A PROOF OF SUDAKOV THEOREM WITH STRICTLY CONVEX NORMS

L. CARAVENNA

ABSTRACT. The paper establishes a solution to the Monge problem in $\mathbb{R}^n$ for a possibly asymmetric norm cost function and absolutely continuous initial measures, under the assumption that the unit ball is *strictly* convex — but not necessarily differentiable nor uniformly convex. The proof follows the strategy initially proposed by Sudakov in 1976, found to be incomplete in 2000; the missing step is fixed in the above case adapting a disintegration technique introduced for a variational problem. By strict convexity, mass moves along rays, and we also investigate the divergence of the vector field of rays.

1. INTRODUCTION

The present paper concerns the existence of deterministic transport plans for the Monge-Kantorovich problem in $\mathbb{R}^n$, with a strictly convex and possibly asymmetric norm cost function

$$c(x, y) = \|y - x\|.$$

Given an initial Borel probability measure μ and a final Borel probability measure ν, the Monge problem deals with the minimization of the functional

$$\tau \mapsto \int_{\mathbb{R}^n} c(x, \tau(x))\, d\mu(x)$$

among the maps τ such that $\tau_\sharp \mu = \nu$. The issue was pushed forward by Monge in 1781 ([Mon]), in the case of the Euclidean norm, for absolutely continuous, compactly supported μ, ν in $\mathbb{R}^3$.
Even existence of solutions is a difficult question, due to the nonlinear dependence on the variable τ and the non-compactness of the set of minimizers in a suitable topology. A natural assumption is the absolute continuity of μ w.r.t. the Lebesgue measure, as shown in Section 8 of [AP]: there are initial measures with dimension arbitrarily close to n such that the transport problem with the Euclidean distance has no solution.

The modern approach passes through the Kantorovich formulation ([Kan1], [Kan2]). Rather than a map $\tau : \mathbb{R}^n \mapsto \mathbb{R}^n$, a transport is defined as a coupling of μ, ν: a probability measure π on the product space $\mathbb{R}^n \times \mathbb{R}^n$ having marginals μ, ν. The family of these couplings, called *transport plans*, is denoted with $\Pi(\mu, \nu)$ and their cost is defined as

$$\pi \mapsto \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y)\, d\pi(x, y).$$

The Kantorovich formulation is the relaxation of the Monge problem in the space of probability measures. A transport map τ induces the transport plan $\pi = (\mathrm{Id}, \tau)_\sharp \mu$, and the cost of the transport map coincides with the one of the induced transport plan. Denoting with $\pi = \int \pi_x \mu(x)$ the disintegration of a transport plan π w.r.t. the projection on the first variable, then the coupling π reduces to a map when the measure π_x, for μ-a.e. x, is concentrated at one point. There corresponds a difference in the model: the mass present at x is not necessarily moved to some point $\tau(x)$: the weaker formulation allows indeed spreading of mass, and the amount of mass at x spread in a region S is $\pi_x(S)$ — where π_x is the conditional measure of the above disintegration of π.
Assuming, more generally, c lower semicontinuous, one can deduce immediately existence of solutions for the relaxed problem by the direct method of calculus of variations. In order to recover solutions in the sense by Monge, then, one proves that some optimal transport plan is deterministic.

The topic has been studied extensively. We focus only on the Monge problem in $\mathbb{R}^n$ with norm cost functions, presenting a partial literature; for a broad overview one can consult for example [Vil], [AKP].

A solution was initially claimed in 1976 by V. N. Sudakov ([Sud]). The idea was to decompose first $\mathbb{R}^n$ into locally affine regions of different dimension invariant for the transport; to reduce then the transport

problem to new transport problems within these regions, by disintegrating the measures to be transported and by considering the transport problems between the conditional probabilities of μ, ν; to recover finally the solution in $\mathbb{R}^n$ by the solutions of the reduced problems.
For the solvability of the new transport problems, one needs however an absolute continuity property of the new initial measures. He thought that this absolute continuity property of the conditional probabilities was granted by Borel measurability properties of his partition, but, instead, at this level of generality the property does not hold — as pointed out in 2000 by Alberti, Kirchheim and Preiss ([AKP]), providing an example where the disintegration of the Lebesgue measure w.r.t. a partition even into disjoint segments with Borel directions has atomic conditional measures; this example is recalled in Example 4.1. Therefore, a gap remains in his proof.

Before this was known, another approach to the Monge problem in $\mathbb{R}^n$, based on partial differential equations, was given in [EG]. Despite some additional regularity on μ, ν, they introduced new interesting ideas. Strategies at least partially in the spirit of Sudakov were instead pursued independently and contemporary in [CFM], [TW], and [AP], improving the result. They achieved the solution to the Monge problem with an absolute continuity hypothesis only on the first marginal μ, requiring that μ, ν have finite first order moments and for cost functions satisfying some kind of uniform convexity property — which allows clever countable partitions of the domain into regions where the direction of the transport is Lipschitz. In [AKP] 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 been solved also in [CP1], with a different technique, focusing on convex bounded domains.

In the context of a different variational problem, a regularity property of the disintegration of the Lebesgue measure w.r.t. a particular partition into segments was proven in 2007 ([Bia], [BG]). They studied the vector fields generated by the directions of maximal growth of a 1-Lipschitz function u defined by the Hopf-Lax formula

$$u(x) := \max\left\{u(\bar{x}) - \|\bar{x} - x\| : \ \bar{x} \in \partial\Omega,\ \alpha x + (1-\alpha)\bar{x} \in \bar{\Omega}\ \forall \alpha \in [0,1]\right\}, \qquad u\restriction_{\partial\Omega} \text{ given}$$

with $\Omega \subset \mathbb{R}^n$ open with compact closure $\bar{\Omega}$ and for a general, not necessarily symmetric norm. In a first part of the work they recover the directions of maximal growth, for $\|\cdot\|$ strictly convex, as the $\mathscr{L}^n$-a.e. pointwise limit of piecewise regular vector fields defined as the directions of maximal growth of

$$u_I(x) := \max\left\{u(\bar{x}_i) - \|\bar{x}_i - x\| : \ i = 1, \dots, I,\ \alpha x + (1-\alpha)\bar{x}_i \in \bar{\Omega}\ \forall \alpha \in [0,1]\right\}, \qquad u\restriction_{\partial\Omega} \text{ given}$$

for a dense sequence of points $\{\bar{x}_i\}_{i=1}^{\infty}$. Notice that this approximating direction vector field points towards finitely many points. By approximation, they prove regularity estimates which allow a smart change of variable determining the disintegration of the Lebesgue measure on Ω w.r.t. the segments of maximal growth of u — and showing the absolute continuity of the conditional measures w.r.t. $\mathcal{H}^1$. The change of variable does not rely on any Lipschitz regularity of the directions of maximal growth, which maybe does not hold because of the irregularity of the norm, and thus does not rely on Area or Coarea formula. It is instead a direct application of Fubini-Tonelli theorem, once one has proved by the regularity estimates that the smart change of variables establishes an isomorphism of $\mathscr{L}^n$-measurable functions on Ω and on a suitable product space which parametrizes Ω. Namely, $\mathscr{L}^n$-a.e. point of Ω lies on a segment of maximal growth, and each segment of maximal growth is detected by its terminal point (on $\partial\Omega$) and its direction (in $\mathbf{S}^1$), while a scalar parameter identifies then the position on the segment.

The aim of this paper is to follow in this more complicated setting the construction relative to the strict convex norm in [BG]. This shows that the strategy started by Sudakov actually works, at least when assuming the unit ball $\{x : \ \|x\| \leq 1\}$ *strictly* convex, but with no smoothness assumption. Moreover, we remove the condition of finite order moments, or bounded support, on μ, ν.

The assumption of strict convexity is of course a simplification on the norm: it is well known that in this case the mass moves along lines. Indeed, under the nontriviality assumption that there exists a transport plan with finite cost, there exists a 1-Lipschitz map $\phi : \mathbb{R}^n \mapsto \mathbb{R}$, the *Kantorovich potential*, such that the transport is possible only towards points where the decrease of ϕ is the maximal allowed; these points by strict convexity are aligned forming transport rays $\{r_z\}_{z\in\mathcal{S}}$ which may intersect only at endpoints, yielding an analogy with the variational problem above.

We will prove the following statement.

Theorem. *Let μ, ν be Borel probability measures on $\mathbb{R}^n$ with μ absolutely continuous w.r.t. the Lebesgue measure $\mathscr{L}^n$. Let $\|\cdot\|$ be a possibly asymmetric norm whose unit ball is strictly convex.*
Suppose there exists a transport plan $\pi \in \Pi(\mu,\nu)$ with finite cost $\int \|y-x\|\,d\pi(x,y)$. Then:

Claim 1. *The family of transport rays $\{r_z\}_{z\in\mathcal{S}}$ can be parametrized with a Borel subset $\mathcal{S}$ of countably many hyperplanes, the* transport set $\mathcal{T}_{\mathtt{e}} = \cup_{z\in\mathcal{S}} r_z$ *is Borel and there exists a Borel function c such that the following disintegration of $\mathscr{L}^n \llcorner \mathcal{T}_{\mathtt{e}}$ holds: $\forall\varphi$ either integrable or positive*

$$\int_{\mathcal{T}_{\mathtt{e}}} \varphi(x)\,d\mathscr{L}^n(x) = \int_{\mathcal{S}} \left\{ \int_{a(z)\cdot d(z)}^{b(z)\cdot d(z)} \varphi(z+(t-z\cdot d(z))d(z))c(t,z)d\mathcal{H}^1(t) \right\} d\mathcal{H}^{n-1}(z).$$

Claim 2. *There exists a unique map τ monotone on each ray r_z solving the Monge-Kantorovich problem*

$$\min_{\tau_\sharp\mu=\nu} \int_{\mathbb{R}^n} \|\tau(x)-x\|\,d\mu(x) = \min_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} \|y-x\|\,d\pi(x,y).$$

Claim 3. *The divergence of the direction vector field of rays is a series of Radon measures, and a Green-Gauss like formula holds on special sets.*

The paper is organized as follows.
Section 2. We study the transport set associated to a 1-Lipschitz function ϕ.

In Subsection 2.1 we recall that, by the strict convexity of the norm, the transport set $\mathcal{T}$ is made of disjoint oriented segments, the transport rays. We prove that the set of points belonging to more rays is $\mathcal{H}^{n-1}$-countably rectifiable. The membership in a ray defines then an equivalence relation on $\mathcal{T}$.

In Subsection 2.2 we partition $\mathcal{T}$, and the transport set with endpoints $\mathcal{T}_{\mathtt{e}}$, into model sets. The model sets are sheaves of rays $\mathcal{Z}$ transversal to some hyperplane and their subsets $\mathcal{K}$ of rays truncated between two transversal hyperplanes.

In Subsection 2.3, we study the disintegration of the Lebesgue measure, on the transport set, w.r.t. the partition in rays, showing the absolute continuity of the conditional measure w.r.t. the 1-dimensional Hausdorff measure on the rays. This is done first on model sets, then in the whole $\mathcal{T}$. Moreover, we show by a density argument that the set of endpoints $\mathcal{T}_{\mathtt{e}} \setminus \mathcal{T}$ is Lebesgue negligible (Lemma 2.20).

In Subsection 2.4 the density of the conditional measures w.r.t. the Hausdorff 1-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.
Section 3. The proof started by Sudakov is completed, in the case $\{\|x\|\le 1\}$ strictly convex.
The transport problem in $\mathbb{R}^n$, with $\mu \ll \mathscr{L}^n$, is reduced to the 1-dimensional case, disintegrating the measures w.r.t. the equivalence relation given by the membership in a transport ray. By the study in Section 2, the conditional measures are still absolutely continuous w.r.t. $\mathcal{H}^1$ on each ray. The 1-dimensional case, well known since old, can then be solved, e.g. with the selection in [AKP]. Putting side by side the 1-dimensional solutions, a global map is constructed.
Section 4. We give some counterexamples.
We first recall the one in [AKP], showing that disintegrating a compact, positive $\mathscr{L}^n$-measure set w.r.t. the membership to disjoint segments, with Borel direction, the conditional probabilities can be atomic. Then, it is shown that the transport set $\mathcal{T}$ in general is just a σ-compact set. In the last examples, one can see how the divergence of the vector field of ray directions, defined as zero out of $\mathcal{T}$, can fail to be a measure.
Appendix A. We recall the disintegration of measures in the form presented in [BC1].
Appendix B. A table of notations is given for the reader's convenience.

Further developments are given in [CP2], [BC2], subsequent to the submission of the present paper.

2. Disintegration in Transport Rays

The present section deals with the following problem: studying the disintegration of the Lebesgue measure on the transport set associated to a potential ϕ w.r.t. the partition induced by the directions of maximal decrease of ϕ. More precisely, in the present section we adopt the following definitions.

Definition 2.1 (Potential). A *potential* is a 1-Lipschitz map $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ such that

$$\phi(x) - \phi(y) \le \|y-x\| \qquad \forall x,y\in\mathbb{R}^n. \tag{2.1}$$

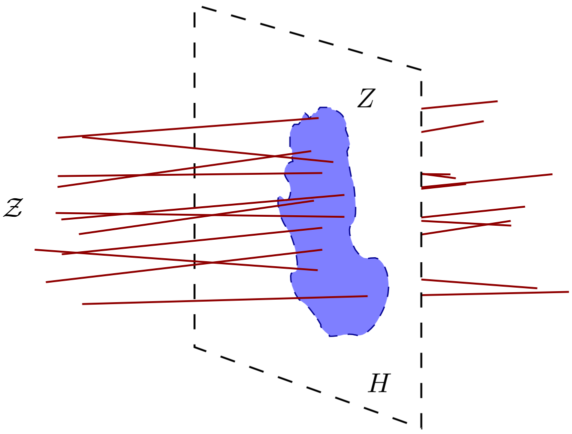

FIGURE 1. A sheaf of rays

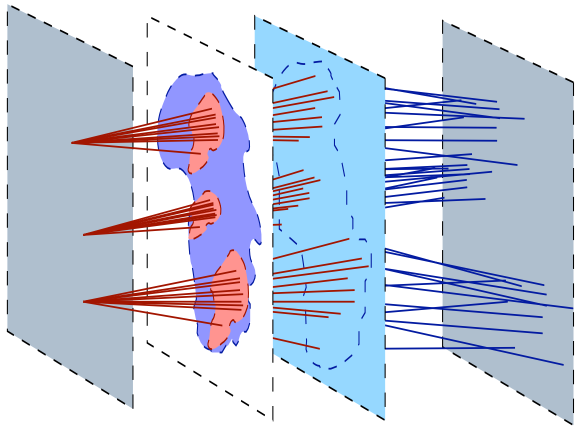

FIGURE 2. Approximation of rays

Definition 2.2 (Transport set). The *transport set* associated to a potential ϕ is the set $\mathcal{T}$ made of the open segments $(\!|x,y|\!)$ for every couple (x,y) such that in (2.1) equality holds:

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

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

We summarize briefly the construction. Due to the strict convexity of the norm, $\mathcal{T}$ is made of disjoint oriented segments — the transport rays — which are the lines of maximal decrease of ϕ. The membership to a transport ray defines then an equivalence relation on $\mathcal{T}$, by identifying points on a same ray. The issue is to show that the conditional measures of $\mathscr{L}^n\llcorner\mathcal{T}$ are absolutely continuous w.r.t. $\mathcal{H}^1$ on the rays, and the fact that the set of endpoints is Lebesgue negligible. We will indeed prove some more regularity.

By the additivity of the measures, the thesis will follow if proved on the elements of a countable partition of $\mathcal{T}$ into Borel sets. In particular, in Subsections 2.1, 2.2 we provide a partition into model sets $\mathcal{Z}$ made of rays transversal to some hyperplane H, let Z be the intersection of $\mathcal{Z}$ with H (Figure 1). Points x belonging to $\mathcal{Z}$ can be parametrized by the point $y(x)\in Z$ where the ray through x intersects H and by the distance $t(x)$ from H, positive if $\phi(y(x))\geq\phi(x)$ or negative otherwise. We prove in Corollary 2.22 that the bijective parameterization

$$\begin{array}{ccc}\mathcal{Z} & \leftrightarrow & \mathrm{Im}((y,t))\subset Z\times\mathbb{R}\\ x & \leftrightarrow & y(x),t(x)\end{array}$$

provides an isomorphism between the $\mathscr{L}^n$-measurable functions on $\mathcal{Z}$ and the $(\mathcal{H}^{n-1}\llcorner Z)\otimes\mathcal{H}^1$-measurable functions on $\mathrm{Im}((y,t))$. The isomorphism implies that the push forward with (y,t) of $\mathscr{L}^n\llcorner\mathcal{Z}$ is absolutely continuous w.r.t. $(\mathcal{H}^{n-1}\llcorner Z)\otimes\mathcal{H}^1$, with density function $\tilde\alpha(t,\cdot)$. By the classical Fubini-Tonelli theorem this proves the disintegration: denoting with $\sigma^t(y)$ the inverse map of (y,t), i.e. $x=\sigma^{t(x)}(y(x))$,

$$\begin{aligned}\int_{\mathcal{Z}}\varphi(x)\,d\mathscr{L}^n(x)&=\int_{(y,t)(\mathcal{Z})}\varphi(\sigma^t(z))\tilde\alpha(t,z)\,d\mathcal{H}^{n-1}(z)\otimes d\mathcal{H}^1(t)\\&=\int_Z\left\{\int_{\inf t(y^{-1}(z))}^{\sup t(y^{-1}(z))}\varphi(\sigma^t(z))\tilde\alpha(t,z)\,d\mathcal{H}^1(t)\right\}d\mathcal{H}^{n-1}(z)\\&=\int_Z\left\{\int_{y^{-1}(z)}\varphi(\sigma^t(z))c(t,z)\,d\mathcal{H}^1(t)\right\}d\mathcal{H}^{n-1}(z),\end{aligned}$$

where c is obtained by an easy change of variables in the 1-dimensional integral.

The isomorphism is derived from the fact that if we consider open rays transversal to two parallel hyperplanes and we consider the bijective map between the two hyperplanes coupling the points on a same ray, then the push forward of the Hausdorff $(n-1)$-dimensional measure on one hyperplane with this map is absolutely continuous w.r.t. the Hausdorff $(n-1)$-dimensional measure on the other hyperplane: positive sections does not shrink to zero if not at endpoints of rays.

This fundamental estimate is proved in Lemma 2.17 by approximating the rays with a sequence of segments starting from a section on one hyperplane and pointing towards finitely many points of a sequence dense in a third section beyond the other hyperplane (see Figure 2), and passing to the limit by the u.s.c. of the Hausdorff measure on compact sets.

The absolute continuity estimate yields more than the existence of the above density c: the distributional divergence of the vector field $\hat{d}$ of the rays on $\mathcal{Z}$, set zero on $\mathbb{R}^n \setminus \mathcal{Z}$, is a Radon measure, and the following formula holds (Lemma 2.30):

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

We see that this implies a Green-Gauss-type formula on special subsets.

2.1. **Elementary structure of the Transport Set.** We define the multivalued functions associating to a point the transport rays through that point — which are the lines of maximal growth of ϕ — the relative directions and endpoints. We then prove that they are Borel multivalued functions (Lemma 2.4).

Definition 2.3. The *outgoing rays* from $x \in \mathbb{R}^n$ are defined as

$$\mathcal{P}(x) := \big\{y : \ \phi(y) = \phi(x) - \|y - x\|\big\}.$$

The *incoming rays* at x are then given by

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

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 $\mathcal{T}_{\mathrm{e}}$ is just 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 $\mathcal{P}^{-1}(x) \neq \{x\}$. The following remarks are in order.

The set $\mathcal{P}(x)$ is a union of closed segments with endpoint x, which we call rays. In fact, ϕ's Lipschitz condition (2.1) implies that, for every $y \in \mathcal{P}(x)$, ϕ must decrease linearly from x to y at the maximal rate allowed:

$$\phi(x + t(y - x)) = \phi(x) - t\|y - x\| \qquad \text{for all } y \in \mathcal{P}(x),\ t \in [0,1]. \tag{2.2}$$

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 if two rays intersect in y, and $x \in \mathcal{P}^{-1}(y)$, $z \in \mathcal{P}(y)$, one has

$$\phi(z) \overset{z\in\mathcal{P}(y)}{=} \phi(y) - \|z - y\| \overset{y\in\mathcal{P}(x)}{=} \phi(x) - \|y - x\| - \|z - y\| \leq \phi(x) - \|z - x\|. \tag{2.3}$$

Again by Lipschitz condition (2.1) equality must hold: then $\|z - y\| = \|y - x\| + \|z - y\|$. Since the norm is strictly convex, this implies that x, y, z must be aligned.

In the following is shown that, at $\mathscr{L}^n$-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:

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

In order to show that in $\mathcal{T}_{\mathrm{e}}$ there exists such a vector field of directions, one has to show that there is at most one transport ray even at $\mathscr{L}^n$-a.e. endpoint. This is not trivial because, up to now, we can't say that the set of endpoints is $\mathscr{L}^n$-negligible, which does not follow from the fact that the set e.g. of initial points is Borel and that from each point starts at least a segment which does not intersect the others, with a Borel direction field — one can see the Example 4.1 (from [Lar], [AKP]).

One should then study before the multivalued map giving the directions of those rays:

$$\mathcal{D}(x) := \left\{\frac{y - x}{|y - x|}\mathbb{1}_{\mathcal{P}(x)}(y) + \frac{x - y}{|y - x|}\mathbb{1}_{\mathcal{P}^{-1}(x)}(y)\right\}_{y\in\mathcal{R}(x)} \qquad \text{for all } x \in \mathcal{T}_{\mathrm{e}}. \tag{2.4}$$

We first show that the above maps $\mathcal{P}$, $\mathcal{D}$ are Borel maps. We remind that a multivalued function F is Borel if the counterimage of an open set is Borel, where the counterimage of a set S is defined as the set of x such that $F(x) \cap S \neq \emptyset$.

Lemma 2.4. *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 $\phi(z_k) = \phi(x_k) - \|z_k - x_k\|$, by continuity we have that $\phi(z) = \phi(x) - \|z - x\|$. 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 $\mathcal{T}$, $\mathcal{T}_{\mathbb{e}}$. 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 2.5. 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 $\mathcal{T}_{\mathbb{e}}$ into segments, up to a $\mathscr{L}^n$-negligible set. This is found as 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 $\mathscr{L}^n$-a.e. point x of $\mathcal{T}_{\mathbb{e}}$, 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.

Before giving this result, we recall the definition of rectifiable set and a rectifiability criterion, that will be used in Lemma 2.8 below in order to show the rectifiability of the set where $\mathcal{D}$ is multivalued.

Definition 2.6 (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 $E \subset \cup f_i(\mathbb{R}^k)$.

Theorem 2.7 (Theorem 2.61, [AFP]). *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 $L(x) \subset \mathbb{R}^n$ such that*

$$S \cap B_{\rho(x)}(x) \subset x + \big\{y \in \mathbb{R}^n : |P_{L(x)^\perp} y| \le m(x)|P_{L(x)} y|\big\},$$

where P_L is the orthogonal projection onto L, $P_{L^\perp}$ onto the orthogonal of L. Then S is contained in the union of countably many Lipschitz k-graphs whose Lipschitz constants do not exceed $2\sup_x m(x)$.

Lemma 2.8. *On $\mathcal{T}_{\mathbb{e}}$, $\mathcal{D}$ is single valued out of a $(n-1)$-countably rectifiable set.*

Proof. We show the rectifiability of the set where $\mathcal{D}$ is multivalued applying Theorem 2.7.

Step 1: Countable covering. By (2.3) $\mathcal{D}$ is single valued where there are both an incoming and outgoing ray. By symmetry, it is then enough to consider the set J were there are more outgoing rays.

Notice that, by strict convexity, for every $d \neq d'$ in the sphere $\mathbf{S}^{n-1}$ there exist $\bar h, \bar\rho > 0$ such that

$$q \cdot d_1 \le -1/h < 1/h \le q \cdot d_2 \qquad \forall (d_1, d_2, q) \in \mathbf{B}_\rho(d) \times \mathbf{B}_\rho(d') \times (\delta D^*(\mathbf{B}_\rho(d')) - \delta D^*(\mathbf{B}_\rho(d)))$$

for $h \ge \bar h$, $\rho \le \bar\rho$, where $\mathbf{B}_\rho(\cdot)$ is the closed ball of radius ρ centered at $\cdot$. One can then extract a countable covering $\{B_1^{hj} \times B_2^{hj}\}_{hj\in\mathbb{N}}$ of $\mathbf{S}^{n-1}\times\mathbf{S}^{n-1}\setminus\{d = d'\}$, with B_1^{hj}, B_2^{hj} balls of radius $1/j$, satisfying

$$q \cdot d_1 \le -1/h < 1/h \le q \cdot d_2 \qquad \forall (d_1, d_2, q) \in B_1^{hj} \times B_2^{hj} \times (\delta D^*(B_2^{hj}) - \delta D^*(B_1^{hj})).$$

Define

$$\begin{aligned} J_{ijp} := \Big\{ x \in \mathcal{T}_{\mathbb{e}} : \exists d_1, d_2 \in \mathcal{D}(x) \text{ s.t. } \|d_1 - d_2\| \ge \frac{1}{p}, \\ d_1 \in B_1^{ip}, d_2 \in B_2^{ip},\ \mathcal{H}^1\big((x + \langle d_i\rangle) \cap \mathcal{P}(x)\big) \ge \frac{1}{p} \Big\}. \end{aligned}$$

It is not difficult to see that $\{J_{ijp}\}_{ijp\in\mathbb{N}}$ provides a countable covering of J.

Step 2: Remarks. Suppose that $x_k \in J_{ijp}$ converges to some x. Then by compactness there is a subsequence such that there exist $d_k^1 \in \mathcal{D}(x_k)\cap B_1^{ip}$ and $d_k^2 \in \mathcal{D}(x_k)\cap B_2^{ip}$ converging respectively to some $d_1 \in B_1^{ip}$, $d_2 \in B_2^{ip}$, and

$$y_k^1 := x_k + d_k^1/p \to y^1 := x + d^1/p \qquad y_k^2 := x_k + d_k^2/p \to y^2 := x + d^2/p.$$

By the continuity of ϕ, since y_k^1, y_k^2 belong to $\mathcal{P}(x_k)$, then y^1, y^2 belong to $\mathcal{P}(x)$ and therefore $x \in J_{ijp}$. In particular, J_{ijp} is closed.

Step 3: Claim. By the previous steps, it suffices to show that each J_{ijp} is $(n-1)$-countably rectifiable. To this purpose, we show that the cone condition of Theorem 2.7 holds: we prove that for every $x \in J_{ijp}$ the relative interior of the cone

$$x + \big\{(\lambda_1 B_1^{ip} - \lambda_2 B_2^{ip}) \cup (-\lambda_1 B_1^{ip} + \lambda_2 B_2^{ip})\big\}_{\lambda_1,\lambda_2 \geq 0}$$

contains no sequence in J_{ijp} converging to x.

Step 4: Claim of the estimate. We prove in the next step that for every sequence of points $x_k \in J_{ijp}$ converging to x, with the notations of Step 2, with $(x_k - x)/|x_k - x|$ converging to some vector ℓ

$$\exists q_1, q_2 \in \delta D^*(d^2/\|d^2\|) - \delta D^*(d^1/\|d^1\|) : \qquad q_1 \cdot \ell \geq 0, \qquad q_2 \cdot \ell \leq 0. \tag{2.5}$$

By definition of B_1^{ip} and B_2^{ip}, if $\ell \in B_1^{ip}$ one would have $q_1 \cdot \ell \leq -1/i$, while $q_2 \cdot \ell \geq 1/i$ would hold if $\ell \in B_2^{ip}$, yielding a contradiction: this means that any possible limit ℓ as above does not belong to $B_1^{ip} \cup B_2^{ip}$. Then (2.5) implies easily that every sequence $\{x_k\}_{k\in\mathbb{N}}$ converging to x definitively does not belong to the relative interior of the cone

$$x + \big\{(\lambda_1 B_1^{ip} - \lambda_2 B_2^{ip}) \cup (-\lambda_1 B_1^{ip} + \lambda_2 B_2^{ip})\big\}_{\lambda_1,\lambda_2 \geq 0}.$$

Step 5: Proof of the estimate (2.5). Let $x_k \in J_{ijp}$ converging to x, up to subsequence as in Step 2 one can assume also that there exist $y_k^1, y_k^2 \in \mathcal{P}(x_k)$ converging respectively to $y^1 = x + d^1/p$, $y^2 = x + d^2/p$ and that $(x_k - x)/|x_k - x|$ converges to some vector ℓ.

Observe first that, given $b \in \mathbb{R}^n$, $\ell \in \mathbf{S}^{n-1}$, there exists a vector v belonging to the subdifferential $\partial^- \|b\|$ of $\|\cdot\|$ at b, and depending on ℓ, such that the equality

$$\|a_k\| = \|b\| + v \cdot (a_k - b) + o(|a_k - b|), \tag{2.6}$$

holds for every $a_k \in \mathbb{R}^n$ converging to $b \in \mathbb{R}^n$ with $\frac{a_k - b}{\|a_k - b\|}$ converging to ℓ.

As a consequence, one can choose vectors $v^2 \in \partial^- \|y^2 - x\|$, $v_k^1 \in \partial^- \|y_k^1 - x\|$ in order to have

$$\begin{aligned}
\phi(x) + v^2 \cdot (x - x_k) + o(|x - x_k|) &\overset{(2.6)}{=} \phi(x) - \|y^2 - x\| + \|y^2 - x_k\| \\
&= \phi(y^2) + \|y^2 - x_k\| \geq \phi(x_k) = \phi(y_k^1) + \|y_k^1 - x_k\| \\
&\geq \phi(x) - \|y_k^1 - x\| + \|y_k^1 - x_k\| \overset{(2.6)}{=} \phi(x) + v_k^1 \cdot (x - x_k) + o(|x - x_k|).
\end{aligned}$$

Consider every subsequence s.t. v_k^1 converges to some v^1, necessarily in $\partial^- \|y^1 - x\|$: this yields

$$(v^2 - v^1) \cdot \ell = \lim_k \left\{ (v^2 - v_k^1) \cdot \frac{x - x_k}{|x - x_k|} \right\} \geq 0$$

proving that

$$\exists q_1 \in \delta D^*(d^2/\|d^2\|) - \delta D^*(d^1/\|d^1\|) : \qquad q_1 \cdot \ell \geq 0.$$

The existence of q_2 can be found by symmetry inverting the roles of d^1, d^2. □

Lemma 2.8 ensures that one can define on a Borel subset of $\mathcal{T}_{\mathrm{e}}$, differing from $\mathcal{T}_{\mathrm{e}}$ for an $\mathscr{L}^n$-negligible set, a vector field giving at each point the direction of the ray passing there:

$$d(x) \quad \text{s.t.} \quad \mathcal{D}(x) := \{d(x)\}. \tag{2.7}$$

On this domain, the function d is Borel, by Lemma 2.4, being just a restriction of the Borel multivalued map $\mathcal{D}$. Since, by the strong triangle inequality in (2.3), rays cannot bifurcate, we are allowed to consider their endpoints, possibly at infinity. After compactifying $\mathbb{R}^n$, define on $\mathcal{T}_{\mathrm{e}}$

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

As we will prove in the following (see resp. Lemma 2.13 and Lemma 2.20), both these functions are Borel, $\mathscr{L}^n$-a.e. single valued and their image is $\mathcal{H}^n$-negligible; in particular, $a(x) \neq x$ for $\mathcal{H}^n$-a.e. $x \in \mathcal{T}$.

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. 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 $\mathcal{T}$.

Definition 2.9 (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 $(\!|a(y), b(y)|\!)$ at one point. The set Z is called 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 $\mathfrak{e}_j \in \mathbf{S}^{n-1}$ such that $\mathbf{S}^{n-1} \subset \cup_{j=1}^{J} \mathbf{B}_\varepsilon(\mathfrak{e}_j)$; define, then, the following finite, disjoint covering $\{S_j\}$ of $\mathbf{S}^{n-1}$:

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

Lemma 2.10. *The following sets are sheaf sets covering $\mathcal{T}$ (resp. $\mathcal{T}_{\mathfrak{e}}$):*

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

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

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

The family $\{\mathcal{Z}_{jk\ell m}\}_{i\ell m}$ is disjoint, it refines and covers a set increasing to $\mathcal{T}$ when k increases. $\mathcal{Z}^{\mathfrak{e}}_{jk\ell m}$ differs from $\mathcal{Z}_{jk\ell m}$ only for endpoints of rays, and the sets $\{\mathcal{Z}^{\mathfrak{e}}_{jk\ell m}\}_{j\ell m}$ can instead intersect each other at points where $\mathcal{D}$ is multivalued. We denote with $Z_{jk\ell m}$ a basis of $\mathcal{Z}_{jk\ell m}$.

A partition of $\mathcal{T}$ is then provided by

$$\mathcal{Z}'_{jk\ell m} = \mathcal{Z}_{jk\ell m} \setminus \bigcup_{k'<k, \ell'<m'} \mathcal{Z}_{jk'\ell' m'}.$$

Proof. Consider a point on a ray. Then $d(x)\in S_j$ for exactly one j. Moreover, since $\mathcal{R}(x)\cdot\mathfrak{e}_j$ is a nonempty interval, for k sufficiently large we can define maximal values of ℓ, m such that $2^{-k}[\ell-1, m+1] \subset \mathcal{R}(x)\cdot\mathfrak{e}_j$. 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 ones $C_{j\alpha\beta p}$ are closed: since S_j is σ-compact, consider a covering of it with compact sets $\mathfrak{S}^p_j$, for $p\in\mathbb{N}$; define then

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

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

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

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

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

Definition 2.11 (d-cylinder). A cylinder subordinated to the vector field d is a σ-compact set of the form

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

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

Lemma 2.12. *With the notations of Lemma 2.10, $\mathcal{T}$ is covered by the d-cylinders*

$$\mathcal{K}_{jk\ell m} = \left\{\sigma^t y = y + \frac{td(y)}{d(y)\cdot \mathfrak{e}_j} \quad \text{with } y \in Z_{jk\ell m} \cap \mathcal{Z}'_{jk\ell m},\ t \in 2^{-k}[\ell, m]\right\}.$$

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

Proof. The proof is similar to the one of Lemma 2.10: 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 2.13. *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 Lemma 2.4). 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 2.10. Define the Borel function pushing, along rays, each point in $\mathcal{V}^-_{jk\ell m}$ to the upper basis:

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

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 2.14. Focus on a sheaf set with axis e_1 and basis $Z \subset \left\{x \cdot \mathrm{e}_1 = 0\right\}$. The composite map

$$\begin{array}{ccccc} \mathcal{Z}(Z) \subset \mathbb{R}^n & \to & \mathbb{R} \times \mathbb{R}^n & \to & Z + (-1,1)\mathrm{e}_1 \subset \mathbb{R}^n \\ z & \to & (z\cdot \mathrm{e}_1, \sigma^{-z\cdot \mathrm{e}_1} z) = (t,x) & \to & \left(x + \frac{t \arctan t}{\arctan(b(x)\cdot \mathrm{e}_1)} \mathrm{e}_1 \mathbb{1}_{t\ge 0} + \frac{t\arctan t}{\arctan(a(x)\cdot \mathrm{e}_1)} \mathrm{e}_1 \mathbb{1}_{t \le 0}\right) \end{array}$$

is a Borel and invertible change of variable from $\mathcal{Z}(Z)$ to the cylinder $Z + (-1,1)\mathrm{e}_1$, with Borel inverse. This will turn out to carry negligible sets into negligible sets (see Corollary 2.22).

Remark 2.15. 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 (2.19) 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 $\mathscr{L}^n$.** 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 Subsection 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 Subsection 2.3.2. In particular, it is proved that the conditional measures are absolutely continuous w.r.t. the Hausdorff 1-dimensional measure on the rays. We recall that this is nontrivial, since some regularity of the field of directions is needed (see Example 4.1).

2.3.1. *Fundamental estimate: the sheaf set $\mathcal{Z}$.* We first show with an example how the vector field d can be approximated with a piecewise radial vector field d_I.

Fix the attention on a sheaf set $\mathcal{Z}_{\mathbb{e}}$ with axis e_1 and a bounded basis $Z \subset \{x : \ \mathrm{e}_1 \cdot x = 0\}$: assume that, for suitable $h^{\pm}$,

$$\mathcal{Z}_{\mathbb{e}} = \cup_{y \in Z} [\![a(y), b(y)]\!], \qquad \mathrm{e}_1 \cdot a \restriction_Z < h^- \leq 0, \qquad \mathrm{e}_1 \cdot b \restriction_Z > h^+ \geq 0.$$

Example 2.16 (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,

$$x \longrightarrow \sigma^t(x) := x + \frac{t}{d(x) \cdot \mathrm{e}_1} d(x).$$

In order to avoid to work with values at infinite, we think to truncate the rays at $\{x \cdot \mathrm{e}_1 = h^-\}$. Choose now a dense sequence $\{\mathsf{a}_i\}$ in $\sigma^{h^-} Z$. Approximate the potential ϕ with the sequence of potentials

$$\phi_I(x) = \max \big\{ \phi(\mathsf{a}_i) - \|x - \mathsf{a}_i\| : \ i = 1, \ldots, I \big\}.$$

Since ϕ is uniformly continuous on $\sigma^{h^-} Z$, as a consequence of the representation formula for ϕ, we see easily that ϕ_I increases to ϕ on the closure of $\mathcal{Z}_{\mathbb{e}} \cap \{x \cdot \mathrm{e}_1 \geq h^-\}$. There, consider now the vector fields of ray directions

$$d_I(x) = \sum_{i=1}^{I} d^i(x) \mathbb{1}_{\overline{\Omega_i^I}}(x) \qquad \text{with} \qquad d^i(x) = \frac{x - \mathsf{a}_i}{|x - \mathsf{a}_i|}, \tag{2.8}$$

where the open sets Ω_i^I are

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

They partition $\mathbb{R}^n$, together with their boundary. Notice that this boundary is $\mathcal{H}^{n-1}$-countably rectifiable: for example apply Lemma 2.8, since it is where the field of ray directions associated to ϕ_I is multivalued. We show that the sequence d_I converges $\mathcal{H}^n$-a.e. to d on $\mathcal{Z}_{\mathbb{e}} \cap \{x \cdot \mathrm{e}_1 > h^-\}$. More precisely, every selection of the d_I converges pointwise to d on $\mathcal{Z}_{\mathbb{e}} \cap \{x \cdot \mathrm{e}_1 > h^-\}$. Consider any sequence $\{d_{I_j}(x)\}_j$ convergent to some $\bar{d}$. The corresponding points a_{i_j} satisfy

$$\phi_{I_j}(\mathsf{a}_{i_j}) = \phi_{I_j}(x) + \|x - \mathsf{a}_{i_j}\|;$$

therefore, they will converge to some point a s.t. $\bar{d} = (x - \mathsf{a})/|x - \mathsf{a}|$ and $\mathsf{a} \cdot \mathrm{e}_1 = h^-$; in particular, $\mathsf{a} \neq x$. Then, taking the limit in the last equation, one gets that $\phi(\mathsf{a}) = \phi(x) + \|x - \mathsf{a}\|$. In particular, where d is single valued, $d = (x - z)/|x - z| = \bar{d}$ follows.

Define the map $\sigma_{d_I}^t$ which, similarly to σ^t, moves points along the rays relative to ϕ_I. Notice that, by (2.8), within Ω_i^I the map $\sigma_{d_I}^t$ moves points towards a_i, for $i \leq I$. As a consequence, for $S \subset \Omega_i^I \cap \{x \cdot \mathrm{e}_1 = h\}$ and $h - h^- > t \geq 0$ the set $\sigma_{d_I}^{-t} S$ is similar to S: precisely

$$\sigma_{d_I}^{-t} S = \mathsf{a}_i + \frac{h - h^-}{h - t - h^-}(S - \mathsf{a}_i).$$

By additivity, also for $S \subset \{x \cdot \mathrm{e}_1 = h\}$ and $h - h^- > t \geq 0$ the following equality holds:

$$\mathcal{H}^{n-1}(\sigma_{d_I}^{-t} S) = \left(\frac{h - h^-}{h - t - h^-} \right)^{n-1} \mathcal{H}^{n-1}(S). \tag{2.9}$$

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

$$\mathcal{K} = \mathcal{Z} \cap \{h^- \leq \mathrm{e}_1 \cdot x \leq h^+\} = \cup_{t \in [h^-, h^+]} \sigma^t Z, \qquad \text{and } a \restriction_{\mathcal{K}} \cdot \mathrm{e}_1 \leq h^-, \ b \restriction_{\mathcal{K}} \cdot \mathrm{e}_1 \geq h^+.$$

Lemma 2.17 (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 Example 2.16. There, we proved pointwise convergence on $\mathcal{Z}_{\mathbb{e}} \cap \{x \cdot \mathrm{e}_1 > h^-\}$. 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

$$\mathcal{H}^{n-1}(A_\eta) \ge \mathcal{H}^{n-1}(\sigma^t S) - \eta. \tag{2.10}$$

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 (2.9) for the regular d_I,

$$\mathcal{H}^{n-1}(A_\eta) \equiv \mathcal{H}^{n-1}(\sigma_{d_I}^0 A_\eta) \le \left(\frac{t - h^-}{s - h^-}\right)^{n-1} \mathcal{H}^{n-1}(\sigma_{d_I}^{s-t} A_\eta). \tag{2.11}$$

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

$$\limsup_{I \to \infty} \mathcal{H}^{n-1}(\sigma_{d_I}^{s-t} A_\eta) \le \mathcal{H}^{n-1}(\sigma_d^{s-t} A_\eta) \le \mathcal{H}^{n-1}(\sigma^s S). \tag{2.12}$$

Collecting (2.10), (2.11) and (2.12) 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 2.18. The consequences of this fundamental formula are given in Subsection 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

$$\left(\frac{h^+ - t}{h^+ - h^-}\right)^{n-1} \mathcal{H}^{n-1}(Z(h^-)) \le \mathcal{H}^{n-1}\big(Z(t)\big) \tag{2.13}$$

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 (Lemma 2.20). As a consequence, we can cover $\mathcal{H}^n$-almost all $\mathcal{T}_{\mathbb{e}}$ with countably many d-cylinders — of positive $\mathcal{H}^n$-measure if $\mathcal{T}_{\mathbb{e}}$ has positive $\mathcal{H}^n$-measure.

2.3.2. *Disintegration of the Lebesgue measure on $\mathcal{Z}$.* We derive now the consequences of the fundamental estimates of Lemma 2.17. We first observe by a density argument that the set of endpoints of transport rays is $\mathcal{H}^n$-negligible (Lemma 2.20). 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 $\mathscr{L}^n$-measurable functions on d-cylinders to $\mathscr{L}^n$-measurable functions on usual cylinders. Some regularity properties of the Jacobian are presented. The fundamental estimate leads then to the explicit disintegration of the Lebesgue measure on the whole transport set $\mathcal{T}_{\mathbb{e}}$ (Theorem 2.25).

Remark 2.19. 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 2.17.

Lemma 2.20. *The set of endpoints of transport rays is negligible: $\mathscr{L}^n(\mathcal{T}_{\mathbb{e}} \setminus \mathcal{T}) = 0$.*

Proof. We analyze just $\mathcal{A} = \cup_x a(x)$, the other case is symmetric. Suppose $\mathcal{H}^n(\mathcal{A}) > 0$. Since we have the decomposition of Subsection 2.2, it is enough to prove the negligibility e.g. of the initial points of the set $\mathfrak{L}$ where $d \in B_\eta(\mathrm{e}_1)$, for some small $\eta > 0$, and $\mathcal{H}^1(\mathcal{P}(x)\cdot \mathrm{e}_1) > 1$. Consider a Lebesgue point of both the sets $\mathcal{A}$ and $\mathfrak{L}$, say the origin. 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 $\lambda \in T$

$$\mathcal{H}^{n-1}(H_\lambda) \geq (1-\varepsilon)r^{n-1} \qquad \text{where } H_\lambda = \mathfrak{L}\cap\mathcal{A}\cap\{x\cdot \mathrm{e}_1 = \lambda,\ |x-\lambda \mathrm{e}_1|_\infty \leq r\}. \tag{2.14}$$

Choose, now, $s < t$, both in T, with $|t-s| < \varepsilon r$. By Lemma 2.17, then

$$\mathcal{H}^{n-1}\big(\sigma^{t-s}H_s\big) \geq \left(\frac{1-t}{1-s}\right)^{n-1}\mathcal{H}^{n-1}\big(H_s\big) \overset{(2.14)}{\geq} (1-2\varepsilon)r^{n-1}.$$

Moreover, since $d \in B_\eta(\mathrm{e}_1)$, we have that $\mathcal{H}^{n-1}\big(\sigma^{t-s}H_s \setminus \{x\cdot \mathrm{e}_1 = t,\ |x-t\mathrm{e}_1|_\infty \leq r\}\big) \leq 2\eta r^{n-1}$. Since points in $\sigma^{t-s}H_s$ do not stay in $\mathcal{A}$, then we reach a contradiction with the estimate (2.14) for $\lambda = t$: we would have

$$r^{n-1} = \big|\{x\cdot \mathrm{e}_1 = t,\ |x-t\mathrm{e}_1|_\infty \leq r\}\big| \geq (1-\varepsilon)r^{n-1} + (1-2\varepsilon-2\eta)r^{n-1} = (2-3\varepsilon-2\eta)r^{n-1}. \qquad \square$$

Lemma 2.21. *With the notations of Lemma 2.17, the push forward of the measure $\mathcal{H}^{n-1}\llcorner Z$ by the map σ^t can be written as*

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

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

$$\left(\frac{h^+-t}{h^+}\right)^{n-1} \leq \frac{1}{\alpha^t} \leq \left(\frac{t-h^-}{-h^-}\right)^{n-1} \qquad \textit{for } t \geq 0,$$
$$\left(\frac{t-h^-}{-h^-}\right)^{n-1} \leq \frac{1}{\alpha^t} \leq \left(\frac{h^+-t}{h^+}\right)^{n-1} \qquad \textit{for } t < 0.$$

Proof. Lemma 2.17 ensures that the measures $\sigma^t_\sharp\mathcal{H}^{n-1}\llcorner Z$ and $\mathcal{H}^{n-1}\llcorner\sigma^t Z$ 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 $\sigma^t_\sharp\mathcal{H}^{n-1}\llcorner Z$ w.r.t. $\mathcal{H}^{n-1}\llcorner\sigma^t Z$. For the inverse mapping σ^{-t}, the Radon-Nikodym derivative is instead $\alpha^t(\sigma^t(y))^{-1}$. The last estimate, then, is straightforward from Lemma 2.17, with $s = 0$. $\square$

Corollary 2.22. *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 $\mathscr{L}^n$-measurable functions on $[h^-,h^+]\times Z$ and on $\mathcal{Z}$, since images and inverse images of $\mathscr{L}^n$-zero measure sets are $\mathscr{L}^n$-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 2.23. *The function $\tilde\alpha(t,y) := \frac{(\sigma^{-t})_\sharp\mathcal{H}^{n-1}\llcorner\sigma^t Z}{\mathcal{H}^{n-1}\llcorner Z}$ 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. $y \in Z$:

$$-\left(\frac{n-1}{\mathsf{b}(y)\cdot \mathrm{e}_1 - t}\right)\tilde{\alpha}(t,y) \le \frac{d}{dt}\tilde{\alpha}(t,y) \le \left(\frac{n-1}{t-\mathsf{a}(y)\cdot \mathrm{e}_1}\right)\tilde{\alpha}(t,y), \tag{2.15}$$

$$\left(\frac{|\mathsf{b}(y)-\sigma^t y|}{|\mathsf{b}(y)-y|}\right)^{n-1}(-1)^{\mathbb{1}_{t<0}} \le \tilde{\alpha}(t,y)(-1)^{\mathbb{1}_{t<0}} \le \left(\frac{|\sigma^t y-\mathsf{a}(y)|}{|y-\mathsf{a}(y)|}\right)^{n-1}(-1)^{\mathbb{1}_{t<0}}. \tag{2.16}$$

Moreover,

$$\int_{\mathsf{a}\cdot \mathrm{e}_1}^{\mathsf{b}\cdot \mathrm{e}_1}\left|\frac{d}{dt}\tilde{\alpha}(t,y)\right| \le 2\left(\frac{|\mathsf{b}-\mathsf{a}|^{n-1}}{|\mathsf{b}|^{n-1}}+\frac{|\mathsf{b}-\mathsf{a}|^{n-1}}{|\mathsf{a}|^{n-1}}-1\right).$$

Proof. The function $\tilde{\alpha}(t,\cdot)$ is by definition $L^1_{\mathrm{loc}}(\mathcal{H}^{n-1}\llcorner Z)$, for each fixed t. We prove that one can take suitable representatives in order to define a function, that we still denote with $\tilde{\alpha}(t,y)$, which is Lipschitz in the t variable, Borel in $y \in Z$ and satisfies the estimates in the statement.

Applying Lemma 2.17 and Corollary 2.22, for $h^- < s < t < h^+$ and every measurable $S \subset Z$, we have

$$\left(\frac{h^+-t}{h^+-s}\right)^{n-1}\int_S \tilde{\alpha}(s,y)d\mathcal{H}^{n-1}(y) \le \int_S \tilde{\alpha}(t,y)d\mathcal{H}^{n-1}(y) \le \left(\frac{t-h^-}{s-h^-}\right)^{n-1}\int_S \tilde{\alpha}(s,y)d\mathcal{H}^{n-1}(y). \tag{2.17}$$

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

$$\left[\left(\frac{h^+-t_j}{h^+-t_i}\right)^{n-1}-1\right]\tilde{\alpha}(t_i,y) \le \tilde{\alpha}(t_j,y)-\tilde{\alpha}(t_i,y) \le \left[\left(\frac{t_j-h^-}{t_i-h^-}\right)^{n-1}-1\right]\tilde{\alpha}(t_i,y). \tag{2.18}$$

One can also redefine $\tilde{\alpha}(t_j,y)$ on a $\mathcal{H}^{n-1}$-negligible set of y in order to have the inequality for all $y \in Z$. Therefore, one can redefine the pointwise values of $\tilde{\alpha}(t,y)$ for $t \notin \{t_i\}_{i\in\mathbb{N}}$ as the limit of $\tilde{\alpha}(t_{i_k},y)$ for any sequence $\{t_{i_k}\}_k$ converging to t: this defines an extension of $\tilde{\alpha}(t_i,y)$ from $\{\cup_{i\in\mathbb{N}} t_i\}\times Z$ to $(h^-,h^+)\times Z$ locally Lipschitz in t. By the above integral estimate this limit function, at any t, must be a representative of the $\mathscr{L}^1(\mathcal{H}^{n-1})$ function $\tilde{\alpha}(t,y)$ — one can see it just taking in (2.17) $t \to s^+$. By the above pointwise estimate (2.18), taking the derivative, we get (2.15). Equation (2.15), moreover, implies the following monotonicity:

$$\frac{d}{dt}\left(\frac{\tilde{\alpha}(t,y)}{(\mathrm{e}_1\cdot\mathsf{b}-t)^{n-1}}\right) \ge 0 \qquad \text{and} \qquad \frac{d}{dt}\left(\frac{\tilde{\alpha}(t,y)}{(t-\mathrm{e}_1\cdot\mathsf{a})^{n-1}}\right) \le 0.$$

Then, since $\frac{\mathrm{e}_1\cdot\mathsf{b}-t}{\mathrm{e}_1\cdot\mathsf{b}} = \frac{|\mathsf{b}-\sigma^t y|}{|\mathsf{b}-y|}$, $\frac{t-\mathrm{e}_1\cdot\mathsf{a}}{-\mathrm{e}_1\cdot\mathsf{a}} = \frac{|\mathsf{a}-\sigma^t y|}{|\mathsf{a}-y|}$ and $\tilde{\alpha}(0,\cdot)\equiv 1$, we obtain exactly (2.16). Furthermore

$$\begin{aligned}
\int_{\mathsf{a}\cdot\mathrm{e}_1}^0\left|\frac{d}{dt}\tilde{\alpha}(t,y)\right| dt &\overset{(2.15)}{\le} \int_{\{\frac{d\tilde{\alpha}(t,y)}{dt}>0\}\cap\{t<0\}}\frac{d}{dt}\tilde{\alpha}(t,y)\,dt + \int_{\mathsf{a}\cdot\mathrm{e}_1}^0 \frac{(n-1)\tilde{\alpha}(t,y)}{\mathsf{b}\cdot\mathrm{e}_1-t}\,dt \\
&\overset{(2.15)}{\le} \int_{\mathsf{a}\cdot\mathrm{e}_1}^0 \frac{d}{dt}\tilde{\alpha}(t,y)\,dt + 2\int_{\mathsf{a}\cdot\mathrm{e}_1}^0 \frac{(n-1)\tilde{\alpha}(t,y)}{\mathsf{b}\cdot\mathrm{e}_1-t}\,dt \\
&\overset{(2.16)}{\le} 1+2\int_{\mathsf{a}\cdot\mathrm{e}_1}^0 \frac{(n-1)(\mathrm{e}_1\cdot\mathsf{b}-t)^{n-2}}{(\mathrm{e}_1\cdot\mathsf{b})^{n-1}}\,dt = 1+2\left(\frac{|\mathsf{b}-\mathsf{a}|^{n-1}}{|\mathsf{b}|^{n-1}}-1\right).
\end{aligned}$$

Summing the symmetric estimate on $(0,\mathsf{b}\cdot\mathrm{e}_1)$, we get

$$\int_{\mathsf{a}\cdot \mathrm{e}_1}^{\mathsf{b}\cdot \mathrm{e}_1}\left|\frac{d}{dt}\tilde{\alpha}(t,y)\right| \le 2\left(\frac{|\mathsf{b}-\mathsf{a}|^{n-1}}{|\mathsf{b}|^{n-1}}+\frac{|\mathsf{b}-\mathsf{a}|^{n-1}}{|\mathsf{a}|^{n-1}}-1\right). \qquad \square$$

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

Lemma 2.24. *On $\mathcal{K} = \{\sigma^t Z\}_{t\in(h^-,h^+)}$, we have the following disintegration of the Lebesgue measure: for all φ such that $\int|\varphi|\,d\mathscr{L}^n < \infty$ the following integrals are well defined and equality holds*

$$\int_{\mathcal{K}}\varphi(x)\,d\mathscr{L}^n(x) = \int_Z\left\{\int_{h^-}^{h^+}\varphi(\sigma^t y)\tilde{\alpha}(t,y)\,d\mathcal{H}^1(t)\right\}d\mathcal{H}^{n-1}(y), \tag{2.19}$$

where $\tilde{\alpha}(t,\cdot)$, strictly positive, is the Radon-Nikodym derivative of $(\sigma^{-t})_\sharp\mathcal{H}^{n-1}\llcorner\sigma^t Z$ w.r.t. $\mathcal{H}^{n-1}\llcorner Z$.

Proof. Consider any integrable function φ. Then, since $(\sigma^{-t})_\sharp \mathcal{H}^{n-1} \llcorner \sigma^t S = \tilde\alpha(t,\cdot)\mathcal{H}^{n-1} \llcorner Z$ and since $\varphi \circ \sigma^t \restriction_Z$ is still $\mathscr{L}^n$-measurable (Corollary 2.22), we have

$$\int_Z \varphi(\sigma^t y)\tilde\alpha(t,y)\, d\mathcal{H}^{n-1}(y) = \int_{\sigma^t Z} \varphi(y)\, d\mathcal{H}^{n-1}(y) = \int_{\mathcal{K}\cap\{x\cdot \mathrm{e}_1 = t\}} \varphi(y)\, d\mathcal{H}^{n-1}(y).$$

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

$$\int_{\mathcal{K}} \varphi(y)\, d\mathscr{L}^n(y) = \int_{h^-}^{h^+} \int_{\mathcal{K}\cap\{x\cdot \mathrm{e}_1 = t\}} \varphi(y)\, d\mathcal{H}^{n-1}(y)\, dt = \int_{h^-}^{h^+} \int_Z \varphi(\sigma^t y)\tilde\alpha(t,y)\, d\mathcal{H}^{n-1}(y)\, dt.$$

Finally, since $\tilde\alpha$ is measurable (Corollary 2.23) and locally integrable, by the above estimate and Tonelli theorem applied to the negative and positive part, Fubini theorem provides the thesis by allowing to exchange the order of the integrals. $\square$

The following is the main theorem of the section. Before it, we set and renew the notation:

- $\{\mathcal{Z}_i\}_{i\in\mathbb{N}}$ is the partition of the transport set $\mathcal{T}_{\mathbb{e}}$ into sheaf sets as in Lemma 2.10;
- Z_i is a section of $\mathcal{Z}_i$ and $\mathfrak{d}_i$ is the relative axis;
- $\mathcal{S}$ is the quotient set of $\mathcal{T}_{\mathbb{e}}$ w.r.t. the membership to transport rays, identified with $\cup_i Z_i$.
- $\sigma^t(x) = x + \frac{t}{d(x)\cdot\mathfrak{d}_i} d(x)$ is the map moving points along rays of $\mathcal{Z}_i$;
- $\tilde\alpha_i(t,\cdot)$ is the Radon-Nikodym derivative of $(\sigma^{-t})_\sharp \mathcal{H}^{n-1} \llcorner \sigma^t Z_i$ w.r.t. $\mathcal{H}^{n-1} \llcorner Z_i$;
- $c(t,y) := \sum_i \tilde\alpha_i\big(d(y)\cdot(t\mathfrak{d}_i - y), y\big) d(y)\cdot \mathfrak{d}_i \mathbb{1}_{\mathcal{Z}_i}(y)$.

Theorem 2.25. *One has then the following disintegration of the Lebesgue measure on $\mathcal{T}_{\mathbb{e}}$*

$$\int_{\mathcal{T}_{\mathbb{e}}} \varphi(x)\, d\mathscr{L}^n(x) = \int_{\mathcal{S}} \bigg\{ \int_{a(y)\cdot d(y)}^{b(y)\cdot d(y)} \varphi(y + (t - y\cdot d(y))d(y)) c(t,y) d\mathcal{H}^1(t) \bigg\} d\mathcal{H}^{n-1}(y), \tag{2.20}$$

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

Remark 2.26. As a consequence of Corollary 2.23, c is measurable in y and locally Lipschitz in t.

Remark 2.27 (Dependence on the partition). Suppose to partition the transport set in a different family of sheaf sets $\{\mathcal{Z}'_i\}_{i\in\mathbb{N}}$, with the quotient space identified with the union $\mathcal{S}'$ of the new basis. Then, one can refine the partitions $\{\mathcal{Z}_i\}_{i\in\mathbb{N}}$ and $\{\mathcal{Z}'_i\}_{i\in\mathbb{N}}$ into a family of sheaf sets $\{\widehat{\mathcal{Z}}_i\}_{i\in\mathbb{N}}$. Consider the change of variables in a single sheaf set $\widehat{\mathcal{Z}}_i$. If we consider $Z \subset \{x\cdot v + q = 0\}$ and $Z' \subset \{x \cdot v' + q'\} = 0$, then points $y \in Z$ and the corresponding $y' \in Z'$ are related as follows

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

Moreover, we have the disintegration formulas

$$\begin{aligned}\int_{\mathcal{Z}} \varphi(x)\, d\mathscr{L}^n(x) &= \int_Z \bigg\{ \int_{a(y)\cdot d(y)}^{b(y)\cdot d(y)} \varphi(y + (t - y\cdot d(y))d(y)) c(t,y) d\mathcal{H}^1(t) \bigg\} d\mathcal{H}^{n-1}(y) \\ &= \int_{Z'} \bigg\{ \int_{a(y)\cdot d(y)}^{b(y)\cdot d(y)} \varphi(y + (t - y\cdot d(y))d(y)) c'(t,y) d\mathcal{H}^1(t) \bigg\} d\mathcal{H}^{n-1}(y),\end{aligned}$$

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 T is the map from Z to Z' and β the following Radon-Nikodym derivative

$$T(y) := y - \frac{q' + y\cdot v'}{d(y)\cdot v'} d(y) \qquad \beta := \frac{dT_\sharp \mathcal{H}^{n-1} \llcorner Z}{d\mathcal{H}^{n-1} \llcorner Z'}.$$

Proof. Forget the set of endpoints of rays, since by Lemma 2.20 they are negligible.

Consider the refinement of the partition $\{\mathcal{Z}_i\}_{i\in\mathbb{N}}$ given in Lemma 2.12, which partitions $\mathcal{T}$ into cylinders subordinated to d: denote these as $\{\mathcal{K}_{ij}\}_{ij\in\mathbb{N}}$ and set $\widehat{Z}_{ij}$, $h^{\pm}_{ij}$ in order to have $Z_i = \cup_{j\in\mathbb{N}}\widehat{Z}_{ij}$ and

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

Since we set

$$c(t,y) = \sum_i \tilde{\alpha}_i\big(d(y)\cdot(t\mathfrak{d}_i - y), y\big)d(y)\cdot\mathfrak{d}_i\mathbb{1}_{\mathcal{Z}_i}(y),$$

the local result of Lemma 2.24, with a translation and the change of variable $t \to \frac{t}{d(y)\cdot\hat{\mathfrak{d}}_i}$, yields

$$\int_{\mathcal{K}_{ij}} \varphi(x)\, d\mathcal{H}^n(x) = \int_{\widehat{Z}_{ij}} \left\{ \int_{\frac{h^-_{ij}}{d(y)\cdot\hat{\mathfrak{d}}_i}}^{\frac{h^+_{ij}}{d(y)\cdot\hat{\mathfrak{d}}_i}} \varphi\big(y + (t - y\cdot d(y))d(y)\big)c(t,y)\, d\mathcal{H}^1(t)\right\} d\mathcal{H}^{n-1}(y).$$

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

$$\begin{aligned}\int_{\mathcal{T}} \varphi(x)\, d\mathcal{H}^n(x) &= \int_{\cup_{ij}\mathcal{K}_{ij}} \varphi(x)\, d\mathcal{H}^n(x) = \sum_{ij}\int_{\mathcal{K}_{ij}} \varphi(x)\, d\mathcal{H}^n(x)\\ &= \sum_i\sum_j \int_{\widehat{Z}_{ij}} \left\{ \int_{\frac{h^-_{ij}}{d(y)\cdot\hat{\mathfrak{d}}_i}}^{\frac{h^+_{ij}}{d(y)\cdot\hat{\mathfrak{d}}_i}} \varphi(y + (t - y\cdot d(y))d(y))c(t,y)\, d\mathcal{H}^1(t)\right\} d\mathcal{H}^{n-1}(y)\\ &= \sum_i \int_{Z_i} \left\{ \int_{a(x)}^{b(x)} \varphi(y + (t - y\cdot d(y))d(y))c(t,y)\, d\mathcal{H}^1(t)\right\} d\mathcal{H}^{n-1}(y)\\ &= \int_{\cup_i Z_i} \left\{ \int_{a(x)}^{b(x)} \varphi(y + (t - y\cdot d(y))d(y))c(t,y)\, d\mathcal{H}^1(t)\right\} d\mathcal{H}^{n-1}(y).\end{aligned}$$

Separating the positive and the negative part of φ, the convergences in the steps above are monotone, if the integrals are thought on $\mathbb{R}^n$ with integrands multiplied by the characteristic function of the domains, and do not give any problem. □

2.4. **Remarks on the divergence of the ray directions.** In this section we extend the function d to be null out of $\mathcal{T}$. We consider then its distributional divergence

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

If the set of initial points, or final points as well, is compact, then it turns out to be a Radon measure concentrated on $\mathcal{T}$. More generally, it is a series of measures (see examples in Section 4). A decomposition of it can be constructed as follows. Consider the countable partition of $\mathcal{T}$ into tuft sets $\{\mathcal{K}_i\}_{i\in\mathbb{N}}$ of Lemma 2.12. 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: in this way one defines a vector field $\hat{d}$ on $\mathbb{R}^n$ which coincides with d on $\mathcal{K}_i$. The i-th addend of the series is then defined as the restriction to $\mathcal{K}_i$ of the divergence of this vector field $\hat{d}$. The absolutely continuous part of this divergences does not depend on the $\{\mathcal{K}_i\}_{i\in\mathbb{N}}$ we have chosen, as well as 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. We notice that it gives a coefficient of an ODE for the density c defined in the previous section.

Definition 2.28. 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 $\mathscr{L}^n$-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

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

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

Lemma 2.29. *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 $\phi(a) - \|x - a\|$ 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 Example 2.16, we see that the potentials

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

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

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

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:

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

It is immediate to estimate the positive part, for $x \notin K$, with

$$(\mathrm{div} d_I)_{\mathrm{a.c.}}(x) = \sum_i \frac{n-1}{|x - \mathtt{a}_i|} \mathbb{1}_{\Omega_i}(x) \le \frac{n-1}{\mathrm{dist}(x, \cup_i \mathtt{a}_i)}.$$

For the negative part, one can observe as in Proposition 4.6 of [BG] that for $B_r(x) \cap K = \emptyset$

$$(\mathrm{div} d_I)^+(B_r(x)) - (\mathrm{div} d_I)^-(B_r(x)) = \mathrm{div} d_I(B_r(x)) = \int_{\partial B_r(x)} d_I(y) \cdot \frac{y}{|y|} d\mathcal{H}^{n-1}(y) \ge -|\partial B_r(0)|.$$

The last two estimates yield

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

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. $\square$

Lemma 2.30. *Let $\mathcal{K}$ be the d-cylinder fixed above for defining $\hat{d}$ (Definition 2.28). Consider any couple $\mathcal{S}, c$ as in the disintegration Theorem 2.25. Then, for any d-sub-cylinder $\mathcal{K}'$ of $\mathcal{K}$, the following formulae hold:*

$$\partial_t c(t,y) - \Big[\big(\mathrm{div}\hat{d}\big)_{\mathrm{a.c.}}(y + (t - d(y)\cdot y)d(y))\Big]c(t,y) = 0 \qquad \mathcal{H}^n\text{-a.e. on } \mathcal{K}. \tag{2.21}$$

$$\int_{\mathcal{K}'} \varphi\, \mathrm{div}\hat{d}_i = \int_{\mathcal{K}'} \varphi\, (\mathrm{div}\hat{d}_i)_{\mathrm{a.c.}} = -\int_{\mathcal{K}'} \nabla\varphi\cdot d + \int_{\partial\mathcal{K}'^{+}-\partial\mathcal{K}'^{-}} \varphi d\cdot \mathrm{e}_1 \qquad \forall \varphi\in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^n). \tag{2.22}$$

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

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

Moreover, $\hat{d}$ is the vector field of directions relative to a potential $\hat{\phi}$: we can apply the disintegration Theorem 2.25, getting (2.20) 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 (Remark 2.26), since we are integrating on a compact support, allows the integration by parts in the t variable

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

after performing this, the above equality becomes

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

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 (2.21) 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 $\hat{c}\,\hat{\sigma}^{+}_{\sharp}\mathcal{H}^{n-1}\llcorner\widehat{\mathcal{S}} - \hat{c}\,\hat{\sigma}^{-}_{\sharp}\mathcal{H}^{n-1}\llcorner\widehat{\mathcal{S}}$. On the other hand that, taking $\varphi\mathbb{1}_{\mathcal{K}'}$, if Z is the relative section and $h^{\pm}$ define the height,

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

Coming back, integrating by parts again, one finds precisely (2.22). □

2.4.2. *Global divergence.* The divergence of the vector field d, generally speaking, is not a measure (see examples of Section 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 Subsection 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 2.31. *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 Theorem 2.25 with relative density c and sections $\mathcal{S}$, one has the relation

$$\partial_t c(t,y) - \Big[\big(\mathrm{div}d\big)_{\mathrm{a.c.}} (y + (t - d(y)\cdot y) d(y)) \Big] c(t,y) = 0 \qquad \mathcal{H}^n \, z\text{-a.e. on } \mathcal{T}, \tag{2.23}$$

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

Remark 2.32. The measurable function $(\mathrm{div}d)_{\mathrm{a.c.}}$ in general does not define a distribution, since it can fail to be locally integrable (Example 4.2, 4.3).

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. 2.30. □

Remark 2.33. Since c does not depend on the construction of the vector fields $\hat{d}_i$, then Equation (2.23) ensures that $(\mathrm{div}d)_{\mathrm{a.c.}}$ is independent of the choices we made to obtain $\hat{d}_i$. Moreover, by Corollary 2.23, 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 2.34. *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, $\forall \varphi \in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^n \setminus \cup_x a(x) \cup b(x))$,

$$\langle \mathrm{div}\hat{d}, \varphi \rangle = \int \varphi \, (\mathrm{div}d)_{\mathrm{a.c.}}. \tag{2.24}$$

Proof. Since $d = \hat{d}_i$ on $\mathcal{K}_i$, Equation (2.22) 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 Lemma 2.25, it follows that the divergence of d is the sum of the above measures: $\forall \varphi \in \mathrm{C}^\infty_{\mathrm{c}}(\mathbb{R}^n)$

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

Equation (2.24) 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 the proof of Sudakov theorem, under the assumption of strict convexity of the norm. We do not assume μ, ν compactly supported, but we require to avoid trivialities that the optimal cost is finite.

The argument is a reduction to dimension 1, since the 1-dimensional theory is well established: we state the known result in Theorem 3.1, which ensures the existence of an optimal transport map for the Monge problem when the initial measure is absolutely continuous.

We observe first the following fact. There exist a closed set Γ containing the support of any optimal transport plan $\pi \in \Pi(\mu,\nu)$ such that function

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

is 1-Lipschitz w.r.t. $\|\cdot\|$ and satisfies $\phi(x) - \phi(y) = \|y - x\|$ for every $(x, y) \in \Gamma$. For completeness we sketch the proof in Remark 3.3.

In the present section we fix the potential ϕ as in (3.1) and we consider the related transport set $\mathcal{T}_{\mathsf{e}}$. The transport rays of $\mathcal{T}_{\mathsf{e}}$ are invariant sets for the transport. In fact, given any optimal transport plan π, by construction if (x, y) belongs to the support of π then $y \in \mathcal{P}(x)$, being $\phi(x) - \phi(y) = \|y - x\|$: therefore, if one disintegrates π w.r.t. the projection onto the first set of n-variables, the conditional probability π_x is concentrated on $\mathcal{P}(x)$ for μ-a.e. $x \in \mathbb{R}^n$. This means that the mass is transported within the rays, in the direction where ϕ decreases.

The analysis performed in Section 2 yields the following information. Up to removing an $\mathscr{L}^n$-negligible set from $\mathcal{T}_{\mathsf{e}}$, the relation which defines the transport rays, precisely

$$x \sim y \qquad \text{if} \qquad \phi(x) - \phi(y) = \|y - x\|,$$

is an equivalence relation. Equivalently, we are saying that the transport rays provide a partition of $\mathcal{T}$ up to a μ-negligible set, into segments since the norm is strictly convex. Moreover, if one disintegrates the Lebesgue measure on $\mathcal{T}_{\mathsf{e}}$ w.r.t. this partition — identifying points on a same transport ray — the conditional measures are absolutely continuous by Theorem 2.25.

This allows to conclude the strategy proposed by Sudakov in [Sud], as we outline here before the formal proof. One disintegrates μ w.r.t. the partition into transport rays, that we can denote with $\{r_y\}_{y \in \mathcal{S}}$ where $\mathcal{S}$ is a σ-compact subset of countably many hyperplanes and $\mathcal{H}^{n-1} \llcorner S$ is absolutely continuous w.r.t. the quotient measure. This is possible by Theorem 2.25 because $\mu \ll \mathscr{L}^n$, and one obtains absolutely continuous conditional probabilities $\{\mu_y\}_{y \in \mathcal{S}}$. If $\nu \ll \mathscr{L}^n$, one can moreover disintegrate also ν in the same way.
As the rays are invariant sets, one obtains that any optimal transport plan $\pi \in \Pi(\mu, \nu)$ can in turn be disintegrated w.r.t. the partition $\{r_y \times r_y\}_{y \in \mathcal{S}}$ and the conditional probabilities $\{\pi_y\}_{y \in \mathcal{S}}$ belong respectively to $\Pi(\mu_y, \nu_y)$: each optimal transport plan is a superposition of transport plans on the rays.
Denoting with m the quotient measure of π and $\gamma := \frac{dm}{d\mathcal{H}^{n-1}}$, the optimal cost can be written as

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \|y - x\| \, d\pi(x, y) = \int_{\mathcal{S}} \left\{ \int_{\mathbb{R}^n \times \mathbb{R}^n} \|y - x\| \gamma(z) \, d\pi_z(x, y) \right\} d\mathcal{H}^{n-1}(z).$$

One can thus obtain other optimal transport plans rearranging the transport between each μ_z, ν_z without increasing the cost realized by π_z w.r.t. $c(x, y) = \|y - x\|$. As the optimal transports from μ_z to ν_z are within the ray r_z, this means that we reduced the original problem to 1-dimensional transport problems, *and by the study of Section 2 we know that the initial measures $\{\mu_y\}_{y \in \mathcal{S}}$ are absolutely continuous.*

An optimal map solving our original Monge problem will then be defined by defining an optimal transport map for the transport problem on each ray r_z: indeed, if one defines a map on each ray r_z, then by juxtaposition a map is defined on μ-almost all of $\mathbb{R}^n$.

While, as already observed in the introduction, the absolute continuity of μ is important, the assumption of absolute continuity of ν is just technical. It was present in Sudakov statement, but it has been removed in subsequent works ([AP]). When ν is singular, then a positive mass can be transported to points belonging to more transport rays: as ν can give positive measure to endpoints, transport rays do not partition any carriage of ν and therefore it is not immediate how to disintegrate ν on $\mathcal{T}_{\mathsf{e}} \setminus \mathcal{T}$ in order to obtain the right conditional measures on the rays. This however is a formal problem, since rays are invariant sets: for any sheaf set $\mathcal{Z}$, one can determine the portion of mass $\nu(\mathcal{Z} \setminus \mathcal{T})$ on the terminal points of $\mathcal{Z}$ which comes from the rays in $\mathcal{Z}$ as the mass initially present in $\mathcal{Z}$ which has not been carried to $\mathcal{Z} \cap \mathcal{T}$; it is just the difference

$$\mu(\mathcal{Z}) - \nu(\mathcal{Z} \cap \mathcal{T}).$$

We recall before the main theorem the 1-dimensional result, that we cite from Theorem 5.1 in [AP].

Theorem 3.1 (1-dimensional theory)**.** *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*

$$t(x) = \sup \left\{ y \in \mathbb{R} : \; F(y) \leq G(x) \right\} \tag{3.2}$$

(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 give now the solution to the Monge problem. For clarity we specify the setting:

- We consider two probability measures μ, ν with μ absolutely continuous w.r.t. $\mathscr{L}^n$.
- The cost function is given by $c(x, y) = \|y - x\|$, and we assume that the optimal cost is finite.
- We fix the potential ϕ in (3.1) and we construct the transport set and the partition in transport rays $\{r_z = [\![a(z), b(z)]\!]\}_{z\in\mathcal{S}}$ as in Section 2.2, so that Theorem 2.25 holds.
- $p : \mathcal{T} \to \mathcal{S}$ denotes the projection onto the quotient, for the partition into transport rays.

We define the following auxiliary measures μ_z, ν_z on the ray r_z:

- Let $\mu = \int_{\mathcal{S}} \mu_z\, m(z)$ be the disintegration of $\mu = \mu \llcorner \mathcal{T}$ w.r.t. the partition in transport rays.
- Let $\nu \llcorner \mathcal{T} = \int_{\mathcal{S}} \tilde\nu_z \tilde m(z)$ be the disintegration of $\nu \llcorner \mathcal{T}$ w.r.t. the partition in transport rays.
- Set $\nu_z = f(z)\delta_{b(z)} + g(z)\tilde\nu_z$, where f, g are the Radon-Nikodym derivatives

$$f = \frac{d(m - \tilde m)}{dm} \qquad \text{and} \qquad g(z) = \frac{d\tilde m}{dm}.$$

Theorem 3.2. *Define on* $\mathcal{T}$ *the two cumulative functions*

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

Then, an optimal transport map for the Monge-Kantorovich problem between μ *and* ν *is given by*

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

Every optimal plan has the form $\pi \llcorner (\mathcal{T} \times \mathcal{T}_{\mathsf{e}}) = \int_{\mathcal{S}} \pi_z\, dm(z)$*, with* $\pi_z \in \Pi(\mu_z, \nu_z)$ *optimal.*

Proof. The proof follows the 1-dimensional reduction argument described in the introduction.

Step 1: Absolute continuity of the conditional probabilities. By assumption there exists a nonnegative integrable function f such that $\mu = f\mathscr{L}^n$. By Theorem 2.25

$$\begin{aligned}\int_{\mathcal{T}_{\mathsf{e}}} \varphi(x)\, d\mu(x) &= \int_{\mathcal{T}_{\mathsf{e}}} \varphi(x) f(x)\, d\mathscr{L}^n(x) \\ &= \int_{\mathcal{S}} \bigg\{ \int_{a(z)\cdot d(z)}^{b(z)\cdot d(z)} \varphi(z + (t - z\cdot d(z))d(z)) f(z + (t - z\cdot d(z))d(z)) c(t, z) d\mathcal{H}^1(t) \bigg\} d\mathcal{H}^{n-1}(z),\end{aligned}$$

Moreover, by the definition of disintegration, and since $m \ll \mathcal{H}^{n-1} \llcorner S$,

$$\int_{\mathcal{T}_{\mathsf{e}}} \varphi\, d\mu = \int_{\mathcal{S}} \bigg\{ \int_{\mathcal{R}(z)} \varphi\, d\mu_z \bigg\} dm(z) = \int_{\mathcal{S}} \bigg\{ \int_{\mathcal{R}(z)} \varphi \frac{dm}{d\mathcal{H}^{n-1}}(z)\, d\mu_z \bigg\} d\mathcal{H}^{n-1}(z).$$

Then, denoting

$$i(z) = \int_{a(z)\cdot d(z)}^{b(z)\cdot d(z)} f(z + (s - z\cdot d(z))d(z)) c(s, z) d\mathcal{H}^1(s),$$

one obtains

$$\mu_z(x) = \frac{f(z + (x - z\cdot d(z))d(z)) c(x, z)}{i(z)} \mathcal{H}^1(x).$$

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

Step 2: Solution of the transport problem on a ray. By the 1-dimensional theory, Theorem 3.1, an optimal transport map from $(\mathcal{R}(y), \mu_y)$ to $(\mathcal{R}(y), \nu_y)$ is given by the restriction of T in (3.3) to $\mathcal{R}(y)$.

Step 3: Measurability of T. 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 2.10. In particular, a subset C of $\mathcal{T}$ is Borel if and only if its intersections with $\mathcal{Z}(Z_k)$ are Borel. Moreover, composing T with the Borel change of variable given in Remark 2.14, from the sheaf set $\mathcal{Z}$ we can reduce to $(0, 1)\mathrm{e}_1 + Z$,

$d(x) = \mathrm{e}_1$ and the map T takes the form $T(y) = y + (T \cdot \mathrm{e}_1 - y \cdot \mathrm{e}_1)\mathrm{e}_1$. One, then, has just to prove that the map $T \cdot \mathrm{e}_1$ is Borel: this map is monotone in the first variable, and Borel in the others; in particular, it is Borel on $(0,1)\mathrm{e}_1 + \cdot Z$.

Step 4: Disintegration of optimal transport plans. Consider any transport plan $\pi \in \Pi(\mu,\nu)$. By construction of ϕ with (3.1), if (x,y) belong to the support of π then $\phi(x) - \phi(y) = \|y - x\|$: the support of π is then contained in $\cup_{y \in S} \mathcal{R}(y) \times \mathcal{R}(y) \cup \{x = y\}$.

Moreover, one can forget of the points out of $\mathcal{T}$, since they stay in place: $\pi(((\mathbb{R}^n \setminus \mathcal{T}) \times \mathbb{R}^n) \setminus \{x = y\}) = 0$; as a consequence $\pi \llcorner (\mathbb{R}^n \setminus \mathcal{T} \times \mathbb{R}^n)$ is already induced by the map T. We assume then for simplicity $\pi(\mathcal{T} \times \mathbb{R}^n) = 1$, eventually considering the transport problem between the marginals of $\pi \llcorner (\mathcal{T} \times \mathbb{R}^n)/\pi(\mathcal{T} \times \mathcal{T}_{\mathrm{e}})$.

As a consequence, one can disintegrate π w.r.t. $\{(\!| a(y), b(y) |\!) \times \mathcal{R}(y)\}_{y \in \mathcal{S}}$ by Theorem A.6.

By the marginal condition $\mu(A) = \pi(A \times \mathbb{R}^n) = \pi(A \times \mathcal{R}(A))$ the quotient measure then is still m:

$$\pi = \int_{y \in \mathcal{S}} \pi_y dm(y), \qquad \pi_y(\mathcal{R}(y) \times \mathcal{R}(y)) = 1.$$

Moreover, for m-a.e y the plan π_y transports μ_y to ν_y: for all measurable $S' \subset \mathcal{S}$, $A \subset \mathbb{R}^n$

$$\begin{aligned}
\int_{S'} \pi_z(A \times \mathbb{R}^n) dm(z) &= \pi((A \cap \mathcal{Z}(S')) \times \mathbb{R}^n) = \mu(A \cap \mathcal{Z}(S')) = \int_{S'} \mu_z(A) dm(z) \\
\int_{S'} \pi_z(\mathbb{R}^n \times A) dm(z) &= \pi(\mathcal{Z}(S') \times (\mathcal{Z}(S') \cap A \setminus \mathcal{T})) + \pi(\mathbb{R}^n \times (\mathcal{Z}(S') \cap A \cap \mathcal{T})) \\
&= [\mu(\mathcal{Z}(S') \cap \mathcal{Z}(A \setminus \mathcal{T})) - \nu(\mathcal{T} \cap \mathcal{Z}(S') \cap \mathcal{Z}(A \cap \mathcal{T}))] + \nu(\mathcal{Z}(S') \cap A \cap \mathcal{T}) \\
&= \int_{S'} \delta_{b(z)}(A)\, d(m(z) - \tilde{m}(z)) + \int_{S'} \tilde{\nu}_z(A)\, d\tilde{m}(z) = \int_{S'} \nu_z(A)\, dm(z).
\end{aligned}$$

Step 5: Optimality of T. Since $T \lceil_{r_z}$ is an optimal transport between μ_z and ν_z (Step 2), then

$$\int_{\mathcal{R}(z) \times \mathcal{R}(z)} \|x - y\|\, d\pi_z(x,y) \geq \int_{\mathcal{R}(z)} \|x - T(x)\|\, d\mu_z(x), \tag{3.4}$$

where $\pi = \int_{\mathcal{S}} \pi_z\, dm(z)$ is any optimal transport plan, as in Step 3. Therefore T is optimal:

$$\begin{aligned}
\int \|x - y\|\, d\pi' \geq \int \|x - y\|\, d\pi &= \int_{\mathcal{S}} \left\{ \int_{\mathcal{R}(z) \times \mathcal{R}(z)} \|x - y\|\, d\pi_z(x,y) \right\} dm(z) \\
&\overset{(3.4)}{\geq} \int_{\mathcal{S}} \left\{ \int_{\mathcal{R}(z) \times \mathcal{R}(z)} x - T(x)\, d\mu_z(x) \right\} dm(z) = \int \|x - T(x)\|\, d\mu
\end{aligned}$$

for every $\pi' \in \Pi(\mu,\nu)$. This yields to the existence of an optimal transport map of the form

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

where T_z is a one-dimensional, optimal transport map from μ_z to ν_y, when $\nu(\cup_x b(x)) = 0$. $\square$

We sketch finally in the following remark the proof of the standard claim in the introduction. One could see that the potential defined by (3.1) with instead of Γ the support of *any* optimal transport plan $\pi \in \Pi(\mu,\nu)$ has the same property: it is 1-Lipschitz and its c-subdifferential contains the support of any other optimal transport plan. We omit it since not needed.

Remark 3.3. We recall the definition of c-monotonicity: given a cost function $c : \mathbb{R}^{2n} \to \mathbb{R}^+$, a set $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ is c-cyclically monotone (briefly c-monotone) if for all $M \in \mathbb{N}$, $(x_i, y_i) \in \Gamma$ one has

$$\sum_{i=1}^{M} c(x_i, y_i) \leq \sum_{i=1}^{M} c(x_{i+1}, y_i), \qquad x_{M+1} := x_1.$$

By Theorem 5.2 in [AP], when the cost function is continuous the support of any optimal transport plan with finite cost is c-monotone. Consider then a sequence $\{\pi_k\}_{k \in \mathbb{N}}$ dense, w.r.t. the weak*-topology, in the set of optimal transport plans for $c(x,y) = \|y - x\|$ and define as Γ the support of the optimal transport plan $\sum_{k \in \mathbb{N}} 2^{-k} \pi_k$. By the u.s.c. of the Borel probability measures on closed sets w.r.t. weak*-convergence, Γ contains the support of any other optimal transport plan.

The function ϕ defined by (3.1),

$$\phi(x)=\inf_{\substack{M,\,x_M:=x,\\ i=1,\dots,M-1,\\ (x_i,y_i)\in\Gamma}}\left\{\sum_{i=0}^{M-1}\big[\|y_i-x_{i+1}\|-\|y_i-x_i\|\big]\right\},\qquad x\in\mathbb{R}^n,\ (x_0,y_0)\in\Gamma\text{ fixed},$$

is trivially 1-Lipschitz, being the infimum of 1-Lipschitz functions. Moreover, if $(x,y)\in\Gamma$, then for every $(M-1)$-uple $(x_i,y_i)\in\Gamma$ one chooses in order to compute $\phi(x)$ one has that $\{(x_i,y_i)\}_{i=1}^{M-1}\cup\{(x,y)\}$ is a M-uple to compute $\phi(y)$, and then

$$\phi(y)\le\|y-y\|-\|y-x\|+\inf_{\substack{M,\,x_M:=x,\\ i=1,\dots,M-1,\\ (x_i,y_i)\in\Gamma}}\sum_{i=0}^{M}\big[\|y_i-x_{i+1}\|-\|y_i-x_i\|\big]=-\|y-x\|+\phi(x).$$

ϕ is real valued as a consequence of c-monotonicity, which implies $\phi(x_0)\ge 0$ and thus $\phi(x_0=0)$.

In general ϕ is not integrable w.r.t. neither μ nor ν. When it is, clearly by the marginal condition $\int c\pi=\int\phi\mu-\int\phi\nu$ for every optimal plan π. Notice however that if $\int c\pi=\int\tilde{\phi}\mu-\int\tilde{\phi}\nu$ with $\tilde{\phi}$ 1-Lipschitz, then necessarily $c(x,y)=\tilde{\phi}(x)-\tilde{\phi}(y)$ for π-a.e. (x,y).

4. Remarks on the Decomposition

We first recall an example from [AKP]. This shows that the absolute continuity of the conditional measures in the disintegration of the Lebesgue measure established in Theorem 2.25 relies on some regularity of the vector field of ray directions, since the Borel measurability is not enough.

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

Theorem. 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.*

This example contradicts Proposition 78 in Sudakov proof ([Sud]): 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 1-dimensional measure on that segments, even if the vector field of directions is Borel. Notice moreover that the set of initial points of the segments from $x\in M_N$ to $(f_2(x),2)$ has $\mathscr{L}^3$ measure one, being the whole M_N.

Another counterexample can be found in [Lar].

In the following two examples we show on one hand that the divergence of the vector field of ray directions can fail to be a Radon measure. On the other hand, we see that in general the transport set is merely a σ-compact subset of $\mathbb{R}^n$: we consider just below, before of the examples, an alternative definition of $\mathcal{T}$, which extends it and has analogous properties; however in dimension $n\ge 2$ even this extension does not fill the space, for any Kantorovich potential of the transport problem — as shown in Example 4.3.

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

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

This means that $\pm\nabla\phi\in D$. Consider now a point where, moreover, there is an outgoing ray. As an immediate consequence of (2.2), just differentiating in the direction of the outgoing ray, we have the relation

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

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

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

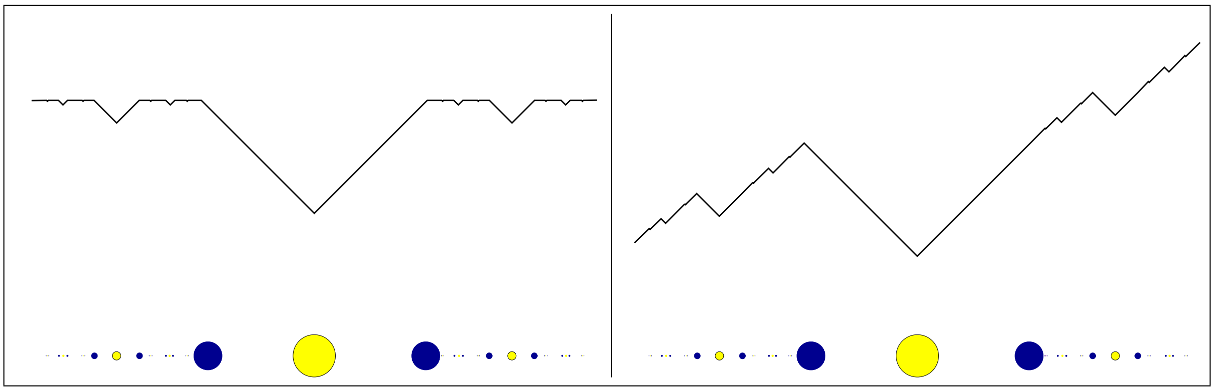

FIGURE 3. Example 4.2. The pictures show the graphs of the Kantorovich potentials ϕ (on the left) and $\tilde{\phi}$ (on the right), for the same transport problem — source masses are the blue ones, destinations the yellow ones. With $\tilde{\phi}$ the vector field of ray directions is defined $\mathscr{L}^1$-a.e., while with ϕ this is not the case, since the gradient vanishes in a $\mathscr{L}^1$-positive measure set.

Equation (4.1) suggests another possible definition of d. Assuming the norm 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 Example 4.2), and has analogous properties. However, even in this case the vector field of direction is not generally defined in positive $\mathscr{L}^n$-measure sets. In fact, in Example 4.3 we find that the gradient of ϕ can vanish on sets with $\mathscr{L}^n$-postive measure.

Example 4.2 (Transport rays do not fill continuously the line). Consider in $[0,1]$ the following transport problem, with $c(x,y) = |y-x|$ (Figure 3).

Fix $\ell \in (0,1/4)$. Construct the following 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 $y_{ik}+\ell^{k+1}(-1,1)$ — where $y_{1k}, \ldots, y_{2^k k}$ 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

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

The map bringing the mass in $y_{ik} \pm \ell^k$ to y_{ik} is induced by the plan

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

and it is easily seen to be the optimal one (e.g. by [AP], checking c-monotonicity).

Clearly, it is not relevant how the map is defined out of $\cup_{ik}\{y_{ik} \pm \ell^k\}$.

We can consider a first Kantorovich potential ϕ given by

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

ϕ is differentiable 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 ray directions by (4.2).
Notice that the divergence of the vector field of ray directions is not a locally finite measure.

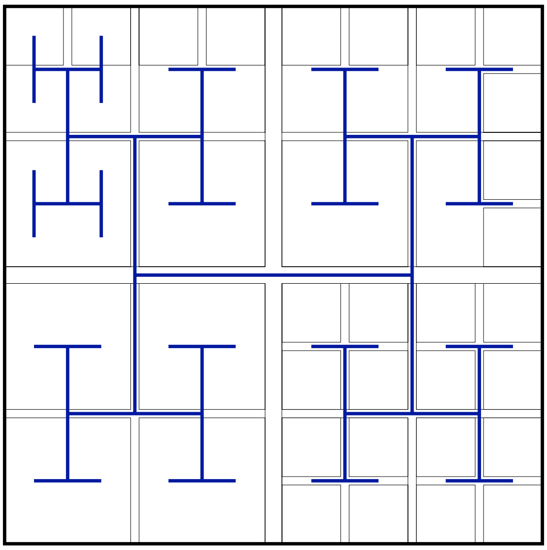

FIGURE 4. Example 4.3. In this case, however one chooses the potential, the vector field of direction is not defined on the whole square. In fact, the potential must be constant in points, belonging to the blue skeleton we begin to draw, dense in a $\mathscr{L}^2$-positive measure set.

Define now another Kantorovich potential $\tilde{\phi}$ as follows. Consider the limit of the functions

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

It is 1-Lipschitz, constant on the intervals we took away. In particular, $\tilde{\phi} := \phi + h$ is again a good potential, which is precisely the one defined in (3.1). Notice that the direction field of rays relative to the potential $\tilde{\phi}$ is defined $\mathscr{L}^1$-almost everywhere, just except in the atoms of μ. It is 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. Again, the divergence of the vector field fails to be a locally finite Radon measure.

Observe that, spreading the atomic measures on suitable small intervals, one gets an analogous example with marginals absolutely continuous w.r.t. $\mathscr{L}^1$.

Example 4.3 ($\mathcal{T}$ does not fill the space). Consider in the unit square $X = Y = [0,1]^2$ the following transport problem (see Figure 4).

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

$$c_1 = \left(\frac{1}{2}, \frac{1}{2}\right), \qquad \Big\{c_h\Big\}_{h=\frac{4^i+2}{3},\dots,\frac{4^{i+1}-1}{3}} = \Big\{c_j \pm (\ell_i + a_i)(\mathrm{e}_1 \pm \mathrm{e}_2)\Big\}_{j=\frac{4^{i-1}+2}{3},\dots,\frac{4^i-1}{3}} \quad \text{for } i \in \mathbb{N}$$

and, finally, the intermediate points

$$z_{i,j,0} = c_i + j r_i \mathrm{e}_1 \quad \text{and} \quad z_{i,j,\pm} = c_i \pm (\ell_1 + a_i) r_i \mathrm{e}_1 + j r_i \mathrm{e}_2 \quad \text{for } i \in \mathbb{N},\ j \in \{-n_i, \dots, n_i\}.$$

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 optimal (e.g. by [AP], checking c-monotonicity).

Let ϕ be any 1-Lipschitz function whose c-subdifferential contains the support of π. Since $\phi(y)-\phi(x) = \|y-x\|$ must hold for all (x,y) in the support of π, 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 of this is example that in general there is no choice of potential ϕ such that the extension of the transport set defined in (4.2) fills the space.

Observe that, spreading the atomic measures on suitable small squares, one gets an analogous example with marginals absolutely continuous w.r.t. $\mathscr{L}^2$.

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 measures are investigated. We present firstly the abstract setting and definition. Then we state an existence and essential uniqueness theorem. Finally, we give two trivial examples in $\mathbb{R}^n$.

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. We will equivalently speak of partitions and functions (whose level set define a partition).

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

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

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

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

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

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

Definition A.2. A *disintegration of ρ consistent with* r is a map $S \ni s \mapsto \rho_s \in \mathcal{P}(R,\mathscr{R})$, where $\mathcal{P}(R,\mathscr{R})$ are the probability measures on $(R,\mathscr{R})$, such that $\rho_\cdot(B)$ is η-measurable for all $B \in \mathscr{R}$ and satisfies

$$\rho(B \cap r^{-1}(S)) = \int_S \rho_s(B)\, dr_\sharp\rho(s) \qquad \forall B \in \mathscr{R}, \forall S \in r_\sharp(\mathscr{R}). \tag{A.1}$$

The measures $\{\rho_s\}_{s\in S}$ are the *conditional probabilities*.
A disintegration is *strongly consistent* if for η-a.e. s the measure ρ_s is carried by $r^{-1}(s)$:

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

Definition A.3. A σ-algebra is *countably generated* if there is a countable basis generating it. A σ-algebra $\mathcal{A}$ is *essentially countably generated*, w.r.t. a measure m, if there is a countably generated σ-algebra $\widehat{\mathcal{A}}$ such that for all $A \in \mathcal{A}$ there exists $\hat{A} \in \widehat{\mathcal{A}}$ satisfying $m(A \vartriangle \hat{A}) = 0$.

Remark A.4. If $\mathscr{R}$ is countably generated or complete, $\rho_\cdot(B)$ is $\mathscr{S}$-measurable, not only η-measurable.

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 [BC1].

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 $s \to \rho_s$ consistent with r.*

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 $(\mathscr{L}^*, \lambda^*)$.*

Clearly, the disintegration of a finite measure is completely analogous. One can moreover extend the notion of disintegration consistent with a map (or, equivalently, with a partition) when $\rho \in \mathcal{M}^+(R, \mathcal{R})$ is a σ-finite nonnegative measure, as a map $S \ni s \mapsto \rho_s \in \mathcal{M}^+(R, \mathscr{R})$ such that $\rho_\cdot(B)$ is η-measurable for all $B \in \mathscr{R}$ and satisfying

$$\rho(B \cap r^{-1}(S)) = \int_S \rho_s(B) \, d\theta(s) \qquad \forall B \in \mathscr{R}, \forall S \in r_\sharp(\mathscr{R}), \tag{A.2}$$

where θ is a σ-finite measure on $(S, \mathscr{R})$ absolutely continuous w.r.t. $r_\sharp\rho$, possibly different from it. The measures $\{\rho_s\}_{s\in S}$ are the conditional measures relative to θ. The disintegration is strongly consistent if $\rho(R \setminus r^{-1}(s)) = 0$ for θ-a.e.s.

The following two examples provide the basic meaning of what a disintegration consistent with a map is. The third example, instead, shows that a disintegration does not need to be strongly consistent, 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.

Example A.7. Let $R = [0,1]^2$. Partition R by slicing it with the parallel hyperplanes $H_\lambda = \{x \cdot \mathrm{e}_1 = \lambda\}$, for $\lambda \in [0,1]$. Then, the quotient image space of $\big(R, \mathcal{B}(R), \mathscr{L}^2 \llcorner R\big)$ is simply $\big([0,1], \mathcal{B}([0,1]), \mathscr{L}^1 \llcorner [0,1]\big)$. Defining the conditional probabilities

$$\rho_\lambda = \mathcal{H}^{n-1} \llcorner H_\lambda, \qquad \lambda \in [0,1]$$

one has the strongly consistent disintegration $\mathscr{L}^2 \llcorner R = \int_{[0,1]} \rho_\lambda \, d\mathscr{L}^1(\lambda)$.

Example A.8. Let $R = \mathbf{B}^{n-1} \setminus \{0\}$. Partition R into rays centered in the origin: set $r_d = (\mathbb{R}d) \cap R$ for every direction $d \in \mathbf{S}^{n-1}$. Then the quotient space of $(R, \mathcal{B}(R), \mathscr{L}^n \llcorner R)$ is $(\mathbf{S}^{n-1}, \mathcal{B}(\mathbf{S}^{n-1}), \mathcal{H}^{n-1} \llcorner \mathbf{S}^{n-1})$ and the conditional measures are $\{\rho_d = t^{n-1}\mathcal{H}^1(t) \llcorner r_d\}_{d\in\mathbf{S}^{n-1}}$.

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 $\rho_s = \mathscr{L}^1 \llcorner [0,1]$ for all $s \in V$. In particular, this disintegration is not strongly consistent.

Appendix B. Notations

The following table lists some notations of this article.

Id	The identity function, $\mathrm{Id}(x) = x$
$\mathbb{1}_S$	The function vanishing out of S, equal to one on S (where $S \subset \mathbb{R}^n$)
$(\!(a, b)\!)$	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
e_k	$\{\mathrm{e}_1, \dots, \mathrm{e}_n\}$ 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
$\mathcal{B}(X)$	The Borel subsets of $X \subset \mathbb{R}^n$
$\mathcal{H}^\alpha$	The α-dimensional Hausdorff measure in $\mathbb{R}^n$
$\mathscr{L}^n$	The Lebesgue measure on $\mathbb{R}^n$
$\ll$	Denotes that a measure is absolutely continuous w.r.t. another one
$\langle\cdot\rangle$	Denotes the linear span
$\langle\cdot,\cdot\rangle$	$\langle\theta, \varphi\rangle = \int \varphi\, d\theta$, where θ is a measure and φ is a θ-integrable function
$\tau_\sharp$	The push forward with a measurable map τ, see Appendix A
$\Pi(\mu, \nu)$	The set of transport plans between two probability measures μ and ν
$f\restriction_S$	The restriction of the function f to a set S
$\theta \llcorner S$	For A θ-measurable, $\theta \llcorner S(A) = \theta(A \cap S)$
$\int \theta_s\, m$	$\theta = \int \theta_s\, dm(s)$ denotes the disintegration of a measure θ, see Appendix A
μ, ν	Probability measures on $\mathbb{R}^n$, $\mu \ll \mathscr{L}^n$
$\lVert\cdot\rVert$	A possibly asymmetric norm on $\mathbb{R}^n$ whose unit ball is strictly convex
$c(x, y)$	The cost function $c(x, y) = \lVert y - x\rVert$
D^*	The unit ball $\{x \in \mathbb{R}^n : \lVert x\rVert \le 1\}$
D	The dual convex set of D^*: $D = \left\{\ell :\ \ell \cdot d \le 1\ \forall d \in D^*\right\}$
∂D	The boundary of D
δD	The support cone of D at $\ell \in \partial D$: $\delta D(\ell) = \left\{d \in \partial D^* :\ d \cdot \ell = 1 = \sup_{\hat\ell \in \partial D} d \cdot \hat\ell\right\}$
ϕ	See Definition 2.1 and (3.1)
$\mathcal{T}$, $\mathcal{T}_{\mathrm{e}}$	See Definition 2.2
$\partial_c\phi$	The c-subdifferential of ϕ: $\partial_c\phi = \left\{(x, y) :\ \phi(x) - \phi(y) = c(x, y)\right\}$
$\partial^-\phi$	The *sub-differential* of a function $\phi : \mathbb{R}^n \mapsto \mathbb{R}$: $\partial^-\phi(x) = \left\{v^* :\ \phi(y) \ge \phi(x) + v^* \cdot (y - x)\ \forall y\right\}$
$\sigma^\cdot(\cdot)$	See Definition 2.11, and also Page 10 for $\sigma^\cdot_{d_I}(\cdot)$
$\mathcal{Z}$, Z	Sheaf set and its basis, see Definition 2.9
$\mathcal{K}$	d-cylinder, see Definition 2.11
$\mathcal{P}$, $\mathcal{R}$	See Definition 2.3
$\mathcal{D}$, d	Directions of the rays, see (2.4) and (2.7)
$\mathcal{S}$	See Theorems 2.25, 3.2
α	See Lemma 2.21
$\tilde\alpha$	See Corollary 2.23, Lemmata 2.21, 2.24
$c(t, z)$	See Theorem 2.25, Lemma 2.30, (2.23)

Acknowledgements. The author wishes to express her gratitude to professor Stefano Bianchini, for his precious advice and many helpful discussions. She thanks moreover the anonymous referee for several suggestions leading to an improvement of the exposition.

References

[AFP] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[AKP] L. Ambrosio, B. Kirchheim, and A. Pratelli. Existence of optimal transport maps for crystalline norms. *Duke Math. J.*, 125(2):207–241, 2004.

[AP] 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.

[BC1] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. *Bull. Inst. Math., Acad. Sin. (N.S.)*, 4(4):353–454, 2009.

[BC2] S. Bianchini and F. Cavalletti. Monge problem in metric spaces. Preprint.

[BG] S. Bianchini and M. Gloyer. On the Euler Lagrange equation for a variational problem: the general case II. To appear on *Math. Z.*

[Bia] S. Bianchini. On the Euler-Lagrange equation for a variational problem. *Discrete Contin. Dyn. Syst.*, 17(3):449–480, 2007.

[CFM] 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.

[CP1] T. Champion and L. De Pascale. The Monge problem for strictly convex norms in $\mathbb{R}^d$. To appear on *Journ. of the Eur. Math. Soc.*

[CP2] T. Champion and L. De Pascale. The Monge problem in $\mathbb{R}^d$. Preprint, 2009.

[EG] L. C. Evans and W. Gangbo. Differential equations methods for the Monge-Kantorovich mass transfer problem. *Current Developments in Mathematics*, pages 65–126, 1997.

[Kan1] L. V. Kantorovich. On the transfer of masses. *Docl. Akad. Nauk. SSSR*, 37:227–229, 1942.

[Kan2] L. V. Kantorovich. On a problem of Monge. *Uskpekhi Mat. Nauk.*, 3:225–226, 1948.

[Lar] 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.

[Mon] G. Monge. Mémoire sur la Theorie des Déblais et des Remblais. *Histoire de l'Acad. de Sciences de Paris*, pages 666–704, 1781.

[Sud] 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).

[TW] N. S. Trudinger and X. J. Wang. On the Monge mass transfer problem. *Calc. Var. PDE*, (13):19–31, 2001.

[Vil] C. Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.

Laura Caravenna, SISSA, via Beirut 4, 34014 Trieste, Italy
E-mail address: `laura.caravenna@sns.it`