)\cdot v'} d(y) \qquad \beta := \frac{dT_\sharp \mathcal{H}^{N-1} \llcorner_Z}{d\mathcal{H}^{N-1} \llcorner_{Z'}}.$$

Proof. Forget the set of initial and end points, since by Lemma 2.16 they are negligible. Consider another partition of the transport set $\mathcal{T}$, the one, given in Lem. 2.8, into cylinders subordinated to d: let

$$\begin{aligned}\mathcal{K}_i &= \left\{ \hat{\sigma}^t(\widehat{Z}_i) : \ t \in [h_i^-, h_i^+] \right\} = \left\{ y + \frac{t d(y)}{d(y)\cdot\hat{\mathfrak{d}}_i} : \quad y \in \widehat{Z}_i, \ t \in [h_i^-, h_i^+] \right\} \\ &= \left\{ y : \quad h^- \le y\cdot\hat{\mathfrak{d}}_i \le h^+ \right\} \bigcap \bigcup_{x\in\widehat{Z}_i} [\![a(x), b(x)]\!].\end{aligned}$$

By construction, moreover, the faces $\sigma^{h_i^{\pm}}(\widehat{Z}_i)$ are artificial, in the sense that, given $\{\mathcal{K}_{jk\ell m}\}$ there is a reordering $\{\mathcal{K}^{\pm}_{j(k+1)\ell m}\}$ of $\{\mathcal{K}^{\pm}_{j(k+1)\ell m}\}$ such that we have

$$\bigcup_\ell \sigma^{h^+}(\widehat{Z}^-_{j(k+1)\ell m}) = \bigcup_\ell \sigma^{h^-}(\widehat{Z}_{jk\ell m}) \qquad \text{and} \qquad \bigcup_\ell \sigma^{h^-}(\widehat{Z}^+_{j(k+1)\ell m}) = \bigcup_\ell \sigma^{h^+}(\widehat{Z}_{jk\ell m})$$
$$\bigcup_{jk\ell m} \mathcal{K}_{j(k+1)\ell m} = \mathcal{T}. \tag{2.17}$$

The original partition into $\mathcal{Z}_i$ essentially, collects families of the above d-cylinders. By the local result of Lem. 2.20, after the translation of the origin and the change of variable $t \to \frac{t}{d(y)\cdot\hat{\mathfrak{d}}_i}$, one has

$$\int_{\mathcal{K}_i} \varphi(x)\, d\mathcal{H}^N(x) = \int_{y\in\widehat{Z}_i} \left\{ \int_{h^-}^{h^+} \varphi\big(y+(t-y\cdot d(y))d(y)\big)\tilde{\alpha}\big(d(y)\cdot(t\hat{\mathfrak{d}}_i - y), y\big) d(y)\cdot\hat{\mathfrak{d}}_i\, d\mathcal{H}^1(t) \right\} d\mathcal{H}^{N-1}(y).$$

Define

$$c(t,y) = \tilde{\alpha}\big(d(y)\cdot(t\hat{\mathfrak{d}}_i - y), y\big) d(y)\cdot\hat{\mathfrak{d}}_i.$$

Trivially, then, one extends the result in the whole domain

$$\begin{aligned}
\int_{\mathcal{T}} \varphi(x)\, d\mathcal{H}^N(x) &= \int_{\cup_i \mathcal{K}_i} \varphi(x)\, d\mathcal{H}^N(x) = \sum_i \int_{\mathcal{K}_i} \varphi(x)\, d\mathcal{H}^N(x) \\
&= \sum_i \int_{y\in\widehat{Z}_i} \left\{ \int_{h^-}^{h^+} \varphi(y+(t-y\cdot d(y))d(y)) c(t,y)\, d\mathcal{H}^1(t) \right\} d\mathcal{H}^{N-1}(x) \\
&= \int_{y\in\cup_i\widehat{Z}_i} \left\{ \int_{h^-}^{h^+} \varphi(y+(t-y\cdot d(y))d(y)) c(t,y)\, d\mathcal{H}^1(t) \right\} d\mathcal{H}^{N-1}(x) \\
&\overset{(2.17)}{=} \int_{y\in\cup_i Z_i} \left\{ \int_{a(x)}^{b(x)} \varphi(y+(t-y\cdot d(y))d(y)) c(t,y)\, d\mathcal{H}^1(t) \right\} d\mathcal{H}^{N-1}(x).
\end{aligned}$$

Separating the positive and the negative part of φ, in the very last step the convergence is monotone and does not give any problem. □

2.4. **Remarks on the divergence of the rays' direction.** In this section we extend the function d to be null out of $\mathcal{T}$. We consider then its divergence, which is defined as the distribution

$$\langle \mathrm{div} d, \varphi\rangle = \int_{\mathcal{T}} \nabla\varphi\cdot d\, dx \qquad \forall \varphi\in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^N).$$

If the set of initial points, or final points as well, is compact, then it turns out to be a Radon measure concentrated on $\mathcal{T}$, without the endpoints. More generally, it is a series of measures (see examples in Sect. 4). A decomposition of it can be constructed as follows. Consider a partition of $\mathcal{T}$ into tuft sets $\{\mathcal{K}_i\}$, defined in 2.2. Fix the attention on one $\mathcal{K}_i$. Truncate the rays with an hyperplane just before they enter $\mathcal{K}_i$, and take that intersection as the new source: you define a vector field $\hat{d}$ on $\mathbb{R}^N$ which coincides with d on $\mathcal{K}_i$. The i-th addend is defined as the divergence of this vector field $\hat{d}$, truncated on $\mathcal{K}_i$. It turns out that the absolutely continuous part of this divergences does not depend on the $\{\mathcal{K}_i\}$ we have chosen, as the limit: this happens, as well, for the distributional limit of the series — which is precisely div$d$.

2.4.1. *Local divergence.* In this subsection, we point out that, if the closure of the set of initial points is a negligible compact K, then the divergence of the vector field of directions, as a distribution on $\mathbb{R}^N\setminus K$, is a locally finite Radon measure. A similar statement holds when the closure of the set of terminal points is a negligible compact. This will then be used to approximate in some sense the divergence of the original vector field d. For the moment, we notice that it gives a coefficient of an ODE for the density c defined in the previous section.

Definition 2.24. Fix the attention on a d-cylinder with bounded basis $\mathcal{K} = \{\sigma^t(Z) : t\in(h^-,h^+)\}$, assume Z compact. Suppose, moreover, that for $\mathcal{L}^N$-a.e. $x\in\mathcal{K}$ the ray $\mathcal{R}(x)$ intersects also the compact

$K = \sigma^{h^- - \varepsilon}(Z)$. Let $\{\mathtt{a}_i\}$ be dense in K. Consider the potential given by

$$\hat{\phi}(x) = \max_{\mathtt{a}\in K} \Big\{ \phi(a) - |x-a|_{D^*} \Big\}$$

and define $\hat{d}$ as the relative vector field of rays' directions.

Lemma 2.25. *The vector field $\hat{d}$ is defined out of K, single valued on $\mathcal{H}^N$-a.a. $\mathbb{R}^N$. Moreover, on $\mathbb{R}^N - K$, its divergence is a locally finite Radon measure.*

Proof. Since K is compact, since the continuous function $\phi(a) - |x-a|_{D^*}$ must attain a minimum on K, then the transport set $\mathcal{T}_{\mathtt{e}}$ is at least $\mathbb{R}^N \setminus K$. Moreover, K is $\mathcal{H}^N$-negligible, being contained in a hyperplane. Therefore the vector field of directions $\hat{d}$ is $\mathcal{H}^N$-a.e. defined and single valued on $\mathbb{R}^N$. Furthermore, by definition it coincides with d on $\mathcal{K}$. The regularity of the divergence, which in general should be only a distribution, is now proved by approximation.

As in Ex. 2.12, we see that the potentials

$$\hat{\phi}_I(x) = \max_{j=i,\dots,I} \Big\{ \phi(\mathtt{a}_j) - |x-\mathtt{a}_j|_{D^*} \Big\}$$

increases to $\hat{\phi}$. Moreover, the corresponding vector field of directions

$$d_I(x) = \sum_{i=1}^{I} d^i(x) \mathbb{1}_{\Omega_i}(x) \qquad d^i(x) = \frac{x-\mathtt{a}_i}{|x-\mathtt{a}_i|},$$

with

$$\Omega_i = \Big\{ x: \ |x-\mathtt{a}_i|_{D^*} > |x-\mathtt{a}_j|_{D^*},\ j\in\{1\dots I\}\setminus i \Big\}, \qquad J_I = \bigcup_i \partial\Omega_i \quad (\mathcal{H}^{N-1} \text{ count. recti.}),$$

converges p.w. $\mathcal{H}^N$-a.e. to $\hat{d}$. By d_I's membership in BV, the distribution $\mathrm{div} d_I$ is a Radon measure: we have thus

$$\langle \mathrm{div} d_I, \varphi \rangle = -\int \nabla\varphi\cdot d_I = \int \varphi \,\mathrm{div} d_I \qquad \forall \varphi \in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^N).$$

By the explicit expression, we have that the singular part is negative and concentrated on J_I:

$$\mathrm{div} d_I = \sum_i \frac{N-1}{|x-\mathtt{a}_i|} \mathcal{L}^2 \llcorner_{\Omega_i}(x) + \left(\frac{x-\mathtt{a}_j}{|x-\mathtt{a}_j|} - \frac{x-\mathtt{a}_i}{|x-\mathtt{a}_i|} \right) \cdot \nu_{ij} \mathcal{H}^1 \llcorner_{\partial\Omega_i\cap\partial\Omega_j}(x).$$

Moreover, integrating on each piece, one can compute directly, as in [7] Prop. 4.6, the estimates

$$(\mathrm{div} d_I)_{\mathrm{a.c.}}(x) \le \frac{N-1}{\mathrm{dist}(x, \cup_i \mathtt{a}_i)},$$

$$|\mathrm{div} d_I|(B_r(x)) \le |\partial B_r(0)| + \frac{2(N-1)|B_r(x)|}{\mathrm{dist}(B_r(x), \cup_i \mathtt{a}_i)} \qquad \text{for } B_r(x)\cap K = \emptyset.$$

In particular, restrict $\mathrm{div} d_I$ on open sets O_k increasing to $\mathbb{R}^N \setminus K$. By compactness, the measures $\mathrm{div} d_I \llcorner_{O_k}$ should converge weakly*, up to subsequence, to a locally finite Radon measure μ. Nevertheless, the whole sequence converges and the limit measure is defined on $\mathbb{R}^N\setminus K$, since μ must coincide with the divergence of the vector field $\hat{d}$: for all $\varphi\in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^N\setminus K)$

$$\langle \mathrm{div}\hat{d}, \varphi\rangle = -\int \nabla\varphi\cdot\hat{d} = \lim_I -\int \nabla\varphi\cdot d_I = \lim_I \int \varphi\,\mathrm{div} d_I = \int \varphi\, d\mu.$$

In particular, this proves that $\mathrm{div}\hat{d}$, in $\mathscr{D}(\mathbb{R}^N\setminus K)$, is a locally finite Radon measure. ☐

Lemma 2.26. *Let $\mathcal{K}$ be the d-cylinder fixed above for defining $\hat{d}$ (Def. 2.24). Consider any couple $\mathcal{S}, c$ as in the disintegration Th. 2.21. Then, for any d-sub-cylinder $\mathcal{K}'$ of $\mathcal{K}$, the following formulae hold:*

$$\partial_t c(t,y) - \Big[\big(\mathrm{div}\hat{d}\big)_{\mathrm{a.c.}}(y + (t - d(y)\cdot y) d(y)) \Big] c(t,y) = 0 \qquad \mathcal{H}^N\text{-a.e. on } \mathcal{K}. \tag{2.18}$$

$$\int_{\mathcal{K}'} \varphi\, \mathrm{div}\hat{d}_i = \int_{\mathcal{K}'} \varphi\,(\mathrm{div}\hat{d}_i)_{\mathrm{a.c.}} = -\int_{\mathcal{K}'} \nabla\varphi\cdot d + \int_{\partial\mathcal{K}'^+ - \partial\mathcal{K}'^-} \varphi d\cdot e_1 \qquad \forall\varphi\in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^N). \tag{2.19}$$

Proof. Since, by the previous lemma, the divergence of $\hat{d}$ is a measure, then we have the equality

$$-\int_{\mathbb{R}^N} \nabla\varphi \cdot \hat{d} = \langle \mathrm{div}\hat{d}, \varphi\rangle = \int_{\mathbb{R}^N} (\mathrm{div}\hat{d})_{\mathrm{a.c.}}\varphi + \int_{\mathbb{R}^N} \varphi(\mathrm{div}\hat{d})_{\mathrm{s}} \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^N \setminus K).$$

Moreover, $\hat{d}$ is the vector field of directions relative to a potential $\hat{\phi}$: we can apply the disintegration Theorem 2.21, getting (2.16) for a couple $\widehat{\mathcal{S}}, \hat{c}$. Notice that on $\mathcal{K}$, being $d = \hat{d}$, one can require $\widehat{\mathcal{S}} \restriction_{\mathcal{K}} \equiv \mathcal{S}$, which will lead to $\hat{c}(t,\cdot) \restriction_{\widehat{\mathcal{S}}} \equiv c(t,\cdot) \restriction_{\mathcal{S}}$. The local Lipschitz estimate on c (Rem. 2.22), since we are integrating on a compact support, allows the integration by parts in the t variable

$$\int_{\hat{a}(x)\cdot\hat{d}(y)}^{\hat{b}(y)\cdot\hat{d}(y)} \hat{c}(t,y)\nabla\varphi\big(y + (t - y\cdot\hat{d}(y))\hat{d}(y)\big)\cdot\hat{d}\,dt$$
$$= \varphi(\hat{b}(y))\hat{c}(b(y)\cdot\hat{d}(y),y) - \varphi(\hat{a}(y))\hat{c}(a(y)\cdot\hat{d}(y),y) - \int_{\hat{a}(x)\cdot\hat{d}(y)}^{\hat{b}(y)\cdot\hat{d}(y)} \varphi(y + (t - y\cdot\hat{d}(y))\hat{d}(y))\partial_t\hat{c}(t,y)\,dt;$$

after performing this, the above equality becomes

$$\int_{\widehat{\mathcal{S}}} \varphi(\hat{b}(y))\hat{c}(\hat{b}(y)\cdot\hat{d}(y),y)\,d\mathcal{H}^{N-1}(y) - \int_{\widehat{\mathcal{S}}} \varphi(\hat{a}(y))\hat{c}(\hat{a}(y)\cdot\hat{d}(y),y)\,d\mathcal{H}^{N-1}(y)$$
$$- \int_{\widehat{\mathcal{S}}}\int_{\hat{a}(x)\cdot\hat{d}(y)}^{\hat{b}(y)\cdot\hat{d}(y)} \Big[\varphi(y + (t - y\cdot\hat{d}(y))\hat{d}(y))\partial_t\hat{c}(t,y)\,dt\Big]\,d\mathcal{H}^{N-1}(y)$$
$$+ \int_{\widehat{\mathcal{S}}}\int_{\hat{a}(x)\cdot\hat{d}(y)}^{\hat{b}(y)\cdot\hat{d}(y)} \Big[\Big((\mathrm{div}\hat{d})_{\mathrm{a.c.}}\varphi\Big)(y + (t - y\cdot\hat{d}(y))\hat{d}(y))\hat{c}(t,y)\,dt\Big]\,d\mathcal{H}^{N-1}(y) + \int_{\mathbb{R}^N}\varphi(\mathrm{div}\hat{d})_{\mathrm{s}} = 0.$$

Moreover, since both $\hat{c}$ and $\partial_t\hat{c}$ are locally bounded, by the dominated convergence theorem this last relation holds also for bounded functions vanishing out of a compact — and in a neighborhood of K, the set of initial points for $\hat{d}$. By the arbitrariness of φ, this relation gives $\mathcal{H}^N$-a.e.

$$\partial_t c(t,y) - \Big[\big(\mathrm{div}\hat{d}\big)_{\mathrm{a.c.}}(y + (t - \hat{d}(y)\cdot y)\hat{d}(y))\Big]\hat{c}(t,y) = 0,$$

which turns out to be (2.18) on $\mathcal{K}$. Furthermore, on one hand we can notice that the singular part is concentrated on $\cup_{y\in K}\hat{b}(y) \cup \cup_{y\in K}\hat{a}(y)$, the endpoints w.r.t. the rays of $\hat{\phi}$. More precisely, denoting with $\hat{\sigma}^\pm$ the maps associating to each point in Z the relative initial or final point, we have that the singular part is given by $\hat{c}\,\hat{\sigma}^+_\sharp\mathcal{H}^{N-1} \restriction_{\widehat{\mathcal{S}}} - \hat{c}\,\hat{\sigma}^-_\sharp\mathcal{H}^{N-1} \restriction_{\widehat{\mathcal{S}}}$. On the other hand that, taking $\varphi = \mathbb{1}_{\mathcal{K}'}$, if Z is the relative section and $h^\pm$ define the height,

$$-\int_Z\int_{h^-}^{h^+} \Big[\varphi(y + (t - y\cdot d(y))d(y))\partial_t c(t,y)\,dt\Big]\,d\mathcal{H}^{N-1}(y)$$
$$+ \int_Z\int_{h^-}^{h^+} \Big[(\mathrm{div}d)_{\mathrm{a.c.}}\varphi(y + (t - y\cdot d(y))d(y))c(t,y)\,dt\Big]\,d\mathcal{H}^{N-1}(y) + \int_{\mathcal{K}'}\varphi(\mathrm{div}d)_{\mathrm{s}} = 0.$$

Coming back, integrating by parts again, one finds precisely (2.19). □

2.4.2. *Global divergence.* The divergence of the vector field d, generally speaking, is not a measure (see examples of Sect. 4). Nevertheless, it is not merely a distribution: it is a series of measures. Consider a covering of d-cylinders $\mathcal{K}_i$, as in Subs. 2.2. Repeat the construction of 2.4.1: one gets measures $\mathrm{div}\hat{d}_i$, which one can cut out of $\mathcal{K}_i$. The finite sum of this sequence of disjoint measures converges to $\mathrm{div}d$, in the sense of distribution. Actually, it turns out to be an absolutely continuous measure on the space of test functions vanishing on $\cup_x a(x) + b(x)$ — $\mathcal{H}^N$-negligible set that, nevertheless, can be dense in $\mathbb{R}^N$...

This construction could depend a priori on the decomposition $\{\mathcal{K}_i\}$ one has chosen. Notwithstanding, it turns out that this is not the case. In fact, the absolutely continuous part $(\mathrm{div}\hat{d})_{\mathrm{a.c.}}$ satisfies the following equation.

Lemma 2.27. *If one, just formally, defines on $\mathcal{T}$ the measurable function*

$$(\mathrm{div}d)_{\mathrm{a.c.}} := \sum_i (\mathrm{div}d_i)_{\mathrm{a.c.}}\mathbb{1}_{\mathcal{K}_i},$$

then, for any partition into d-cylinders as in Th. 2.21 with relative density c and sections $\mathcal{S}$, one has the relation

$$\partial_t c(t,y) - \Big[\big(\mathrm{div} d\big)_{\mathrm{a.c.}}(y+(t-d(y)\cdot y)d(y))\Big]c(t,y) = 0 \qquad \mathcal{H}^N\, z\text{-a.e. on } \mathcal{T}, \tag{2.20}$$

where $z = y + (t - d(y)\cdot y)d(y)$ with $y \in S$.

Remark 2.28. The measurable function $(\mathrm{div} d)_{\mathrm{a.c.}}$ in general does not define a distribution, since it can fail to be locally integrable (Ex. 4.3, 4.4).

Proof. Since the $\mathcal{K}_i$ are a partition of $\mathcal{T}$, and their bases are $\mathcal{H}^N$-negligible, then the statement — which is a pointwise relation — is a direct consequence of Lemm. 2.26. □

Remark 2.29. Since c does not depend on the construction of the vector fields $\hat{d}_i$, then Equation 2.20 ensures that $(\mathrm{div} d)_{\mathrm{a.c.}}$ is independent of the choices we made to obtain $\hat{d}_i$. Moreover, by Cor. 2.19, one has the bounds $-\frac{N-1}{\mathtt{b}(y)\cdot d(y)-t} \leq \big(\mathrm{div} d\big)_{\mathrm{a.c.}}(y+(t-d(y)\cdot y)d(y)) \leq \frac{N-1}{t-\mathtt{a}(y)\cdot d(y)}$.

Lemma 2.30. *We have the equality*

$$\mathrm{div} d = \sum_i \big(\mathrm{div}\hat{d}_i\big)_{\mathrm{a.c.}} \llcorner_{\mathcal{K}_i} - \mathcal{H}^{N-1} \llcorner_{\partial\mathcal{K}_i^+} + \mathcal{H}^{N-1} \llcorner_{\partial\mathcal{K}_i^-}.$$

Therefore, $\forall \varphi \in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^N \setminus \cup_x a(x) \cup b(x))$,

$$\langle \mathrm{div}\hat{d}, \varphi\rangle = \int \varphi\, (\mathrm{div} d)_{\mathrm{a.c.}}. \tag{2.21}$$

Proof. Since $d = \hat{d}_i$ on $\mathcal{K}_i$, Eq. (2.19) can be rewritten as

$$\int_{\mathcal{K}_i} \nabla\varphi\cdot d = -\int_{\mathcal{K}_i} \varphi(\mathrm{div}\hat{d}_i)_{\mathrm{a.c.}} + \int_{\partial\mathcal{K}_i^+ - \partial\mathcal{K}_i^-} \varphi \qquad \forall\varphi\in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^N).$$

Using the partition in the proof of Lem. 2.21, it follows that the divergence of d is the sum of the above measures: $\forall\varphi \in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^N)$

$$-\langle \mathrm{div} d, \varphi\rangle = \int_{\mathcal{T}} \nabla\varphi\cdot d = \sum_i \int_{\mathcal{K}_i}\nabla\varphi\cdot d = -\sum_i\int_{\mathcal{K}_i}\varphi(\mathrm{div}\hat{d}_i)_{\mathrm{a.c.}} + \int_{\partial\mathcal{K}_i^+ - \partial\mathcal{K}_i^-}\varphi.$$

Eq. 2.21 follows from the fact that one can choose a partition $\mathcal{K}_i$ whose bases are outside of the support of φ. □

3. Sudakov Proof Completed

In the present section we deal with Sudakov theorem, with the additional assumption D^* strictly convex. Up to now, the transport set — which is really the set where mass can be moved — is partitioned in segments, the transport rays. By D^* strictly convexity, mass can just be rearranged inside transport rays. Moreover, repeting [7] we found in Sect. 2 an explicit disintegration of measures absolutely continuous w.r.t. the Lebesgue measure: from it, follows that the disintegrated measures are absolutely continuous w.r.t. $\mathcal{H}^1$ on the segments. Consequently, it is possible to show that Sudakov argument works, in this case: disintegrate the given measures μ, ν w.r.t. the transport rays. If $\mu \ll \mathcal{L}^N$, as just stated, μ's disintegrated measures are absolutely continuous w.r.t. $\mathcal{H}^1$. The well established one dimensional theory, then, leads to an optimal transport map on the rays, between the disintegrated measures. Gluing together these maps we get an optimal transport map for the original problem.

We start this section recalling basic theorems of optimal mass transport theory (see [4], [18], [15]).

Theorem 3.1. *The minimum of the Kantorovich problem is equal to*

$$\max\left\{\int_X \phi(x)\,d\mu(x) - \int_Y \phi(y)\,d\nu(y)\right\},$$

where the supremum runs among all pairs $\phi \in \mathrm{Lip}(\mathbb{R}^N)$ such that $\phi(x)-\phi(y) \leq |y-x|_{D^*}$ for all $x, y \in \mathbb{R}^N$. Any maximizer ϕ is called Kantorovich potential. Moreover, if π is an optimal transport plan with finite *cost and ϕ is a maximizer, then the equality $\phi(x)-\phi(y)=|y-x|_{D^*}$ holds on its support. Finally, a maximizer, for example, is given by*

$$\phi(x):=\inf_{\substack{n,\, x_n:=x,\\ i=1,\dots,n-1,\\ (x_i,y_i)\in\Gamma}}\left\{\sum_{i=0}^{n-1}\left[|y_i-x_{i+1}|_{D^*}-|y_i-x_i|_{D^*}\right]\right\},\qquad x\in\mathbb{R}^N,\ (x_0,y_0)\in\Gamma \text{ fixed.} \tag{3.1}$$

Corollary 3.2. *Consider the decomposition in rays relative to a Kantorovich potential ϕ. By Th. 3.1, we have $\phi(x)-\phi(y)=|y-x|_{D^*}$ for all (x,y) in the support of π, denoted by Γ . This means exactly that, if $(x,y)\in\Gamma$, then $y\in\mathcal{P}(x)$. In particular, $P_Y(\Gamma\cap\{x\}\times\mathbb{R}^N)\subset\mathcal{P}(x)$, while $P_X(\Gamma\cap\mathbb{R}^N\times\{y\})\subset\mathcal{P}^{-1}(y)$. Then*

$$\Gamma\cap\mathcal{R}(x)\times\mathcal{R}(x)=\Gamma\cap\mathcal{R}(x)\times\mathbb{R}^N=\Gamma\cap\mathbb{R}^N\times\mathcal{R}(x)$$

which implies, for all subsets $S\subset\mathcal{T}$,

$$\mu(\mathcal{Z}(S))=\pi(\mathcal{Z}(S)\times\mathbb{R}^N)=\pi(\mathcal{Z}(S)\times\mathcal{Z}_{\mathfrak{e}}(S))=\pi(\mathbb{R}^N\times\mathcal{Z}_{\mathfrak{e}}(S))=\nu(\mathcal{Z}_{\mathfrak{e}}(S)).$$

In the case ν is absolutely continuous, too, then, in the last equation, we can neglect the endpoints.

Theorem 3.3 (1-dimensional theory)**.** *([4], Th. 5.1) Let μ, ν be probability measures on $\mathbb{R}$, μ without atoms, and let*

$$G(x)=\mu((-\infty,x)),\qquad F(x)=\nu((-\infty,x))$$

be respectively the distribution functions of μ, ν. Then

- *the nondecreasing function $t:\mathbb{R}\mapsto\overline{\mathbb{R}}$ defined by*
$$t(x)=\sup\left\{y\in\mathbb{R}:\ F(y)\le G(x)\right\} \tag{3.2}$$
(with the convention $\sup\emptyset=-\infty$) maps μ into ν. Any other nondecreasing map t' such that $t'_\sharp\mu=\nu$ coincides with t on the support of μ up to a countable set.
- *If $\phi:[0,+\infty]\to\mathbb{R}$ is nondecreasing and convex, then t is an optimal transport relative to the cost $c(x,y)=\phi(|y-x|)$. Moreover, t is the unique optimal transport map, in the case ϕ is strictly convex.*

We present now the main statement, which corresponds to Sudakov theorem plus the assumption of D^* strictly convex.

Theorem 3.4. *Consider two probability measures $\mu,\nu\in\mathcal{P}(\mathbb{R}^N)$, suppose μ is absolutely continuous w.r.t. $\mathcal{L}^N$. Assume that the optimal cost of the transport is finite, for the cost function $c(x,y)=|y-x|_{D^*}$; consider then a Kantorovich potential ϕ. Construct the partition in transport rays given in Sect. 2.2, in order to have the disintegration Th. 2.21, and let p be the projection onto the quotient $\mathcal{S}$. Define the following one dimensional measures on the rays.*

- *In the case $\nu(\cup_x b(x))=0$, disintegrate μ and $\nu\restriction_{\mathcal{T}}$ w.r.t. the membership to a transport ray, let μ_y, ν_y be the disintegrated measures, for $y\in\mathcal{S}$.*
- *In the case $\nu(\cup_x b(x))\neq 0$, disintegrate μ, ν before w.r.t. the partition $\{\mathcal{P}^{-1}(b)\}_{b\in b(\mathcal{T})}$, let $\hat\mu_\alpha$ and $\hat\nu_\alpha\restriction_{\mathcal{T}}$ be the disintegrated measures. Disintegrate, then, each μ_α and $\nu_\alpha\restriction_{\mathcal{T}}$ w.r.t. the membership to a ray, neglecting the point $b(x_\alpha)$. This gives quotient measures $m(\alpha,y)$, $\tilde m(\alpha,y)$ and disintegrated measures μ_y, $\tilde\nu_y$, for $y\in\mathcal{S}$. Consider now the Radon-Nikodym derivatives*
$$f(\alpha,y)=\frac{m(\alpha,y)-\tilde m(\alpha,y)}{m(\alpha,y)}\qquad\text{and}\qquad g(\alpha,y)=\frac{\tilde m(\alpha,y)}{m(\alpha,y)}.$$
Let $\nu_y=f(\alpha,y)\delta_{b(x_\alpha)}+g(\alpha,y)\tilde\nu_{\alpha,y}$.

Define, on $\mathcal{T}$, two functions as follows.

$$F(z)=\mu_{p(z)}\big(\llparenthesis a(z),z\rrparenthesis\big),\qquad G(z)=\nu_{p(z)}\big(\llparenthesis a(z),z\rrparenthesis\big).$$

Then, an optimal transport map is given by

$$T:\ z\mapsto\begin{cases} z & \text{if } z\notin\mathcal{T}\\ x+td(x) & \text{where } t=\sup\{s:\ F(x+sd(x))\le G(z)\},\ \text{if } z\in\mathcal{T}\end{cases}. \tag{3.3}$$

Proof. Denote with $(\!|a(x), b(x)|\!)$ the transport ray passing through x. Let $\mathcal{S}$ be the quotient space and p the projection onto it.

`Step 1.` Suppose $\nu(\cup_x b(x)) = 0$. This, since μ is absolutely continuous, means also that $\mu(\cup_x a(x)) = \mu(\cup_x b(x)) = \nu(\cup_x a(x)) = 0$. In fact, whenever (x, y) belongs to the c-monotone set of a plan π, then there is a ray containing $(\!|x, y|\!)$. This is just the fact that we are transporting μ to ν, so mass is moved along the rays *from a to b*. In this case, we have that μ and ν are concentrated on $\mathcal{T}$, where the membership to a ray defines an equivalence relation. The disintegration theorem Th. A.6 gives then

$$\mu = \int_{y\in\mathcal{S}} \mu_y \, dm(y) \qquad \nu = \int_{y\in\mathcal{S}} \nu_y \, dm(y),$$

with

- the probability measures μ_y, ν_y concentrated on the one dimensional segment $(\!|a(y), b(y)|\!)$,
- $y \to \mu_y(K)$, $y \to \nu_y(K)$ Borel functions for all Borel $K \subset \mathbb{R}^N$,
- $m(S) = \mu(\mathcal{Z}(S)) = \nu(\mathcal{Z}(S)) = \pi(\mathcal{Z}(S) \times \mathcal{Z}(S))$ for all $S \subset \mathcal{S}$, for all $\pi \in \Pi(\mu, \nu)$ (Cor. 3.2).

Moreover, we are assuming that there exists an integrable function f such that

$$\mu = f\mathcal{L}^N.$$

By the explicit disintegration Th. 2.21, then, denoting $i(y) = \int_{a(y)\cdot d(y)}^{b(y)\cdot d(y)} f(y+(s-y\cdot d(y))d(y))c(s, y)d\mathcal{H}^1(s)$, we get

$$\mu_y = f_y\mathcal{H}^1 = \frac{f(y + (t - y\cdot d(y))d(y))c(t, y)}{i(y)}\mathcal{H}^1(t).$$

This was exactly the missing step in Sudakov proof, since one has to prove that the disintegrated measures of μ are absolutely continuous w.r.t. $\mathcal{H}^1$.

`Step 2.` By the one dimensional theory, Th. 3.3, an optimal transport map from $(\mathcal{R}(y), \mu_y)$ to $(\mathcal{R}(y), \nu_y)$ is exactly given by (3.3), restricted to $z \in (\!|a(y), b(y)|\!)$. The map T in (3.3) is Borel, not only on the rays, but in the whole $\mathcal{T}$. To see it, consider the countable partition of $\mathcal{T}$ into σ-compact sets $\mathcal{Z}(Z_k)$, by Lemma 2.6. In particular, a subset C of $\mathcal{T}$ is Borel if and only if its intersections with the $\mathcal{Z}(Z_k)$ are Borel. Moreover, composing T with the Borel change of variable given in Rem. 2.10, from the sheaf set $\mathcal{Z}$ we can reduce to $(0, 1)\times Z$, $d(x) = e_1$ and the map T takes the form $T(y) = y+(T\cdot e_1 - y\cdot e_1)e_1$. One, then, has just to prove that the map $T \cdot e_1$ is Borel: this map is monotone in the first variable, and Borel in the second; in particular, it is Borel on $(0, 1) \cdot Z$.

`Step 3.` Consider now any other transport plan $\pi \in \Pi(\mu, \nu)$. By Cor. 3.2, then, its support Γ is contained in $\cup_{y\in S}\mathcal{R}(y) \times \mathcal{R}(y) \cup \{x = y\}$. Moreover, we can forget about the points which stay in place, since they do not contribute to the cost of the transport. As a consequence, then, one applies again the disintegration theorem A.6, and disintegrates π in

$$\pi = \int_{y\in S} \pi_y dm(y),$$

with π_y concentrated on $\mathcal{R}(y) \times \mathcal{R}(y)$. Notice that the quotient measure m is the same as the one in the disintegration of μ and ν, by Cor. 3.2. Moreover, for m-a.e y the plan π_y necessarily transports μ_y to ν_y: for all measurable $S' \subset \mathcal{S}$, $A \subset \mathbb{R}^N$

$$\int_{S'} \pi_y(A \times \mathbb{R}^N)dm(y) = \pi(A \times \mathbb{R}^N \cap \mathcal{Z}(S') \times \mathcal{Z}(S')) \overset{\text{Cor. 3.2}}{=} \mu(A \cap \mathcal{Z}(S')) = \int_{S'} \mu_y(A)dm(y).$$

In particular, by the one-dimensional result,

$$\int_{\mathcal{R}(y)\times\mathcal{R}(y)} |x - y|_{D^*} \, d\pi_y \geq \int_{\mathcal{R}(y)} |x - T(x)|_{D^*} \, d\mu_y.$$

Therefore, one can conclude the optimality of T: in fact, for every $\pi \in \Pi(\mu, \nu)$

$$\int |x - y|_{D^*} \, d\pi = \int\left\{ \int |x - y|_{D^*} \, d\pi_y \right\} dm \geq \int\left\{ \int |x - T(x)|_{D^*} \, d\mu_y \right\} dm = \int |x - T(x)|_{D^*} \, d\mu.$$

This yields to the existence of an optimal transport map of the form

$$T = \mathrm{Id}_{\mathbb{R}^N\setminus\mathcal{T}} + \sum_{y\in S} T_y \mathbb{1}\lceil_{(\!|a(y),b(y)|\!)},$$

where T_y is a one-dimensional, optimal transport map from μ_y to ν_y, when $\nu(\cup_x b(x)) = 0$. Every other plan has the form $\pi = \int_{\mathcal{S}} \pi_y \, dm(y)$, with $\pi_y \in \Pi(\mu_y, \nu_y)$.

`Step 4.` Allow now $\nu(\cup_x b(x)) > 0$. In that case, ν is not supported on $\mathcal{T}$, thus the terminal points have to be taken into account. A partition of $\cup_x \llparenthesis a(x), b(x)\rrbracket$ is the following: $\{\mathcal{P}^{-1}(b)\}_{b\in b(\mathcal{T})}$. This partition does not identify just the points on a same ray, but all the points whose ray has the same terminal point. In this way, precisely as in the previous step, one gets a first disintegration $\mu_\alpha, \nu_\alpha, m'(\alpha), \pi_\alpha$. Do now a second step. Focus on a single equivalence class $\mathcal{P}^{-1}(b(x_\alpha))$; since the disintegrated measure μ_α is not carried by the terminal points, for m'-a.e. α, it is possible to disintegrate it w.r.t. the membership to a ray, neglecting the point $b(x_\alpha)$. This gives a quotient measure $m(\alpha, y)$, where α can be determined from y. One can do the same with $\nu_\alpha \restriction_{\mathcal{T}}$, getting a quotient measure $\tilde{m}(\alpha, y)$ and the disintegrated measures $\tilde{\nu}_\alpha$. The difference between the two quotient measures, $m(\alpha, y) - \tilde{m}(\alpha, y)$, must be the part of mass coming from the ray (α, y) which goes to the terminal point $b(x_\alpha)$. Consider now the Radon-Nikodym derivatives

$$f(\alpha, y) = \frac{m(\alpha, y) - \tilde{m}(\alpha, y)}{m(\alpha, y)} \qquad \text{and} \qquad g(\alpha, y) = \frac{\tilde{m}(\alpha, y)}{m(\alpha, y)}.$$

Therefore, define the following new 'disintegration' of ν, in order to have m as quotient measure:

$$\nu_{\alpha,y} = f(\alpha, y)\delta_{b(x_\alpha)} + g(\alpha, y)\tilde{\nu}_{\alpha,y}.$$

One can now repeat Steps 2 and 3. Coming back to the old notation, index the rays with $y \in \mathcal{S}$ instead of (α, y). Consider T_y optimal transport map from μ_y to ν_y and define $T = \mathrm{Id}_{\mathbb{R}^N \setminus \mathcal{T}} + \sum_{y\in S} T_y \mathbb{1}\restriction_{\llparenthesis a(y), b(y)\rrparenthesis}$. The measurability is analogous as above. Any other plan $\pi \in \Pi(\mu, \nu)$ can be disintegrated w.r.t. the partition $\{\llparenthesis a(y), b(y)\rrparenthesis \times \mathbb{R}^N\}_x$, in maps $\pi_y \in \Pi(\mu_y, \nu_y)$ and the same quotient measure m. The same inequalities as above show that $\int |x - y|_{D^*} \, d\pi \geq \int |x - T(x)|_{D^*} \, d\mu$, and thus the optimality of T. □

4. Remarks on the decomposition

Since ϕ is lipschitz, then it is $\mathcal{H}^N$-a.e. differentiable. At each point x where ϕ is differentiable, the Lipschitz inequality, just by differentiating along the segment from x to $x + d$, implies that

$$|\nabla\phi(x) \cdot d| \leq 1 \qquad \text{for all } d \in \partial D^*.$$

This means that $\pm\nabla\phi \in D$. Consider now a point where, moreover, there is an outgoing ray. As an immediate consequence of (2.1), just differentiating in the direction of the outgoing ray, we have the relation

$$-\nabla\phi(x) \cdot \frac{d(x)}{|d(x)|_{D^*}} = 1.$$

This means that $-\nabla\phi(x) \in \partial D$, and moreover

$$d \in \mathcal{D}(x) \quad \text{satisfies} \quad \frac{d(x)}{|d(x)|_{D^*}} \in \delta D(-\partial\phi(x)) \qquad (d(x) \in \partial\phi(x) \text{ if } D \text{ is a ball}). \tag{4.1}$$

Equation (4.1) suggests another possible definition of d. Assuming D^* strictly convex, $\delta D(-\nabla\phi(x))$ is single-valued. Therefore one could define for example

$$d(x) = \delta D(-\nabla\phi(x)), \qquad \text{where } -\nabla\phi(x) \in \partial D. \tag{4.2}$$

This, generally, extends the vector field we analyzed (see Ex. 4.3), and has analogous properties. However, even in this case the vector field of direction is not generally defined in positive $\mathcal{L}^N$-measure sets. In fact, in Ex. 4.4 we find that the gradient of ϕ can vanish in non negligible sets.

Before showing that, we recall the example in [3] where that the disintegration of the Lebesgue measure given in 2.21 is not for free, and requires some additional regularity of the vector field.

Example 4.1 (A Nikodym set in $\mathbb{R}^3$). In [3], Section 2, it is proved the following theorem.

Theorem 4.2. *There exist a Borel set $M_N \subset [-1, 1]^3$ with $|[-1, 1]^3 \setminus M_N| = 0$ and a Borel map $f : M_N \to [-2, 2]^2 \times [-2, 2]^2$ such that the following holds. If we define for $x \in M_N$ the open segment l_x connecting $(f_1(x), -2)$ to $(f_2(x), 2)$, then*

- *$\{x\} = l_x \cap M_N$ for all $x \in M_N$,*
- *$l_x \cap l_y = \emptyset$ for all $x, y \in M_N$ different.*

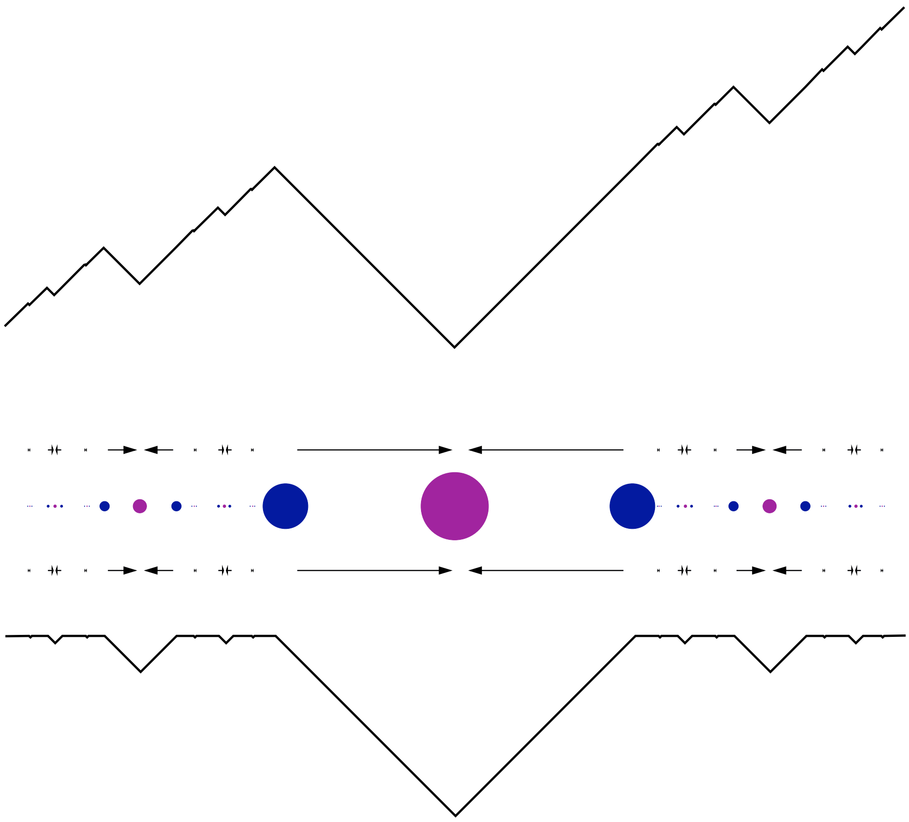

Figure 1: Example 4.3. With the potential above, the vector field of rays' direction is defined $\mathcal{L}^1$-a.e.. With the potential below, this is not the case, since the gradient vanishes in a $\mathcal{L}^1$-positive measure set.

Firstly, we recall that this example contradicts Prop. 78 in Sudakov proof ([16]): the disintegration of the Lebesgue measure on $[0,1]^3$ w.r.t. the segments l_x cannot be absolutely continuous w.r.t. the Hausdorff one dimensional measure on that segments. Secondly, we notice that the set of initial points of the segments from $x \in M_N$ to $(f_2(x), 2)$ has $\mathcal{L}^3$ measure one, being the whole M_N.

Another counterexample can be found in [13].

Example 4.3 (Transport rays do not fill continuously the line). Consider in $[0,1]$ the following transport problem, with $c(x,y) = |y-x|$ (Fig. 1). Fix $\ell \in (0,1/4)$ Construct a Cantor set of positive measure: remove from the interval $[0,1]$ first the subinterval $\left(\frac{1}{2}-\ell,\, \frac{1}{2}+\ell\right)$; then, in each of the remaining intervals, the central subinterval of length $2\ell^2$, and so on: at the step $k+1$ remove the subintervals $y_{i,k}+\ell^{k+1}(-1,1)$ — where $y_{1,k}, \dots, y_{2^k,k}$ are the centers of the intervals remaining at the step k. The measure of the set we remove is $\sum_{k=1}^{\infty}(2\ell)^k = \frac{2\ell}{1-2\ell} \in (0,1)$. Consider then the transport problem between

$$\mu = \sum_{k=1}^{+\infty}\sum_{i=1}^{2^k} 2^{-k-2^k-1}\big(\delta_{y_{i,k}+\ell^k} + \delta_{y_{i,k}-\ell^k}\big) \qquad \text{and} \qquad \nu = \sum_{k=1}^{+\infty}\sum_{i=1}^{2^k} 2^{-k-2^k}\delta_{y_{i,k}}.$$

The map bringing the mass in $y_{i,k} \pm \ell^k$ to $y_{i,k}$ is optimal,by [4], since it gives a c-monotone plan

$$\sum_{k=1}^{+\infty}\sum_{i=1}^{2^k} 2^{-k-2^k-1}\big(\delta_{(y_{i,k}+\ell^k,y_{i,k})} + \delta_{(y_{i,k}-\ell^k,y_{i,k})}\big).$$

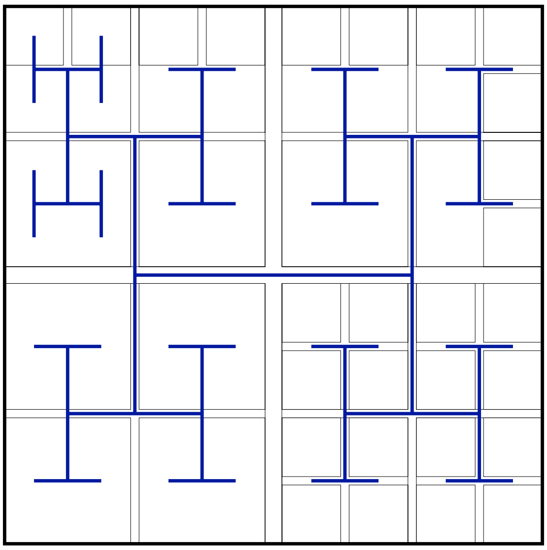

Figure 2: Example 4.4. In this case, however one chooses the potential, the vector field of direction is not defined on the whole square. In fact, the potential must be constant in points, belonging to the blue skeleton we begin to draw, dense in a $\mathcal{L}^2$-positive measure set.

Out of $\cup_{i,k}\{y_{i,k} \pm \ell^k\}$, clearly it is not relevant how the map is defined. The Kantorovich potential is not unique, up to constants. In particular, we can take

$$\phi(x) = \begin{cases} |\lambda| - \ell^k & \text{if } x = y_{i,k} + \lambda, \text{ with } \lambda \in (-\ell^k, \ell^k), \\ 0 & \text{on the Cantor set and out of } [-1,1]. \end{cases}$$

We have differentiability exactly in the points where no mass is set. In the points of the Cantor set the differential of ϕ vanishes: its gradient does not help in defining the field of rays' directions. Moreover, in every neighborhood we have points where the direction is towards right and others where it is towards left. In particular, the field of directions cannot be defined on the whole $\mathbb{R}$ with continuity. Consider now the limit of the functions

$$h_k(x) = x - \sum_{h=1}^{\infty} \sum_{i=1}^{2^k} \Big[(x - y_{i,k} + \ell^k) \mathbb{1}_{(y_{i,k}-\ell^k, y_{i,k}+\ell^k)} + \ell^k \mathbb{1}_{[y_{i,k}+\ell^k, +\infty)} \Big].$$

It is 1-Lipschitz, constant on the intervals we took away. In particular, $\tilde{\psi} = \psi + h$ is again a good potential. It is precisely the one defined in (3.1). Notice that, with the potential $\tilde{\phi}$, the maximal directions defined as in (4.2) are defined almost everywhere, just except in the atoms of μ. They are an extension of the previous vector field of directions. Notwithstanding, there is no continuity of this vector field on the Cantor set, which has positive measure. Continuity is recovered in open sets not containing the atoms of μ, ν. Observe that, spreading the atomic measures on suitable small intervals, one gets an analogous example with marginals absolutely continuous w.r.t. $\mathcal{L}^1$. Notice, finally, that the divergence of the vector field fails to be a locally finite Radon measure.

Example 4.4 ($\mathcal{T}$ does not fill the space). Consider in the unit square $X = Y = [0,1]^2$ the following transport problem (see Fig. 2).

Fix $\lambda \in (0,1)$. Define, recursively, the half edge $\ell_0 = 1/2$ and then, for $i \in \mathbb{N}$,

$$\ell_i = \frac{\lambda^{\frac{1}{2^{i+1}}} \ell_{i-1}}{2} = \lambda^{\sum_{j=2}^{i+1} 2^{-j}} 2^{-i-1}, \qquad a_i = \ell_{i-1} - 2\ell_i, \qquad n_i \text{ maximum in } 2\mathbb{N} \text{ s.t. } r_i := \frac{\ell_i + a_i}{n_i} < a_i.$$

Define moreover the sequence of centers

$$c_1 = \left(\frac{1}{2}, \frac{1}{2}\right), \qquad \left\{c_h\right\}_{h=\frac{4^i+2}{3},\dots,\frac{4^{i+1}-1}{3}} = \left\{c_j \pm (\ell_i + a_i)(e_1 \pm e_2)\right\}_{j=\frac{4^{i-1}+2}{3},\dots,\frac{4^i-1}{3}} \text{ for } i \in \mathbb{N}$$

and, finally, the intermediate points

$$z_{i,j,0} = c_i + jr_ie_1 \quad \text{and} \quad z_{i,j,\pm} = c_i \pm (\ell_1 + a_i)r_ie_1 + jr_ie_2 \quad \text{for } i \in \mathbb{N},\ j \in \{-n_i, \dots, n_i\}.$$

Then, the marginal measures be given by

$$\mu = \sum_{i=1}^{\infty} \frac{2^{-i}}{3(n_i+1)} \sum_{k=0,\pm,\ j \text{ even}} \delta_{z_{i,j,k}} \qquad \nu = \sum_{i=1}^{\infty} \frac{2^{-i}}{3n_i} \sum_{k=0,\pm} \delta_{z_{i,j,k},\ j \text{ odd}}.$$

One can immediately verify that the transport plan

$$\pi = \sum_{i=1}^{\infty} \frac{2^{-i}}{3} \sum_{j=1\dots n_i,\ k=0,\pm} \Bigg[\left(\frac{j}{n_i+1} - \frac{j-1}{n_i}\right) \delta_{(z_{i(-n_i+2j-2)k}, z_{i(-n_i+2j-1)k})} + \left(\frac{j}{n_i} - \frac{j}{n_i+1}\right) \delta_{(z_{i(-n_i+2j)k}, z_{i(-n_i+2j-1)k})}\Bigg]$$

is c-monotone, thus, in particular, optimal ([4]). Therefore, since $\phi(y) - \phi(x) = |y-x|_{D^*}$ for all (x, y) in a c-monotone carriage of π (Th. 3.1), we have that ϕ is constant on the set of points $\{z_{i(2j)k}\}_{ijk}$, say null. Moreover, these points are dense in the region

$$K = \bigcap_{i\in\mathbb{N}} \bigcup_{j=2^i+1}^{2^{i+1}} c_j + [-\ell_i, \ell_i]^2,$$

therefore ϕ must vanish on K. In the Lebesgue points of K, in particular, $\nabla\phi$ must vanish, too. This implies, by (4.1), that K is in the complementary of $\mathcal{T}$. The measure of this compact set is

$$\lim_{i\to\infty} 2^{2i}(2\ell_i)^2 = \lim_{i\to\infty} \lambda^{\sum_{j=1}^{i} 2^{-j}} = \lambda \in (0, 1).$$

The conclusion is that the transport set, in general, does not fill the space. Observe that, spreading the atomic measures on suitable small squares, one gets an analogous example with marginals absolutely continuous w.r.t. $\mathcal{L}^2$.

Appendix A. The Disintegration Theorem

The disintegration of a measure is a tool in order to decompose, and localize further into prescribed regions, a measure given on a space. In a measure theoretic environment, it goes in the opposite direction of the usual Fubini theorem, where product measure are investigated. We present firstly the abstract setting and definition, then we show this in two basic examples in $\mathbb{R}^N$. Finally, an existence and essential uniqueness theorem is enunciated.

Be given a measurable space $(R, \mathscr{R})$ and a function $r : R \mapsto S$, for some set S. A typical case is when one has a partition of R: this defines in R an equivalence relation and one considers the projection onto the quotient.

Definition A.1. The set S can be endowed with the *push forward σ-algebra* $\mathscr{S}$ of $\mathscr{R}$: it is defined as the biggest σ-algebra such that r is measurable, and it is explicitly given by

$$Q \in \mathscr{S} = r_\sharp(\mathscr{R}) \quad \Longleftrightarrow \quad r^{-1}(Q) \in \mathscr{R}.$$

Given a measure space $(R, \mathscr{R}, \rho)$, the *push forward measure* η is then defined as

$$\eta(Q) = r_\sharp\rho(Q) = \rho(r^{-1}(Q)) \qquad \forall Q \in r_\sharp(\mathscr{R}).$$

Consider, then, a probability space ind its push forward with a given map:

$$r : (R, \mathscr{R}, \rho) \mapsto (S, \mathscr{S}, \eta).$$

Definition A.2. A *disintegration* of ρ relative to r is a map $\rho_s(B) : \mathscr{R} \times S \to [0, 1]$ such that

- $\rho_s(\cdot)$ is a probability measure on $(R, \mathscr{R})$, for all $s \in S$
- $\rho.(B)$ is η-measurable for all $B \in \mathscr{R}$

that satisfies for all $B \in \mathscr{R}$, $S \in r_\sharp R$

$$\rho(B \cap r^{-1}(S)) = \int_S \rho_v(B)\, dr_\sharp \rho(v). \tag{A.1}$$

A disintegration is *subordinated to* r if for η-a.e. s the measure ρ_s is carried by $r^{-1}(s)$:

$$\rho_s(r^{-1}(s)) = 1 \qquad \text{for } \eta\text{-a.e. } s.$$

Remark A.3. If $\mathscr{R}$ is countably generated or complete, $\rho.(B)$ is $\mathscr{S}$-measurable, not only η-measurable.

Definition A.4. A σ-algebra is *countably generated* if there is a countable basis generating it. A σ-algebra is *essentially countably generated*, w.r.t. a measure m, if there is a countably generated σ-algebra $\mathcal{A}$ such that for all $A \in \mathcal{A}$ there exists $\hat{A} \in \hat{A}$ satisfying $m(A \vartriangle \hat{A}) = 0$.

Remark A.5. A *measure algebra* is a couple $(\mathscr{X}^*, \chi^*)$ where

- $\mathscr{X}^*$ is a σ-algebra,
- χ^* is a countably additive functional from $\mathscr{X}^*$ to $\mathbb{R}^+$.

Given a measure space $(X, \mathscr{X}, \chi)$, there is a natural measure algebra associated to it. Define the following equivalence relation in $\mathscr{X}$: for all $S_1, S_2 \in \mathscr{X}$

$$S_1 \sim S_2 = S^* \in \mathscr{X}^* \qquad \Longleftrightarrow \qquad \chi(S_1 \vartriangle S_2) = 0.$$

The quotient σ-algebra $\mathscr{X}^*$ is a measure algebra with $\chi^*(S^*) = \chi(S)$, for all $S \in \mathscr{X}^*$. Passing to the measure algebras, one can see that a measure space is essentially countably generated if and only if the associated measure algebra is countably generated.

We state a synthesis of the disintegration theorem in the form of [6].

Theorem A.6 (Disintegration theorem)**.** *Assume $(R, \mathscr{R}, \rho)$ is a countably generated probability space, $R = \cup_s R_s$ a decomposition of R, $r : R \to S$ the quotient map. Let $(S, \mathscr{S}, \eta)$ the quotient measure space defined by $\mathscr{S} = r_\sharp \mathscr{R}$, $\eta = r_\sharp \rho$. Then there exists a unique disintegration $s \to \rho_s$.*

Moreover, $\mathscr{S}$ is essentially countably generated w.r.t. η, say by the family $\{S_n\}_{n\in\mathbb{N}}$ generating $\widehat{\mathscr{S}}$. Identify furthermore the atoms of $\widehat{\mathscr{S}}$: define the equivalence relation

$$s \sim s' \qquad \text{if} \qquad \forall \widehat{S} \in \widehat{\mathscr{S}} \quad \{\ s \in \widehat{S} \ \Rightarrow\ s' \in \widehat{S}\ \}.$$

Denoting with p the quotient map and with $(L, \mathscr{L}, \lambda)$ the quotient space, the following properties hold:

- *$R_\ell := \cup_{s \in p^{-1}(\ell)} R_s \equiv (p \circ r)^{-1}(\ell)$ is ρ-measurable, and $R = \cup_{\ell \in L} R_\ell$;*
- *the disintegration $\rho = \int_L \rho_\ell\, d\eta(\ell)$ satisfies $\rho_\ell(R_\ell) = 1$;*
- *the disintegration $\rho = \int_S \rho_s\, d\eta(s)$ satisfies $\rho_s = \rho_{p(s)}$;*
- *the measure algebra $(\mathscr{S}^*, \eta^*)$ is isomorphic to the measure algebra $(\mathscr{L}^*, \lambda^*)$.*

The following two examples provide the basic meaning of what a disintegration subordinated to a map is. The third example, instead, shows that a disintegration does not need to be subordinated, and that the quotient space $(S, \mathscr{S}, \eta)$ in general is not countably generated. In that case, it is an object which does not succeed in localizing the measure, and does not carry many information. A less trivial example of a useful disintegrations is essentially the present article, as well as [7], [6].

Example A.7. Partition the unit square of $\mathbb{R}^N \setminus \{0\}$ into the coordinate hyperplanes $H_\lambda = \{x \cdot e_1 = \lambda\}$, for $\lambda \in [0, 1]$. Then, the quotient space is simply $([0, 1], \mathscr{L}^1([0, 1]), \mathcal{L}^1)$. Moreover, defining

$$\rho_\lambda = \mathcal{H}^{N-1} \llcorner_{H_\lambda},$$

the family $\{\rho_\lambda\}_{\lambda \in [0,1]}$ is a disintegration subordinated to the projection onto the line $\mathbb{R}e_1$.

Example A.8. Partition the unit ball of $\mathbb{R}^N \setminus \{0\}$ into rays centered in the origin: $r_d = (0, 1]d$ for every direction $d \in S^{N-1}$. Then the quotient space is $(S^{N-1}, \mathscr{L}(S^{N-1}), \mathcal{H}^{N-1} \llcorner_{S^{N-1}})$. This time, the disintegration subordinated to the projection onto S^{N-1} is given by $\{\rho_d = t^{N-1} \mathcal{H}^1(t) \llcorner_{r_d}\}_{d \in S^{N-1}}$. One could normalize in order to have probability measures.

Example A.9. Consider the following partition of $[0,1]$: $x \sim y$ if $x - y \in \mathbb{Z}\alpha$, with α irrational. The quotient set, by definition, is the Vitali set V. One can verify that the quotient σ-algebra of the Lebesgue one contains just sets of either full, or null quotient measure. Consequently, (A.1) implies that the only disintegration of the Lebesgue measure is given by $\rho_s = \mathcal{L}^1 \llcorner_{[0,1]}$ for all $s \in V$. In particular, this disintegration is not subordinated to the projection onto the quotient.

Acknowledgements. The author wishes to express her gratitude to professor Stefano Bianchini, for his advice and many helpful discussions.

References

[1] G. Alberti, B. Kirchheim, and D. Preiss. Personal communication in [3].
[2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford Clarendon Press, 2000.
[3] L. Ambrosio, B. Kircheim, and A. Pratelli. Existence of optimal transport maps for crystalline norms. *Duke Math. J.*, 125(2):207–241, 2004.
[4] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. In *Optimal transportation and applications (Martina Franca, 2001)*, volume 1813 of *Lecture Notes in Math.*, pages 123–160. Springer, Berlin, 2003.
[5] S. Bianchini. On the Euler-Lagrange equation for a variational problem. *Discrete Contin. Dyn. Syst.*, 17(3):449–480, 2007.
[6] S. Bianchini and L. Caravenna. Sufficient conditions for optimality of c-cyclically monotone transference plans. In preparation.
[7] S. Bianchini and M. Gloyer. On the Euler Lagrange equation for a variational problem: the general case II. Preprint, 2007.
[8] L. Caffarelli, M. Feldman, and R. McCann. Constructing optimal mass for Monge's transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.*, (15):1–26, 2002.
[9] T. Champion and L. De Pascale. The Monge problem for strictly convex norms in $\mathbb{R}^d$. Preprint, 2008.
[10] L. C. Evans and W. Gangbo. Differential equations methods for the Monge-Kantorovich mass transfer problem. *Current Developments in Mathematics*, pages 65–126, 1997.
[11] L. V. Kantorovich. On the transfer of masses. *Docl. Akad. Nauk. SSSR*, 37:227–229, 1942.
[12] L. V. Kantorovich. On a problem of Monge. *Uskpekhi Mat. Nauk.*, 3:225–226, 1948.
[13] D. G. Larman. A compact set of disjoint line segments in $\mathbb{R}^3$ whose end set has positive measure. *Mathematika*, 18:112–125, 1971.
[14] G. Monge. Mémoire sur la Theorie des Déblais et des Remblais. *Histoire de l'Acad. de Sciences de Paris*, pages 666–704, 1781.
[15] L. Rüschendorf. On c-optimal random variables. *Statist. Probab. Lett.*, 27(3):267–270, 1996.
[16] V. N. Sudakov. Geometric problems in the theory of infinite-dimensional probability distributions. *Proc. Steklov Inst. Math.*, 2:1–178, 1979. Number in Russian series statements: t. 141 (1976).
[17] N. S. Trudinger and X. J. Wang. On the Monge mass transfer problem. *Calc. Var. PDE*, (13):19–31, 2001.
[18] C. Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften.* Springer, 2009.

Laura Caravenna, S.I.S.S.A., via Beirut 4, 34014 Trieste, Italy
E-mail address: `l.caravenna@sissa.it`