Scuola Internazionale Superiore di Studi Avanzati
Settore di Analisi Funzionale e Applicazioni

Ph.D. Thesis

The Disintegration Theorem and Applications to Optimal Mass Transport

Candidate:
Laura Caravenna

Advisor:
Prof. Stefano Bianchini

Accademic Year 2008/09

Dichiarazione

Il presente lavoro costituisce la tesi presentata da Laura Caravenna, sotto la direzione di ricerca del Prof. Stefano Bianchini, al fine di ottenere l'attestato di ricerca post-universitaria di *"Doctor Philosophiae" in Analisi Matematica* presso la Scuola Internazionale Superiore di Studi Avanzati, settore di Analisi Funzionale e Applicazioni. Ai sensi dell' art. 18, comma 3 dello Statuto della Sissa pubblicato sulla G.U. no 62 del 15.03.2001, il predetto attestato è equipollente al titolo di *Dottore di Ricerca in Matematica*.

Trieste, Settembre 2009.

Declaration

The present work constitutes the thesis presented by Laura Caravenna, with the supervisory of Prof. Stefano Bianchini, in order to obtain the research title *"Doctor Philosophiae" in Mathematical Analysis* at the International School of Advanced Studies (SISSA), sector of Functional Analysis and Applications. By the article 18, comma 3 of Sissa Statute, published on G.U. number 62, 15 March 2001, this research title is equipollent to the Italian title *Dottore di Ricerca in Matematica*.

Trieste, Sptember 2009.

Abstract

The thesis begins recalling a fundamental result in Measure Theory: the Disintegration Theorem in countably generated measure spaces. Given a partition $X = \cup_{\alpha \in A} X_\alpha$ of the probability space (X, Ω, μ), when possible it endows the equivalence classes $(X_\alpha, \Omega \llcorner X_\alpha)$ with probability measures μ_α suitable to reconstruct the original measure, by integrating over the quotient probability space $(A, \mathcal{A}, m)$:

$$\mu = \int_A \mu_\alpha m(d\alpha).$$

We consider two specific affine partitions of $\mathbb{R}^n$: the equivalence classes are respectively the 1D rays of maximal growth of a 1-Lipschitz function ϕ and the projection of the faces of a convex function f. We establish the nontrivial fact that for this two specific locally affine partitions of $\mathbb{R}^n$ the equivalence classes are equivalent as measure spaces to themselves with the Hausdorff measure of the proper dimension, and the same holds for the relative quotient space. More concretely, this result is a regularity property of the graphs of functions whose 'faces' define the partitions, once projected. The remarkable fact is that a priori the directions of the rays and of the faces are just Borel and no Lipschitz regularity is known. Notwithstanding that, we also prove that a Green-Gauss formula for these directions holds on special sets.

The above study is then applied to some standard problem in Optimal Mass Transportation, namely

- if a transport plan is extremal in $\Pi(\mu, \nu)$;
- if a transport plan is the unique measure in $\Pi(\mu, \nu)$ concentrated on a given set A;
- if a transport plan is a solution to the Kantorovich problem;
- if there exists a solution to the Monge problem in $\mathbb{R}^n$, with a strictly convex norm.

We face these problems with a common approach, decomposing the space into suitable invariant regions for the transport and 'localizing' the study in the equivalence classes and in the quotient space by means of the Disintegration Theorem. Explicit procedures are provided in the cases above, which are fulfilled depending on regularity properties of the disintegrations we consider.

As by sides results, we study the Disintegration Theorem w.r.t. family of equivalence relations, the construction of optimal potentials, a natural relation obtained from c-cyclical monotonicity.

The Thesis collects the results obtained in [BC1], [Car2], [CD].

Contents

1 Introduction

Given a space with some structure, a fundamental abstract operation in Mathematics is to define an equivalence relation somehow compatible with the structure and to pass to the quotient. A basic problem is to study the behavior of the structure w.r.t. this operation, for suitably chosen partitions in equivalence classes. This operation can be the basis of reduction arguments, spitting up a problem in simpler ones, or simply a way to sweep out obstacles or unessential informations, focussing on what remains.

This is the key operation all throughout the thesis, in the basic framework of Measure Spaces both as general as countably generated ones and as specific as $\mathbb{R}^n$ with the Lebesgue measure.

A fundamental result in Measure Theory, the Disintegration Theorem, deals with the probability structure of the equivalence classes X_t, $t \in Q$, in a probability space (X, Ω, μ). While the quotient probability space $(Q, \mathcal{Q}, m)$ is easily defined, one needs some regularity in order to have unique probabilities μ_t on X_t satisfying

$$\mu(B \cap h^{-1}(A)) = \int_A \mu_t(B) m(dt) \qquad \text{for all } B \in \Omega,\ A \in \mathcal{Q}.$$

In a first part, we recall these basic results. We study moreover with these tools the disintegration w.r.t. a family of equivalence relation closed under countable intersection, finding that the family contains a sharpest equivalence relation for the measure.

We then consider separately two specific partitions of $\mathbb{R}^n$. The equivalence classes are in one case the lines of maximal growth of a 1-Lipschitz potential, w.r.t. a strictly convex norm. In the second case, they correspond to the faces of a convex function — they are the relative interior of maximal convex sets where the convex function is linear. Existence and uniqueness of a disintegration are provided by measurability properties of the partitions, applying the Disintegration Theorem. We establish the nontrivial fact that for these two particular affine partitions of $\mathbb{R}^n$ the conditional measures on the equivalence classes are equivalent to the Hausdorff measure of the proper dimension, and the same holds for the quotient spaces. This is a regularity property of the graphs of the functions defining the partitions, as any BV regularity of the direction field in general does not hold. Notwithstanding that, we also prove that a Green-Gauss formula for these directions holds on special sets.

In a second part, the study above is applied to some standard problem in Optimal Mass Transportation. As a common strategy we decompose the space into invariant regions for the transport, and we 'localize' the questions by means of a disintegration.

We consider two countably generated probability spaces (X, Ω, μ), (Y, Σ, ν), and a cost function $c : X \times Y \to [0, +\infty]$. By an isomorphism theorem, we are allowed to fix $X = Y = [0, 1]$ and μ, ν Borel probability measures. We assume that $c : [0, 1]^2 \to [0, +\infty]$ is coanalytic. The family of transference plan between μ and ν is defined as the subset of probability measures satisfying the marginal conditions $(P_1)_\sharp\pi = \mu$, $(P_2)_\sharp\pi = \nu$, where $P_1(x, y) = x$, $P_2(x, y) = y$ are the projection on X, Y:

$$\Pi(\mu, \nu) := \Big\{\pi \in \mathcal{P}([0, 1]^2) : (P_1)_\sharp\pi = \mu, (P_2)_\sharp\pi = \nu\Big\}.$$

Namely, we study

- if a transport plan is extremal in $\Pi(\mu, \nu)$;
- if a transport plan is the unique measure in $\Pi(\mu, \nu)$ concentrated on a given set A;
- if a transport plan π is a solution to the Kantorovich problem;
- if there exist potentials relative to an optimal transport plan π;
- if there exists a solution to the Monge problem in $\mathbb{R}^n$, with strictly convex norms.

The main issue is to define the suitable partitions where the question can be answered. Then in order to continue with the investigation we will need a regularity property of the disintegration. The procedure will be finally accomplished when the resulting problem in the quotient space will be well posed.

In the following we provide a more detail description of each topic.
The content of Chapters 4, 6.1 and 13 is from [Car2]. The content of Chapters 5, 6.2 and C is from [CD]. The content of Chapters 2, 3, 7, 8, 9,10, 11, 12, B is from [BC1].

Part I

Part I is devoted to the Disintegration Theorem and to further regularity of specific disintegrations in $\mathbb{R}^n$.

Chapter 2 — · —

All the results can be found in Section 452 of [Fre2]; we refer also to [HJ, Fre1, AFP, Sri].

Given a probability space (X, Ω, μ) and a partition $X = \cup_{t \in Q} X_t$, the quotient is itself a probability space with the push forward σ-algebra and measure defined by

$$S \in q_\sharp\Omega \quad \Longleftrightarrow \quad q^{-1}(S) \in \Omega, \qquad (q_\sharp\mu)(S) := \mu(q^{-1}(S))$$

where q is the quotient map.
The equivalence classes $\{X_t\}_{t\in Q}$ inherit the restricted σ-algebra $\Omega \llcorner_{X_t} = \{S \cap X_t : S \in \Omega\}$. It would be suitable to have probabilities μ_t on $(X_t, \Omega\llcorner_{X_t})$, $t \in Q$, satisfying

$$\mu(B \cap h^{-1}(S)) = \int_S \mu_t(B) m(dt) \qquad \text{for all } B \in \Omega,\ S \in q_\sharp \Omega. \tag{1.1}$$

Such a family $\{\mu_t\}_{t\in Q}$ is called *disintegration* of the probability measure μ *strongly consistent* with the partition $X = \cup_{t\in Q} X_t$, and $\{\mu_t\}_{t\in Q}$ *conditional probabilities*.

The Disintegration Theorem deals with existence and uniqueness of a disintegration.

Under the assumption that μ is determined by the values on a countable subset of Ω, then two such families $\{\mu_t\}_{t\in Q}$, $\{\tilde{\mu}_t\}_{t\in Q}$ coincide for $(q_\sharp\mu)$-a.e. t by usual arguments — the disintegration is thus *unique*.

The *existence* is instead a compatibility condition between the partition and the measure structure — equivalently, it is a regularity property of the quotient space. When μ is determined by the values on countably many sets, a first result is that even the quotient measure is determined by the values on a countable family $\{B_n\}_{n\in\mathbb{N}}$ of saturated sets, where by a *saturated set* we mean a subset B of X such that $B = q^{-1} \circ q(B)$. In measure theory, this is a corollary of Maharam Theorem.
If $\{B_n\}_{n\in\mathbb{N}}$ separate equivalence classes, except for a μ-negligible set in Ω, then one can identify the quotient with a Borel probability space on $[0,1]$ (see Appendix A) and existence is guaranteed.
If not, then one can consider a poorer partition, whose equivalence classes are the union of the ones in the previous partition which are not separated by $\{B_n\}_{n\in\mathbb{N}}$. The quotient space L w.r.t. this second equivalence relation is smaller and there is a natural map $p : Q \to L$. From the previous case one obtains the existence of a unique family of probabilities $\{\mu_t\}_{t\in Q}$ *on* (X, Ω) satisfying (1.1), but they are concentrated on the equivalence classes $\cup_{t\in p^{-1}(\ell)} X_t$, $\ell \in L$, instead of X_t, $t \in Q$. Precisely, $\mu_t \equiv \mu_{p(t)}$ for $(q_\sharp\mu)$-a.e. t. In this case the *disintegration* is called *consistent with the partition*, but it is not *strongly consistent*.
For example, the disintegration of $([0,1], \mathcal{B}, \mathcal{L}^1)$ w.r.t. the equivalence classes $x^\bullet := \{x + k\alpha \mod 1, k \in \mathbb{N}\}$ is strongly consistent only if $\alpha \in \mathbb{Q}$, otherwise one obtains the Vitali set as a quotient space.

One can identify Q with a subset of X, by the Axiom of Choice. Moreover, in the case considered above of (X, Ω, μ) *essentially countably generated*, by Appendix A one can assume that $X = [0,1]$, μ a Borel probability measure. The condition of strong consistency of the disintegration is then equivalent to the possibility of choosing $q(\Gamma) \subset X = [0,1]$ in such a way that the quotient σ-algebra contains Borel sets, for a $\Gamma \in \Omega$ with $\mu(\Gamma) = 1$.

The main result can then be summarized as follows.

Theorem 1.1. *Let* (X, Ω, μ) *be an essentially countably generated probability space and* $X = \cup_{t\in Q} X_t$ *a partition of* X. *Then there exists a unique disintegration consistent with the partition.*

The disintegration is strongly consistent if and only if its quotient space is isomorphic to a Borel probability space on $[0,1]$.

The disintegration results extend to σ-finite spaces instead of probability ones.

Chapter 3 — · —

Motivated by Theorem 1.1, we restrict to (X, Ω, μ) essentially countably generated.
We consider a family $\mathfrak{E}$ of equivalence relations E_e, $e \in \mathcal{E}$, which is closed under countable intersection: for every countable family $\{e_n\}_{n\in\mathbb{N}} \subset \mathcal{E}$

$$\bigcap_n E_{e_n} \in \mathfrak{E}.$$

Then we prove that there exists an equivalence relation $E_{\hat{e}} \in \mathfrak{E}$ defining the finest partition for the family, in a measure theoretic sense.
We denote the disintegration w.r.t. an equivalence relation E_e with

$$\mu = \int \mu_{E_e,\beta} m_{E_e}(d\beta).$$

Theorem 1.2. *There exists $E_{\hat{e}} \in \mathfrak{E}$ such that for all E_e on X, $e \in \mathcal{E}$, the following holds:*

- *Every saturated set w.r.t. the equivalence relation E_e differs from a saturated set w.r.t. $E_{\hat{e}}$ for a μ-negligible set.*
- *There exists a measure preserving projection p from the quotient space w.r.t. $E_{\hat{e}}$ to the quotient space w.r.t. E_e. Denote the disintegration consistent with p as*

$$m_{E_{\hat{e}}} = \int m_{E_{\hat{e}},\alpha} m_{E_e}(d\alpha).$$

- *The disintegration consistent with the equivalence relation E_e is determined by disintegrating μ w.r.t. $E_{\hat{e}}$ and the quotient measure w.r.t. the level set of p of the previous step:*

$$\mu = \int \mu_{E_e,\alpha} m_{E_e}(d\alpha) \quad \textit{with} \quad \mu_{E_e,\alpha} \equiv \int \mu_{E_{\hat{e}},\beta} m_{E_{\hat{e}}\alpha}(d\beta).$$

Chapter 4 — · —

We turn the attention to $\mathbb{R}^n$ with the Lebesgue measure.
Let $\tilde{\|} \cdot \tilde{\|}$ be a possibly asymmetric norm whose unit ball is *strictly* convex and $\phi : \mathbb{R}^n \to \mathbb{R}$ a 1-Lipschitz function w.r.t. this norm, that we call *potential*.
We define as *transport set* the set of rays of maximal growth of ϕ, with or without the endpoints: denoting $\partial_c\phi := \Big\{(x, y) : \ \phi(x) - \phi(y) = \tilde{\|} y - x \tilde{\|}\Big\} \subset \mathbb{R}^n \times \mathbb{R}^n$, then

$$\bar{\mathcal{T}} := \bigcup_{(x,y)\in\partial_c\phi} (\!|x, y|\!) \qquad\qquad \bar{\mathcal{T}}_e := \bigcup_{(x,y)\in\partial_c\phi\setminus\{y=x\}} [\![x, y]\!].$$

Due to the strict triangular inequality, holding by the strict convexity of the unit ball, it is well known that two rays of maximal growth may intersect only at a common starting or final point. In particular, one can partition $\bar{\mathcal{T}}$ in rays with the equivalence relation

$$x \sim y \quad \Longleftrightarrow \quad \phi(x) - \phi(y) = \tilde{\|} y - x \tilde{\|}.$$

Lemma 1.3. *Both $\bar{\mathcal{T}}$ and $\bar{\mathcal{T}}_e$ are σ-compact sets.*

Lemma 1.4. *There exists a countably $(n-1)$-rectifiable set $\mathcal{N}$ such that $\sim$ is an equivalence relation on $\bar{\mathcal{T}}_e \setminus \mathcal{N}$.*

We deal with the disintegration of the Lebesgue measure on $\bar{\mathcal{T}}_e \setminus \mathcal{N}$, or, as well, on $\bar{\mathcal{T}}$, w.r.t. this partition $\{r(y)\}_{y \in \mathcal{S}}$. Existence and uniqueness of a strongly consistent disintegration are easily provided by Theorem 1.1. We establish that the quotient space is equivalent to $\oplus_{k \in \mathbb{N}}(\mathbb{R}^{n-1}, \mathscr{L}^{n-1}(\mathbb{R}^{n-1}), \mathscr{L}^{n-1})$, while the conditional measures are equivalent to $(\mathbb{R}, \mathscr{L}^1(\mathbb{R}), \mathscr{L}^1)$. As a by sides result, we find that $\mathscr{L}^n \llcorner \bar{\mathcal{T}} = \mathscr{L}^n \llcorner \bar{\mathcal{T}}_e$: the set of endpoints of rays is a Borel set of zero Lebesgue measure.

More precisely, the main theorem of the section is summarized as follows.

Theorem 1.5. *The following formula holds: for every integrable function $\varphi \in L(\mathscr{L}^n)$*

$$\int_{\bar{\mathcal{T}}_e} \varphi(x)\, d\mathscr{L}^n(x) = \int_{\mathcal{S}} \left\{ \int_{r(y)} \varphi(z)\gamma(z)\, d\mathcal{H}^1(z) \right\} d\mathcal{H}^{n-1}(y)$$

with $\mathcal{S}$ a countable union of σ-compact subsets of hyperplanes and γ a Borel positive function.

We stress that the statement is nontrivial, as any Lipschitz regularity of the direction field of rays is not known. Indeed, there are σ-compact sets of positive Lebesgue measure which can be partitioned into disjoint segments, with Borel direction field, and such that the conditional measures of the disintegration are deltas (see Example 4.3 taken from [AKP1], improving [Lar1]). The theorem above is thus a regularity property of the partition we are considering.

The proof follows the technique introduced in [Bia], [BG], where the role of the potential ϕ was belonging to the solution u to the variational problem

$$\inf_{\bar{u} + W_0^{1,\infty}(\Omega)} \int_\Omega \big(\chi_D(\nabla u) + g(u)\big)\, dx,$$

with D convex closed with nonempty interior, $g : \mathbb{R} \to \mathbb{R}$ strictly monotone increasing and differentiable, Ω open bounded, $\nabla \bar{u} \in D$ a.e. in Ω.

Idea of the argument

The main steps are the following.

One first observes that, by additivity of the measures, it suffices to determine the disintegration of the Lebesgue measure on the elements of a countable partition of $\bar{\mathcal{T}}$. In particular, one considers sheaf of rays transversal to an hyperplane (Figure 1).

Varying the hyperplanes in a dense set one finds a covering and then constructs the partition. The quotient set $\mathcal{S}$ is identified with the intersection of the rays of these model sets with the relative transversal hyperplane.

One then studies the map which moves points along rays within a model set. More specifically, fix two transversal parallel hyperplanes, restrict the model set to those rays which intersect both and consider the bijective function between the two sections coupling points belonging to a same ray (Figure 3). Then one proves that the push forward of the Hausdorff $(n-1)$-dimensional measure on one section is absolutely continuous w.r.t. the Hausdorff $(n-1)$-dimensional measure on the other section, and viceversa.

This is proved by approximation, by the upper semicontinuity of the Hausdorff $(n-1)$-dimensional measure. Indeed, being 1-Lipschitz, the potential satisfies the Hopf-Lax formula

$$\phi(x) = \max_{z\in A} \left\{ \phi(z) - \tilde{\|} x - z \tilde{\|} \right\}$$

on the model set between the two hyperplanes, where A is the section corresponding to the hyperplane with the higher values of ϕ. One then approximates ϕ with the locally uniformly converging sequence

$$\phi_I(x) = \max \left\{ \phi(a_i) - \tilde{\|} x - a_i \tilde{\|} : \ i = 1, \dots, I \right\}, \qquad \text{with } \{a_i\}_{i\in\mathbb{N}} \text{ dense in } A,$$

and in turn the rays of ϕ joining the two sections A, B with the rays of ϕ_I passing through B; the approximating section obtained by the intersection with any other third parallel hyperplane between the two converges Hausdorff to the approximated one (Figure 2). By the simple expression of the approximating direction field, pointing towards finitely many points, one can easily prove area estimates which pass to the limit.

Let $\bar{\mathcal{Z}}$ be a sheaf of rays transversal to $\{x \cdot e = 0\}$, $e \in \mathbb{R}^n$: by Fubini Theorem

$$\int_{\bar{\mathcal{Z}}} \varphi(x)\, dx = \int_{-\infty}^{+\infty} \int_{\bar{\mathcal{Z}}\cap\{x\cdot e = t\}} \varphi(z)\, dt \times d\mathcal{H}^{n-1}(z).$$

Denoting as $(\sigma^t)^{-1} : \bar{\mathcal{Z}} \cap \{x \cdot e = t\} \to \bar{\mathcal{Z}} \cap \{x \cdot e = 0\}$ the map coupling points on a same ray, the absolutely continuous estimate yields the existence of a function $\tilde{\alpha}(t, y)$ s.t. $(\sigma^t)^{-1}_{\sharp} \mathcal{H}^{n-1} \llcorner_{\bar{\mathcal{Z}}\cap\{x\cdot e=t\}} = \tilde{\alpha}(t, y) \mathcal{H}^{n-1} \llcorner_{\bar{\mathcal{Z}}\cap\{x\cdot e=0\}}$. Then

$$\int_{\bar{\mathcal{Z}}\cap\{x\cdot e=t\}} \varphi(z)\, dt \times d\mathcal{H}^{n-1}(z) = \int_{\bar{\mathcal{Z}}\cap\{x\cdot e=0\}} \varphi(\sigma^t y) \tilde{\alpha}(t, y)\, dt \times d\mathcal{H}^{n-1}(y).$$

Substituting in the former, one applies again Fubini-Tonelli Theorem and derives the disintegration.

In particular, denoting $Z_0 := \bar{\mathcal{Z}} \cap \{x \cdot e = 0\}$, we construct a change of variable $\psi : \bar{\mathcal{Z}} \to \mathbb{R} \times Z_0$, one to one to its image, which brings Lebesgue measurable function on $\bar{\mathcal{Z}}$ to $\mathcal{H}^1 \otimes (\mathcal{H}^{n-1} \llcorner_{Z_0})$-measurable functions, since 0-measure sets are mapped to 0-measure sets and viceversa. In the language of measure theory of the first chapter, it induces an isomorphism between the relative measure algebras.

The negligibility of $\bar{\mathcal{T}}_e \setminus \bar{\mathcal{T}}$ is proved by a density argument involving the push forward estimate.

Chapter 5 — · —

We disintegrate the Lebesgue measure on the graph of a convex function w.r.t. the partition given by its faces. As the graph of a convex function naturally supports the Lebesgue measure, its faces, being convex, have a well defined linear dimension, and then they naturally support a proper dimensional Hausdorff measure. Our main result is that the conditional measures induced by the disintegration are equivalent to the Hausdorff measure on the faces on which they are concentrated.

Theorem 1.6. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let $\mathscr{H}^n$ be the Hausdorff measure on its graph. Define a face of f as the convex set obtained by the intersection of its graph with a supporting hyperplane and consider the partition of the graph of f into the relative interiors of the faces $\{F_\alpha\}_{\alpha\in A}$.*
Then, the Lebesgue measure on the graph of the convex function admits a unique disintegration

$$\mathscr{H}^n = \int_A \lambda_\alpha \, dm(\alpha)$$

w.r.t. this partition and the conditional measure λ_α which is concentrated on the relative interior of the face F_α is equivalent to $\mathscr{H}^k \llcorner F_\alpha$, where k is the linear dimension of F_α.

This apparently intuitive fact does not always hold, as even for a partition given by a Borel measurable collection of segments in $\mathbb{R}^3$ (1-dimensional convex sets). If any Lipschitz regularity of the directions of these segments is not known, it may happen that the conditional measures induced by the disintegration of the Lebesgue measure are Dirac deltas (as for Chaper 4, see the Couterexamples 4.3 taken from [AKP2], improving [Lar1]). Also in our case, the directions of the faces of a convex function are just Borel measurable (Example 3.9 in [Bia]). Therefore, our result, other than answering a quite natural question, enriches the regularity properties of the faces of a convex function, which have been intensively studied for example in [ELR], [KM], [Lar1], [LR], [AKP2], [PZ]. As a byproduct, we recover the Lebesgue negligibility of the set of relative boundary points of the faces, which was first obtained in [Lar2].

Idea of the argument

The proof of our theorem does not rely on Area or Coarea Formula, which in several situations allow to obtain in one step both the existence and the absolute continuity of the disintegration (in applications to optimal mass transport problem, see for example [TW], [FM1], [AP]). The basis of the technique has been described for the previous chapter, applied to 1D rays of a 1-Lipschitz function. For notational convenience, we work as there with the projections of the faces on $\mathbb{R}^n$, and we neglect the set where the convex function is not differentiable.

Just to give an idea, focus as there on a collection of 1-dimensional faces $\mathscr{C}$ which are transversal to a fixed hyperplane $H_0 = \{x \in \mathbb{R}^n : x \cdot e = 0\}$ and such that the projection of each face on the line spanned by the fixed vector e contains the interval $[h^-, h^+]$, with $h^- < 0 < h^+$ (Figure 3). The notations are analogous to the previous ones.

The core of the proof, as before, is to show the absolute continuity

$$\mathscr{H}^{n-1} \llcorner (\mathscr{C} \cap H_t) \ll \sigma^t_\sharp(\mathscr{H}^{n-1} \llcorner (\mathscr{C} \cap H_0)).$$

We prove in particular the following quantitative estimate: for all $0 \leqslant t \leqslant h^+$ and $S \subset \mathscr{C} \cap H_0$

$$\mathscr{H}^{n-1}(\sigma^t(S)) \leqslant \left(\frac{t-h^-}{-h^-}\right)^{n-1} \mathscr{H}^{n-1}(S). \tag{1.2}$$

This fundamental estimate is proved again approximating the 1-dimensional faces with a sequence of finitely many cones with vertex in $\mathscr{C} \cap H_{h^-}$ and basis in $\mathscr{C} \cap H_t$. At this step of the technique, the construction of such approximating sequence heavily depends on the nature of the partition one has to deal with. In this case, our main task is to find the suitable cones relying on the fact that we are approximating the faces of a convex function. The strategy, roughly, is to approximate f with the function f_I having as epigraph the convex envelope of $f_{\llcorner H_t} \cup \{\tilde{y}_i\}_{i=1,\dots,I}$, where $\tilde{y}_i$, $i \in \mathbb{N}$, constitute a dense sequence in $(\mathbb{I}, f)(\mathscr{C} \cap H_0)$ (Figure 6). In turn, this allows to approximate the rays of f, that we are supposing to be 1D, with the ones of f_I — which point towards finitely many points.

We obtain the disintegration of the Lebesgue measure on the k-dimensional faces, with $k > 1$, from a reduction argument to this case. Indeed, consider a model set $\mathscr{W}^k$ made of k-dimensional faces transversal to a fixed $(n-k)$-dimensional hyperplane W^{n-k} (Figure 4). In particular, the faces we consider have positive projection on the k-dimensional subspace V^k orthogonal to W^{n-k}. Let $V^k = \langle e_1, \dots, e_k \rangle$. Then by Fubini Theorem, denoting the plane $\{q + \langle e_1 \rangle + W^{n-k}\}$ as $\mathfrak{W}(q)$,

$$\int_{\mathscr{W}^k} \varphi(x)\, d\mathscr{L}^n(x) = \int_{\langle e_2,\dots,e_k\rangle} \int_{\mathscr{W}^k \cap \mathfrak{W}(q)} \varphi(z)\, d\mathcal{H}^{n-k+1}(z) \otimes \mathcal{H}^{k-1}(q),$$

and $\mathscr{W}^k \cap \mathfrak{W}(q)$ is a model set made of 1D faces of the convex function $f_{\llcorner \mathfrak{W}(q)}$.

Chapter 6 — · —

In Chapters 4, 5 we studied a regularity property for two affine partitions in $\mathbb{R}^n$, classifying the conditional and quotient measures: they were equivalent to the Hausdorff measure of the proper dimension. In both cases, after a countable partition in model sets, the result was based on the absolute continuity push forward estimate (1.2). The fundamental estimate (1.2) implies moreover a Lipschitz continuity and BV regularity of $\alpha^t(z)$ w.r.t t: this yields a further improvement of the regularity of the partition that now we are going to describe.

In the case of a 1-Lipschitz function the density of the conditional measures w.r.t. the 1D Hausdorff 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 $\bar{\mathfrak{T}}$, is the coefficient for an ODE for the above density.

Consider for both the cases a vector field v which at each point x is parallel to the face (ray) through that point x.

If we restrict the vector field to an open Lipschitz set Ω which does not contain points in the relative boundaries of the faces, then we prove that its distributional divergence is the sum of two terms: an absolutely continuous measure, and a $(n-1)$-rectifiable measure representing the flux of v through the boundary of Ω. The density $(\operatorname{div} v)_{\mathrm{a.c.}}$ of the absolutely continuous part is related to the density of the conditional measures defined by the disintegration above.
In the case of the set $\mathscr{C}$ previously considered, if the vector field is such that $v \cdot e = 1$, the expression of the density of the absolutely continuous part of the divergence is

$$\partial_t \alpha^t = (\operatorname{div} v)_{\mathrm{a.c.}} \alpha^t.$$

No piecewise BV regularity of the vector field v of faces and rays directions hold in general (Example 3.9 in [Bia]). Therefore, it is a remarkable fact that a divergence formula holds.

The divergence of the whole vector field v is the limit, in the sense of distributions, of the sequence of measures which are the divergence of truncations of v on the elements $\{\bar{\mathcal{K}}_\ell\}_{\ell \in \mathbb{N}}$ of a suitable partition of $\mathbb{R}^n$. However, in general, it fails to be a measure (Examples 13.7, 13.8).

This additional regularity is proved first for the rays of a 1-Lipschitz function, where v is basically the direction field. The argument follows the one in [BG], and it is based on the cone approximation of rays described above: we consider the divergence of the approximating direction field and by uniform estimates on their total variation we pass to the limit.

It is then proved for the faces of a convex function. We give an alternative prove based on the Disintegration Theorem 1.6 and on regularity properties of the density function proved in Chapter 5.

In the last part, we change point of view: instead of looking at vector fields constrained to the faces of the convex function, we describe the faces as an $(n+1)$-uple of currents, the k-th one corresponding to the family of k-dimensional faces, for $k = 0, \dots, n$. The regularity results obtained for the vector fields can be rewritten as regularity results for these currents. More precisely, we prove that they are locally flat chains. When truncated on a set Ω as above, they are locally normal, and we give an explicit formula for their border; the $(n+1)$-uple of currents is the limit, in the flat norm, of the truncations on the elements of a partition.

An application of this kind of further regularity is presented in Section 8 of [BG]. Given a vector field v constrained to live on the faces of f, the divergence formula we obtain allows to reduce the transport equation

$$\operatorname{div} \rho v = g$$

to a PDE on the faces of the convex function. We do not pursue this issue.

Part II

In Part II we turn the attention to basic problems in Optimal Mass Transportation, that we face as applications of the Disintegration Theorem.

Chapter 7 — · —

In Chapter 7 we explain the approach we follow for studying extremality, uniqueness and optimality of a transference plan between two countably generated probability spaces (X, Ω, μ), (Y, Σ, ν). By Appendix A, we directly assume that they are Borel probability spaces on $[0, 1]$.

The issue is in particular the following: we have necessary conditions for extremality, uniqueness and optimality, but they are not sufficient in general. How one can test if a given transport plan $\pi \in \Pi(\mu, \nu)$ satisfying the necessary condition is actually extremal, unique or optimal?

The idea of the general scheme is described by the following theorem. The set $\Pi^f(\mu, \nu)$ denotes the family of transport plans which have to be compared to π.

Theorem 1.7. *Assume that there are partitions $\{X_\alpha\}_{\alpha\in[0,1]}$, $\{Y_\beta\}_{\beta\in[0,1]}$ of X, Y such that*

1. *for all $\pi' \in \Pi^f(\mu, \nu)$ it holds $\pi'(\cup_\alpha X_\alpha \times Y_\alpha) = 1$,*
2. *the disintegration $\pi = \int \pi_\alpha m(d\alpha)$ of π w.r.t. the partition $\{X_\alpha \times Y_\alpha\}_{\alpha\in[0,1]}$ is strongly consistent,*
3. *in each equivalence class $X_\alpha \times Y_\alpha$ the measure π_α is extremal/unique/optimal in $\Pi(\mu_\alpha, \nu_\alpha)$, where*

$$\mu_\alpha := (P_1)_\sharp \pi_\alpha, \quad \nu_\alpha := (P_2)_\sharp \pi_\alpha.$$

Then π is extremal/unique/optimal.

Since the necessary conditions we consider, specified later, are pointwise properties to be verified by a suitable carriage Γ of π, we propose in the following explicit procedures.

We define crosswise equivalence relations in Γ, meaning that the equivalence classes are

$$\Gamma_\alpha = \Gamma \bigcap X_\alpha \times [0, 1] = \Gamma \bigcap [0, 1] \times Y_\alpha = \Gamma \bigcap X_\alpha \times Y_\alpha$$

for partitions $\{X_\alpha\}_{\alpha\in[0,1]}$, $\{Y_\alpha\}_{\alpha\in[0,1]}$ of X, Y. In particular, $\Gamma \subset \cup_\alpha X_\alpha \times Y_\alpha$. They are chosen in such a way that each probability measure concentrated on Γ_α is always extremal, unique or optimal as a transport map between its marginals. This is a consequence of the fact that the necessary conditions become sufficient if joint with other assumptions, and we specify equivalence classes satisfying for free the necessary conditions and these further requirements.

In order to apply Theorem 1.7, one needs then the following assumptions:

- the strong consistency of the disintegration, treated in Chapter 2;
- $\hat{\pi}(\cup_{\alpha\in[0,1]} X_\alpha \times Y_\alpha) = 1$ for every other transport plan $\hat{\pi} \in \Pi^f(\mu, \nu)$.

The last condition is for free in the study of extremality, as if π is a linear combination of π_1, π_2 then $\pi_1, \pi_2 \ll \pi$. In the other two cases, it translates to a transport problem in the quotient space.
Indeed, on one hand under the first assumption the crosswise structure implies, together with the marginal conditions, that the quotient measure of π w.r.t. the product partition $\{X_\alpha \times Y_\beta\}_{\alpha,\beta\in[0,1]}$ is of the form $(\mathbb{I}, \mathbb{I})_\sharp m$. On the other hand, the marginal conditions force that the quotient measure of any other transport plan $\hat{\pi}$ w.r.t. the product partition is a transport plan from m to itself: as $\hat{\pi}$ is concentrated on Γ if and only if its quotient measure is $(\mathbb{I}, \mathbb{I})_\sharp m$, the second assumption is then a problem of uniqueness in the quotient space.

In Section 7.1 we give a further interpretation of the assumptions in Theorem 1.7.

In Section 7.2 we formalize the setting and the strategy, that we prove without specifying the actual necessary conditions and relative partitions under consideration. We collect 4 steps encoding the method which will be used to obtain the results in the next chapters.

Chapter 8 — · —

We address the problem of extremality. The results obtained with our approach are already known in the literature: this part can be seen as an exercise to understand how the procedure works. The difficulties of both approaches are the same: in fact the existence of a Borel rooting set up to negligible sets is equivalent to the strong consistency of the disintegration, and to the existence of a Borel limb numbering system ([HW]) on which π is concentrated — the equivalence classes are Borel limbs.

The necessary condition we consider is the acyclicity (Th. 3 of [HW]). We say that $\Gamma \subset [0,1]^2$ is *acyclic* if for all finite sequences $(x_i, y_i) \in \Gamma$, $i = 1, \dots, n$, with $x_i \neq x_{i+1 \mod n}$ and $y_i \neq y_{i+1 \mod n}$ it holds

$$\Big\{(x_{i+1}, y_i), i = 1, \dots, n, x_{n+1} = x_1\Big\} \not\subset \Gamma.$$

A measure is *acyclic* if it is concentrated on an acyclic set Γ.

Given such an acyclic carriage Γ, one has the following crosswise partition.

Definition 1.8 (Axial equivalence relation)**.** We define $(x, y)E(x', y')$ if there are $(x_i, y_i) \in \Gamma$, $0 \leqslant i \leqslant I$ finite, such that

$$(x, y) = (x_0, y_0), (x', y') = (x_I, y_I) \quad \text{and} \quad (x_{i+1} - x_i)(y_{i+1} - y_i) = 0.$$

Theorem 1.7 leads to the following statement.

Corollary 1.9 (Extremality (Theorem 8.8))**.** *Let π concentrated on a σ-compact acyclic set Γ.*

If we partition the set Γ into axial equivalence classes, then π is extremal in $\Pi(\mu, \nu)$ if the disintegration is strongly consistent.

We moreover recall the following result ([Dou, Lin]), which we can prove by means of duality.

Proposition 1.10. *The transference plan* $\pi \in \Pi(\mu, \nu)$ *is extremal if and only if* $L^1(\mu) + L^1(\nu)$ *is dense in* $L^1(\pi)$.

CHAPTER 9 — · —

We consider the problem of verifying if an analytic set A can carry more than one transference plan — if not, we say that A is a *set of uniqueness*.

The necessary condition we consider is the A-acyclicity. A set $\Gamma \subset A$ is A-*acyclic* if for all finite sequences $(x_i, y_i) \in \Gamma$, $i = 1, \dots, n$, with $x_i \neq x_{i+1 \text{ mod } n}$ and $y_i \neq y_{i+1 \text{ mod } n}$ it holds

$$\Big\{(x_{i+1}, y_i), i = 1, \dots, n, x_{n+1} = x_1\Big\} \not\subset A.$$

A measure is A-*acyclic* if it is concentrated on an A-acyclic set Γ.

As an A-acyclic set is in particular acyclic, then we consider once more the axial equivalence relation on Γ. In this case, not only the disintegration should be strongly consistent, but we must verify also Condition (1) of Theorem 7.1. Condition (2) of the following corollary implies this fact.

Corollary 1.11 (Uniqueness (Page 119))**.** *Let* π *concentrated on a* σ*-compact* A*-acyclic set* Γ*.*

If we partition the set Γ *into axial equivalence classes, then* π *is the unique measure in* $\Pi(\mu, \nu)$ *concentrated on* A *if*

1. *the disintegration is strongly consistent,*
2. A' *is a subset of* $\{\alpha \leqslant \beta\}$*, up to measure preserving maps.*

Essentially, we are just showing that in the quotient space the uniqueness problem can be reduced to the trivial one

$$\{\alpha = \beta\} \subset A' \subset \{\alpha \leqslant \beta\}, \quad \pi \in \Pi(m, m).$$

We use the following easy observation:

Lemma 1.12. *If* $\pi = (\mathbb{I}, f)_\sharp \mu$ *and* $A = \mathrm{epi}(f)$*, then* A *is a set of uniqueness of* $\Pi(\mu, \nu)$*, where* $\nu = f_\sharp(\mu)$*.*

We remark that the diagonal is A'-acyclic, and thus A' is the graph of a partial order relation. In the case m is atomic, then A' can be completed to a Borel linear order on $[0, 1]$, and A is a set of uniqueness. We are studying if also in the general case A is a set of uniqueness if and only if A' can be completed to a Borel linear order. As a partial order can be always completed to a linear order by the Axiom of Choice, then this would be again a measurability assumption.

CHAPTER 10 — · —

We consider the optimality of a transference plan w.r.t. a coanalytic cost $c : [0,1]^2 \to [0,+\infty]$: we ask whether

$$\int_{[0,1]^2} c\pi = \min_{\hat{\pi} \in \Pi(\mu,\nu)} \int_{[0,1]^2} c\hat{\pi};$$

The necessary condition we consider is the well-known c-monotonicity. A subset Γ of $[0,1]^2$ is c-*cyclically monotone* when for all I, $i = 1,\dots,I$, $(x_i,y_i) \in \Gamma$, $x_{I+1} := x_1$ we have

$$\sum_{i=1}^{I} \big[c(x_{i+1},y_i) - c(x_i,y_i)\big] \geqslant 0.$$

A transference plan $\pi \in \Pi(\mu,\nu)$ is c-*cyclically monotone* if there exists a π-measurable c-cyclically monotone set Γ such that $\pi(\Gamma) = 1$.

The easiest equivalence relation is the cycle equivalence relation, introduced also in [BGMS].

Definition 1.13 (Closed cycles equivalence relation)**.** We say that $(x,y)\bar{E}(x',y')$ or (x,y) *is equivalent to* (x',y') *by closed cycles* if there is a closed cycle with finite cost passing through them: there are $(x_i,y_i) \in \Gamma$ such that $(x_0,y_0) = (x,y)$ and $(x_j,y_j) = (x',y')$ for some $j \in \{0,\dots,I\}$ such that

$$\sum_{i=1}^{I} c(x_i,y_i) + c(x_{i+1},y_i) < +\infty, \qquad x_{I+1} := x_0.$$

The optimality within each class is immediate from the fact that there exists a couple of $\mathcal{A}$-optimal potentials ϕ,ψ, and after the discussion of the above two problems the statement of the following Corollary should be clear.

Corollary 1.14 (Optimality (Theorem 10.6))**.** *Let* π *concentrated on a* σ*-compact* c*-cyclically monotone set* Γ *and partition* Γ *w.r.t. the cycle equivalence relation (Definition 10.1).*

Then, π c*-cyclically monotone is optimal if*

1. *the disintegration is strongly consistent,*
2. *the image set* $A' := (h_X \otimes h_Y)_\sharp\{c < +\infty\}$ *is a set of uniqueness.*

If one chooses the existence of optimal potentials as sufficient criterion, or even more general criteria, it is in general possible to construct other equivalence relations, such that in each class the conditional probabilities π_α are optimal. This is done extending the construction in Section 10.1 with a different equivalence realtion.

The above result generalizes the previous known cases:

1. if μ or ν are atomic ([Pra]): clearly m must be atomic;

2. if $c(x,y) \leqslant a(x) + b(y)$ with $a \in L^1(\mu)$, $b \in L^1(\nu)$ ([RR]): m is a single δ;
3. if $c : [0,1]^2 \to \mathbb{R}$ is real valued and satisfies the following assumption ([AP])
$$\nu\left(\left\{y : \int c(x,y)\mu(dx) < +\infty\right\}\right) > 0, \quad \mu\left(\left\{x : \int c(x,y)\nu(dy) < +\infty\right\}\right) > 0 :$$
in this case m is a single δ;
4. If $\{c < +\infty\}$ is an open set O minus a $\mu \otimes \nu$ negligible set N ([BGMS]): in this case every point in $\{c < +\infty\}$ has a squared neighborhood satisfying condition (10.5) below.

In each case the equivalence classes are countably many Borel sets, so that the disintegration is strongly consistent and the acyclic set A' is a set of uniqueness (Lemma 9.9).

Section 10.2 is more set theoretical: its aim is just to show that there are other possible decompositions for which our procedure can be applied, and in particular situations where a careful analysis may give the validity of Theorem 7.1 for this new decomposition, but not of the above Corollary for the cycle decomposition. The main result is that under PD and CH we can construct a different equivalence relation satisfying Condition (3) of Theorem 7.1 and the crosswise structure w.r.t. Γ.

Chapter 11 — · —

We give several examples: for historical reasons, we restrict to examples concerning the optimality of π, but trivial variations can be done in order to adapt to the other two problems. We split the section into 2 parts. In Section 11.1 we study how the choice of Γ can affect our construction: it turns out that in pathological cases a wrong choice of Γ may lead to situations for which either the disintegration is not strongly consistent or in the quotient space there is no uniqueness. This may happen both for optimal or not optimal transference plans. In Section 11.2 instead we consider if one can obtain conditions on the problem in the quotient space less strict than the uniqueness condition: the examples show that this is not the case in general.

When the assumption is not satisfied, then one can modify the cost in order to have the same quotient measure but both c-cyclically monotone optimal and c-cyclically non optimal transport plans (Example 11.5, Proposition 11.9).

Chapter 12 — · —

We address the natural question: if we have optimal potentials in each set $X_\alpha \times Y_\alpha$, is it possible to construct an optimal couple (ϕ, ψ) in $\cup_\alpha X_\alpha \times Y_\alpha$? We show that under the assumption of strong consistency this is the case. The main tool is Von Neumann's Selection Theorem, and the key point is to show that the set

$$\Big\{(\alpha, \phi_\alpha, \psi_\alpha) : \phi_\alpha, \psi_\alpha \text{ optimal couple in } X_\alpha \times Y_\alpha\Big\}$$

is analytic in a suitable Polish space. The Polish structure on the family of optimal couples is obtained identifying each μ-measurable function ϕ with the sequence of measures $\{(\phi \vee (-M)) \wedge M)\mu\}_{M\in\mathbb{N}}$, which is shown to be a Borel subset of $\mathcal{M}^{\mathbb{N}}$.

Chapter 13

This chapter 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

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

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 raised 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]), already considered in the previous chapters. 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 $\pi = (\mathbb{I},\tau)_\sharp\mu$, 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], [AKP2].

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 ([AKP2]), 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 13.7. 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], providing the first complete proof of existence for the Euclidean norm. 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 [AKP2] 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, in convex bounded domains. The case of a general norm is treated in [CP2], and is also being considered separately, with different approaches, by myself and Daneri.

In [BC2] the disintegration approach is applied in metric spaces, with non branching geodesics.

We work under the assumption of strict convexity of the unit ball. This 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. We apply the Disintegration Theorem 1.5 in order to conclude Sudakov proof in the case the unit ball is strictly convex.

We will prove the following statement.

Theorem 1.15. *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 $\tilde{\|} \cdot \tilde{\|}$ 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 $\int \tilde{\|} y - x \tilde{\|} \, d\pi(x,y)$. Then:

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

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

Claim 2. *There exists a unique map τ monotone on each ray $\mathbf{r}_z$ solving the Monge-Kantorovich problem*

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

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.*

We also give an expression of the transport density in terms of the conditional probabilities of μ, ν w.r.t. the ray equivalence relation (see (13.5)).
In Section 13.1 we give some counterexamples.
It is shown that the transport set $\bar{\mathcal{T}}$ in general is just a σ-compact set. One can see how the divergence of the vector field of ray directions, defined as zero out of $\bar{\mathcal{T}}$, can fail to be a measure.

Appendix

The Thesis closes with a table of notations, a list of figures and, of course, the bibliography. We collect moreover the following auxiliary chapters.

Chapter A — · —

We briefly recall the isomorphism of measure algebras between a countably generated Probability space and a Borel probability space on $[0,1]$.

Chapter B — · —

We formalize the concepts of cyclic perturbations and acyclic perturbations. After recalling the properties of projective sets in Polish spaces in Section B.1 and the duality results of [Kel] (Section B.2), we show how to define the n-cyclic part of a signed measure λ with 0 marginals: this is the largest measure $\lambda_n \ll \lambda$ which can be written as $\lambda_n = \lambda_n^+ - \lambda_n^-$ with

$$\lambda_n^+ = \frac{1}{n} \int_{C_n} \sum_{i=1}^{n} \delta_{P_{(2i-1,2i)} w} m(dw) \qquad \lambda_n^- = \frac{1}{n} \int_{C_n} \sum_{i=1}^{n} \delta_{P_{(2i+1,2i \mod 2n)} w} m(dw)$$

where $C_n \subset [0,1]^{2n}$ is the set of n-closed cycles and $m \in \mathcal{M}^+(C_n)$. This approach leads to the definition of cyclic perturbations λ: these are the signed measures with 0 marginals which can be written as sum (without cancellation) of cyclic measures. The acyclic measures are those measures for which there are not n-cyclic measures $\lambda_n \ll \lambda$ for all $n \geqslant 2$: in particular they are concentrated on an acyclic set. This approach leads naturally to the well known results on the properties of sets on which extremal/unique/optimal measures are concentrated: in fact, in all cases we ask that there are not cyclic perturbations which either are concentrated on the carriage set Γ, or on the set of uniqueness A, or diminish the cost of the measure π. One then deduces the well known criteria that Γ is acyclic, Γ is A-acyclic and Γ is c-cyclically monotone.

Chapter C — · —

We give very essential recalls on tensors and currents, in view of the applications to Chapter 6. We refer mainly to Chapter 4 of [Mor] and Sections 1.5.1, 4.1 of [Fed].

Part I

The Disintegration Theorem

2 The Disintegration Theorem

In this section we prove the Disintegration Theorem for measures in countably generated σ-algebras, with some applications. The results of this sections can be deduced from Section 452 of [Fre2]; for completeness we give here self-contained proofs.

We first recall the next definition. Let (X, Ω, ν) a generic measure space.

Definition 2.1. The σ-algebra $\mathcal{A}$ is *essentially countably generated* if there is a countable family of sets $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, such that for all $A \in \mathcal{A}$ there exists $\hat{A} \in \mathfrak{A}$, where $\mathfrak{A}$ is the σ-algebra generated by A_n, $n \in \mathbb{N}$, which satisfies $m(A \bigtriangleup \hat{A}) = 0$.

We consider now the following objects:

1. (X, Ω, μ) a countably generated probability space;
2. $X = \cup_{\alpha \in \mathtt{A}} X_\alpha$ a partition of X;
3. $\mathtt{A} = X/\sim$ the quotient space, where $x_1 \sim x_2$ if and only if there exists α such that $x_1, x_2 \in X_\alpha$;
4. $h : X \to \mathtt{A}$ the quotient map $h(x) = x^\bullet = \{\alpha : x \in X_\alpha\}$.

We can give to $\mathtt{A}$ the structure of probability space as follows:

1. define the σ-algebra $\mathcal{A} = h_\sharp(\Omega)$ on $\mathtt{A}$ as
$$S \in \mathcal{A} \quad \Longleftrightarrow \quad \cup_{\alpha \in S}\{x : h(x) = \alpha\} = h^{-1}(S) \in \Omega;$$
2. define the probability measure $m = h_\sharp \mu$.

We can rephrase (1) by saying that $\mathcal{A}$ is the largest σ-algebra such that $h : X \to \mathtt{A}$ is measurable: it can be considered as the subalgebra of Ω made of all saturated measurable sets.

The first result of this section is the structure of $\mathcal{A}$ as a σ-algebra.

Proposition 2.2. *The σ-algebra $\mathcal{A}$ is essentially countably generated.*

Notice that we cannot say that the σ-algebra $\mathcal{A}$ is countably generated: for example, $x \in [0, 1]$ and $x^\bullet = \{x + \mathbb{Q}\} \cap [0, 1]$. We are stating that the measure algebra $\mathcal{A}/\mathcal{N}_m$, where $\mathcal{N}_m$ is the σ-ideal of m-negligible sets, is countably generated.

This proposition is a consequence of Maharam Theorem, a deep result in measure theory, and can be found in [Fre1], Proposition 332T(b). We give a direct proof of this proposition. The fundamental observation is the following lemma.

Lemma 2.3. *Let* f_n *be a countable sequence of measurable functions on* $\mathtt{A}$*. Then there is a countably generated* σ*-subalgebra* $\mathfrak{A}$ *of* A *such that each* f_n *is measurable.*

Proof. The proof is elementary, since this σ-algebra is generated by the countable family of sets

$$\left\{f_n^{-1}(q_m, +\infty), q_m \in \mathbb{Q}, m \in \mathbb{N}\right\}.$$

This is actually the smallest σ-algebra such that all f_n are measurable. $\square$

Proof of Proposition 2.2. The proof will be given in 3 steps.

Step 1. Define the map $\Omega \ni B \to f_B \in L^\infty(m)$ by

$$h_\sharp \mu \llcorner_B = \int f_B m.$$

The map is well defined by Radon-Nykodým theorem, and $0 \leqslant f_B \leqslant 1$ m-a.e..

Given an increasing sequence of $B_i \in \Omega$, then

$$\int_A f_{\cup_i B_i} m = \mu(h^{-1}(A) \cap \cup_i B_i) = \lim_i \mu(h^{-1}(A) \cap B_i) = \lim_i \int_A f_{B_i} m = \int_A \lim_i f_{B_i} m,$$

where we have used twice the Monotone Convergence Theorem and the fact that f_{B_i} is increasing m-a.e.. Hence $f_{\cup_i B_i} = \lim_i f_{B_i}$. By repeating the same argument and using the fact that m is a probability measure, the same formula holds for decreasing sequences of sets, and for disjoint sets one obtains in the same way $f_{\cup_i B_i} = \sum_i f_{B_i}$.

Step 2. Let $\mathcal{B} = \{B_n, n \in \mathbb{N}\}$, be a countable family of sets σ-generating Ω: without any loss of generality, we can assume that $\mathcal{B}$ is closed under complementation and

$$B_m \cap B_n \in \{\emptyset, B_n, B_m\}.$$

In particular, $\mathcal{B}$ is closed under finite intersection: in other words, we are considering the Boolean algebra generated by a σ-family, which is countable.

We now recall that if $B \in \Omega$, then there exists a sequence sequence of sets $B_{nm} \in \mathcal{B}$ such that:

1. B_{nm}, $n \in \mathbb{N}$, is disjoint for m fixed: in fact, if B_{n_i} is a sequence in $\mathcal{B}$, one considers the sequence defined by $\tilde{B}_{n_i} = B_{n_i} \setminus \cup_{j<i} B_{n_j}$, which is in $\mathcal{B}$ because of the closures w.r.t. complementation and finite intersection and satisfies $\cup_i \tilde{B}_{n_i} = \cup_i B_{n_i}$;

2. $\cup_n B_{nm}$ is decreasing w.r.t. m and $E \subset \cup_n B_{nm}$ for all $m \in \mathbb{N}$;

3. $\mu(\cap_m \cup_n B_{nm} \setminus E) = 0$.

The last two properties follow from the elementary fact that the outer measure

$$\theta(B) = \inf\left\{\sum_n \mu(B_n), A \subset \cup_n B_n, B_n \in \mathcal{B}\right\}$$

coincides with μ on the σ-algebra generated by $\mathcal{B}$, because $\mathcal{B}$ is a Boolean algebra and $\theta = \mu$ on $\mathcal{B}$ implies $\theta = \mu$ on the σ-algebra generated by $\mathcal{B}$.

We conclude that

$$\begin{aligned}\int_A f_B m = \mu(h^{-1}(A) \cap B) &= \lim_m \sum_n \mu(h^{-1}(A) \cap B_{nm}) \\ &= \lim_m \sum_n \int_A f_{B_{nm}} m = \int_A \lim_m \sum_n f_{B_{nm}} m.\end{aligned}$$

and then $f_B = \lim_m \sum_n f_{B_{nm}}$ m-a.e..

Step 3. Let $\mathfrak{A}$ be a countably generated σ-algebra such that the functions f_{B_n}, $B_n \in \mathcal{B}$, are measurable; it is provided by Lemma 2.3.

Applying the last equality to the set $B = h^{-1}(A)$ with $A \in \mathcal{A}$, we obtain that there exists a function f in $L^\infty(m)$, measurable w.r.t. the σ-algebra $\mathfrak{A}$ such that $\chi_A = f$ m-a.e., and this concludes the proof, because up to negligible set f is the characteristic function of a measurable set in $\mathfrak{A}$. $\square$

Remark 2.4. We observe there that the result still holds if Ω is the μ-completion of a countably generated σ algebra: this is easily implied by Step 2 of the previous proof.

More generally, the same proof shows that every σ-algebra $\mathcal{A} \subset \Omega$ is essentially countably generated.

In general, the atoms of $\mathfrak{A}$ are larger than the atoms of $\mathcal{A}$. It is then natural to introduce the following quotient space.

Definition 2.5. Let $(\mathsf{A}, \mathcal{A}, m)$ be a measure space, $\mathfrak{A} \subset \mathcal{A}$ a σ-subalgebra. We define the quotient $(\mathsf{L}, \mathfrak{L}, \ell)$ as the image space by the equivalence relation

$$\alpha_1 \sim_1 \alpha_2 \quad \Longleftrightarrow \quad \big[\alpha_1 \in A \iff \alpha_2 \in A \ \forall A \in \mathfrak{A}\big].$$

We note that $(\mathfrak{L}, \ell)$ is isomorphic as a measure algebra to $(\mathfrak{A}, m)$, so that in the following we will not distinguish the σ-algebras and the measures, but just the spaces A and $\mathsf{L} = \mathsf{A}/\sim_1$. The quotient map will be denoted by $p : \mathsf{A} \to \mathsf{L}$.

We next define a disintegration of μ *consistent* with the partition $X = \cup_\alpha X_\alpha$ ([Fre2], Definition 452E).

Definition 2.6 (Disintegration)**.** The *disintegration* of the probability measure μ *consistent* with the partition $X = \cup_{\alpha \in \mathsf{A}} X_\alpha$ is a map $\mathsf{A} \ni \alpha \mapsto \mu_\alpha \in \mathcal{P}(X, \Omega)$ such that

1. for all $B \in \Omega$, $\mu_\alpha(B)$ is m-measurable;
2. for all $B \in \Omega$, $A \in \mathcal{A}$,

$$\mu(B \cap h^{-1}(A)) = \int_A \mu_\alpha(B) m(d\alpha), \tag{2.1}$$

where $h : X \to \mathsf{A}$ is the quotient map.

We say that the disintegration is *unique* if for all two measure valued functions $\alpha \mapsto \mu_{1,\alpha}$, $\alpha \mapsto \mu_{2,\alpha}$ which satisfy points (1), (2) it holds $\mu_{1,\alpha} = \mu_{2,\alpha}$ m-a.e. α.
The measures μ_α, $\alpha \in \mathsf{A}$, are called *conditional probabilities*.
We say that the disintegration is *strongly consistent* if for m-a.e. α $\mu_\alpha(X \setminus X_\alpha) = 0$.

We make the following observations.

1. At this level of generality, we do not require $\mu_\alpha(X_\alpha) = 1$, i.e. that μ_α is concentrated on the class X_α: in fact, we are not even requiring X_α to be μ-measurable.

2. The choice of the σ-algebra $\mathcal{A}$ in A is quite arbitrary: in our choice it is the largest σ-algebra which makes point (2) of Definition 2.6 meaningful, but one can take smaller σ-algebras, for example Λ considered in Definition 2.5.

3. If $A \in \mathcal{A}$ is an atom of the measure space $(\mathsf{A}, \mathcal{A}, m)$, then the measurability of $\mu_h(B)$ implies that $\mu_h(B)$ is constant m-a.e. on A for all $B \in \Omega$. In particular, if we want to have μ_h concentrated on the smallest possible set, we need to check μ_h with the largest σ-algebra on A: equivalently, this means that the atoms of the measure space $(\mathsf{A}, \mathcal{A}, m)$ are as small as possible. However, negligible sets are useless to this extent.

4. The formula (2.1) above does not require to have Ω countably generated, and in fact there are disintegration results in general probability spaces (see Section 452 of [Fre2] for general results). However, no general uniqueness result can be expected in that case.

5. The formula (2.1) can be easily extended to integrable functions by means of monotone convergence theorem: for all μ-integrable functions f, f is μ_α-integrable for m-a.e. α, $\int f \mu_\alpha$ is m-integrable and it holds
$$\int f\mu = \int \left(\int f\mu_\alpha \right) m(d\alpha). \tag{2.2}$$

We are ready for proving the general disintegration theorem.

Theorem 2.7 (Disintegration theorem). *Assume* (X, Ω, μ) *countably generated probability space,* $X = \cup_{\alpha \in \mathsf{A}} X_\alpha$ *a decomposition of* X, $h : X \to X_\alpha$ *the quotient map. Let* $(\mathsf{A}, \mathcal{A}, m)$ *the measure space defined by* $\mathcal{A} = h_\sharp \Omega$, $m = h_\sharp \mu$.

Then there exists a unique disintegration $\alpha \mapsto \mu_\alpha$ *consistent with the partition* $X = \cup_{\alpha \in \mathsf{A}} X_\alpha$.

Moreover, if $\mathfrak{A}$ *is a countably generated* σ*-algebra such that Proposition 2.2 holds, and* L *is the quotient space introduced in Definition 2.5,* $p : \mathsf{A} \to \mathsf{L}$ *the quotient map, then the following properties hold:*

1. $X = X_\lambda$ *is* μ*-measurable, and* $X = \cup_{\lambda \in \mathsf{L}} X_\lambda$*;*

2. *the disintegration* $\mu = \int_{\mathsf{L}} \mu_\lambda m(d\lambda)$ *satisfies* $\mu_\lambda(X_\lambda) = 1$*;*

3. *the disintegration* $\mu = \int_A \mu_\alpha m(d\alpha)$ *satisfies* $\mu_\alpha = \mu_{p(\alpha)}$ m*-a.e.*.

The last point means that the disintegration $\mu = \int_A \mu_\alpha m(d\alpha)$ has conditional probabilities μ_α constant on each atom of $\mathcal{L}$ in A, precisely given by $\mu_\alpha = \mu_\lambda$ for $\alpha = p^{-1}(\lambda)$ m-a.e.: i.e. μ_α is the pullback of the measure μ_λ.

Proof. We base the proof on well known disintegration theorem for measurable functions from $\mathbb{R}^d$ into $\mathbb{R}^{d-k}$, see for example [AFP], Theorem 2.28.

Step 1: Uniqueness. To prove uniqueness, let $\mathcal{B} = \{B_n\}_{n\in\mathbb{N}}$ be a countable family of sets generating Ω. We observe that the $L^\infty(m)$ functions given by $\int_A f_n(\alpha)m(d\alpha) = \mu(h^{-1}(A) \cap B_n)$ are uniquely defined up to a m-negligible set. This means that $\mu_\alpha(B_n)$ is uniquely defined on the algebra $\mathcal{B}$ m-a.e., so that it is uniquely determined on the σ-algebra Ω generated by $\mathcal{B}$.

Step 2: Existence. By measurable space isomorphisms (see for example the proof of the last theorem of [HJ]), we can consider $(X, \Omega) = (L, \mathfrak{A}) = ([0, 1], \mathcal{B})$, so that there exists a unique strongly consistent disintegration $\mu = \int_L \mu_\lambda m(d\lambda)$ by Theorem 2.28 of [AFP] and Step 1 of the present proof.

Step 3: Point (3) Again by the uniqueness of Step 1, we are left in proving that $\int \mu_{p(\alpha)} m(d\alpha)$ is a disintegration on $X = X_\alpha$.

Since $p : A \to L$ is measurable and p is measure preserving, $\alpha \mapsto \mu_{p(\alpha)}(B)$ is m-measurable for all $B \in \Omega$. By Proposition 2.2, for all $A \in \mathcal{A}$ there exists $\hat{A} \in \mathcal{L}$ such that $m(A \triangle \hat{A}) = \mu(h^{-1}(A) \triangle h^{-1}(\hat{A})) = 0$: then

$$\begin{aligned}\int_A \mu_{p(\alpha)}(B)m(d\alpha) &= \int_{\hat{A}} \mu_{p(\alpha)}(B)m(d\alpha) = \int_{\hat{A}} \mu_\lambda(B)m(d\lambda) \\ &= \mu(h^{-1}(\hat{A}) \cap B) = \mu(h^{-1}(A) \cap B).\end{aligned}$$

□

The final result concerns the existence of a section S for the equivalence relation $X = \cup_\alpha X_\alpha$.

Definition 2.8. We say that S is a *section* for the equivalence relation $X = \cup_{\alpha\in A} X_\alpha$ if for $\alpha \in A$ there exists a unique $x_\alpha \in S \cap X_\alpha$.

We say that S_μ is a μ-*section* for the equivalence relation induced by the partition $X = \cup_{\alpha\in A} X_\alpha$ if there exists a Borel set $\Gamma \subset X$ of full μ-measure such that the decomposition

$$\Gamma = \bigcup_{\alpha\in A} \Gamma_\alpha = \bigcup_{\alpha\in A} \Gamma \cap X_\alpha$$

has section S_μ.

Clearly from the axiom of choice, there is certainly a rooting set S, and by pushing forward the σ-algebra Ω on S we can make $(S, \mathcal{S})$ a measurable space. The following result is a classical application of selection principles.

Proposition 2.9. *The disintegration of* μ *consistent with the partition* $X = \cup_{\alpha\in A} X_\alpha$ *is strongly consistent if and only if there exists a Borel measurable* μ*-section* S *such that the* σ*-algebra* $\mathcal{S}$ *contains* $\mathcal{B}(S)$.

Proof. Since we are looking for a μ-section, we can replace (X, Ω) with $([0,1], \mathcal{B})$ by a measurable injection.

If the disintegration is strongly consistent, then the map $x \to \{\alpha : x \in X_\alpha\}$ is a μ-measurable map by definition, where the measurable space $(A, \mathcal{A})$ can be taken to be $([0,1], \mathcal{B})$ (Step 2 of Theorem 2.7). By removing a set of μ-measure 0, we can assume that h is Borel, so that by Proposition 5.1.9 of [Sri] it follows that there exists a Borel section.

The converse is a direct consequence of Theorem 2.7 and the Isomorphism Theorem among Borel spaces, Theorem 3.3.13 of [Sri]. $\square$

Remark 2.10 (Disintegration of σ-finite measures). If the total variation of μ is not finite, the quotient measure $h_\sharp\mu$ is in general infinite valued (take for example $X = \mathbb{R}^n$, $\Sigma = \mathscr{B}(\mathbb{R}^n)$, $\mu = \mathscr{L}^n$ and $X_\alpha = \{x : x \cdot z = \alpha\}$, where z is a fixed vector in $\mathbb{R}^n$ and $\alpha \in \mathbb{R}$).

Nevertheless, if μ is σ-finite and (X, Σ), $(A, \mathscr{A})$ satisfies the hypothesis of Theorem 2.7, replacing the possibly infinite-valued measure $\nu = p_\#\mu$ with an equivalent σ-finite measure m on $(A, \mathscr{A})$ one can find a family of σ-finite measures $\{\tilde{\mu}_\alpha\}_{\alpha \in A}$ on X such that

$$\mu = \int \tilde{\mu}_\alpha \, dm(\alpha) \tag{2.3}$$

and

$$\tilde{\mu}_\alpha(X \setminus X_\alpha) = 0 \quad \text{for } m\text{-a.e. } \alpha \in A. \tag{2.4}$$

Take for example $m = p_\#\theta$ for a finite measure θ equivalent to μ.

We recall that two measures μ_1 and μ_2 are *equivalent* if and only if

$$\mu_1 \ll \mu_2 \quad \text{and} \quad \mu_2 \ll \mu_1. \tag{2.5}$$

Moreover, if λ and $\{\tilde{\lambda}_\alpha\}_{\alpha \in A}$ satisfy (2.3) and (2.4) as well as m and $\{\tilde{\mu}_\alpha\}_{\alpha \in A}$, then λ is equivalent to m and

$$\tilde{\lambda}_\alpha = \frac{dm}{d\lambda}(\alpha)\tilde{\mu}_\alpha,$$

where $\frac{dm}{d\lambda}$ is the Radon-Nikodym derivative of m w.r.t. λ.

By *disintegration* of a σ-finite measures μ *strongly consistent* with a given partition we mean any family of σ-finite measures $\{\tilde{\mu}_\alpha\}_{\alpha \in A}$ which satisfies the above properties. Whenever μ has finite total variation we choose the quotient measure on the quotient space.

Finally, we recall that any disintegration of a σ-finite measure μ can be recovered by the disintegrations of the finite measures $\{\mu \llcorner K_n\}_{n \in \mathbb{N}}$, where $\{K_n\}_{n \in \mathbb{N}} \subset X$ is a partition of X into sets of finite μ-measure.

3 Characterization of the disintegration for a family of equivalence relations

Consider a family of equivalence relations on X,

$$\mathfrak{E} = \Big\{ E_e \subset X \times X : \ E_e \text{ equivalence relation}, \ e \in \mathcal{E} \Big\}$$

closed under countable intersection. By Theorem 2.7, to each E we can associate the disintegration

$$X = \bigcup_{\alpha \in A} X_\alpha, \quad \mu = \int_A \mu_\alpha m(d\alpha).$$

The key point of this section is the following easy lemma. For simplicity we will use the language of measure algebras: their elements are the equivalence classes of measurable sets w.r.t. the equivalence relation

$$A \sim A' \quad \Longleftrightarrow \quad \mu(A \vartriangle A') = 0.$$

Let $\mathfrak{Z} = \{C_z, z \in \mathcal{Z}\}$ be a family of countably generated σ-algebras such that $C_z \subset \mathcal{A}$, where $\mathcal{A}$ is a given countably generated σ-algebra on X. Let $\mathcal{C}$ be the σ-algebra generated by $\cup \mathfrak{Z} = \cup_{z \in \mathcal{Z}} C_z$.

Lemma 3.1. *There is a countable subfamily $\mathfrak{Z}' \subset \mathfrak{Z}$ such that the measure algebra generated by $\mathfrak{Z}'$ coincides with the measure algebra of $\mathcal{C}$.*

Proof. The proof follows immediately by observing that $\mathcal{C}$ is essentially countably generated because it is a σ-subalgebra of $\mathcal{A}$: one can repeat the proof of Proposition 2.2, see also Remark 2.4, or [Fre1], Proposition 332T(b).

Let A_n, $n \in \mathbb{N}$, be a generating family for $\mathcal{C}$: it follows that there is a countable subfamily $C_n \in \mathfrak{Z}$ such that A_n belongs to the σ-algebra generated by $\cup_n C_n$. Let A_{nm}, $m \in \mathbb{N}$, be the countable family of sets generating $C_n \in \mathfrak{Z}$: it is straightforward that $\{A_{nm}, n, m \in \mathbb{N}\}$ essentially generates $\mathcal{C}$. □

We can then state the representation theorem.

Theorem 3.2. *Assume that the family $\mathfrak{E}$ of equivalence relations is closed w.r.t. countable intersection: if $E_{e_n} \in \mathfrak{E}$ for all $n \in \mathbb{N}$, then*

$$\bigcap_n E_{e_n} \in \mathfrak{E}. \tag{3.1}$$

Then there exists $E_{\bar{e}} \in \mathfrak{E}$ such that for all E_e, $e \in \mathcal{E}$, the following holds:

1. *if $\mathcal{A}_e$, $\mathcal{A}_{\bar{e}}$ are the σ-subalgebra of Ω made of the saturated sets for E_e, $E_{\bar{e}}$ respectively, then for all $A \in \mathcal{A}_e$ there is $A' \in \mathcal{A}_{\bar{e}}$ such that $\mu(A \triangle A') = 0$;*

2. *if m_e, $m_{\bar{e}}$ are the restrictions of μ to $\mathcal{A}_e$, $\mathcal{A}_{\bar{e}}$ respectively, then $\mathcal{A}_e$ can be embedded (as measure algebra) in $\mathcal{A}_{\bar{e}}$ by point (1): let*

$$m_{\bar{e}} = \int m_{\bar{e},\alpha} m_e(d\alpha)$$

be the unique consistent disintegration of $m_{\bar{e}}$ w.r.t. the equivalence classes of $\mathcal{A}_e$ in $\mathcal{A}_{\bar{e}}$.

3. *If*

$$\mu = \int \mu_{e,\alpha} m_e(d\alpha), \quad \mu = \int \mu_{\bar{e},\beta} m_{\bar{e}}(d\beta)$$

are the unique consistent disintegration w.r.t. E_e, $E_{\bar{e}}$ respectively, then

$$\mu_{e,\alpha} = \int \mu_{\bar{e},\beta} m_{\bar{e},\alpha}(d\beta). \tag{3.2}$$

for m_e-a.e. α.

The last point essentially tells us that the disintegration w.r.t. $E_{\bar{e}}$ is the sharpest one, the others being obtained by integrating the conditional probabilities $\mu_{\bar{e},\beta}$ w.r.t. the probability measures $m_{\bar{e},\alpha}$.

Note that the result is useful but it can lead to trivial result if $E = \{(x,x), x \in X\}$ belongs to $\mathfrak{E}$: in this case

$$\mu_{\bar{e},\beta} = \delta_\beta, \quad m_{\bar{e},\alpha} = \mu_{e,\alpha}.$$

Proof. Point (1). We first notice that if $E_{e_1}, E_{e_2} \in \mathfrak{E}$ and $A \in \mathcal{A}_{e_1}$, the σ-algebra of saturated sets generated by E_{e_1}, then $A \in \mathcal{A}_{e_{12}}$, the σ-algebra of saturated sets generated by $E_{e_1} \cap E_{e_2}$. Hence, if $\mathfrak{E}$ is closed under countable intersection, then for every family of equivalence relations $E_{e_n} \in \mathfrak{E}$ there exists $E_{\bar{e}} \in \mathfrak{E}$ such that the σ-algebras $\mathcal{A}_{e_n}$ made of the saturated measurable sets w.r.t. E_{e_n} are subalgebras of the σ-algebra $\mathcal{A}_{\bar{e}}$ made of the saturated measurable sets w.r.t. the equivalence relation $E_{\bar{e}}$.

By Lemma 3.1 applied to the family $\mathfrak{Z} = \{\mathcal{A}_e | e \in \mathcal{E}\}$, we can take a countable family of equivalence relations such that the σ-algebra of saturated sets w.r.t. their intersection satisfies Point (1).

Point (2). This point is a consequence of the Disintegration Theorem 2.7, using the embedding $\mathcal{A}_e \ni A \mapsto A' \in \mathcal{A}_{\bar{e}}$ given by the condition $\mu(A \triangle A') = 0$, and the map in Appendix A.

Point (3). Since consistent disintegrations are unique, it is enough to show that (3.2) is a disintegration for E_e. By definition, for all $C \in \Omega$, $\mu_{\bar{e},\beta}(C)$ is a $m_{\bar{e}}$-measurable function, so that by (2.2) it is also $m_{\bar{e},\alpha}$-measurable for m_e-a.e. α and

$$\alpha \mapsto \int \mu_{\bar{e},\beta}(C) m_{\bar{e},\alpha}(d\beta)$$

is m_e-measurable. Denoting with h_e the equivalence map for E_e, for all $A \in \mathcal{A}_e$ we have

$$\mu(C \cap h_e^{-1}(A)) = \int_A \mu_{\bar{e},\beta}(C) m_{\bar{e}}(d\beta) = \int_A \left(\int \mu_{\bar{e},\beta}(C) m_{\bar{e},\alpha}(d\beta) \right) m_e(d\alpha),$$

where we used the definition of $\mu_{\bar{e},\beta}$ in the first equality and (2.2) in the second one. $\square$

4 Disintegration on 1D rays of a Lipschitz function

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

$$\phi(x) - \phi(y) \leqslant |\!|\!| y - x |\!|\!| \qquad \forall x, y \in \mathbb{R}^n. \tag{4.1}$$

Definition 4.2 (Transport set)**.** The *transport set* associated to a potential ϕ is the set $\bar{\mathcal{T}}$ made of the open segments $]\!]x, y[\![$ for every couple (x, y) such that in (4.1) equality holds:

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

Similarly, we will also consider the transport set with all the endpoints:

$$\bar{\mathcal{T}}_e = \bigcup_{(x,y)\in\partial_c\phi\setminus\{y=x\}} [\![x, y]\!].$$

We summarize briefly the construction. Due to the strict convexity of the norm, $\bar{\mathcal{T}}$ 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 $\bar{\mathcal{T}}$, by identifying points on a same ray. The issue is to show that the conditional measures of $\mathcal{L}^n \llcorner \bar{\mathcal{T}}$ 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.

The following example shows that the absolute continuity of the conditional measures in the disintegration of the Lebesgue measure established in Theorem 4.26 relies on some regularity of the vector field of ray directions, since the Borel measurability is not enough.

Example 4.3 (A Nikodym set in $\mathbb{R}^3$)*.* In [AKP2], 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 $x \in M_N$,

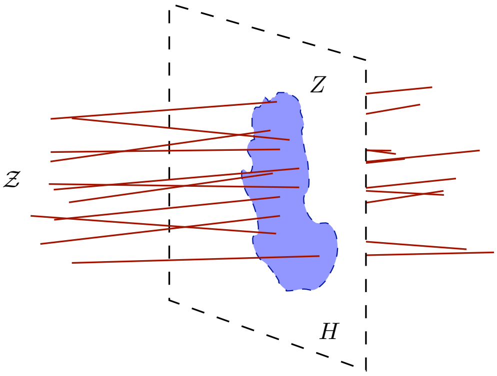


Figure 1: A sheaf of rays

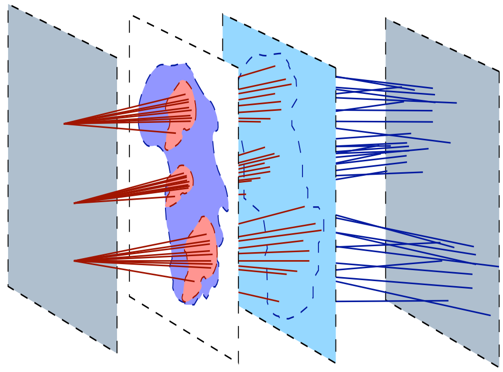


Figure 2: Approximation of rays

- $l_x \cap l_y = \emptyset$ 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 one 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 [Lar1].

By the additivity of the measures, the thesis will follow if proved on the elements of a countable partition of $\bar{\mathcal{T}}$ into Borel sets. In particular, in Subsections 4.1, 4.2 we provide a partition into model sets $\bar{\mathcal{Z}}$ made of rays transversal to some hyperplane H, let Z be the intersection of $\bar{\mathcal{Z}}$ with H (Figure 1). Points x belonging to $\bar{\mathcal{Z}}$ 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 $\phi(y(x)) \geqslant \phi(x)$ or negative otherwise. We prove in Corollary 4.23 that the bijective parameterization

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

provides an isomorphism between the $\mathscr{L}^n$-measurable functions on $\bar{\mathcal{Z}}$ 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 $\mathscr{L}^n \llcorner \bar{\mathcal{Z}}$ is absolutely continuous w.r.t. the measure $(\mathcal{H}^{n-1} \llcorner Z) \otimes \mathcal{H}^1$, 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. $x = \sigma^{t(x)}(y(x))$,

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

where c is obtained by an easy change of variables in the one 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 4.18 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 $\bar{\mathcal{Z}}$, set zero on $\mathbb{R}^n \setminus \bar{\mathcal{Z}}$, is a Radon measure, and the following formula holds (Lemma 6.3):

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

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

4.1 Elementary structure of the Transport Set $\bar{\mathcal{T}}$

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 4.5).

Definition 4.4. The *outgoing rays* from $x \in \mathbb{R}^n$ are defined as

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

The *incoming rays* at x are then given by

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

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

The transport set with the endpoints $\bar{\mathcal{T}}_e$ is just the subset of $\mathbb{R}^n$ where there is some non degenerate transport ray: those x such that $\bar{\mathcal{R}}(x) \neq \{x\}$. Similarly, $\bar{\mathcal{T}}$ is the set where both $\bar{\mathcal{P}}(x) \neq \{x\}$ and $\bar{\mathcal{P}}^{-1}(x) \neq \{x\}$. The following remarks are in order.

The set $\bar{\mathcal{P}}(x)$ is a union of closed segments with endpoint x, which we call rays. In fact, ϕ's Lipschitz condition (4.1) implies that, for every $y \in \bar{\mathcal{P}}(x)$, ϕ must decrease linearly from x to y at the maximal rate allowed:

$$\phi(x + t(y-x)) = \phi(x) - t\tilde{|\!|}y - x\tilde{|\!|} \qquad \text{for all } y \in \bar{\mathcal{P}}(x),\ t \in [0,1]. \tag{4.2}$$

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 $x \in \bar{\mathcal{P}}^{-1}(y)$, $z \in \bar{\mathcal{P}}(y)$, one has

$$\phi(z) \overset{z\in\bar{\mathcal{P}}(y)}{=} \phi(y) - \tilde{\|}z-y\tilde{\|} \overset{y\in\bar{\mathcal{P}}(x)}{=} \phi(x) - \tilde{\|}y-x\tilde{\|} - \tilde{\|}z-y\tilde{\|} \leqslant \phi(x) - \tilde{\|}z-x\tilde{\|}. \quad (4.3)$$

Again by Lipschitz condition (4.1) equality must hold: then $\tilde{\|}z-y\tilde{\|} = \tilde{\|}y-x\tilde{\|} + \tilde{\|}z-y\tilde{\|}$. 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 $x \in \bar{\mathcal{T}}_e$, it is possible to define a vector field giving the direction of the ray through x:

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

In order to show that in $\bar{\mathcal{T}}_e$ 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 4.3 (from [Lar1], [AKP2]).

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

$$\bar{\mathcal{D}}(x) := \left\{ \frac{y-x}{|y-x|}\chi_{\bar{\mathcal{P}}(x)}(y) + \frac{x-y}{|y-x|}\chi_{\bar{\mathcal{P}}^{-1}(x)}(y) \right\}_{y\in\bar{\mathcal{R}}(x)} \qquad \text{for all } x \in \bar{\mathcal{T}}_e. \quad (4.4)$$

We first show that the above maps $\bar{\mathcal{P}}$, $\bar{\mathcal{D}}$ 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 4.5. *The multivalued functions $\bar{\mathcal{P}}$, $\bar{\mathcal{P}}^{-1}$, $\bar{\mathcal{R}}$, $\bar{\mathcal{D}}$ 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 $\bar{\mathcal{T}}$ and $\bar{\mathcal{T}}_e$ are σ-compact.*

Proof. Firstly, consider the graph of $\bar{\mathcal{P}}$: it is closed. In fact, take a sequence (x_k, z_k), with $z_k \in \bar{\mathcal{P}}(x_k)$, converging to a point (x, z). Then, since $\phi(z_k) = \phi(x_k) - \tilde{\|}z_k - x_k\tilde{\|}$, by continuity we have that $\phi(z) = \phi(x) - \tilde{\|}z-x\tilde{\|}$. Therefore the limit point (x, z) belongs to $\mathrm{Graph}(\bar{\mathcal{P}}(x))$. Since the graph is closed, then both the image and the counterimage of a closed set are σ-compact. In particular, this means that $\bar{\mathcal{P}}$, $\bar{\mathcal{P}}^{-1}$ and $\bar{\mathcal{R}}$ are Borel. Secondly, since the graph of $\bar{\mathcal{P}}$ is closed, both the graphs of $\bar{\mathcal{P}} \setminus \mathbb{I}$ and $\bar{\mathcal{P}}^{-1} \setminus \mathbb{I}$ are still σ-compact. In particular, the intersection and the union of their images must be σ-compact. These are, respectively, the transport sets $\bar{\mathcal{T}}$, $\bar{\mathcal{T}}_e$. Finally, the map $\bar{\mathcal{D}}$ is exactly the composite map $x \in \bar{\mathcal{T}} \to \mathrm{dir}(x, \bar{\mathcal{R}}^{-1}(x) \setminus \{x\})$, 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 4.6. 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 $\bar{\mathcal{T}}_e$ 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 $\bar{\mathcal{T}}_e$, 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 13.1). 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 4.9 below in order to show the rectifiability of the set where $\bar{\mathcal{D}}$ is multivalued.

Definition 4.7 (Rectifiable set)**.** Let $E \subset \mathbb{R}^n$ be an $\mathcal{H}^k$-measurable set. We say that E is *countably* k*-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 4.8 (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*

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

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 4.9. *On $\bar{\mathcal{T}}_e$, $\bar{\mathcal{D}}$ is single valued out of a countably $(n-1)$-rectifiable set.*

Proof. We show the rectifiability of the set where $\bar{\mathcal{D}}$ is multivalued applying Theorem 4.8.

Step 1: Countable covering. By (4.3) $\bar{\mathcal{D}}$ 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

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

for $h \geqslant \bar{h}$, $\rho \leqslant \bar{\rho}$, 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

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

Define

$$\begin{aligned} J_{ijp} := \Big\{ &x \in \bar{\mathcal{T}}_e : \exists d_1, d_2 \in \bar{\mathcal{D}}(x) \text{ s.t. } [\![d_1 - d_2]\!] \geqslant \frac{1}{p}, \\ &\qquad d_1 \in B_1^{ip}, d_2 \in B_2^{ip},\ \mathcal{H}^1\big((x + \langle d_i \rangle) \cap \bar{\mathcal{P}}(x)\big) \geqslant \frac{1}{p} \Big\}. \end{aligned}$$

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 $d^1_k \in \bar{\mathcal{D}}(x_k) \cap B^{ip}_1$ and $d^2_k \in \bar{\mathcal{D}}(x_k) \cap B^{ip}_2$ converging respectively to some $d_1 \in B^{ip}_1$, $d_2 \in B^{ip}_2$, and

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

By the continuity of ϕ, since y^1_k, y^2_k belong to $\bar{\mathcal{P}}(x_k)$, then y^1, y^2 belong to $\bar{\mathcal{P}}(x)$ 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 countably $(n-1)$-rectifiable. To this purpose, we show that the cone condition of Theorem 4.8 holds: we prove that for every $x \in J_{ijp}$ the relative interior of the cone

$$x + \big\{(\lambda_1 B^{ip}_1 - \lambda_2 B^{ip}_2) \cup (-\lambda_1 B^{ip}_1 + \lambda_2 B^{ip}_2)\big\}_{\lambda_1,\lambda_2\geqslant 0}$$

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

$$\exists q_1, q_2 \in \delta D^*(d^2/|\!|\!|d^2|\!|\!|) - \delta D^*(d^1/|\!|\!|d^1|\!|\!|): \qquad q_1 \cdot \ell \geqslant 0, \qquad q_2 \cdot \ell \leqslant 0. \tag{4.5}$$

By definition of B^{ip}_1 and B^{ip}_2, if $\ell \in B^{ip}_1$ one would have $q_1 \cdot \ell \leqslant -1/i$, while $q_2 \cdot \ell \geqslant 1/i$ would hold if $\ell \in B^{ip}_2$, yielding a contradiction: this means that any possible limit ℓ as above does not belong to $B^{ip}_1 \cup B^{ip}_2$. Then (4.5) 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

$$x + \big\{(\lambda_1 B^{ip}_1 - \lambda_2 B^{ip}_2) \cup (-\lambda_1 B^{ip}_1 + \lambda_2 B^{ip}_2)\big\}_{\lambda_1,\lambda_2\geqslant 0}.$$

Step 5: Proof of the estimate (4.5). Let $x_k \in J_{ijp}$ converging to x, up to subsequence as in Step 2 one can assume also that there exist $y^1_k, y^2_k \in \bar{\mathcal{P}}(x_k)$ 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 ν belonging to the subdifferential $\partial^-|\!|\!|b|\!|\!|$ of $|\!|\!|\cdot|\!|\!|$ at b, and depending on ℓ, such that the equality

$$|\!|\!|a_k|\!|\!| = |\!|\!|b|\!|\!| + \nu \cdot (a_k - b) + o(|a_k - b|), \tag{4.6}$$

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 $\nu^2 \in \partial^-|\!|\!|y^2 - x|\!|\!|$, $\nu^1_k \in \partial^-|\!|\!|y^1_k - x|\!|\!|$ in order to have

$$\begin{aligned}
\phi(x) + \nu^2 \cdot (x - x_k) + o(|x - x_k|) &\overset{(4.6)}{=} \phi(x) - |\!|\!|y^2 - x|\!|\!| + |\!|\!|y^2 - x_k|\!|\!| \\
&= \phi(y^2) + |\!|\!|y^2 - x_k|\!|\!| \geqslant \phi(x_k) = \phi(y^1_k) + |\!|\!|y^1_k - x_k|\!|\!| \\
&\geqslant \phi(x) - |\!|\!|y^1_k - x|\!|\!| + |\!|\!|y^1_k - x_k|\!|\!| \overset{(4.6)}{=} \phi(x) + \nu^1_k \cdot (x - x_k) + o(|x - x_k|).
\end{aligned}$$

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

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

proving that

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

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

Lemma 4.9 ensures that one can define on a Borel subset of $\bar{\mathcal{T}}_e$, differing from $\bar{\mathcal{T}}_e$ for an $\mathscr{L}^n$-negligible set, a vector field giving at each point the direction of the ray passing there:

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

On this domain, the function d is Borel, by Lemma 4.5, being just a restriction of the Borel multivalued map $\bar{\mathcal{D}}$. Since, by the strong triangle inequality in (4.3), rays cannot bifurcate, we are allowed to consider their endpoints, possibly at infinity. After compactifying $\mathbb{R}^n$, define on $\bar{\mathcal{T}}_e$

$$a(x) = \big\{x + td, \text{ where } t \text{ minimal value s.t. } \phi(x) = \phi(x + td) + td,\ d \in \bar{\mathcal{D}}(x)\big\},$$
$$b(x) = \big\{x + td, \text{ where } t \text{ maximal value s.t. } \phi(x) = \phi(x + td) + td,\ d \in \bar{\mathcal{D}}(x)\big\}.$$

As we will prove in the following (see resp. Lemma 4.14 and Lemma 4.21), 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. $x \in \bar{\mathcal{T}}$.

4.2 Partition of $\bar{\mathcal{T}}$ into model sets

Here we decompose the transport set $\bar{\mathcal{T}}$ 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 $\bar{\mathcal{T}}$.

Definition 4.10 (Sheaf set)**.** The sheaf sets $\bar{\mathcal{Z}}$, $\bar{\mathcal{Z}}_e$ are defined to be σ-compact subsets of $\bar{\mathcal{T}}$ of the form

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

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 $\bar{\mathcal{T}}$ (resp. $\bar{\mathcal{T}}_e$) with countably many possibly disjoint sets $\bar{\mathcal{Z}}_i$ (resp. $\bar{\mathcal{Z}}_{ei}$). 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 $\mathbf{S}^{n-1}$:

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

Lemma 4.11. *The following sets are sheaf sets covering $\bar{\mathcal{T}}$ (resp. $\bar{\mathcal{T}}_e$):*

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

$$\begin{aligned}
\bar{\mathcal{Z}}_{jk\ell m} &= \Big\{ x \in \bar{\mathcal{T}} : \ d(x) \in S_j, \ \ell, m \text{ extremal values s.t.} \\
&\qquad\qquad 2^{-k}[\ell - 1, m + 1] \subset \bar{\mathcal{R}}(x) \cdot \mathfrak{e}_j \Big\} \\
\bar{\mathcal{Z}}^e_{jk\ell m} &= \Big\{ x \in \bar{\mathcal{T}}_e : \ \exists d \in S_j \cap \bar{\mathcal{D}}(x), \ \ell, m \text{ extremal values s.t.} \\
&\qquad\qquad 2^{-k}[\ell - 1, m + 1] \subset \big(\bar{\mathcal{R}}(x) \cap \{x + \mathbb{R}d\} \big) \cdot \mathfrak{e}_j \Big\}.
\end{aligned}$$

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

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

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

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

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

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

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

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_h^p, clearly depending also on j, k, ℓ, m. □

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

Definition 4.12 (d-cylinder)**.** A cylinder subordinated to the vector field d is a σ-compact set of the form

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

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 4.13. *With the notations of Lemma 4.11, $\bar{\mathcal{T}}$ is covered by the d-cylinders*

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

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

Proof. The proof is similar to the one of Lemma 4.11: just cut the sets $\bar{\mathcal{Z}}'_{jk\ell m}$ 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 $\bar{\mathcal{Z}}'_{jk\ell m}$ 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 4.14. *The (multivalued) functions a, b are Borel on the transport set with endpoints $\bar{\mathcal{T}}_e$.*

Proof. A first way could be to show that their graph is σ-compact (as for Lemma 4.5). Define instead the following intermediate sets between a d-cylinder and a sheaf set:

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

where $\{\bar{\mathcal{Z}}^{\mathfrak{e}}_{jk\ell m}\}$ is the partition defined in Lemma 4.11. Define the Borel function pushing, along rays, each point in $\mathcal{V}^-_{jk\ell m}$ to the upper basis:

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

Then, the Borel functions $\cup_{j\ell m} \sigma^+ \chi_{\mathcal{V}^-_{jk\ell m}}(x)$, 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 $\cup_{j\ell m} \sigma^- \chi_{\mathcal{V}^+_{jk\ell m}}(x)$. □

Remark 4.15. Focus on a sheaf set with axis e_1 and basis $Z \subset \{x \cdot e_1 = 0\}$. The composite of the following two maps

$$\begin{array}{ccc} \bar{\mathcal{Z}}(Z) \subset \mathbb{R}^n & \to & \mathbb{R} \times \mathbb{R}^n \\ z & \to & (z \cdot e_1, \sigma^{-z \cdot e_1} z) = (t, x) \end{array}$$

$$\begin{array}{ccc} \mathbb{R} \times \mathbb{R}^n & \to & Z + (-1,1)e_1 \subset \mathbb{R}^n \\ (z \cdot e_1, \sigma^{-z \cdot e_1} z) = (t, x) & \to & \left(x + \frac{t \arctan t}{\arctan(b(x) \cdot e_1)} e_1 \chi_{t \geqslant 0} + \frac{t \arctan t}{\arctan(a(x) \cdot e_1)} e_1 \chi_{t \leqslant 0} \right) \end{array}$$

is a Borel and invertible change of variable from $\bar{\mathcal{Z}}(Z)$ to the cylinder $Z + (-1,1)e_1$, with Borel inverse. This will turn out to carry negligible sets into negligible sets (see Corollary 4.23).

Remark 4.16. Consider a d-cylinder of the above partition

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

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 (4.18) 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.

4.3 Fundamental estimate: the sheaf set $\mathcal{Z}$

In this section we arrive to the explicit disintegration of the Lebesgue measure on $\bar{\mathcal{T}}$, w.r.t. the partition in rays when 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 the present section. 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 4.4. In particular, it is proved that the conditional measures are absolutely continuous w.r.t. the Hausdorff one dimensional measure on the rays. We recall that this is nontrivial, since some regularity of the field of directions is needed (see Example 4.3).

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 $\bar{\mathcal{Z}}_e$ with axis e_1 and a bounded basis $Z \subset \{x : e_1 \cdot x = 0\}$: assume that, for suitable $h^{\pm}$,

$$\bar{\mathcal{Z}}_e = \cup_{y \in Z} [\![a(y), b(y)]\!], \qquad e_1 \cdot a\llcorner_Z < h^- \leqslant 0, \qquad e_1 \cdot b\llcorner_Z > h^+ \geqslant 0.$$

Example 4.17 (Local approximation of the vector field d). Suppose $h^- < 0$. Consider the Borel functions moving points along rays, parametrized with the projection on the e_1 axis,

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

In order to avoid to work with values at infinite, we think to truncate the rays at $\{x \cdot e_1 = h^-\}$. Choose now a dense sequence $\{a_i\}$ in $\sigma^{h^-}Z$. Approximate the potential ϕ with the sequence of potentials

$$\phi_I(x) = \max\left\{\phi(a_i) - \tilde{\|}x - a_i\tilde{\|} : i = 1, \dots, I\right\}.$$

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 $\tilde{\mathcal{Z}}_e \cap \{x \cdot e_1 \geqslant h^-\}$. There, consider now the vector fields of ray directions

$$d_I(x) = \sum_{i=1}^{I} d^i(x)\chi_{\overline{\Omega_i^I}}(x) \qquad \text{with} \qquad d^i(x) = \frac{x - a_i}{|x - a_i|}, \tag{4.7}$$

where the open sets Ω_i^I are

$$\begin{aligned}\Omega_i^I &= \left\{x : \ \phi(a_i) - \tilde{\|}x - a_i\tilde{\|} > \phi(a_j) - \tilde{\|}x - a_j\tilde{\|}, \ j \in \{1 \dots I\} \setminus i\right\} \\ &= \text{interior of } \left\{x : \ \phi(a_i) = \phi_I(x) + \tilde{\|}x - a_i\tilde{\|}\right\}.\end{aligned}$$

They partition $\mathbb{R}^n$, together with their boundary. Notice that this boundary is $\mathcal{H}^{n-1}$-countably rectifiable: for example apply Lemma 4.9, 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 $\tilde{\mathcal{Z}}_e \cap \{x \cdot e_1 > h^-\}$. More precisely, every selection of the d_I converges pointwise to d on $\tilde{\mathcal{Z}}_e \cap \{x \cdot e_1 > h^-\}$. Consider any sequence $\{d_{I_j}(x)\}_j$ convergent to some $\bar{d}$. The corresponding points a_{i_j} satisfy

$$\phi_{I_j}(a_{i_j}) = \phi_{I_j}(x) + \tilde{\|}x - a_{i_j}\tilde{\|};$$

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

Define the map $\sigma^t_{d_I}$ which, similarly to σ^t, moves points along the rays relative to ϕ_I. Notice that, by (4.7), within Ω_i^I the map $\sigma^t_{d_I}$ moves points towards a_i, for $i \leqslant I$. As a consequence, for $S \subset \Omega_i^I \cap \{x \cdot e_1 = h\}$ and $h - h^- > t \geqslant 0$ the set $\sigma^{-t}_{d_I} S$ is similar to S: precisely

$$\sigma^{-t}_{d_I} S = a_i + \frac{h - h^-}{h - t - h^-}(S - a_i).$$

By additivity, also for $S \subset \{x \cdot e_1 = h\}$ and $h - h^- > t \geqslant 0$ the following equality holds:

$$\mathcal{H}^{n-1}(\sigma^{-t}_{d_I} S) = \left(\frac{h - h^-}{h - t - h^-}\right)^{n-1} \mathcal{H}^{n-1}(S). \tag{4.8}$$

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

$$\bar{\mathcal{K}} = \bar{\mathcal{Z}} \cap \{h^- \leqslant e_1 \cdot x \leqslant h^+\} = \cup_{t \in [h^-,h^+]} \sigma^t Z, \qquad \text{and } a_{\llcorner \bar{\mathcal{K}}} \cdot e_1 \leqslant h^-,\ b_{\llcorner \bar{\mathcal{K}}} \cdot e_1 \geqslant h^+.$$

Lemma 4.18 (Absolutely continuous push forward)**.** *For $h^- < s \leqslant t < h^+$ the following estimate holds:*

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

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

Proof. Fix $h^- < s \leqslant t \leqslant h^+$. Consider $S \subset Z$ and assume firstly that $\mathcal{H}^{n-1}(\sigma^t S) > 0$. Approximate the vector field d as in Example 4.17. There, we proved pointwise convergence on $\bar{\mathcal{Z}}_e \cap \{x \cdot e_1 > h^-\}$. Choose any $\eta > 0$. By Egoroff theorem, the convergence of d_I to d is uniform on a compact subset $A_\eta \subset \sigma^t S$ such that

$$\mathcal{H}^{n-1}(A_\eta) \geqslant \mathcal{H}^{n-1}(\sigma^t S) - \eta. \tag{4.9}$$

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 $\sigma^{s-t}_{d_I}(A_\eta)$ converges in Hausdorff metric to $\sigma^{s-t}_{d}(A_\eta)$. Moreover, by the explicit formula (4.8) for the regular d_I,

$$\mathcal{H}^{n-1}(A_\eta) \equiv \mathcal{H}^{n-1}(\sigma^0_{d_I} A_\eta) \leqslant \left(\frac{t - h^-}{s - h^-}\right)^{n-1} \mathcal{H}^{n-1}(\sigma^{s-t}_{d_I} A_\eta). \tag{4.10}$$

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

$$\limsup_{I \to \infty} \mathcal{H}^{n-1}(\sigma^{s-t}_{d_I} A_\eta) \leqslant \mathcal{H}^{n-1}(\sigma^{s-t}_{d} A_\eta) \leqslant \mathcal{H}^{n-1}(\sigma^s S). \tag{4.11}$$

Collecting (4.9), (4.10) and (4.11) 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 $h^- \leqslant s \leqslant t < h^+$. 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 4.19*.* The consequences of this fundamental formula are given in Subsection 4.4. 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

$$\left(\frac{h^+ - t}{h^+ - h^-}\right)^{n-1} \mathcal{H}^{n-1}(Z(h^-)) \leqslant \mathcal{H}^{n-1}(Z(t)) \tag{4.12}$$

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 4.21). As a consequence, we can cover $\mathcal{H}^n$-almost all $\bar{\mathcal{T}}_e$ with countably many d-cylinders — of positive $\mathcal{H}^n$-measure if $\bar{\mathcal{T}}_e$ has positive $\mathcal{H}^n$-measure.

4.4 Explicit disintegration of $\mathscr{L}^n$

We derive now the consequences of the fundamental estimates of Lemma 4.18. We first observe by a density argument that the set of endpoints of transport rays is $\mathcal{H}^n$-negligible (Lemma 4.21). 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 $\bar{\mathcal{T}}_e$ (Theorem 4.26).

Remark 4.20. We underline that the results of this section 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 4.18.

Lemma 4.21. *The set of endpoints of transport rays is negligible: $\mathscr{L}^n(\bar{\mathcal{T}}_e \setminus \bar{\mathcal{T}}) = 0$.*

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 4.2, it is enough to prove the negligibility e.g. of the initial points of the set $\mathfrak{L}$ where $d \in B_\eta(e_1)$, for some small $\eta > 0$, and $\mathcal{H}^1(\bar{\mathcal{P}}(x) \cdot e_1) > 1$. 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 $\lambda \in T$

$$\mathcal{H}^{n-1}(H_\lambda) \geqslant (1-\varepsilon)r^{n-1} \qquad \text{where } H_\lambda = \mathfrak{L} \cap \mathcal{A} \cap \{x \cdot e_1 = \lambda,\ |x - \lambda e_1|_\infty \leqslant r\}. \tag{4.13}$$

Choose, now, $s < t$, both in T, with $|t - s| < \varepsilon r$. By Lemma 4.18, then

$$\mathcal{H}^{n-1}\left(\sigma^{t-s} H_s\right) \geqslant \left(\frac{1-t}{1-s}\right)^{n-1} \mathcal{H}^{n-1}(H_s) \overset{(4.13)}{\geqslant} (1-2\varepsilon) r^{n-1}.$$

Moreover, since $d \in B_\eta(e_1)$, we have that $\mathcal{H}^{n-1}\left(\sigma^{t-s} H_s \setminus \{x \cdot e_1 = t,\ |x - te_1|_\infty \leqslant r\}\right) \leqslant 2\eta r^{n-1}$. Since points in $\sigma^{t-s} H_s$ do not stay in $\mathcal{A}$, then we reach a contradiction with the estimate (4.13) for $\lambda = t$: we would have

$$\begin{aligned} r^{n-1} &= \left|\{x \cdot e_1 = t,\ |x - te_1|_\infty \leqslant r\}\right| \geqslant \\ &(1-\varepsilon) r^{n-1} + (1 - 2\varepsilon - 2\eta) r^{n-1} = (2 - 3\varepsilon - 2\eta) r^{n-1}. \end{aligned}$$

□

Lemma 4.22. *With the notations of Lemma 4.18, the push forward of the measure $\mathcal{H}^{n-1} \llcorner Z$ by the map σ^t can be written as*

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

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

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

Proof. Lemma 4.18 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 4.18, with $s = 0$. □

Corollary 4.23. *The map $\sigma^t(x) : [h^-, h^+] \times Z \mapsto \bar{\mathcal{Z}}$ 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 $\bar{\mathcal{Z}}$, 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 $N \subset \bar{\mathcal{Z}}$ is $\mathcal{H}^n$-negligible, then

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

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

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

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 4.24. *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* a, 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.* $y \in Z$*:*

$$-\left(\frac{n-1}{b(y)\cdot e_1 - t}\right)\tilde{\alpha}(t,y) \leqslant \frac{d}{dt}\tilde{\alpha}(t,y) \leqslant \left(\frac{n-1}{t - a(y)\cdot e_1}\right)\tilde{\alpha}(t,y), \tag{4.14}$$

$$\left(\frac{|b(y)-\sigma^t y|}{|b(y)-y|}\right)^{n-1}(-1)^{\chi_{t<0}} \leqslant \tilde{\alpha}(t,y)(-1)^{\chi_{t<0}} \leqslant \left(\frac{|\sigma^t y - a(y)|}{|y-a(y)|}\right)^{n-1}(-1)^{\chi_{t<0}}. \tag{4.15}$$

Moreover,

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

Proof. The function $\tilde{\alpha}(t,\cdot)$ is by definition $L^1_{loc}(\mathcal{H}^{n-1}\llcorner Z)$, 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 4.18 and Corollary 4.23, for $h^- < s < t < h^+$ and every measurable $S \subset Z$, we have

$$\begin{aligned}\left(\frac{h^+-t}{h^+-s}\right)^{n-1}\int_S \tilde{\alpha}(s,y)d\mathcal{H}^{n-1}(y) &\leqslant \int_S \tilde{\alpha}(t,y)d\mathcal{H}^{n-1}(y)\\ &\leqslant \left(\frac{t-h^-}{s-h^-}\right)^{n-1}\int_S \tilde{\alpha}(s,y)d\mathcal{H}^{n-1}(y).\end{aligned} \tag{4.16}$$

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 ($t_j \geqslant t_i$):

$$\left[\left(\frac{h^+-t_j}{h^+-t_i}\right)^{n-1} - 1\right]\tilde{\alpha}(t_i,y) \leqslant \tilde{\alpha}(t_j,y) - \tilde{\alpha}(t_i,y) \leqslant \left[\left(\frac{t_j-h^-}{t_i-h^-}\right)^{n-1} - 1\right]\tilde{\alpha}(t_i,y). \tag{4.17}$$

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 (4.16) $t \to s^+$. By the above pointwise estimate (4.17), taking the derivative, we get (4.14). Equation (4.14), moreover, implies the following monotonicity:

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

Then, since $\frac{e_1\cdot b - t}{e_1 \cdot b} = \frac{|b-\sigma^t y|}{|b-y|}$, $\frac{t-e_1\cdot a}{-e_1\cdot a} = \frac{|a-\sigma^t y|}{|a-y|}$ and $\tilde{\alpha}(0,\cdot) \equiv 1$, we obtain exactly (4.15). Furthermore

$$\begin{aligned}
\int_{a\cdot e_1}^0 \left|\frac{d}{dt}\tilde{\alpha}(t,y)\right| dt &\overset{(4.14)}{\leqslant} \int_{\left\{\frac{d\tilde{\alpha}(t,y)}{dt}>0\right\}\cap\{t<0\}} \frac{d}{dt}\tilde{\alpha}(t,y)\, dt + \int_{a\cdot e_1}^0 \frac{(n-1)\tilde{\alpha}(t,y)}{b\cdot e_1 - t}\, dt \\
&\overset{(4.14)}{\leqslant} \int_{a\cdot e_1}^0 \frac{d}{dt}\tilde{\alpha}(t,y)\, dt + 2\int_{a\cdot e_1}^0 \frac{(n-1)\tilde{\alpha}(t,y)}{b\cdot e_1 - t}\, dt \\
&\overset{(4.15)}{\leqslant} 1 + 2\int_{a\cdot e_1}^0 \frac{(n-1)(e_1\cdot b - t)^{n-2}}{(e_1\cdot b)^{n-1}}\, dt = 1 + 2\left(\frac{|b-a|^{n-1}}{|b|^{n-1}} - 1\right).
\end{aligned}$$

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

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

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

Lemma 4.25. *On $\bar{\mathcal{K}} = \{\sigma^t Z\}_{t\in(h^-,h^+)}$, we have the following disintegration of the Lebesgue measure: $\forall \phi : \int |\phi|\, d\mathscr{L}^n < \infty$*

$$\int_{\bar{\mathcal{K}}} \varphi(x)\, d\mathscr{L}^n(x) = \int_{y\in Z} \left\{ \int_{h^-}^{h^+} \varphi(\sigma^t y)\tilde{\alpha}(t,y)\, d\mathcal{H}^1(t) \right\} d\mathcal{H}^{n-1}(y), \tag{4.18}$$

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 $\varphi \circ \sigma^t \llcorner_Z$ is still $\mathscr{L}^n$-measurable (Corollary 4.23), we have

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

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

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

Finally, since $\tilde{\alpha}$ is measurable (Corollary 4.24) and locally integrable, by the above estimate and Tonelli theorem applied to the negative and positive part, Fubini theorem provides the thesis. $\square$

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

- $\{\bar{\mathcal{Z}}_i\}_{i\in\mathbb{N}}$ is the partition of the transport set $\bar{\mathcal{T}}$ into sheaf sets as in Lemma 4.11;
- Z_i is a section of $\bar{\mathcal{Z}}_i$ and $\mathfrak{d}_i$ is the relative axis;
- $\mathcal{S}$ is the quotient set of $\bar{\mathcal{T}}_e$ 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 $\bar{\mathcal{Z}}_i$;
- $\tilde{\alpha}_i(t,\cdot)$ is the Radon-Nikodym derivative of $(\sigma^{-t})_\sharp \mathcal{H}^{n-1} \llcorner \sigma^t Z_i$ w.r.t. $\mathcal{H}^{n-1} \llcorner Z_i$;
- $c(t,y) := \sum_i \tilde{\alpha}_i\big(d(y)\cdot(t\mathfrak{d}_i - y), y\big) d(y)\cdot \mathfrak{d}_i \chi_{\bar{\mathcal{Z}}_i}(y)$.

Theorem 4.26. *One has then the following disintegration of the Lebesgue measure on $\bar{\mathcal{T}}_e$*

$$\int_{\bar{\mathcal{T}}_e} \varphi(x)\, d\mathscr{L}^n(x) = \int_{\mathcal{S}} \left\{ \int_{a(y)\cdot d(y)}^{b(y)\cdot d(y)} \varphi(y + (t - y\cdot d(y))d(y)) c(t,y) d\mathcal{H}^1(t) \right\} d\mathcal{H}^{n-1}(y), \tag{4.19}$$

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

Remark 4.27. As a consequence of Corollary 4.24, c is measurable in y and locally Lipschitz in t.

Remark 4.28 (Dependence on the partition)*.* Suppose to partition the transport set in a different family of sheaf sets $\bar{\mathcal{Z}}'_i$, with the quotient space identified with the union $\mathcal{S}'$ of the new basis. Then, one can refine the partitions $\{\bar{\mathcal{Z}}_i\}_i$ and $\{\bar{\mathcal{Z}}'_i\}_i$ into a family of sheaf sets $\{\hat{\bar{\mathcal{Z}}}_i\}_i$. Consider the change of variables in a single sheaf set $\hat{\bar{\mathcal{Z}}}_i$. If we consider $Z \subset \{x\cdot\nu + c = 0\}$ and $Z' \subset \{x\cdot\nu' + c'\} = 0$, then

$$y + \big(t - y\cdot d(y)\big)d(y) = y' + \big(t' - y'\cdot d(y)\big)d(y)$$

with $y' = y - \frac{c'+y\cdot\nu'}{d(y)\cdot\nu'} d(y)$ and $t' = t$. Moreover, we have the disintegration formulas

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

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

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

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

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

Proof. Forget the set of endpoints of rays, since by Lemma 4.21 they are negligible.

Consider the refinement of the partition $\{\bar{Z}_i\}_{i\in\mathbb{N}}$ given in Lemma 4.13, which partitions $\bar{\mathcal{T}}$ into cylinders subordinated to d: denote these as $\{\bar{\mathcal{K}}_{ij}\}_{ij\in\mathbb{N}}$ and set $\widehat{Z}_{ij}$, $h^{\pm}_{ij}$ in order to have $Z_i = \cup_{j\in\mathbb{N}}\widehat{Z}_{ij}$ and

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

Since we set

$$c(t,y)=\sum_i\tilde{\alpha}_i\big(d(y)\cdot(t\partial_i-y),y\big)d(y)\cdot\partial_i\chi_{\bar{Z}_i}(y),$$

the local result of Lemma 4.25, with a translation and the change of variable $t\to\frac{t}{d(y)\cdot\partial_i}$, yields

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

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

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

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. $\square$

5 The disintegration of the Lebesgue measure on the faces of a convex body

In this section, after setting the notation and some basic definitions, we apply Theorem 2.7 to get the existence, uniqueness and strong consistency of the disintegration of the Lebesgue measure on the faces of a convex function. Then, we give a rigorous formulation of the problem we are going to deal with and state our main theorem.

5.1 Setting and statement

Let us consider the **ambient space**

$$(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n), \mathscr{L}^n \llcorner K),$$

where $\mathscr{L}^n$ is the Lebesgue measure on $\mathbb{R}^n$, $\mathscr{B}(\mathbb{R}^n)$ is the Borel σ-algebra, K is any set of finite Lebesgue measure and $\mathscr{L}^n \llcorner K$ is the restriction of the Lebesgue measure to the set K. Indeed, the disintegration of the Lebesgue measure w.r.t. a given partition is determined by the disintegrations of the Lebesgue measure restricted to finite measure sets.

Then, let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function.
We recall that the subdifferential of f at a point $x \in \mathbb{R}^n$ is the set $\partial^- f(x)$ of all $r \in \mathbb{R}^n$ such that

$$f(w) - f(x) \geqslant r \cdot (w - x), \quad \forall w \in \mathbb{R}^n.$$

From the basic theory of convex functions, as f is real-valued and is defined on all $\mathbb{R}^n$, $\partial^- f(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$ and it consists of a single point if and only if f is differentiable at x. Moreover, in that case, $\partial^- f(x) = \{\nabla f(x)\}$, where $\nabla f(x)$ is the differential of f at the point x.
We denote by $\operatorname{dom} \nabla f$ a σ-compact set where f is differentiable and such that $\mathbb{R}^n \setminus \operatorname{dom} \nabla f$ is Lebesgue negligible. $\nabla f : \operatorname{dom} \nabla f \to \mathbb{R}$ denotes the differential map and $\operatorname{Im} \nabla f$ the image of $\operatorname{dom} \nabla f$ with the differential map.

The **partition** of $\mathbb{R}^n$ on which we want to decompose the Lebesgue measure is given by the sets

$$\nabla f^{-1}(y) = \{x \in \mathbb{R}^n : \nabla f(x) = y\}, \quad y \in \operatorname{Im} \nabla f,$$

along with the set $\Sigma^1(f) = \mathbb{R}^n \setminus \operatorname{dom} \nabla f$.
By the convexity of f, we can moreover assume w.l.o.g. that the intersection of $\nabla f^{-1}(y)$ with $\operatorname{dom} \nabla f$ is convex

Since ∇f is a Borel map and $\Sigma^1(f)$ is a $\mathscr{L}^n$-negligible Borel set (see e.g. [AAC], [AA]), we can assume that the quotient map p of Definition 2.6 is given by ∇f and that the quotient space is given by $(\operatorname{Im}\nabla f, \mathscr{B}(\operatorname{Im}\nabla f))$, which is measurably included in $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$.

Then, this partition satisfies the hypothesis of **Theorem 2.7** and there exists a family

$$\{\mu_y\}_{y\in\operatorname{Im}\nabla f}$$

of probability measures on $\mathbb{R}^n$ such that

$$\mathscr{L}^n \llcorner K(B\cap\nabla f^{-1}(A)) = \int_A \mu_y(B)\, d\nabla f_\#(\mathscr{L}^n\llcorner K)(y), \quad \forall A, B\in\mathscr{B}(\mathbb{R}^n).$$

In the following we give the formal definition of **face of a convex function** and relate this object to the sets $\nabla f^{-1}(y)$ of our partition.

Definition 5.1. A *tangent hyperplane* to the graph of a convex function $f:\mathbb{R}^n\to\mathbb{R}$ is a subset of $\mathbb{R}^{n+1}$ of the form

$$H_y = \{(z, h_y(z)) : z\in\mathbb{R}^n, \text{ and } h_y(z) = f(x) + y\cdot(z-x)\}, \tag{5.1}$$

where $x\in\nabla f^{-1}(y)$.

We note that, by convexity, the above definition is independent of $x\in\nabla f^{-1}(y)$.

Definition 5.2. A *face of a convex function* $f:\mathbb{R}^n\to\mathbb{R}$ is a set of the form

$$H_y\cap\operatorname{graph} f_{|\operatorname{dom}\nabla f}. \tag{5.2}$$

It is easy to check that, $\forall\, y\in\operatorname{Im}\nabla f$ and $\forall\, z$ such that $(z, f(z))\in H_y\cap\operatorname{graph} f_{|\operatorname{dom}\nabla f}$, we have that $y = \nabla^{-1} f(z)$.

If we denote by $\pi_{\mathbb{R}^n}:\mathbb{R}^{n+1}\to\mathbb{R}^n$ the projection map on the first n coordinates, one can see that, for all $y\in\operatorname{Im}\nabla f$,

$$\nabla f^{-1}(y) = \pi_{\mathbb{R}^n}(H_y\cap\operatorname{graph} f_{|\operatorname{dom}\nabla f}).$$

For notational convenience, the set $\nabla f^{-1}(y)$ will be denoted as F_y.
We also write F^k_y instead of F_y whenever we want to emphasize the fact that the latter has dimension k, for $k = 0,\dots,n$ (where the dimension of a convex set C is the dimension of its affine hull $\operatorname{aff}(C)$) and we set

$$F^k = \bigcup_{\{y:\,\dim(F_y)=k\}} F_y. \tag{5.3}$$

5.1.1 *Absolute continuity of the conditional probabilities*

Since the measure we are disintegrating ($\mathscr{L}^n$) has the same Hausdorff dimension of the space on which it is concentrated ($\mathbb{R}^n$) and since the sets of the partition on which the conditional probabilities are concentrated have a well defined linear dimension, we address the problem of whether this absolute continuity property of the initial measure is still satisfied by the conditional probabilities produced by the disintegration: we want to see if

$$\dim(F_y) = k \quad \Rightarrow \quad \mu_y \ll \mathscr{H}^k \llcorner F_y. \tag{5.4}$$

The answer to this question is not trivial. Indeed, when $n \geqslant 3$ one can construct sets of full Lebesgue measure in $\mathbb{R}^n$ and Borel partitions of those sets into convex sets such that the conditional probabilities of the corresponding disintegration do not satisfy property (5.4) for $k = 1$ (see e.g. [AKP2]).

However, for the partition given by the faces of a convex function, we show that the absolute continuity property is preserved by the disintegration. Our main result is the following:

Theorem 5.3. *Let $\{\mu_y\}_{y \in \mathrm{Im}\,\nabla f}$ be the family of probability measures on $\mathbb{R}^n$ such that*

$$\mathscr{L}^n \llcorner K(B \cap \nabla f^{-1}(A)) = \int_A \mu_y(B)\, d\nabla f_\#(\mathscr{L}^n \llcorner K)(y), \quad \forall A, B \in \mathscr{B}(\mathbb{R}^n). \tag{5.5}$$

Then, for $\nabla f_\#(\mathscr{L}^n \llcorner K)$-a.e. $y \in \mathrm{Im}\,\nabla f$, the conditional probability μ_y is equivalent to the k-dimensional Hausdorff measure $\mathscr{H}^k$ restricted to $F^k_y \cap K$, i.e.

$$\mu_y \ll \mathscr{H}^k \llcorner (F^k_y \cap K) \quad \textit{and} \quad \mathscr{H}^k \llcorner (F^k_y \cap K) \ll \mu_y. \tag{5.6}$$

Remark 5.4. The result for $k = 0, n$ is trivial. Indeed, for all y such that $F_y \cap K \neq \emptyset$ and $\dim(F_y) = 0$ we must put $\mu_y = \delta_{\{F_y\}}$, where δ_{x_0} is the Dirac mass supported in x_0, whereas if $\dim(F_y \cap K) = n$ we have that $\mu_y = \frac{\mathscr{L}^n \llcorner F_y}{|\mathscr{L}^n \llcorner F_y|}$.

Remark 5.5. Since the map

$$\begin{aligned} \mathrm{id} \times f : & \mathbb{R}^n \to \mathbb{R}^{n+1} \\ & x \mapsto (x, f(x)) \end{aligned}$$

is locally Lipschitz and preserves the Hausdorff dimension of sets, Theorem 5.3 holds also for the disintegration of the $(n+1)$-dimensional Lebesgue measure over the partition of the graph of f given by the faces defined in (5.2). We have chosen to deal with the disintegration of the Lebesgue measure over the projections of the faces on $\mathbb{R}^n$ only for notational convenience.

Theorem 5.3 will be proved in Section 5.8, where we provide also an explicit expression for the conditional probabilities.

If we knew some Lipschitz regularity for the field of directions of the faces of a convex function, we could try to apply the Area or Coarea Formula in order to obtain within a single step the disintegration of the Lebesgue measure and the absolute continuity property (5.6).

However, such regularity is presently not known and for this reason we have to follow a different approach.

5.2 A disintegration technique

In this paragraph we give an outline of the technique we use in order to prove Theorem 5.3.

This kind of strategy was first used in order to disintegrate the Lebesgue measure on a collection of disjoint segments in [BG], and then in [Car2].

For simplicity, we focus on the disintegration of the Lebesgue measure on the 1-dimensional faces and, in the end, we give an idea of how we will extend this technique in order to prove the absolute continuity of the conditional probabilities on the faces of higher dimension.

The disintegration on model sets: Fubini-Tonelli theorem and absolute continuity estimates on affine planes which are transversal to the faces. First of all, let us suppose that the projected 1-dimensional faces of f are given by a collection of disjoint segments $\mathscr{C}$ whose projection on a fixed direction $e \in \mathbf{S}^{n-1}$ is equal to a segment $[h^-e, h^+e]$ with $h^- < 0 < h^+$, more precisely

$$\mathscr{C} = \bigcup_{z \in Z_t} [a(z), b(z)], \tag{5.7}$$

where Z_t is a compact subset of an affine hyperplane of the form $\{x \cdot e = t\}$ for some $t \in \mathbb{R}$ and $a(z) \cdot e = h^-$, $b(z) \cdot e = h^+$. Any set of the form (5.7) will be called a *model set* (see also Figure 3).

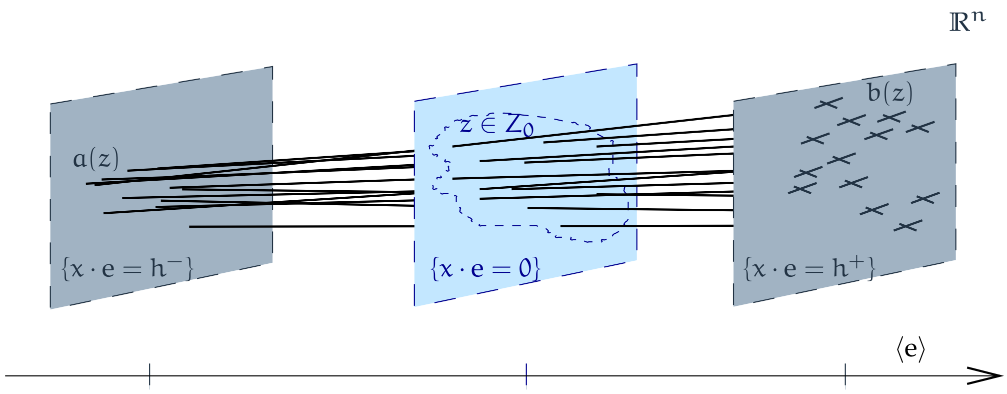


Figure 3: A model set of one dimensional projected faces. Given a subset Z_0 of the hyperplane $\{x \cdot e = 0\}$, the above model set is made of the one dimensional faces of f passing through some $z \in Z_0$, truncated between $\{x \cdot e = h^-\}$, $\{x \cdot e = h^+\}$ and projected on $\mathbb{R}^n$.

We want to find the conditional probabilities of the disintegration of the Lebesgue measure on the segments which are contained in the model set $\mathscr{C}$ and see if they are absolutely continuous w.r.t. the $\mathscr{H}^1$ measure.

The idea of the proof is to obtain the required disintegration by a Fubini-Tonelli argument, that reverts the problem of absolute continuity w.r.t. $\mathscr{H}^1$ of the conditional probabilities on the projected 1-dimensional faces to the absolute continuity w.r.t. $\mathscr{H}^{n-1}$ of the push forward by the flow induced by the directions of the faces of the $\mathscr{H}^{n-1}$-measure on transversal hyperplanes.

First of all, we cut the set $\mathscr{C}$ with the affine hyperplanes which are perpendicular to the segment $[h^-e, h^+e]$, we apply Fubini-Tonelli theorem and we get

$$\int_{\mathscr{C}} \varphi(x)\, d\mathscr{L}^n(x) = \int_{h^-}^{h^+} \int_{\{x\cdot e = t\}\cap\mathscr{C}} \varphi \, d\mathscr{H}^{n-1}\, dt, \quad \forall\, \varphi \in C_c^0(\mathbb{R}^n). \tag{5.8}$$

Then we observe the following: for every $s, t \in [h^-, h^+]$, the points of $\{x \cdot e = t\}\cap\mathscr{C}$ are in bijective correspondence with the points of the section $\{x\cdot e = s\}\cap\mathscr{C}$ and a bijection is obtained by pairing the points that belong to the same segment $[a(z), b(z)]$, for some $z \in Z_t$.
For example, a map which sends the transversal section $Z = \{x\cdot e = 0\}\cap\mathscr{C}$ into the section $Z_t = \{x\cdot e = t\}\cap\mathscr{C}$ (for any $t \in [h^-, h^+]$) is given by

$$\begin{aligned} \sigma^t : & Z \to \sigma^t(Z) = \{x\cdot e = t\}\cap\mathscr{C} \\ & z \mapsto z + t\frac{v_e(z)}{|v_e(z)\cdot e|} = \{x\cdot e = t\}\cap[a(z), b(z)], \end{aligned}$$

where $[a(z), b(z)]$ is the segment of $\mathscr{C}$ passing through the point z and $v_e(z) = \frac{b(z)-a(z)}{|b(z)-a(z)|}$.

Therefore, as soon as we fix a transversal section of $\mathscr{C}$, say for e.g. $Z = \{x\cdot e = 0\}\cap\mathscr{C}$, we can try to rewrite the inner integral in the r.h.s. of (5.8) as an integral of the function $\varphi\circ\sigma^t$ w.r.t. to the $\mathscr{H}^{n-1}$ measure of the fixed section Z.
This can be done if

$$(\sigma^t)^{-1}_{\#}(\mathscr{H}^{n-1} \llcorner \sigma^t(Z)) \ll \mathscr{H}^{n-1} \llcorner Z. \tag{5.9}$$

Indeed,

$$\int_{\sigma^t(Z)} \varphi(y)\, d\mathscr{H}^{n-1}(y) = \int_Z \varphi(\sigma^t(z))\, d(\sigma^t)^{-1}_{\#}(\mathscr{H}^{n-1}\llcorner \sigma^t(Z))(z) \tag{5.10}$$

and if (5.9) is satisfied for all $t \in [h^-, h^+]$, then

$$(5.8) = \int_{h^-}^{h^+} \int_Z \varphi(\sigma^t(z))\, \alpha(t,z)\, d\mathscr{H}^{n-1}(z)\, dt,$$

where $\alpha(t,z)$ is the Radon-Nikodym derivative of $(\sigma^t)^{-1}_{\#}(\mathscr{H}^{n-1}\llcorner \sigma^t(Z))$ w.r.t. $\mathscr{H}^{n-1}\llcorner Z$.

Having turned the r.h.s. of (5.8) into an iterated integral over a product space isomorphic to $Z + [h^{-}e, h^{+}e]$, the final step consists in applying Fubini-Tonelli theorem again so as to exchange the order of the integrals and get

$$\int_{\mathscr{C}} \varphi(x)\, d\mathscr{L}^n(x) = \int_Z \int_{h^-}^{h^+} \varphi(\sigma^t(z))\, \alpha(t,z)\, dt\, d\mathscr{H}^{n-1}(z). \tag{5.11}$$

This final step can be done if α is Borel-measurable and locally integrable in (t,z).

By the uniqueness of the disintegration stated in Theorem 2.7 we have that

$$d\mu_z(t) = \frac{\alpha(t,z)\cdot d\mathscr{H}^1 \llcorner [a(z), b(z)](t)}{\int_{h^-}^{h^+} \alpha(s,z)\, ds}, \quad \text{for } \mathscr{H}^{n-1}\text{-a.e. } z \in Z. \tag{5.12}$$

The same reasoning can be applied to the case $k > 1$. Indeed, let us consider a collection $\mathscr{C}^k$ of k-dimensional faces whose projection on a certain k-plane $\langle e_1, \dots, e_k\rangle$ is given by a rectangle $\prod_{i=1}^{k}[h_i^- e_i, h_i^+ e_i]$, with $h_i^- < 0 < h_i^+$ for all $i = 1, \dots, k$ (see Figure 4).

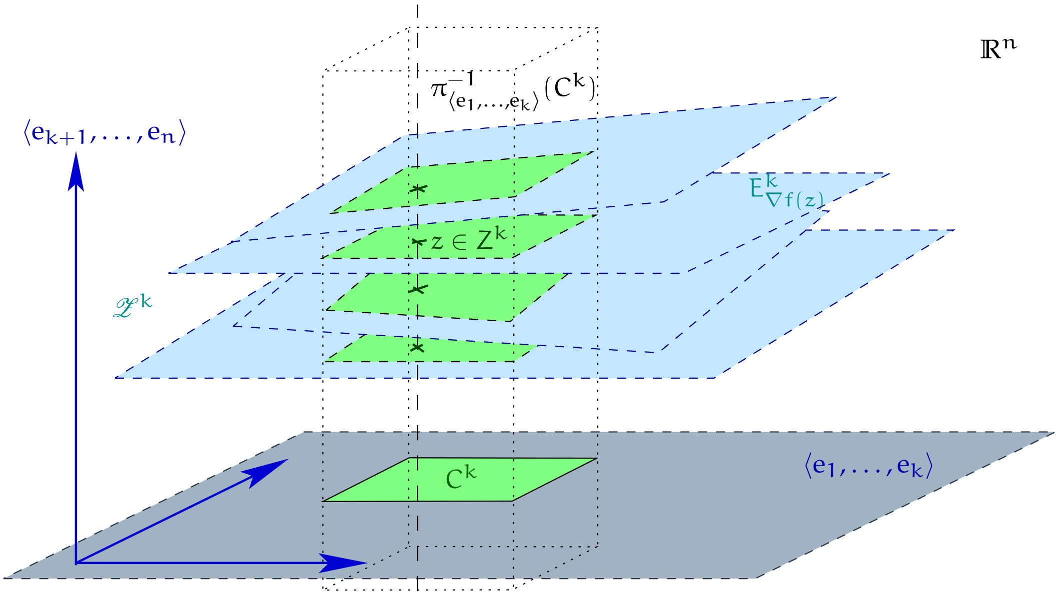


Figure 4: Sheaf sets and $\mathcal{D}$-cylinders (Definitions 5.8, 5.10). Roughly, a sheaf set $\mathscr{Z}^k$ is a collection of k-faces of f, projected on $\mathbb{R}^n$, which intersect exactly at one point some set Z^k contained in a $(n-k)$-dimensional plane. A $\mathcal{D}$-cylinder $\mathscr{C}^k$ is the intersection of a sheaf set with $\pi^{-1}_{\langle e_1,\dots,e_k\rangle}(C^k)$, for some rectangle $C^k = \mathrm{conv}(\{t_i^- e_i, t_i^+ e_i\}_{i=1,\dots,k})$, where $\{e_1, \dots, e_n\}$ are an orthonormal basis of $\mathbb{R}^n$. Such sections Z^k are called basis, while the k-plane $\langle e_1, \dots, e_k\rangle$ is an axis.

Then, as soon as we fix an affine $(n-k)$-dimensional plane which is perpendicular to the k-plane $\langle e_1, \dots, e_k\rangle$, as for example $H^k = \bigcap_{i=1}^{k}\{x \cdot e_i = 0\}$, and we denote by $\pi_{\langle e_1,\dots,e_k\rangle} : \mathbb{R}^n \to \langle e_1,\dots,e_k\rangle$ the projection map on the k-plane $\langle e_1,\dots,e_k\rangle$, the k-dimensional faces in $\mathscr{C}^k$ can be parametrized with the map

$$\sigma^{te}(z) = z + t\frac{\nu_e(z)}{|\pi_{\langle e_1,\dots,e_k\rangle}(\nu_e(z))|}, \tag{5.13}$$

where $z \in Z^k = H^k \cap \mathscr{C}^k$, e is a unit vector in the k-plane $\langle e_1,\dots,e_k\rangle$, $t \in \mathbb{R}$ satisfies $te \cdot e_i \in [h_i^-, h_i^+]$ for all $i = 1,\dots,k$ and $\nu_e(z)$ is the unit direction contained in the face passing through z which is such that $\frac{\pi_{\langle e_1,\dots,e_k\rangle}(\nu_e(z))}{|\pi_{\langle e_1,\dots,e_k\rangle}(\nu_e(z))|} = e$.

If we cut the set $\mathscr{C}^k$ with affine hyperplanes which are perpendicular to e_i for $i = 1,\dots,k$ and apply k-times the Fubini-Tonelli theorem, the main point is again to show that, for every e and t as above,

$$(\sigma^{te})^{-1}_{\#}(\mathscr{H}^{n-k} \llcorner Z^k) \ll \mathscr{H}^{n-k} \llcorner Z^k \tag{5.14}$$

and, after this, that the Radon-Nikodym derivative between the above measures satisfies proper measurability and integrability conditions.

Then, to prove Theorem 5.3 on model sets that are, up to translations and rotations, like the set $\mathscr{C}^k$, it is sufficient to prove (5.14) and some weak properties of the related density function, such as Borel-measurability and local integrability. Actually, the properties of this function will follow immediately from our proof of (5.14), which is given in a stronger form in Lemma 5.37.

Partition of $\mathbb{R}^n$ into model sets and the global disintegration theorem. In the next section we show that the set F^k defined in (5.3), for $k = 1,\dots,n-1$, can be partitioned, up to a negligible set, into a countable collection of Borel-measurable model sets like $\mathscr{C}^k$. After proving the disintegration theorem on the model sets we will see how to glue the "local" results in order to obtain a global disintegration theorem for the Lebesgue measure over the whole faces of the convex function (restricted to a set of $\mathscr{L}^n$-finite measure).

5.3 Measurability of the directions of the k-dimensional faces

The aim of this subsection is to show that the set of the projected k-dimensional faces of a convex function f can be parametrized by a $\mathscr{L}^n$-measurable (and multivalued) map. This will allow us to decompose $\mathbb{R}^n$ into a countable family of Borel model sets on which to prove Theorem 5.3.

First of all we give the following definition, which generalizes Definition 5.1.

Definition 5.6. A *supporting hyperplane* to the graph of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ is an affine hyperplane in $\mathbb{R}^{n+1}$ of the form

$$H = \{w \in \mathbb{R}^{n+1} : w \cdot b = \beta\},$$

where $b \neq 0$, $w \cdot b \leqslant \beta$ for all $w \in \operatorname{epi} f = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : t \geqslant f(x)\}$ and $w \cdot b = \beta$ for at least one $w \in \operatorname{epi} f$. As f is defined and real-valued on all $\mathbb{R}^n$, every supporting hyperplane is of the form

$$H_y = \{(z, h_y(z)) : z \in \mathbb{R}^n, h_y(z) = f(x) + y \cdot (z - x)\}, \tag{5.15}$$

for some $y \in \partial^- f(x)$. Whenever $y \in \operatorname{Im} \nabla f$, H_y is a tangent hyperplane to the graph of f according to Definition 5.1.

Then we define the map

$$x \mapsto \mathcal{P}(x) = \left\{ z \in \mathbb{R}^n : \exists\, y \in \partial^- f(x) \text{ such that } f(z) - f(x) = y \cdot (z - x) \right\}. \quad (5.16)$$

By definition, $\mathcal{P}(x) = \bigcup_{y \in \partial^- f(x)} \pi_{\mathbb{R}^n} (H_y \cap \operatorname{graph}(f))$.

Moreover, the map

$$\operatorname{dom} \nabla f \ni x \mapsto \mathcal{R}(x) := \mathcal{P}(x) \cap \operatorname{dom} \nabla f,$$

gives precisely the set F_y of our partition that passes through the point x.

As the disintegration over the 0-dimensional faces is trivial, we will restrict our attention to the set

$$\mathcal{T} = \{x \in \operatorname{dom} \nabla f : \mathcal{R}(x) \neq \{x\}\}.$$

For all such points there is at least one maximal segment $[w, z] \subset \mathcal{R}(x)$ such that $w \neq z$.

We can also define the multivalued map giving the unit directions contained in the faces passing through the set $\mathcal{T}$, that is

$$\mathcal{T} \ni x \mapsto \mathcal{D}(x) = \left\{ \frac{z - x}{|z - x|} : z \in \mathcal{R}(x),\ z \neq x \right\}. \quad (5.17)$$

We recall that a multivalued map is defined to be Borel measurable if the counter-image of any open set is Borel.
The measurability of the above maps is proved in the following lemma:

Lemma 5.7. *The graph of the multivalued function $\mathcal{P}$ is a closed set in $\mathbb{R}^n \times \mathbb{R}^n$. As a consequence, $\mathcal{P}$, $\mathcal{R}$ and $\mathcal{D}$ are Borel measurable multivalued maps and $\mathcal{T}$ is a Borel set.*

Proof. The closedness of the graph of $\mathcal{P}$ follows immediately from the continuity of f and from the upper-semicontinuity of its subdifferential. Then, the graph of $\mathcal{P}$ is σ-compact in $\mathbb{R}^n \times \mathbb{R}^n$ and, due to the continuity of the projections from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$. It follows then that the map is Borel.

Moreover, since we chose $\operatorname{dom} \nabla f$ to be σ-compact, also the graph of $\mathcal{R}$ is σ-compact, thus $\mathcal{R}$ is a Borel map.

The same reasoning that is made for the map $\mathcal{P}$ can be applied to the multifunction $\mathcal{P} \setminus \mathcal{I}$ (where $\mathcal{I}$ denotes the identity map), thus giving the mesurability of the set $\mathcal{T}$, since

$$\mathcal{T} = \pi\big(\operatorname{graph}(\mathcal{P} \setminus \mathcal{I})\big) \cap \operatorname{dom} \nabla f,$$

where $\pi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ denotes the projection on the first n coordinates.

The measurability of $\mathcal{D}$ follows by the continuity of the map $\mathbb{R}^n \times \mathbb{R}^n \ni (x, z) \mapsto \frac{z-x}{|z-x|}$ out of the diagonal. □

5.4 Partition into model sets

First of all, we introduce some preliminary notation.
If $K \subset \mathbb{R}^d$ is a convex set and $\mathrm{aff}(K)$ is its affine hull, we denote by $\mathrm{ri}(K)$ the relative interior of K, which is the interior of K in the topology of $\mathrm{aff}(K)$, and by $\mathrm{rb}(K)$ its relative boundary, which is the boundary of K in $\mathrm{aff}(K)$.

In order to find a countable partition of F^k into model sets like the set $\mathscr{C}^k$ which was defined in Section 5.2, we have to neglect the points that lie on the relative boundary of the k-dimensional faces.
More precisely, from now onwards we look for the disintegration of the Lebesgue measure over the sets

$$E_y = \mathrm{ri}(F_y), \quad y \in \mathrm{Im}\, \nabla f. \tag{5.18}$$

As we did for the sets F_y, we set

$$E^k_y = E_y, \quad \text{if } \dim(E_y) = k$$

and

$$E^k = \bigcup_{\{y \in \mathrm{Im}\, \nabla f :\, \dim(E_y) = k\}} E^k_y. \tag{5.19}$$

This restriction will not affect the characterization of the conditional probabilities because, as we will prove in Lemma 5.24, the set

$$\mathcal{T} \setminus \bigcup_{k=1}^{n} E^k$$

is Lebesgue negligible.

Now we can start to build the partition of E^k into model sets.

Definition 5.8. For all $k = 1, \dots, n$, we call *sheaf set* a σ-compact subset of E^k of the form

$$\mathscr{Z}^k = \bigcup_{z \in Z^k} \mathrm{ri}(\mathcal{R}(z)), \tag{5.20}$$

where Z^k is a σ-compact subset of E^k which is contained in an affine $(n-k)$-plane in $\mathbb{R}^n$ and is such that

$$\mathrm{ri}(\mathcal{R}(z)) \cap Z^k = \{z\}, \quad \forall z \in Z^k. \tag{5.21}$$

We call *sections* of $\mathscr{Z}^k$ all the sets Y^k that satisfy the same properties of Z^k in the definition.
A *subsheaf* of a sheaf set $\mathscr{Z}^k$ is a sheaf set $\mathscr{W}^k$ of the form

$$\mathscr{W}^k = \bigcup_{w \in W^k} \mathrm{ri}(\mathcal{R}(w)),$$

where W^k is a σ-compact subset of a section of the sheaf set $\mathscr{Z}^k$.

Similarly to Lemma 2.6 in [Car2], we prove that the set E^k can be covered with countably many disjoint sets of the form (5.20).
First of all, let us take a dense sequence $\{V_i\}_{i\in\mathbb{N}} \subset \mathbf{G}(k,n)$, where $\mathbf{G}(k,n)$ is the compact set of all the k-planes in $\mathbb{R}^n$ passing through the origin, and fix, $\forall\, i \in \mathbb{N}$, an orthonormal set $\{e_{i_1},\dots,e_{i_k}\}$ in $\mathbb{R}^n$ such that

$$V_i = \langle e_{i_1},\dots,e_{i_k}\rangle. \tag{5.22}$$

Denoting by $\mathbf{S}^{n-1}\cap V$ the k-dimensional unit sphere of a k-plane $V\subset\mathbb{R}^n$ w.r.t. the Euclidean norm and by $\pi_i = \pi_{V_i} : \mathbb{R}^n \to V_i$ the projection map on the k-plane V_i, for every fixed $0<\varepsilon<1$ the following sets form a disjoint covering of the k-dimensional unit spheres in $\mathbb{R}^n$:

$$\mathbf{S}_i^{k-1} = \big\{\mathbf{S}^{n-1}\cap V :\ V\in\mathbf{G}(k,n),\ \inf_{x\in\mathbf{S}^{n-1}\cap V}|\pi_i(x)| \geqslant 1-\varepsilon\big\}\setminus\bigcup_{j=1}^{i-1}\mathbf{S}_j^{k-1}, \qquad i=1,\dots,I, \tag{5.23}$$

where $I\in\mathbb{N}$ depends on the ε we have chosen.

In order to determine a countable partition of E^k into sheaf sets we consider the k-dimensional rectangles in the k-planes (5.22) whose boundary points have dyadic coordinates. For all

$$l=(l_1,\dots,l_k),\ m=(m_1,\dots,m_k)\in\mathbb{Z}^k \quad \text{with } l_j<m_j\ \ \forall\, j=1,\dots,k \tag{5.24}$$

and for all $i=1,\dots,I,\ p\in\mathbb{N}$, let C^k_{iplm} be the rectangle

$$C^k_{iplm} = 2^{-p}\prod_{j=1}^{k}[l_j\, e_{i_j}, m_j\, e_{i_j}]. \tag{5.25}$$

Lemma 5.9. *The following sets are sheaf sets covering E^k: for $i=1,\dots,I$, $p\in\mathbb{N}$, and $S\subset\mathbb{Z}^k$ take*

$$\mathscr{Z}^k_{ipS} = \Big\{x\in E^k : \mathcal{D}(x)\subset\mathbf{S}_i^{k-1} \text{ and } S\subset\mathbb{Z}^k \text{ is the maximal set such that } \bigcup_{l\in S} C^k_{ipl(l+1)} \subset \pi_i[\mathrm{ri}(\mathcal{R}(x))]\Big\}. \tag{5.26}$$

Moreover, a disjoint family of sheaf sets that cover E^k is obtained in the following way: in case $p=1$ we consider all the sets $\mathscr{Z}^k_{ipS}$ as above, whereas for all $p>1$ we take a set $\mathscr{Z}^k_{ipS}$ if and only if the set $\bigcup_{l\in S} C^k_{ipl(l+1)}$ does not contain any rectangle of the form $C^k_{ip'l(l+1)}$ for every $p'<p$.

As soon as a nonempty sheaf set $\mathscr{Z}^k_{ipS}$ belongs to this partition, it will be denoted by $\mathscr{Z}^k_{ipS}$.

For the proof of this lemma we refer to the analogous Lemma 2.6 in [Car2].

Then, we can refine the partition into sheaf sets by cutting them with sections which are perpendicular to fixed k-planes.

Definition 5.10. (See Figure 4) A k-dimensional $\mathcal{D}$-cylinder is a σ-compact set of the form

$$\mathscr{C}^k = \mathscr{Z}^k \cap \pi^{-1}_{\langle e_1,\dots,e_k\rangle}(C^k), \tag{5.27}$$

where $\mathscr{Z}^k$ is a k-dimensional sheaf set, $\langle e_1,\dots,e_k\rangle$ is any fixed k-dimensional subspace which is perpendicular to a section of $\mathscr{Z}^k$ and C^k is a rectangle in $\langle e_1,\dots,e_k\rangle$ of the form

$$C^k = \prod_{i=1}^{k}[t_i^- e_i, t_i^+ e_i],$$

with $-\infty < t_i^- < t_i^+ < +\infty$ for all $i = 1,\dots,k$, such that

$$C^k \subset \pi_{\langle e_1,\dots,e_k\rangle}[\mathrm{ri}(\mathcal{R}(z))] \quad \forall z \in \mathscr{Z}^k \cap \pi^{-1}_{\langle e_1,\dots,e_k\rangle}(C^k). \tag{5.28}$$

We set $\mathscr{C}^k = \mathscr{C}^k(\mathscr{Z}^k, C^k)$ when we want to refer explicitily to a sheaf set $\mathscr{Z}^k$ and to a rectangle C^k that can be taken in the definition of $\mathscr{C}^k$.
The k-plane $\langle e_1,\dots,e_k\rangle$ is called the *axis* of the $\mathcal{D}$-cylinder and every set Z^k of the form

$$\mathscr{C}^k \cap \pi^{-1}_{\langle e_1,\dots,e_k\rangle}(w), \quad \text{for some } w \in \mathrm{ri}(C^k)$$

is called a *section* of the $\mathcal{D}$-cylinder.

We also define the border of $\mathscr{C}^k$ transversal to $\mathcal{D}$ and its outer unit normal as

$$\begin{aligned} \mathfrak{d}\mathscr{C}^k &= \mathscr{C}^k \cap \pi^{-1}_{\langle e_1,\dots,e_k\rangle}(\mathrm{rb}(C^k)), \\ \hat{n}_{\llcorner \mathfrak{d}\mathscr{C}^k}(x) &= \text{ outer unit normal to } \pi^{-1}_{\langle e_1,\dots,e_k\rangle}(C^k) \text{ at } x, \quad \text{for all } x \in \mathfrak{d}\mathscr{C}^k. \end{aligned} \tag{5.29}$$

Lemma 5.11. *The set E^k can be covered by the $\mathcal{D}$-cylinders*

$$\mathscr{C}^k(\mathscr{Z}^k_{ipS}, C^k_{ipl(l+1)}), \tag{5.30}$$

where $S \subset \mathbb{Z}^k$, $l \in S$ and $\mathscr{Z}^k_{ipS}$, $C^k_{ipl(l+1)}$ are the sets defined in (5.26),(5.25).

Moreover, there exists a countable covering of E^k with $\mathcal{D}$-cylinders of the form (5.30) *such that*

$$\pi_i\left[\mathscr{C}^k(\mathscr{Z}^k_{ipS}, C^k_{ipl(l+1)}) \cap \mathscr{C}^k(\mathscr{Z}^k_{ip'S'}, C^k_{ip'l'(l'+1)})\right] \subset \mathrm{rb}[C^k_{ipl(l+1)}] \cap \mathrm{rb}[C^k_{ip'l'(l'+1)}] \tag{5.31}$$

for any couple of $\mathcal{D}$-cylinders which belong to this countable family (if $i \neq i'$, it follows from the definition of sheaf set that $\mathscr{C}^k(\mathscr{Z}^k_{ipS}, C^k_{ipl(l+1)}) \cap \mathscr{C}^k(\mathscr{Z}^k_{i'p'S'}, C^k_{i'p'l'(l'+1)})$ must be empty).

Proof. The fact that the $\mathcal{D}$-cylinders defined in (5.30) cover E^k follows directly from Definitions 5.8 and 5.10 as in [Car2].

Our aim is then to construct a countable covering of E^k with $\mathcal{D}$-cylinders wich satisfy property (5.31).
First of all, let us fix a nonempty sheaf set $\bar{\mathscr{Z}}^k_{ipS}$ which belongs to the countable partition of E^k given in Lemma 5.9.
In the following we will determine the $\mathcal{D}$-cylinders of the countable covering which are contained in $\bar{\mathscr{Z}}^k_{ipS}$; the others can be selected in the same way starting from a different sheaf set of the partition given in Lemma 5.9.
Then, the $\mathcal{D}$-cylinders that we are going to choose are of the form

$$\mathscr{C}^k\big(\mathscr{Z}^k_{i\hat{p}\hat{S}}, C^k_{i\hat{p}\hat{l}(\hat{l}+1)}\big),$$

where $\mathscr{Z}^k_{i\hat{p}\hat{S}}$ is a subsheaf of the sheaf set $\bar{\mathscr{Z}}^k_{ipS}$.

The construction is done by induction on the natural number $\hat{p}$ which determines the diameter of the squares $C^k_{i\hat{p}\hat{l}(\hat{l}+1)}$ obtained projecting the $\mathcal{D}$-cylinders contained in $\bar{\mathscr{Z}}^k_{ipS}$ on the axis $\langle e_{i_1}, \dots, e_{i_k}\rangle$. Then, as the induction step increases, the diameter of the k-dimensional rectangles associated to the $\mathcal{D}$-cylinders that we are going to add to our countable partition will be smaller and smaller (see Figure 5).

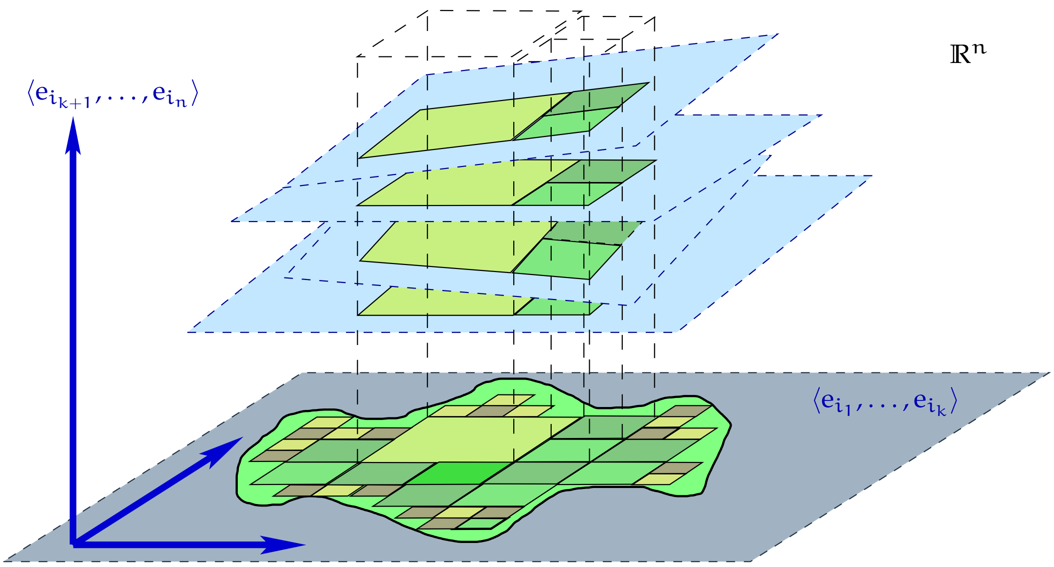


Figure 5: Partition of E^k into $\mathcal{D}$-cylinders (Lemma 5.11).

By definition (5.26) and by the fact that $\bar{\mathscr{Z}}^k_{ipS}$ is a nonempty element of the partition defined in Lemma 5.9, the smallest natural number $\hat{p}$ such that there exists a k-dimensional rectangle of the form $C^k_{i\hat{p}\hat{l}(\hat{l}+1)}$ which is contained in $\pi_i(\bar{\mathscr{Z}}^k_{ipS})$ is exactly p; then, w.l.o.g., we can assume in our induction argument that $p = 1$.

For all $\hat{p} \in \mathbb{N}$, we call $Cyl_{\hat{p}}$ the collection of the $\mathcal{D}$-cylinders which have been chosen up to step $\hat{p}$.

When $\hat{p} = 1$ we set

$$\mathcal{C}yl_1 = \{\mathscr{C}^k(\mathscr{Z}^k_{i1S}, C^k_{i1l(l+1)}) : l \in S\}.$$

Now, let us suppose to have determined the collection of $\mathcal{D}$-cylinders $\mathcal{C}yl_{\hat{p}}$ for some $\hat{p} \in \mathbb{N}$.

Then, we define

$$\begin{aligned}\mathcal{C}yl_{\hat{p}+1} = &\mathcal{C}yl_{\hat{p}} \bigcup \Big\{\mathscr{C}^k = \mathscr{C}^k(\mathscr{Z}^k_{i(\hat{p}+1)\tilde{S}}, C^k_{i(\hat{p}+1)\tilde{l}(\tilde{l}+1)}) : \mathscr{Z}^k_{i(\hat{p}+1)\tilde{S}} \text{subsheaf of } \mathscr{Z}^k_{i\hat{p}S}, \\ &\mathscr{C}^k \nsubseteq \mathscr{C}^k(\mathscr{Z}^k_{ip'S'}, C^k_{ip'l'(l'+1)}) \text{ for all } \mathscr{C}^k(\mathscr{Z}^k_{ip'S'}, C^k_{i'p'l'(l'+1)}) \in \mathcal{C}yl_{\hat{p}}\Big\}.\end{aligned} \tag{5.32}$$

□

As we did in (5.13), any k-dimensional $\mathcal{D}$-cylinder $\mathscr{C}^k = \mathscr{C}^k(\mathscr{Z}^k, C^k)$ can be parametrized in the following way: if we fix $w \in \mathrm{ri}(C^k)$, then

$$\begin{aligned}\mathscr{C}^k = \Big\{&\sigma^{w+te}(z) : z \in Z^k = \pi^{-1}_{\langle e_1,\ldots,e_k\rangle}(w) \cap \mathscr{C}^k,\ e \in \mathbf{S}^{n-1} \cap \langle e_1, \ldots, e_k\rangle \\ &\text{and } t \in \mathbb{R} \text{ is such that } (w+te)\cdot e_j \in [t_j^-, t_j^+] \quad \forall j = 1, \ldots, k\Big\},\end{aligned} \tag{5.33}$$

where

$$\sigma^{w+te}(z) = z + t\frac{v_e(z)}{|\pi_{\langle e_1,\ldots,e_k\rangle}(v_e(z))|}, \tag{5.34}$$

and $v_e(z) \in \mathcal{D}(z)$ is the unit vector such that $\frac{\pi_{\langle e_1,\ldots,e_k\rangle}(v_e(z))}{|\pi_{\langle e_1,\ldots,e_k\rangle}(v_e(z))|} = e$.

We observe that, according to our notation,

$$(\sigma^{w+te})^{-1} = \sigma^{(w+te)-te}. \tag{5.35}$$

5.5 An absolute continuity estimate

According to the strategy outlined in Section 5.2, in order to prove Theorem 5.3 for the disintegration of the Lebesgue measure on the $\mathcal{D}$-cylinders we have to show that, for every $\mathcal{D}$-cylinder $\mathscr{C}^k$ parametrized as in (5.33)

$$(\sigma^{w+te})^{-1}_{\#}(\mathscr{H}^{n-k} \llcorner \sigma^{w+te}(Z^k)) \ll \mathscr{H}^{n-k} \llcorner Z^k. \tag{5.36}$$

This will allow us to make a change of variables between the measure spaces

$$(\sigma^{w+te}(Z^k), \mathscr{H}^{n-k} \llcorner (\sigma^{w+te}(Z^k)) \to (Z^k, \alpha \cdot \mathscr{H}^{n-k} \llcorner Z^k),$$

where α is an integrable function w.r.t. $\mathscr{H}^{n-k} \llcorner Z^k$ (see Section 5.2).

It is clear that the domain of the parameter t, which can be interpreted as a time parameter for a flow σ^{w+te} that moves points along the k-dimensional projected

faces of a convex function, depends on the section Z^k which has been chosen for the parametrization of $\mathscr{C}^k$ and on the direction e.
Then, if $\langle e_1, \dots, e_k\rangle$ is the axis of a $\mathcal{D}$-cylinder $\mathscr{C}^k$, for every $w \in \mathrm{ri}(C^k)$ and for every $e \in \mathbf{S}^{n-1} \cap \langle e_1, \dots, e_k\rangle$, we define the numbers

$$h^-(w,e) = \inf\{t \in \mathbb{R} : w + te \in C^k\}, \quad h^+(w,e) = \sup\{t \in \mathbb{R} : w + te \in C^k\}.$$

We observe that, as $w \in \mathrm{ri}(C^k)$, $h^-(w,e) < 0 < h^+(w,e)$.

We obtain (5.36) in Corollary 5.20 as a consequence of the following fundamental lemma.

Lemma 5.12 (Absolutely continuous push forward). *Let $\mathscr{C}^k$ be a k-dimensional $\mathcal{D}$-cylinder parametrized as in (5.33). Then, for all $S \subset Z^k$ the following estimate holds:*

$$\begin{aligned}\left(\frac{h^+(w,e)-t}{h^+(w,e)-s}\right)^{n-k} \mathscr{H}^{n-k}(\sigma^{w+se}(S)) &\leqslant \mathscr{H}^{n-k}(\sigma^{w+te}(S)) \\ &\leqslant \left(\frac{t-h^-(w,e)}{s-h^-(w,e)}\right)^{n-k} \mathscr{H}^{n-k}(\sigma^{w+se}(S)),\end{aligned} \tag{5.37}$$

where $h^-(w,e) < s \leqslant t < h^+(w,e)$.
Moreover, if $s = h^-(w,e)$ *the left inequality in* (5.37) *still holds and if* $t = h^+(w,e)$ *the right one.*

Lemma 5.12 will be proven at page 71.
The idea to prove this lemma, as in [BG] and [Car2], is to get the estimate (5.37) for the flow σ_j^{w+te} induced by simpler vector fields $\{v_j\}_{j\in\mathbb{N}}$ and then to show that they approximate the initial vector field v_e in such a way that the inequalities in (5.37) pass to the limit.
The main problem in our proof is then to find a suitable sequence of vector fields $\{v_j\}_{j\in\mathbb{N}}$ that approximate, in a certain region, the geometry of the projected k-dimensional faces of a convex function in the direction e, which is described by the vector field v_e.
For the construction of this family of vector fields we strongly rely on the fact that the sets on which we want to disintegrate the Lebesgue measure are, other than disjoint, the projections of the k-dimensional faces of a convex function.

For simplicity, we first prove the estimate (5.37) for 1-dimensional $\mathcal{D}$-cylinders. In this case, if $\langle e\rangle$ is the axis of a 1-dimensional $\mathcal{D}$-cylinder $\mathscr{C}$, there are only two possible directions $\pm e$ that can be chosen to parametrize it. Up to translations by a multiple of the same vector, we can assume that $w = 0$. Moreover, since choosing $-e$ instead of e in the definition of the parametrization map (5.34) simply reverses the order of s and t in (5.37), in order to prove (5.37) it is sufficient to show that, for all $0 \leqslant t \leqslant h^+$ and for all $S \subset \sigma^t(Z)$

$$\mathscr{H}^{n-1}(S) \leqslant \left(\frac{t-h^-}{-h^-}\right)^{n-1} \mathscr{H}^{n-1}((\sigma^t)^{-1}(S)), \tag{5.38}$$

where $\sigma^t = \sigma^{0+te}$ and $h^\pm = h^\pm(0,e)$.

In our construction we first approximate the 1-dimensional faces that lie on the graph of f restricted to the given $\mathcal{D}$-cylinder and then we get the approximating vector fields $\{v_j\}_{j\in\mathbb{N}}$ simply projecting the directions of those approximations on the first n coordinates.

Before giving the details we recall and introduce some useful notation:

$$\mathbf{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\};$$

$$e \in \mathbf{S}^{n-1} \quad \text{a fixed vector;}$$

$$H_t := \{x \in \mathbb{R}^n : x \cdot e = t\}, \quad \text{where} \quad t \in [h^-, h^+] \quad \text{and} \quad h^-, h^+ \in \mathbb{R} : h^- < 0 < h^+;$$

$$\mathbf{B}^{n-1}_R(x) = \{z \in H_{\{x\cdot e\}} : |z - x| \leqslant R\};$$

$$Z \subset Z_0 \quad \sigma\text{-compact section of the 1-dimensional } \mathcal{D}\text{-cylinder } \mathscr{C};$$

$$v_e(x) \in \mathcal{D}(x) \quad \text{is the unit vector such that} \quad \pi_e(v_e(x)) = |\pi_e(v_e(x))|\, e,\ \forall\, x \in \mathscr{C};$$

$$\mathscr{C} = \{\sigma^t(z) : z \in Z,\ t \in [h^-, h^+]\}, \quad \sigma^t(z) = z + t\frac{v_e(z)}{|\pi_e(v_e(z))|};$$

$$\mathscr{C}_t = \bigcup_{s\in[h^-,t]} H_s \cap \mathscr{C};$$

$$l_t(x) = \mathcal{R}(x) \cap \mathscr{C}_t, \quad \forall\, x \in \mathscr{C}_t;$$

$$\forall\, x \in \mathbb{R}^n, \quad \tilde{x} := (x, f(x)) \in \mathbb{R}^{n+1} \quad \text{and} \quad \forall\, A \subset \mathbb{R}^n, \quad \tilde{A} := \text{graph } f_{|_A}.$$

Moreover, we recall the following definitions:

Definition 5.13. The *convex envelope* of a set of points $X \subset \mathbb{R}^n$ is the smaller convex set $\mathrm{conv}(X)$ that contains X. The following characterization holds:

$$\mathrm{conv}(X) = \left\{ \sum_{j=1}^{J} \lambda_j\, x_j : x_j \in X,\ 0 \leqslant \lambda_j \leqslant 1,\ \sum_{j=1}^{J} \lambda_j = 1,\ J \in \mathbb{N} \right\}. \tag{5.39}$$

Definition 5.14. The *graph of a compact convex set* $C \subset \mathbb{R}^{n+1}$, that we denote by $\mathrm{graph}(C)$, is the graph of the function $g : \pi_{\mathbb{R}^n}(C) \to \mathbb{R}$ which is defined by

$$g(x) = \min\{t \in \mathbb{R} : (x,t) \in C\}. \tag{5.40}$$

Definition 5.15. A *supporting* k*-plane* to the graph of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ is an affine k-dimensional subspace of a supporting hyperplane to the graph of f (see Definition 5.6) whose intersection with graph f is nonempty.

Definition 5.16. An R-face of a convex set $C \subset \mathbb{R}^d$ is a convex subset C' of C such that every closed segment in C with a relative interior point in C' has both endpoints in C'. The zero-dimensional R-faces of a convex set are also called *extreme points* and the set of all extreme points in a convex set C will be denoted by $\mathrm{ext}(C)$.

The definition of R-face corresponds to the definition of face of a convex set in [Roc].

We also recall the following propositions, for which we refer to Section 18 of [Roc].

Proposition 5.17. *Let* $C = \mathrm{conv}(D)$*, where* D *is a set of points in* $\mathbb{R}^d$*, and let* C' *be a nonempty* R*-face of* C*. Then* $C' = \mathrm{conv}(D')$*, where* D' *consists of the points in* D *which belong to* C'*.*

Proposition 5.18. *Let* C *be a bounded closed convex set. Then* $C = \mathrm{conv}(\mathrm{ext}(C))$*.*

The key to get fundamental estimate (5.38) is contained in the following lemma:

Lemma 5.19 (Construction of regular approximating vector fields)**.** *For all* $0 \leqslant t \leqslant h^+$*, there exists a sequence of* $\mathscr{H}^{n-1}$*-measurable vector fields*

$$\{v_j^t\}_{j\in\mathbb{N}}, \quad v_j^t : \sigma^t(Z) \to \mathbf{S}^{n-1}$$

such that

1. v_j^t *converges* $\mathscr{H}^{n-1}$*-a.e. to* v_e *on* $\sigma^t(Z)$*;* (5.41)
2. $$\mathscr{H}^{n-1}(S) \leqslant \left(\frac{t-h^-}{-h^-}\right)^{n-1} \mathscr{H}^{n-1}((\sigma^t_{v_j^t})^{-1}(S)), \quad \forall S \subset \sigma^t(Z),$$
   where $\sigma^t_{v_j^t}$ *is the flow map associated to the vector field* v_j^t*. .* (5.42)

Indeed, if we have such a sequence of vector fields, the proof of the estimate (5.38) follows as in [Car2].

Proof. Step 1: Preliminary considerations.

First of all, let us fix $t \in [0, h^+]$.

Eventually partitioning $\mathscr{C}$ into a countable collection of sets, we can assume that $\sigma^t(Z)$ and $\sigma^{h^-}(Z)$ are bounded, with $\sigma^t(Z) \subset \mathbf{B}^{n-1}_{R_1}(x_1) \subset H_t$ and $\sigma^{h^-}(Z) \subset \mathbf{B}^{n-1}_{R_2}(x_2) \subset H_{h^-}$. Then, if we call K_t the convex envelope of $\mathbf{B}^{n-1}_{R_1}(x_1) \cup \mathbf{B}^{n-1}_{R_2}(x_2)$, the function $f_{|K_t}$ is uniformly Lipschitz with a certain Lipschitz constant L_f.

Step 2: Construction of approximating functions. (see Figure 6)

Now we define a sequence of functions $\{f_j\}_{j\in\mathbb{N}}$ whose 1-dimensional faces approximate, in a certain sense, the pieces of the 1-dimensional faces of f which are contained in $\mathscr{C}_t$. The directions of a properly chosen subcollection of the 1-dimensional faces of f_j will give, when projected on the first n coordinates, the approximate vector field v_j^t.

First of all, take a sequence $\{\tilde{y}_i\}_{i\in\mathbb{N}} \subset \tilde{\sigma}^{h^-}(Z)$ such that the collection of segments $\{\tilde{l}_t(y_i)\}_{i\in\mathbb{N}}$ is dense in $\bigcup_{y\in\sigma^{h^-}(Z)} \tilde{l}_t(y)$.

For all $j \in \mathbb{N}$, let C_j be the convex envelope of the set

$$\{\tilde{y}_i\}_{i=1}^j \quad \cup \quad \mathrm{graph}\, f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}} \tag{5.43}$$

and call $f_j : \pi_{\mathbb{R}^n}(C_j) \to \mathbb{R}$ the function whose graph is the graph of the convex set C_j.

We note that $\pi_{\mathbb{R}^n}(C_j) \cap H_{h^-} = \mathrm{conv}(\{y_i\}_{i=1}^j)$ and

$$\mathrm{graph}\, f_{j|_{\mathrm{conv}(\{y_i\}_{i=1}^j)}} = \mathrm{graph}(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j)).$$

We claim that the graph of f_j is made of segments that connect the points of $\mathrm{graph}(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j))$ to the graph of $f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ (indeed, by convexity and by the fact that $\tilde{y}_i = (y_i, f(y_i))$, $f_j = f$ on $\mathbf{B}^{n-1}_{R_1}(x_1)$).
In order to prove this, we first observe that, by definition, all segments of this kind are contained in the set C_j. On the other hand, by (5.39), all the points in C_j are of the form

$$w = \sum_{i=1}^{J} \lambda_i w_i, \tag{5.44}$$

where $\sum_{i=1}^{J} \lambda_i = 1$, $0 \leqslant \lambda_i \leqslant 1$ and $w_i \in \{\tilde{y}_i\}_{i=1}^j \cup \mathrm{graph}\, f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$. In particular, we can write

$$w = \alpha z + (1-\alpha) r, \quad \text{where } 0 \leqslant \alpha \leqslant 1, \quad z \in \mathrm{conv}\big(\{\tilde{y}_i\}_{i=1}^j\big) \quad \text{and } r \in \mathrm{epi}\, f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}. \tag{5.45}$$

Moreover, if we take two points $z' \in \mathrm{graph}(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j))$, $r' \in \mathrm{graph}\, f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ such that $\pi_{\mathbb{R}^n}(z') = \pi_{\mathbb{R}^n}(z)$ and $\pi_{\mathbb{R}^n}(r') = \pi_{\mathbb{R}^n}(r)$, we have that the point

$$w' = \alpha z' + (1-\alpha) r' \tag{5.46}$$

belongs to C_j, lies on a segment which connects $\mathrm{graph}(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j))$ to $\mathrm{graph}\, f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ and its $(n+1)$ coordinate is less than the $(n+1)$ coordinate of w.

The graph of f_j contains also all the pieces of 1-dimensional faces $\{\tilde{l}_t(y_i)\}_{i=1}^j$, since by construction it contains their endpoints and it lies over the graph of $f_{|_{\pi_{\mathbb{R}^n}(C_j)}}$.

Step 3: Construction of approximating vector fields. (see Figure 6)

Among all the segments in the graph of f_j that connect the points of $\mathrm{graph}(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j))$ to the graph of $f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$, we select those of the form $[\tilde{x}, \tilde{y_k}]$, where $x \in \sigma^t(Z)$, $y_k \in \{y_i\}_{i=1}^j$, and we show that for $\mathscr{H}^{n-1}$-a.e. $x \in \sigma^t(Z)$ there exists only one segment within this class which passes through $\tilde{x}$. The approximating vector field will be given by the projection on the first n coordinates of the directions of these segments.

First of all, we claim that for all $x \in \mathbf{B}^{n-1}_{R_1}(x_1)$ the graph of f_j contains at least a segment of the form $[\tilde{x}, \tilde{y}_i]$ for some $i \in \{1, \dots, j\}$.
Indeed, we show that if $\tilde{x}$ is the endpoint of a segment of the form $[\tilde{x}, (y, f_j(y))]$ where y belongs to $\mathrm{conv}\big(\{y_i\}_{i=1}^j\big)$ but $(y, f_j(y)) \notin \mathrm{ext}\big(\mathrm{conv}\big(\{y_i\}_{i=1}^j\big)\big)$, then there are at least two segments of the form $[\tilde{x}, \tilde{y}_k]$ with $\tilde{y_k} \in \mathrm{ext}(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j)) \subset \{\tilde{y}_i\}_{i=1}^j$ (here we assume that $j \geqslant 2$).
In order to prove this, take a point $(z, f_j(z))$ in the open segment $(\tilde{x}, (y, f_j(y)))$ and a supporting hyperplane $H(z)$ to the graph of f_j that contains that point. By definition, $H(z)$ contains the whole segment $[\tilde{x}, (y, f_j(y))]$ and the set $H(z) \cap \big(H_{h^-} \times$

$\mathbb{R}$) is a supporting hyperplane to the set graph$(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j))$ that contains the point $(y, f_j(y))$.
Now, take the smallest R-face C of $\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j))$ which is contained in

$$\mathrm{graph}(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j)))$$

and contains the point $(y, f_j(y))$, that is given by the intersection of all R-faces which contain $(y, f_j(y))$.
By Propositions 5.17 and 5.18, $C = \mathrm{conv}\left[\mathrm{ext}\left(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j)\right) \cap C\right]$ and as $(y, f_j(y))$ does not belong to $\mathrm{ext}\left(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j)\right)$, $\dim(C) \geqslant 1$ and the set $\mathrm{ext}\left(\mathrm{conv}(\{\tilde{y}_i\}_{i=1}^j)\right) \cap C$ contains at least two points $\tilde{y}_k, \tilde{y}_l$.
In particular, since both C and $\tilde{x}$ belong to $H(z) \cap \mathrm{graph}(f_j)$, by definition of supporting hyperplane we have that the graph of f_j contains the segments $[\tilde{x}, \tilde{y}_k]$, $[\tilde{x}, \tilde{y}_l]$ and our claim is proved.

Now, for each $j \in \mathbb{N}$, we define the (possibly multivalued) map $\mathcal{D}_j^t : \mathbf{B}_{R_1}^{n-1}(x_1) \to \mathbb{R}^n$ as follows:

$$\mathcal{D}_j^t : x \mapsto \left\{ \frac{y_i - x}{|y_i - x|} : [\tilde{x}, \tilde{y_i}] \subset \mathrm{graph}(f_j) \right\} \tag{5.47}$$

and we prove that the set

$$B_j := \sigma^t(Z) \cap \{x \in \mathbf{B}_{R_1}^{n-1}(x_1) : \mathcal{D}_j^t(x) \text{ is multivalued } \} \tag{5.48}$$

is $\mathscr{H}^{n-1}$–negligible, $\forall j \in \mathbb{N}$.
Thus, if we neglect the set $B = \bigcup_{j \in \mathbb{N}} B_j$, we can define our approximating vector field as

$$v_j^t(x) = \{\mathcal{D}_j^t(x)\}, \quad \forall x \in \sigma^t(Z) \backslash B, \quad \forall j \in \mathbb{N}. \tag{5.49}$$

In order to show that $\mathscr{H}^{n-1}(B_j) = 0$ we first prove that, for $\mathscr{H}^{n-1}$-a.e. $x \in \mathbf{B}_{R_1}^{n-1}(x_1)$, whenever $\mathcal{D}_j^t(x)$ contains the directions of two segments, f_j must be linear on their convex envelope.
Indeed, suppose that the graph of f_j contains two segments $[\tilde{x}, \tilde{y}_{i_k}]$, where $i_k \in \{1, \dots, j\}$ and $k = 1, 2$, and consider two points $(z_k, f_j(z_k)) \subset [\tilde{x}, \tilde{y}_{i_k}]$ such that

$$\begin{aligned} z_1 &= x + se + a_1 v_1, \quad s \in [h^- - t, 0),\ v_1 \in H_0; \\ z_2 &= x + se + a_2 v_2, \quad s \in [h^- - t, 0),\ v_2 \in H_0. \end{aligned} \tag{5.50}$$

As f_j is linear on $[x, y_{i_k}]$, we have that

$$f_j(z_k) = f_j(x) + r_k \cdot (se + a_k v_k), \tag{5.51}$$

where $r_k \in \partial^- f_j(x), \quad k = 1, 2$.
Moreover, since

$$\pi_{H_0}(\partial^- f_j(x)) = \partial^- f_{|_{\mathbf{B}_{R_1}^{n-1}(x_1)}}(x) \tag{5.52}$$

and the set where $\partial^- f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ is multivalued is $\mathscr{H}^{n-2}$-rectifiable (see for e.g. [Zaj, AA]), we have that, for $\mathscr{H}^{n-1}$-a.e. $x \in \mathbf{B}^{n-1}_{R_1}(x_1)$

$$r \cdot v = \nabla(f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}})(x) \cdot v, \quad \forall\, r \in \partial^- f_j(x),\ \forall\, v \in H_0. \tag{5.53}$$

Then, if we put $w = \nabla(f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}})(x)$, (5.51) becomes

$$f_j(z_k) = f_j(x) + r_k \cdot se + w \cdot a_k v_k. \tag{5.54}$$

If $z_\lambda = (1-\lambda)z_1 + \lambda z_2$, we have that

$$\begin{aligned} f_j(z_\lambda) &\leqslant (1-\lambda)f_j(z_1) + \lambda f_j(z_2) \\ &\overset{(5.54)}{=} f_j(x) + s((1-\lambda)r_1 + \lambda r_2) \cdot e + w \cdot ((1-\lambda)a_1 v_1 + \lambda a_2 v_2). \end{aligned} \tag{5.55}$$

As $((1-\lambda)r_1 + \lambda r_2) \in \partial^- f_j(x)$, we also obtain that

$$\begin{aligned} f_j(z_\lambda) &\geqslant f_j(x) + s((1-\lambda)r_1 + \lambda r_2) \cdot e \\ &\quad + ((1-\lambda)r_1 + \lambda r_2) \cdot ((1-\lambda)a_1 v_1 + \lambda a_2 v_2) = \\ &= f_j(x) + s((1-\lambda)r_1 + \lambda r_2) \cdot e + w \cdot ((1-\lambda)a_1 v_1 + \lambda a_2 v_2) = \\ &\overset{(5.54)}{=} (1-\lambda)f_j(z_1) + \lambda f_j(z_2). \end{aligned} \tag{5.56}$$

Thus, we have that $f_j((1-\lambda)z_1 + \lambda z_2) = (1-\lambda)f_j(z_1) + \lambda f_j(z_2)$ and our claim is proved.

In particular, there exists a supporting hyperplane to the graph of f_j which contains the affine hull of the convex envelope of $\{[\tilde{x}, \tilde{y}_{i_k}]\}_{k=1,2}$ and then this affine hull must intersect $H_t \times \mathbb{R}$ into a supporting line to the graph of $f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ which is parallel to the segment $[\tilde{y}_{i_1}, \tilde{y}_{i_2}]$.

Thus, if all the supporting lines to the graph of $f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ which are parallel to a segment $[\tilde{y}_k, \tilde{y}_m]$ (with $k, m \in \{1, \dots, j\}$, $k \neq m$) are parametrized as

$$l_{k,m} + w, \tag{5.57}$$

where $l_{k,m}$ is the linear subspace of $\mathbb{R}^{n+1}$ which is parallel to $[\tilde{y}_k, \tilde{y}_m]$ and $w \in W_{k,m} \subset H_t \times \mathbb{R}$ is perpendicular to $l_{k,m}$, we have that

$$B_j = \sigma^t(Z) \cap \Bigg[\bigcup_{\substack{k,m \in \{1,\dots,j\} \\ k<m}} \ \bigcup_{w \in W_{k,m}} \pi_{\mathbb{R}^n}(l_{k,m} + w) \Bigg]. \tag{5.58}$$

By this characterization of the set B_j and by Fubini theorem on H_t w.r.t. the partition given by the lines which are parallel to $\pi_{\mathbb{R}^n}(l_{k,m})$ for every k and m, in order to show that $\mathscr{H}^{n-1}(B_j) = 0$ it is sufficient to prove that, $\forall\, w \in W_{k,m}$,

$$\mathscr{H}^{n-1}(\sigma^t(Z) \cap \pi_{\mathbb{R}^n}(l_{k,m} + w)) = 0. \tag{5.59}$$

Finally, (5.59) follows from the fact that a supporting line to the graph of $f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ cannot contain two distinct points of $\tilde{\sigma}^t(Z)$, because otherwise they would be contained in a higher dimensional face of the grapf of f contraddicting the definition of $\tilde{\sigma}^t(Z)$.

Then, the vector field defined in (5.49) is defined $\mathscr{H}^{n-1}$-a.e..

Step 4: Convergence of the approximating vector fields

Here we prove the convergence property of the vector field defined in (5.49) as stated in (5.41).

This result is obtained as a consequence of the uniform convergence of the approximating functions f_j to the function $\hat{f}$ which is the graph of the set

$$\hat{C} = \mathrm{conv}\big(\{\tilde{l}_t(y_i)\}_{i\in\mathbb{N}}\big). \tag{5.60}$$

First of all we observe that, since $C_j \nearrow \hat{C}$,

$$\mathrm{dom}\, f_j = \pi_{\mathbb{R}^n}(C_j) \nearrow \mathrm{dom}\, \hat{f} = \pi_{\mathbb{R}^n}(\hat{C}) \quad \text{and} \quad f_j(x) \searrow \hat{f}(x) \quad \forall\, x \in \mathrm{ri}(\pi_{\mathbb{R}^n}(\hat{C})), \tag{5.61}$$

where $f_j(x)$ is defined $\forall\, j \geqslant j_0$ such that $x \in \pi_{\mathbb{R}^n}(C_{j_0})$.

In order to prove that $f_j(x) \searrow \hat{f}(x)$ uniformly, we show that the functions f_j are uniformly Lipschitz on their domain, with uniformly bounded Lipschitz constants. We recall that the graph of f_j is made of segments that connect the points of graph $f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ to the points of graph$(\mathrm{conv}\big(\{\tilde{y}_i\}_{i=1}^j\big))$.

In order to find and upper bound for the incremental ratios between points $z, w \in \mathrm{dom}\, f_j$, we distinguish two cases.

<u>Case 1</u>: $[z,w] \subset [x, y_k]$, where $x \in \mathbf{B}^{n-1}_{R_1}(x_1)$, $y_k \in \{y_i\}_{i=1}^j$ and $[\tilde{x}, \tilde{y}_k] \subset \mathrm{graph}(f_j)$.
In this case we have that

$$\frac{|f_j(z) - f_j(w)|}{|z-w|} = \frac{|f_j(x) - f_j(y_k)|}{|x-y_k|} = \frac{|f(x) - f(y_k)|}{|x-y_k|} \leqslant L_f, \tag{5.62}$$

where L_f is tha Lipschitz constant of f on K_t.

<u>Case 2</u>: Otherwise we observe that, since f_j is convex,

$$|f_j(z) - f_j(w)| \leqslant \sup_{r \in \partial^- f_j(z) \cup \partial^- f_j(w)} |r \cdot (z-w)|. \tag{5.63}$$

Let then $r \in \partial^- f_j(z) \cup \partial^- f_j(w)$ be a maximizer of the r.h.s. of (5.63) and let us suppose, without loss of generality, that $r \in \partial^- f_j(z)$. If $x \in \mathbf{B}^{n-1}_{R_1}(x_1)$ is such that $(z, f_j(z)) \subset [(y, f_j(y)), \tilde{x}] \subset \mathrm{graph}(f_j)$ for some $y \in \mathrm{conv}(\{y_i\}_{i=1}^j)$, we have the following unique decomposition

$$w - z = \beta_j(z,w)\left(\frac{x-z}{|x-z|}\right) + \gamma_j(z,w) q, \tag{5.64}$$

where $q \in \mathbf{S}^{n-1} \cap H_0$ and $\beta_j(z,w)$, $\gamma_j(z,w) \in \mathbb{R}$.

Then,

$$r \cdot (w - z) = \beta_j(z, w)\left(r \cdot \frac{x - z}{|x - z|}\right) + \gamma_j(z, w)(r \cdot q). \tag{5.65}$$

The first scalar product in (5.65) can be estimated as in Case 1.

As for the second term, we note that the supporting hyperplane to the graph of f_j given by the graph of the affine function $h(p) = f_j(z) + r \cdot (p - z)$ contains the segment $[(z, f_j(z)), \tilde{x}]$ and its intersection with the hyperplane $H_t \times \mathbb{R}$ is given by a supporting hyperplane to the graph of $f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$ which contains the point $\tilde{x}$.

Moreover, as $q \in H_0$, we have that

$$r \cdot q = \pi_{H_0}(r) \cdot q, \tag{5.66}$$

and we know that $\pi_{H_0}(r) \in \partial^- f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}(x)$.

By definition of subdifferential, for all $s \in \partial^- f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}(x)$ and for all $\lambda > 0$ such that $x + \lambda q$, $x - \lambda q \in \mathbf{B}^{n-1}_{R_1}(x_1)$,

$$\frac{f(x) - f(x - \lambda q)}{\lambda} \leqslant s \cdot q \leqslant \frac{f(x + \lambda q) - f(x)}{\lambda} \tag{5.67}$$

and so the term $|r \cdot q|$ is bounded from above by the Lipschitz constant of f.
As the scalar products $\beta_j(z, w), \gamma_j(z, w)$ are uniforlmly bounded w.r.t. j on $\operatorname{dom} f_j \subset \operatorname{dom} \hat{f}$, we conclude that the functions $\{f_j\}_{j \in \mathbb{N}}$ are uniformly Lipschitz on the sets $\{\operatorname{dom} f_j\}_{j \in \mathbb{N}}$ and their Lipschitz constants are uniformly bounded by some positive constant $\hat{L}$.
If we call $\hat{f}_j$ a Lipschitz extension of f_j to the set $\operatorname{dom} \hat{f}$ which has the same Lipschitz constant (Mac Shane lemma), by Ascoli-Arzelá theorem we have that

$$\hat{f}_j \to \hat{f} \quad \text{uniformly on } \operatorname{dom} \hat{f}.$$

Now we prove that, for $\mathscr{H}^{n-1}$-a.e. $x \in \sigma^t(Z) \backslash B$, $v^t_j(x) \to v_e(x)$.
Given a point $x \in \sigma^t(Z) \backslash B$, we call $\tilde{y}_{j(x)}$, where $j \in \mathbb{N}$, the unique point $\tilde{y}_k \in \{\tilde{y}_i\}_{i=1}^j$ such that

$$v^t_j(x) = \frac{y_k - x}{|y_k - x|}.$$

By compactness of $\operatorname{graph}(\operatorname{conv}(\{\tilde{y}_i\}_{i \in \mathbb{N}}))$, there is a subsequence $\{j_n\}_{n \in \mathbb{N}} \subset \mathbb{N}$ such that

$$\tilde{y}_{j_n(x)} \to \hat{y} \in \operatorname{graph} f,$$

hence

$$v^t_{j_n}(x) \to \hat{v} = \frac{\hat{y} - x}{|\hat{y} - x|}.$$

As the functions f_j converge to $\hat{f}$ uniformly, the point $\hat{y}$ and the whole segment $[\tilde{x}, \hat{y}]$ belong to the graph of $\hat{f}$.

So, there are two segments $\tilde{l}_t(x)$ and $[\tilde{x}, \hat{y}]$ which belong to the graph of $\hat{f}$ and pass through the point $\tilde{x}$.

Since $\hat{f}_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}} = f_{|_{\mathbf{B}^{n-1}_{R_1}(x_1)}}$, we can apply the same reasoning we made in order to prove that the set (5.48) was $\mathscr{H}^{n-1}$-negligible to conclude that the set

$$\sigma^t(Z) \cap \Big\{ x \in \mathbf{B}^{n-1}_{R_1}(x_1) : \exists \text{ more than two segments in the graph of } \hat{f} \\ \text{that connect } \tilde{x} \text{ to a point of graph}(\operatorname{conv}(\{\tilde{y}_i\}_{i\in\mathbb{N}})) \Big\}$$

has zero $\mathscr{H}^{n-1}$-measure.

Then, $[\tilde{x}, \hat{y}] = \tilde{l}_t(x)$ and $\hat{v} = v_e(x)$ for $\mathscr{H}^{n-1}$-a.e. $x \in \sigma^t(Z)$, so that property (5.41) is proved.

Step 5: Proof of the estimate (5.42). (see Figure 7)

The estimate for the map $\sigma^t_{v^t_j}$ induced by the approximating vector fields v^t_j follows as in [BG] and [Car2] from the fact that the collection of segments with directions given by v^t_j and endpoints in $\operatorname{dom} v^t_j$, $\sigma^{h^-}(Z)$ form a finite union of cones with bases in $\operatorname{dom} v^t_j$ and vertex in $\{y_i\}_{i=1}^j$.

Indeed, if we define the sets

$$\Omega_{ij} = \left\{ x \in \sigma^t(Z) : \mathcal{D}^t_j(x) = \{v^t_j(x)\} \text{ and } v^t_j(x) = \frac{y_i - x}{|y_i - x|} \right\}, \quad j \in \mathbb{N}, \quad i = 1, \dots, j, \tag{5.68}$$

for all $S \subset \sigma^t(Z) \setminus B$ we have that

$$\mathscr{H}^{n-1}(S) = \sum_{i=1}^{j} \mathscr{H}^{n-1}(S \cap \Omega_{ij})$$

and

$$\mathscr{H}^{n-1}((\sigma^t_{v^t_j})^{-1}(S)) = \sum_{i=1}^{j} \mathscr{H}^{n-1}((\sigma^t_{v^t_j})^{-1}(S \cap \Omega_{ij})).$$

Then it is sufficient to prove (5.42) when the vector field v^t_j is defined as

$$v^t_j(x) = \frac{y_i - x}{|y_i - x|}.$$

After these preliminary considerations, (5.42) follows from the fact that the set

$$\bigcup_{s \in [0, t-h^-]} \sigma^{-s}_{v^t_j}(S) \tag{5.69}$$

is a cone with with base $S \subset H_t$ and vertex $y_i \in H_{h^-}$ and $\sigma^{-t}_{v^t_j}(S)$ is the intersection of this cone with the hyperplane H_0. $\square$

Proof of Lemma 5.12. Given a k-dimensional $\mathcal{D}$-cylinder $\mathscr{C}^k$ parametrized as in (5.33), the collection of segments

$$\bigcup_{z\in Z^k}\{\sigma^{w+te}(z) : t \in [h^-(w,e), h^+(w,e)]\} \tag{5.70}$$

is a 1-dimensional $\mathcal{D}$-cylinder of the convex function f restricted to the $(n-k+1)$-dimensional set

$$\pi^{-1}_{\langle e_1,\dots,e_k\rangle}(\{w+te:\ t\in[h^-(w,e),h^+(w,e)]\}). \tag{5.71}$$

Then, as in Lemma 5.19, we can construct a sequence of approximating vector fields also for the directions of the segments (5.70). The only difference with respect to the approximation of the 1-dimensional faces of f is that the domain of the approximating vector fields will be a subset of an $(n-k)$-dimensional affine plane of the form $\pi^{-1}_{\langle e_1,\dots,e_k\rangle}(w)$ and so the measure involved in the estimate (5.38) will be $\mathscr{H}^{n-k}$ instead of $\mathscr{H}^{n-1}$. Finally, we pass to the limit as with the approximating vector fields given in Lemma 5.19 and we obtain the fundamental estimate (5.37) for the k-dimensional $\mathcal{D}$-cylinders. □

5.6 Properties of the density function

In this subsection, we show that the quantitative estimates of Lemma 5.19 allow not only to derive the absolute continuity of the push forward with σ^{w+te}, but also to find regularity estimates on the density function. This regularity properties will be used in Section 6.2.

Corollary 5.20. *Let $\mathscr{C}^k$ be a k-dimensional $\mathcal{D}$-cylinder parametrized as in (5.33) and let $\sigma^{w+se}(Z^k)$, $\sigma^{w+te}(Z^k)$ be two sections of $\mathscr{C}^k$ with s and t as in (5.37). Then, if we put $s = w+se$ and $t = w+te$, we have that*

$$\sigma^{t-|s-t|e}_{\#}\big(\mathscr{H}^{n-k}\llcorner\sigma^t(Z^k)\big) \ll \mathscr{H}^{n-k}\llcorner\sigma^s(Z^k) \tag{5.72}$$

and by the Radon-Nikodym theorem there exists a function $\alpha(t,s,\cdot)$ which is $\mathscr{H}^{n-k}$-a.e. defined on $\sigma^s(Z^k)$ and is such that

$$\sigma^{t-|s-t|e}_{\#}\big(\mathscr{H}^{n-k}\llcorner\sigma^t(Z^k)\big) = \alpha(t,s,\cdot)\cdot\mathscr{H}^{n-k}\llcorner\sigma^s(Z^k). \tag{5.73}$$

Proof. Without loss of generality we can assume that $s=0$. If $\mathscr{H}^{n-k}(A)=0$ for some $A\subset Z^k$, by definition of push forward of a measure we have that

$$(\sigma^{w+te})^{-1}_{\#}\big(\mathscr{H}^{n-k}\llcorner\ \sigma^{w+te}(Z^k)\big)(A) = \mathscr{H}^{n-k}(\sigma^{w+te}(A)) \tag{5.74}$$

and taking $s=0$ in (5.37) we find that $\mathscr{H}^{n-k}(A)=0$ implies that $\mathscr{H}^{n-k}(\sigma^{w+te}(A)) = 0$. □

Remark 5.21. The fuction $\alpha = \alpha(t,s,y)$ defined in (5.73) is measurable w.r.t. y and, for $\mathscr{H}^{n-k}$-a.e. $y'\in\sigma^{w+te}(Z^k)$, we have that

$$\alpha(s,t,y') = \alpha(t,s,\sigma^{t-|s-t|e}(y'))^{-1}. \tag{5.75}$$

Moreover, from Lemma 5.12 we immediately get the uniform bounds:

$$\begin{aligned}
\left(\frac{h^+(t,e)-u}{h^+(t,e)}\right)^{n-k} &\leqslant \alpha(t+ue,t,\cdot) \leqslant \left(\frac{u-h^-(t,e)}{-h^-(t,e)}\right)^{n-k} && \text{if } u\in[0,h^+(t,e)],\\
\left(\frac{u-h^-(t,e)}{-h^-(t,e)}\right)^{n-k} &\leqslant \alpha(t+ue,t,\cdot) \leqslant \left(\frac{h^+(t,e)-u}{h^+(t,e)}\right)^{n-k} && \text{if } u\in[h^-(t,e),0].
\end{aligned} \tag{5.76}$$

We conclude this section with the following proposition:

Proposition 5.22. *Let $\mathscr{C}^k(\mathscr{Z}^k, C^k)$ be a k-dimensional $\mathcal{D}$-cylinder parametrized as in (5.33) and assume without loss of generality that $w = \pi_{\langle e_1,\dots,e_k\rangle}(Z^k) = 0$. Then, the function $\alpha(t,0,z)$ defined in (5.73) is locally Lipschitz in $t \in \mathrm{ri}(C^k)$ (and so jointly measurable in (t,z)). Moreover, for $\mathscr{H}^{n-k}$-a.e. $y \in \sigma^t(Z)$ the following estimates hold:*

1. *Derivative estimate*

$$-\left(\frac{n-k}{h^+(t,e)-u}\right)\alpha(t+ue,t,y) \leqslant \frac{d}{du}\alpha(t+ue,t,y) \leqslant \left(\frac{n-k}{u-h^-(t,e)}\right)\alpha(t+ue,t,y); \tag{5.77}$$

2. *Integral estimate*

$$\left(\frac{|h^+(t,e)-u|}{|h^+(t,e)|}\right)^{n-k}(-1)^{\chi_{\{u<0\}}} \leqslant \alpha(t+ue,t,y)\,(-1)^{\chi_{\{u<0\}}} \tag{5.78}$$

$$\leqslant \left(\frac{|h^-(t,e)-u|}{|h^-(t,e)|}\right)^{n-k}(-1)^{\chi_{\{u<0\}}}; \tag{5.79}$$

3. *Total variation estimate*

$$\int_{h^-(t,e)}^{h^+(t,e)}\left|\frac{d}{du}\alpha(t+ue,0,z)\right| du \leqslant 2\alpha(t,0,z)\left[\frac{|h^+-h^-|^{n-k}}{|h^+|^{n-k}}+\frac{|h^+-h^-|^{n-k}}{|h^-|^{n-k}}-1\right], \tag{5.80}$$

where h^+, h^- stand for $h^+(t,e), h^-(t,e)$.

Proof. Lipschitz regularity estimate First we prove the local Lipschitz regularity of $\alpha(t,0,z)$ w.r.t. $t \in \mathrm{ri}(C^k)$.
Given $s,t \in C^k$, we set $e = \frac{s-t}{|s-t|}$.
As

$$\sigma^{s-|s|\frac{s}{|s|}} = \sigma^{t-|t|\frac{t}{|t|}} \circ \sigma^{s-|s-t|e},$$

then

$$\begin{aligned}
\sigma_\#^{s-|s|\frac{s}{|s|}}(\mathscr{H}^{n-k}\llcorner\sigma^s(Z)) &= \sigma_\#^{t-|t|\frac{t}{|t|}}(\sigma_\#^{s-|s-t|e}\mathscr{H}^{n-k}\llcorner\sigma^s(Z))\\
&= \sigma_\#^{t-|t|\frac{t}{|t|}}\Big(\alpha(s,t,y)\cdot\mathscr{H}^{n-k}\llcorner\sigma^t(Z)\Big)\\
&= \alpha(t,0,z)\cdot\alpha(s,t,\sigma^t(z))\cdot\mathscr{H}^{n-k}_{|Z}.
\end{aligned} \tag{5.81}$$

By definition of α it follows that

$$\alpha(s,0,z)-\alpha(t,0,z)=\alpha(t,0,z)[\alpha(s,t,\sigma^t(z))-1]. \tag{5.82}$$

Now we want to estimate the term $[\alpha(s,t,\sigma^t(z))-1]$ with the lenght $|s-t|$ times a constant which is locally bounded w.r.t. t. In order to do this, we proceed as in the Corollary 2.19 of [Car2] using the estimate

$$\begin{aligned}\left(\frac{h^+(t,e)-u_2}{h^+(t,e)-u_1}\right)^{n-k}&\mathscr{H}^{n-k}(\sigma^{t+u_1e}(S))\leqslant\mathscr{H}^{n-k}(\sigma^{t+u_2e}(S))\\&\leqslant\left(\frac{u_2-h^-(t,e)}{u_1-h^-(t,e)}\right)^{n-k}\mathscr{H}^{n-k}(\sigma^{t+u_1e}(S)),\end{aligned} \tag{5.83}$$

which holds $\forall\ h^-(t,e)<u_1\leqslant u_2<h^+(t,e)$ and $\forall\, S\subset\sigma^t(Z)$.

Indeed, (5.83) can be rewritten in the following way:

$$\begin{aligned}\left(\frac{h^+(t,e)-u_2}{h^+(t,e)-u_1}\right)^{n-k}&\int_S\alpha(t+u_1e,t,y)\,d\mathscr{H}^{n-k}(y)\leqslant\int_S\alpha(t+u_2e,t,y)\,d\mathscr{H}^{n-k}(y)\\&\leqslant\left(\frac{u_2-h^-(t,e)}{u_1-h^-(t,e)}\right)^{n-k}\int_S\alpha(t+u_1e,t,y)\,d\mathscr{H}^{n-k}(y).\end{aligned} \tag{5.84}$$

Therefore, there is a dense sequence $\{u_i\}_{i\in\mathbb{N}}$ in $(h^-(t,e),h^+(t,e))$ such that for $\mathscr{H}^{n-k}$-a.e. $y\in S$ and for all $u_i\leqslant u_j$, $i,j,\in\mathbb{N}$ the following inequalities hold

$$\begin{aligned}\left[\left(\frac{h^+(t,e)-u_j}{h^+(t,e)-u_i}\right)^{n-k}-1\right]&\alpha(t+u_ie,t,y)\leqslant\alpha(t+u_je,t,y)-\alpha(t+u_ie,t,y)\\&\leqslant\left[\left(\frac{u_j-h^-(t,e)}{u_i-h^-(t,e)}\right)^{n-k}-1\right]\alpha(t+u_ie,t,y).\end{aligned} \tag{5.85}$$

Thanks to the uniform bounds (5.76), for all $y\in\sigma^t(Z)$ such that (5.85) holds, the function $\alpha(t+\cdot e,t,y)$ is locally Lipschitz on $\{u_i\}_{i\in\mathbb{N}}$ and for every $[a,b]\subset(h^-(t,e),h^+(t,e))$ the Lipschitz constants of α on $\{u_i\}_{i\in\mathbb{N}}\cap[a,b]$ are uniformly bounded w.r.t. y.
Then, on every compact interval $[a,b]\subset(h^-(t,e),h^+(t,e))$ there exists a Lipschitz extension $\tilde\alpha(t+\cdot e,t,y)$ of $\alpha(t+\cdot e,t,y)$ which has the same Lipschitz constant.
By the dominated convergence theorem, whenever $\{u_{j_n}\}_{n\in\mathbb{N}}\subset\{u_j\}_{j\in\mathbb{N}}$ converges to some $u\in[a,b]$ we have

$$\int_S\alpha(t+u_{j_n}e,t,y)\,d\mathscr{H}^{n-k}(y)\quad\longrightarrow\quad\int_S\tilde\alpha(t+ue,t,y)\,d\mathscr{H}^{n-k}(y),\quad\forall\,S\subset\sigma^t(Z).$$

However, the integral estimate (5.84) implies that

$$\int_S\alpha(t+u_{j_n}e,t,y)\,d\mathscr{H}^{n-k}(y)\quad\longrightarrow\quad\int_S\alpha(t+ue,t,y)\,d\mathscr{H}^{n-k}(y),$$

so that the Lipschitz extension $\tilde{\alpha}$ is an $L^1(\mathscr{H}^{n-k})$ representative of the original density α for all $u \in [a, b]$. Repeating the same reasoning for an increasing sequence of compact intervals $\{[a_n, b_n]\}_{n\in\mathbb{N}}$ that converge to $(h^-(t,e), h^+(t,e))$, we can assume that the density function $\alpha(t+ue, t, y)$ is locally Lipschitz in u with a Lipschitz constant that depends continuously on t and on e.
Then, by (5.82), the local Lipschitz regularity in t of the function $\alpha(t, 0, z)$ is proved.

Derivative estimate If we derive w.r.t. u the pointwise estimate (5.85) (which holds for all $u \in (h^-(t,e), h^+(t,e))$ by the first part of the proof) we obtain the derivative estimate (5.77).

Integral estimate (5.77) implies the monotonicity of the following quantities:

$$\frac{d}{du}\left(\frac{\alpha(t+ue,t,y)}{(h^+(t,e)-u)^{n-k}}\right) \geqslant 0, \qquad \frac{d}{du}\left(\frac{\alpha(t+ue,t,y)}{(u-h^-(t,e))^{n-k}}\right) \leqslant 0.$$

Integrating the above inequalities from $u \in (h^-(t,e), h^+(t,e))$ to 0 we obtain (5.79).

Total variation estimate In order to prove (5.80) we proceed as in Corollary 2.19 of [Car2].

$$\begin{aligned}
\int_{h^-(t,e)}^{0} \left|\frac{d}{du}\alpha(t+ue,0,z)\right| du \leqslant & \qquad\qquad (5.86)\\
& \int_{\{\frac{d}{du}\alpha(t+ue,0,z)>0\}\cap\{u\in(h^-(t,e),0)\}} \frac{d}{du}\alpha(t+ue,0,z)\, du \\
& \quad + \int_{h^-(t,e)}^{0} \frac{(n-k)\alpha(t+ue,0,z)}{|h^+(t,e)-u|}\, du \\
& \leqslant \int_{h^-(t,e)}^{0} \frac{d}{du}\alpha(t+ue,0,z)\, du + +2\int_{h^-(t,e)}^{0} \frac{(n-k)\alpha(t+ue,0,z)}{|h^+(t,e)-u|}\, du \\
& \leqslant \alpha(t,0,z) + 2\int_{h^-(t,e)}^{0} \frac{(n-k)\alpha(t+ue,0,z)}{|h^+(t,e)-u|}\, du. \qquad\qquad (5.87)
\end{aligned}$$

From (5.82) we know that $\alpha(t+ue,0,z) = \alpha(t,0,z)\,\alpha(t+ue,t,\sigma^t(z))$.
Moreover, since $u < 0$

$$\alpha(t+ue,t,\sigma^t(z)) \underset{(5.79)}{\leqslant} \left(\frac{|h^+(t,e)-u|}{|h^+(t,e)|}\right)^{n-k}.$$

If we substitute this inequality in (5.87) we find that

$$\begin{aligned}
(5.87) &\leqslant \alpha(t,0,z) + 2\alpha(t,0,z)\int_{h^-(t,e)}^{0} \frac{(n-k)|h^+(t,e)-u|^{n-k-1}}{|h^+(t,e)|^{n-k}}\, du \\
&= -\alpha(t,0,z) + 2\alpha(t,0,z)\frac{|h^+(t,e)-h^-(t,e)|^{n-k}}{|h^+(t,e)|^{n-k}}. \qquad\qquad (5.88)
\end{aligned}$$

Adding the symmetric estimate on $(0, h^+(t,e))$ we obtain (5.80). □

5.7 The disintegration on model sets

Now we conclude the proof of Theorem 5.3 on the model sets, giving also an explicit formula for the conditional probabilities.

We consider a k-dimensional $\mathcal{D}$-cylinder $\mathscr{C}^k = \mathscr{C}^k(\mathscr{Z}^k, C^k)$ parametrized as in (5.33) and we assume, without loss of generality, that $\pi_{\langle e_1,\dots,e_k\rangle}(Z^k) = 0 \in \mathbb{R}^n$. We also set $h_j^{\pm} = h^{\pm}(0, e_j)$, $\forall\, j = 1, \dots, k$, and we omit the point $w = 0$ in the notation for the map (5.34).

Theorem 5.23. *Let $\mathscr{C}^k$ be a k-dimensional $\mathcal{D}$-cylinder parametrized as in (5.33). Then, $\forall\, \varphi \in L^1_{loc}(\mathbb{R}^n)$,*

$$\int_{\mathscr{C}^k} \varphi \, d\mathscr{L}^n = \int_{Z^k} \int_{h_k^-}^{h_k^+} \dots \int_{h_1^-}^{h_1^+} \alpha(t_1 e_1 + \dots + t_k e_k, 0, z) \cdot \varphi(\sigma^{(t_1 e_1 + \dots + t_k e_k)}(z)) \, dt_1 \dots dt_k \, d\mathscr{H}^{n-k}(z).$$

Then, as $(Z^k, \mathscr{B}(Z^k))$ is isomorphic to the quotient space determined by the map ∇f on $\mathscr{C}^k$, by the uniqueness of the disintegration the conditional probabilities of the disintegration of the Lebesgue measure on the pieces of k-dimensional faces of f which are contained in $\mathscr{C}^k$ are given by

$$\mu_z(dt_1 \dots dt_k) = \frac{\alpha(t_1 e_1 + \dots + t_k e_k, 0, z)\, \mathscr{H}^k \llcorner [\mathrm{ri}(\mathcal{R}(z)) \cap \mathscr{C}^k](dt_1 \dots dt_k)}{\int_{h_k^-}^{h_k^+} \dots \int_{h_1^-}^{h_1^+} \alpha(s_1 e_1 + \dots + s_k e_k, 0, z)\, ds_1 \dots ds_k}, \tag{5.89}$$

for $\mathscr{H}^{n-k}$-a.e. $z \in Z^k$.

Proof. We proceed using the disintegration technique which was presented in Section 5.2.

$$\begin{aligned}
\int_{\mathscr{C}^k} \varphi(x)\, d\mathscr{L}^n(x) &= \int_{h_k^-}^{h_k^+} \dots \int_{h_1^-}^{h_1^+} \int_{\mathscr{C}^k \cap \{x\cdot e_k = t_k\} \cap \dots \cap \{x \cdot e_1 = t_1\}} \varphi \, d\mathscr{H}^{n-k} \\
&\underset{(5.72)}{=} \int_{h_k^-}^{h_k^+} \dots \int_{h_1^-}^{h_1^+} \int_{Z^k} \alpha(t_k e_k, 0, z) \dots \alpha(t_1 e_1 + \dots + t_k e_k, t_2 e_2 + \dots + t_k e_k, \sigma^{(t_2 e_2 + \dots + t_k e_k)}(z)) \\
&\qquad \cdot \varphi(\sigma^{(t_1 e_1 + \dots + t_k e_k)}(z))\, d\mathscr{H}^{n-k}(z)\, dt_1 \dots dt_k \\
&\underset{(5.81)}{=} \int_{h_k^-}^{h_k^+} \dots \int_{h_1^-}^{h_1^+} \int_{Z^k} \alpha(t_1 e_1 + \dots + t_k e_k, 0, z)\, \varphi(\sigma^{(t_1 e_1 + \dots + t_k e_k)}(z))\, d\mathscr{H}^{n-1}(z)\, dt_1 \dots dt_k \\
&\underset{\substack{(5.76)\\ \text{Prop } 5.22}}{=} \int_{Z^k} \int_{h_k^-}^{h_k^+} \dots \int_{h_1^-}^{h_1^+} \alpha(t_1 e_1 + \dots + t_k e_k, 0, z)\, \varphi(\sigma^{(t_1 e_1 + \dots + t_k e_k)}(z))\, dt_1 \dots dt_k \, d\mathscr{H}^{n-1}(z).
\end{aligned} \tag{5.90}$$

□

5.8 The global disintegration

In this section we prove Theorem 5.3, concerning the disintegration of the Lebesgue measure (restricted to a set of finite Lebesgue measure $K \subset \mathbb{R}^n$) on the whole k-dimensional faces of a convex function.

The idea is to put side by side the disintegrations on the model $\mathcal{D}$-cylinders which belong to the countable family defined in Lemma 5.11, so as to obtain a global disintegration.

What will remain apart will be set $\mathcal{T} \setminus \cup_{k=1}^{n} E^k$, projection of those points which do not belong to the relative interior of any face. Nevertheless, the following lemma ensures that this set is $\mathscr{L}^n$-negligible. Indeed, the union of the borders of the n-dimensional faces has zero Lebesgue measure by convexity and by the fact that the n-dimensional faces of f are at most countable.

For faces of dimension k, with $1 < k < n$, the proof is by contradiction: one considers a Lebesgue point of suitable subsets of $\cup_y F_y^k$ and applies the fundamental estimate (5.37) in order to show that the complementary is too big.

Equation (5.91) below was first proved using a different technique in [Lar2] — where it was shown that the union of the relative boundaries of the R-faces (see Definition 5.16) of an n-dimensional convex body C which have dimension at least 1 has zero $\mathscr{H}^{n-1}$-measure.

Lemma 5.24. *The set of points which do not belong to the relative interior of any face is $\mathscr{L}^n$-negligible:*

$$\mathscr{L}^n\left(\mathcal{T} \setminus \bigcup_{k=1}^{n} E^k\right) = 0, \qquad \text{where } E^k = \bigcup_y \operatorname{ri}\left(F_y^k\right). \tag{5.91}$$

Proof. Consider any n-dimensional face F_y^n. Being convex, it has nonempty interior. As a consequence, since two different faces cannot intersect, there are at most countably many n-dimensional faces $\{F_{y_i}^n\}_{i\in\mathbb{N}}$; moreover, by convexity, each $F_{y_i}^n$ has an $\mathscr{L}^n$-negligible boundary. Thus

$$\mathscr{L}^n\left(\bigcup_i \operatorname{rb}\left(F_{y_i}^n\right)\right) = 0.$$

Since $\mathcal{T} \subset \bigcup\limits_{k=1}^{n} F^k$, the thesis is reduced to showing that, for $0 < k < n$,

$$\mathscr{L}^n\left(F^k \setminus E^k\right) = 0. \tag{5.92}$$

Given a k dimensional subspace $V \in \mathbf{G}(k,n)$, a unit direction $e \in \mathbf{S}^{n-1} \cap V$, and $p \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, define the set $\mathcal{A}^{p,e,V}$ of those $x \in \mathcal{T} \setminus \operatorname{ri}(F^k_{\nabla f(x)})$ which satisfy the two relations

$$\inf_{d\in\mathcal{D}(x)} |\pi_V(d)| \geqslant 1/\sqrt{2} \tag{5.93}$$

$$\pi_V(F^k_{\nabla f(x)}) \supset \operatorname{conv}\left(\{\pi_V(x)\} \cup \pi_V(x) + 2^{-p+1}e + 2^{-p}(\mathbf{S}^{n-1} \cap V)\right). \tag{5.94}$$

Choosing (p, e, V) in a sequence $\{(p_i, e_i, V_i)\}_{i\in\mathbb{N}}$ which is dense in $\mathbb{N}_0 \times (\mathbf{S}^{n-1} \cap V) \times \mathbf{G}(k, n)$, the family $\{\mathcal{A}^{p_i,e_i,V_i}\}_{i\in\mathbb{N}}$ provides a countable covering of $F^k \setminus E^k$ with measurable sets. The measurability of each $\mathcal{A}^{p,e,V}$ can be deduced as follows. The set defined by (5.93) is exactly

$$\mathcal{D}^{-1} \circ \pi_V^{-1}\left(V \setminus \mathrm{ri}\left(\frac{1}{\sqrt{2}}\mathbf{B}^n\right)\right).$$

Moreover, (5.94) is equivalent to

$$\pi_V\big(\mathcal{R}(x) - x\big) \supset \mathrm{conv}(2^{-p+1}e + 2^{-p}\big(\mathbf{S}^{n-1} \cap V)\big).$$

Since $\mathcal{R}$ and $\mathcal{D}$ are measurable (Lemma 5.7), then the measurability of $\mathcal{A}^{p,e,V}$ follows.

In particular, if by absurd (5.92) does not hold, then there exists a subset $\mathcal{A}^{p,e,V}$ of $F^k \setminus E^k$ with positive Lebesgue measure. Up to rescaling, one can assume w.l.o.g. that $p = 0$, $V = \langle e_1, \ldots, e_k\rangle$, where $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $\mathbb{R}^n$, and $e = e_1$. Moreover, we will denote $\mathcal{A}^{p,e,V}$ simply with $\mathcal{A}$.

Before reaching the contradiction $\mathscr{L}^n(\mathcal{A}) = 0$, we need the following remarks. First of all we notice that, for $0 \leqslant h \leqslant 3$ and $t \in \pi_V(\mathcal{A})$, one can prove the fundamental estimate

$$\mathscr{H}^{n-k}\big(\sigma^{t+he}(S)\big) \geqslant \left(\frac{3-h}{3}\right)^{n-k} \mathscr{H}^{n-k}(S) \qquad \forall S \subset \mathcal{A} \cap \pi_V^{-1}(t) \quad (5.95)$$

exactly as in Lemma 5.12, with the approximating vector field given in Step 3, Page 65. Indeed, the $(n-k+1)$-plane $\pi_V^{-1}(\mathbb{R}e)$ cuts the face of each $z \in \mathcal{A} \cap \pi_V^{-1}(t)$ into exactly one line l; this line has projection on V containing at least $[t, t+3e]$. Notice moreover that, by (5.94), each point $x \in l$, with $\pi_V(x) \in \mathrm{ri}\,\big([t, t+3e]\big)$, is a point in the relative interior of the face. In particular, it does not belong to $\mathcal{A}$.

Let us now prove the claim, assuming by contradiction that $\mathscr{L}^n(\mathcal{A}) > 0$ (see also Figure 8). Fix any $\varepsilon > 0$ small enough. w.l.o.g. one can suppose the origin to be a Lebesgue point of $\mathcal{A}$. Therefore, for every $0 < r < \bar{r}(\varepsilon) < 1$, there exists $T \subset \prod_{i=1}^k [0, re_i]$, with $\mathscr{H}^k(T) > (1-\varepsilon)r^k$, such that

$$\mathscr{H}^{n-k}\big(\mathcal{A} \cap \pi_V^{-1}(t) \cap [0, r]^n\big) \geqslant (1-\varepsilon)r^{n-k} \qquad \text{for all } t \in T. \qquad (5.96)$$

Moreover, there is a set $Q \subset [0, re]$, with $\mathscr{H}^1(Q) > (1-2\varepsilon)r$, such that

$$\mathscr{H}^{k-1}\big(T \cap \pi_{\langle e\rangle}^{-1}(q)\big) > (1-\varepsilon)r^{k-1} \qquad \text{for } q \in Q. \qquad (5.97)$$

Consider two points $q, s := q + 2\varepsilon re \in Q$, and take $t \in T \cap \pi_{\langle e\rangle}^{-1}(q)$. By the fundamental estimate (5.95), one has

$$\mathscr{H}^{n-k}\big(\sigma^{t+2\varepsilon re}(S_{t,r})\big) \geqslant (1-\varepsilon)^{n-k}\mathscr{H}^{n-k}(S_{t,r}) \qquad \text{where } S_{t,r} := \mathcal{A} \cap \pi_V^{-1}(t) \cap [0, r]^n.$$

Furthermore, condition (5.92) implies that $|x + 2\varepsilon re - \sigma^{t+2\varepsilon re}(x)| \leqslant 2\varepsilon r$ for each $x \in A \cap \pi_V^{-1}(t)$. Moving points within $\pi_V^{-1}(t) \cap [0, r]^n$ by means of the map $\sigma^{t+2\varepsilon re}$, they can therefore reach only the square $\pi_V^{-1}(s) \cap [-2\varepsilon r, (1+2\varepsilon)r]^n$. Notice that for ε small, since our proof is needed for $n \geqslant 3$ and $k \geqslant 1$,

$$\begin{aligned}\mathscr{H}^{n-k}([-2\varepsilon r, (1+2\varepsilon)r]^n \setminus [0,r]^n) &= (1+4\varepsilon)^{n-k} r^{n-k} - r^{n-k} \\ &\leqslant 4(n-k)\varepsilon r^{n-k} + o(\varepsilon) < n2^n \varepsilon r^{n-k}.\end{aligned}$$

As a consequence, the portion which exceeds $\pi_V^{-1}(s) \cap [0,r]^n$ can be estimated as follows:

$$\mathscr{H}^{n-k}\big(\sigma^{t+2\varepsilon re}(S_{t,r}) \cap [0,r]^n\big) \geqslant \mathscr{H}^{n-k}\big(\sigma^{t+2\varepsilon re}(S_{t,r})\big) - n2^n \varepsilon r^{n-k}.$$

As notice before, condition (5.94) implies that the points $\sigma^{t+2\varepsilon re}(S_{t,r}) \cap [0,r]^n$ belong to the complementary of $\mathcal{A}$. By the above inequalities we obtain then

$$\begin{aligned}\mathscr{H}^{n-k}\big(\mathcal{A}^c \cap \pi_V^{-1}(t+2\varepsilon re) \cap [0,r]^n\big) &\geqslant \mathscr{H}^{n-k}\big(\sigma^{t+2\varepsilon re}(S_{t,r}) \cap [0,r]^n\big) \\ &\geqslant (1-\varepsilon)^{n-k} \mathscr{H}^{n-k}(S_{t,r}) - n2^n \varepsilon r^{n-k} \\ &\overset{(5.96)}{\geqslant} (1-\varepsilon)^{n-k+1} r^{n-k} - n2^n \varepsilon r^{n-k} \\ &\geqslant \frac{1}{2} r^{n-k}.\end{aligned}$$

The last estimate shows that, for each $t \in T \cap \pi_{\langle e\rangle}^{-1}(q)$, the point $s = t + 2\varepsilon re$ does not satisfy the inequality in (5.96): thus $\big(T \cap \pi_{\langle e\rangle}^{-1}(q)\big) + 2\varepsilon re$ lies in the complementary of T. In particular

$$\mathscr{H}^{k-1}\big(T \cap \pi_{\langle e\rangle}^{-1}(s)\big) < r^{k-1} - \mathscr{H}^{k-1}\big(T \cap \pi_{\langle e\rangle}^{-1}(q)\big).$$

However, by construction both t and s belong to Q. This yields the contradiction, by definition of Q:

$$\frac{1}{2} r^{k-1} \overset{(5.97)}{<} \mathscr{H}^{k-1}\big(T \cap \pi_{\langle e\rangle}^{-1}(s)\big) < r^{k-1} - \mathscr{H}^{k-1}\big(T \cap \pi_{\langle e\rangle}^{-1}(t)\big) \overset{(5.97)}{<} \frac{1}{2} r^{k-1}. \qquad \square$$

Proof of Theorem 5.3. As we observed in Remark 5.4, it is sufficient to prove the theorem for the disintegration of the Lebegue measure on the set F^k when $k \in \{1, \ldots, n-1\}$.

Thanks to Lemma 5.24, we can further restrict the disintegration to the set E^k defined in (5.19); moreover, by (6.8), for all $k = 1, \ldots, n-1$ there exists a $\mathscr{L}^n$-negligible set N^k such that

$$E^k \setminus N^k = \bigcup_{j \in \mathbb{N}} \mathscr{C}_j^k \setminus \partial \mathscr{C}_j^k,$$

where $\{\mathscr{C}_j^k\}_{j \in \mathbb{N}}$ is the countable collection of k-dimensional $\mathcal{D}$-cylinders covering E^k which was constructed in Lemma 5.11, so that the sets $\hat{\mathscr{C}}_j^k = \mathscr{C}_j^k \setminus \partial \mathscr{C}_j^k$ are disjoint.

The fundamental observation is the following:

$$\bigcup_{j\in\mathbb{N}} \hat{\mathscr{C}}_j^k = \bigcup_{j\in\mathbb{N}} \bigcup_{y\in \operatorname{Im}\nabla f_{|_{E^k}}} E^k_{y,j} = \bigcup_{y\in \operatorname{Im}\nabla f_{|_{E^k}}} \bigcup_{j\in\mathbb{N}} E^k_{y,j} = \bigcup_{y\in \operatorname{Im}\nabla f_{|_{E^k}}} E^k_y \setminus N^k, \tag{5.98}$$

where $E^k_{y,j} = E^k_y \cap \hat{\mathscr{C}}_j^k$.

For all $j \in \mathbb{N}$, we set

$$Y_j = \{y \in \operatorname{Im}\nabla f_{|_{E^k}} : E^k_{y,j} \neq \emptyset\}, \tag{5.99}$$

we denote by $p_j : \hat{\mathscr{C}}_j^k \to Y_j$ the quotient map corresponding to the partition

$$\hat{\mathscr{C}}_j^k = \bigcup_{y\in \operatorname{Im}\nabla f_{|_{E^k}}} E^k_{y,j}$$

and we set $\nu_j = p_{j\#}\mathscr{L}^n \llcorner \hat{\mathscr{C}}_j^k$.

Since the quotient space $(Y_j, \mathscr{B}(Y_j))$ is isomorphic to $(Z^k_j, \mathscr{B}(Z^k_j))$, where Z^k_j is a section of $\hat{\mathscr{C}}_j^k$, by Theorem 5.23 we have that

$$\mathscr{L}^n \llcorner \hat{\mathscr{C}}_j^k(E_j \cap p_j^{-1}(F_j)) = \int_{F_j} \mu^j_y(E_j)\, d\nu_j(y), \quad \forall E_j \in \mathscr{B}(\hat{\mathscr{C}}_j^k),\ F_j \in \mathscr{B}(Y_j), \tag{5.100}$$

where μ^j_y is equivalent to $\mathscr{H}^k \llcorner E^k_{y,j}$ for ν_j-a.e. $y \in Y_j$.

Moreover, for every $E \in \mathscr{B}(\mathbb{R}^n) \cap E^k$ there exist sets $E_j \in \mathscr{B}(\hat{\mathscr{C}}_j^k)$ such that

$$E = \bigcup_{j\in\mathbb{N}} E_j$$

and for all $F \in \mathscr{B}(Y)$, where $Y = \bigcup_{j\in\mathbb{N}} Y_j = \operatorname{Im}\nabla f_{|_{E^k}}$, there exist sets $F_j \in \mathscr{B}(Y_j)$ such that

$$F = \bigcup_{j\in\mathbb{N}} F_j \quad \text{and} \quad \nabla f^{-1}(F) = \bigcup_{j\in\mathbb{N}} p_j^{-1}(F_j).$$

Then,

$$\begin{aligned}
\mathscr{L}^n \llcorner K(E \cap \nabla f^{-1}(F)) &= \sum_{j=1}^{+\infty} \mathscr{L}^n \llcorner \hat{\mathscr{C}}_j^k(E_j \cap p_j^{-1}(F_j)) \\
&\underset{(5.100)}{=} \sum_{j=1}^{+\infty} \int_{F_j} \mu^j_y(E_j)\, d\nu_j(y) \\
&= \sum_{j=1}^{+\infty} \int_{Y_j} \chi_{F_j}(y)\mu^j_y(E_j)\, d\nu_j(y) \\
&= \sum_{j=1}^{+\infty} \int_{Y} \chi_{F_j}(y)\mu^j_y(E_j) f_j(y)\, d\nu(y),
\end{aligned} \tag{5.101}$$

where f_j is the Radon-Nikodym derivative of ν_j w.r.t. the measure ν on Y given by $\nabla f_{\#}\mathscr{L}^n \llcorner K$.

Since, as we proved in Section 5.1, there exists a unique disintegration $\{\mu_y\}_{y\in \mathrm{Im}\,\nabla f_{|_{E^k}}}$ such that

$$\mathscr{L}^n \llcorner K(E\cap \nabla f^{-1}(F)) = \int_F \mu_y(E)\, d\nu(y) \quad \text{for all} \quad E\in \mathscr{B}(\mathbb{R}^n),\, F\in \mathscr{B}(Y),$$

we conclude that the last term in (5.101) converges and

$$\mu_y = \sum_{j=1}^{+\infty} f_j(y)\, \mu_y^j \quad \text{for } \nu\text{-a.e. } y\in Y, \tag{5.102}$$

so that the Theorem is proved. □

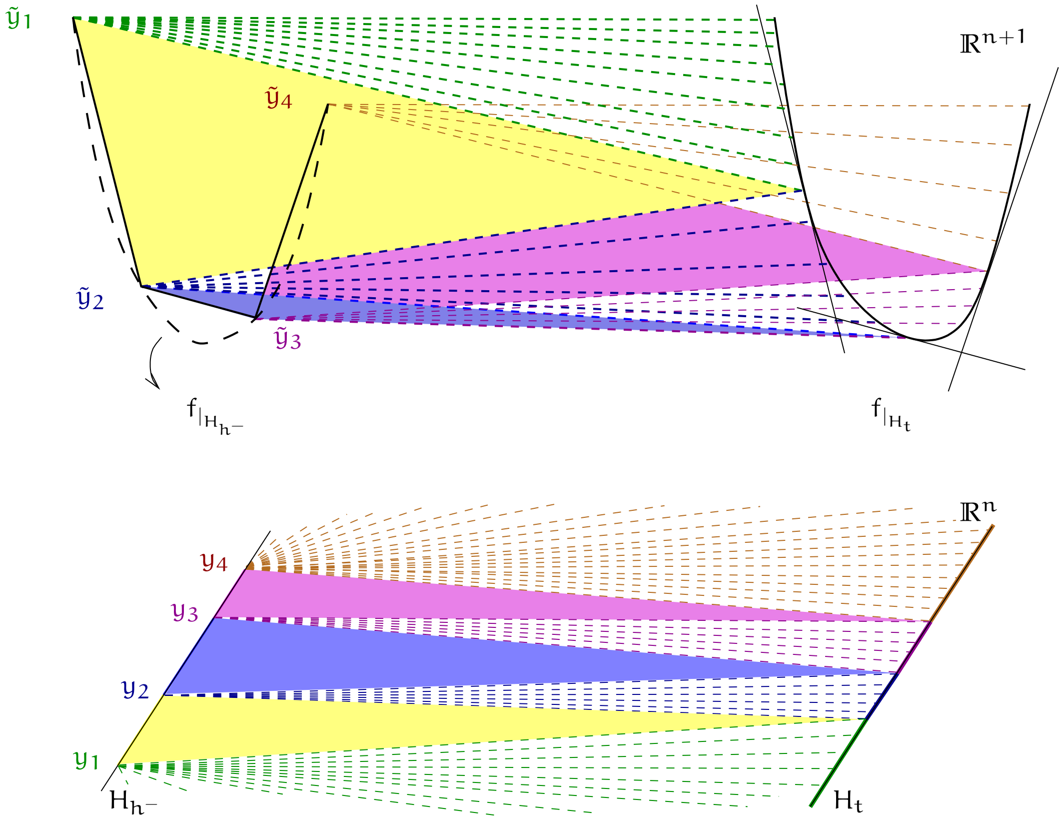


Figure 6: Illustration of a vector field approximating the one dimensional faces of f (Lemma 5.19). One can see in the picture the graph of f_4, which is the convex envelope of $\{\tilde{y}_i\}_{i=1,\dots,4}$ and $f_{\llcorner H_t}$. The faces of f_j connect $\mathscr{H}^{n-1}$-a.e. point of H_t to a single point among the $\{\tilde{y}_i\}_i$, while the remaining points of H_t correspond to some convex envelope $\mathrm{conv}(\{\tilde{y}_{i_\ell}\}_\ell)$ — here represented by the segments $[\tilde{y}_i, \tilde{y}_{i+1}]$. The region where the vector field v_4^t, giving the directions of the faces of f_j, is multivalued corresponds to the 'planar' faces of f_4. The affine span of these planar faces, restricted to suitable planes contained in H_t, provides a supporting hyperplane for the restriction of f to these latter planes — in the picture they are depicted as tangent lines. The intersection of $\sigma^t(Z) \subset H_t$ with any supporting plane to the graph of $f_{\llcorner H_t}$ must contain just one point, otherwise $\mathcal{D}$ would be multivalued at some point of $\sigma^t(Z)$.

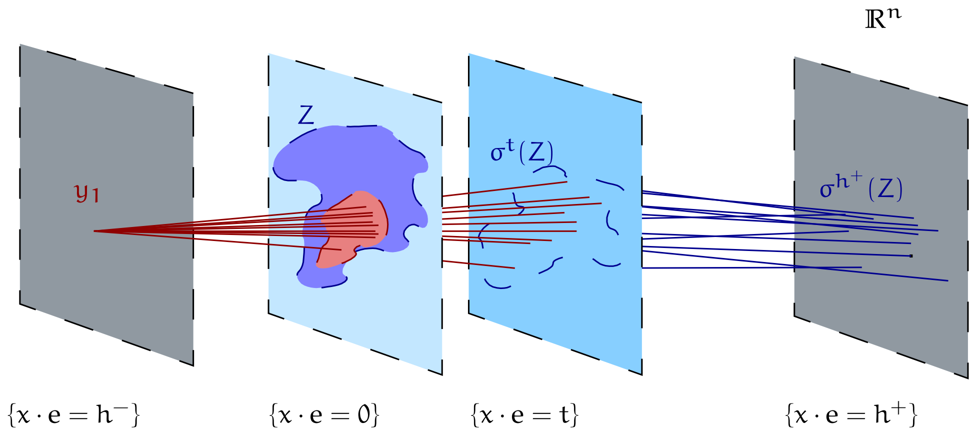


Figure 7: The vector field v_e is approximated by directions of approximating cones, in the picture one can see the first one. At the same time, Z is approximated by the push forward of $\sigma^t(Z)$ with the approximating vector field: compare the blue area with the red one.

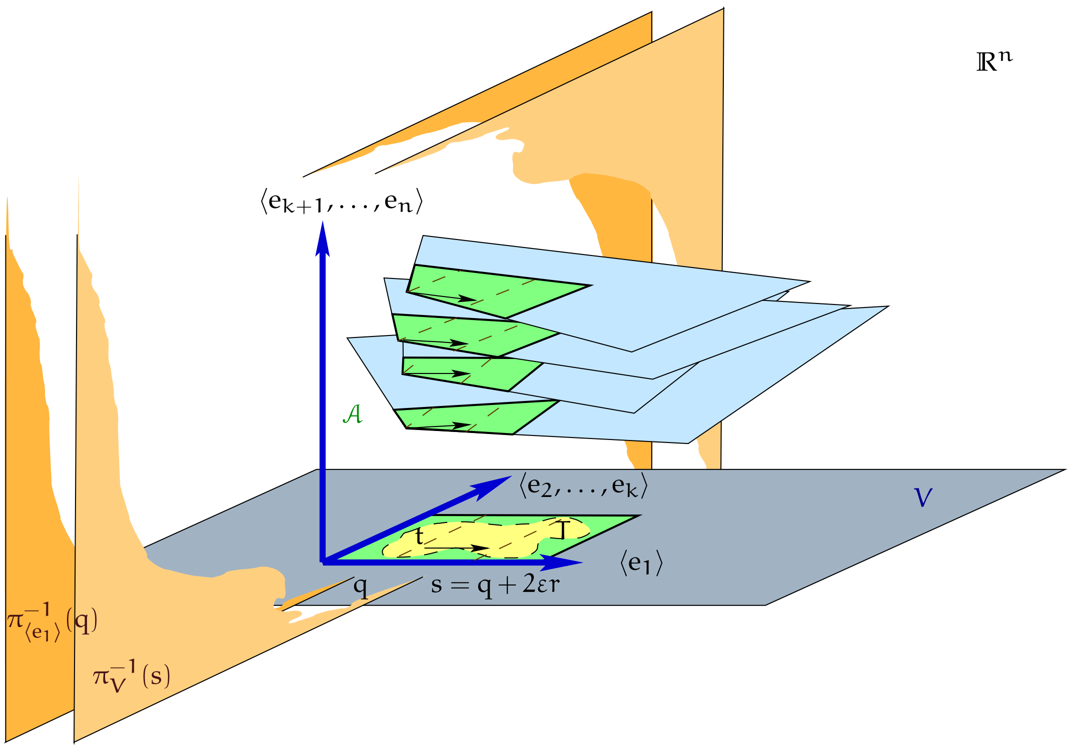


Figure 8: Illustration of the construction in the proof of Lemma 5.24. $\mathcal{A}$ is the set of points on the border of k-faces of f, projected on $\mathbb{R}^n$, having directions close to $V = \langle e_1, \dots, e_k \rangle$ and such that, for each point $x \in \mathcal{A}$, $\pi_V(F^k_{\nabla f(x)})$ contains a fixed half k-cone centered at x with direction e_1. T is a subset of the square $\prod_{i=1}^k [0, re_i]$ such that, for every $t \in T$, $\pi_V^{-1}(t) \cap \mathcal{A}$ is 'big'. Finally, q, $s = q + 2\varepsilon re_1$ are points on $[0, re_1]$ such that the intersection of T with the affine hyperplanes $\pi^{-1}_{\langle e_1 \rangle}(q)$, $\pi^{-1}_{\langle e_1 \rangle}(s)$ is 'big'. The absurd arises from the following. Due to the fundamental estimate, translating by $2\varepsilon re_1$ the points $T \cap \pi^{-1}_{\langle e_1 \rangle}(q)$, one finds points in the complementary of T. Since $T \cap \pi^{-1}_{\langle e_1 \rangle}(q)$ was 'big', then $T \setminus \pi^{-1}_{\langle e_1 \rangle}(s)$ should be big, contradicting the fact that $T \cap \pi^{-1}_{\langle e_1 \rangle}(s)$ is 'big'.

6 Further regularity: a divergence formula

6.1 One dimensional rays

In this section we extend the function d to be null out of $\bar{\mathcal{T}}$. We consider then its distributional divergence

$$\langle \operatorname{div} d, \varphi \rangle = \int_{\bar{\mathcal{T}}} \nabla\varphi \cdot d \, dx \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

If the set of initial points, or final points as well, is compact, then it turns out to be a Radon measure concentrated on $\bar{\mathcal{T}}$. More generally, it is a series of measures (see examples in Section 13.1). A decomposition of it can be constructed as follows. Consider the countable partition of $\bar{\mathcal{T}}$ into tuft sets $\{\bar{\mathcal{K}}_i\}_{i\in\mathbb{N}}$ of Lemma 4.13. Fix the attention on one $\bar{\mathcal{K}}_i$. Truncate the rays with an hyperplane just before they enter $\bar{\mathcal{K}}_i$, 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 $\bar{\mathcal{K}}_i$. The i-th addend of the series is then defined as the restriction to $\bar{\mathcal{K}}_i$ of the divergence of this vector field $\hat{d}$. The absolutely continuous part of this divergences does not depend on the $\{\bar{\mathcal{K}}_i\}_{i\in\mathbb{N}}$ we have chosen, as well as the distributional limit of the series — which is precisely $\operatorname{div} d$.

6.1.1 *Local divergence*

In this section, 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 6.1. Fix the attention on a d-cylinder with bounded basis $\bar{\mathcal{K}} = \{\sigma^t(Z) : t \in (h^-, h^+)\}$, assume Z compact. Suppose, moreover, that for $\mathscr{L}^n$-a.e. $x \in \bar{\mathcal{K}}$ the ray $\bar{\mathcal{R}}(x)$ intersects also the compact $K = \sigma^{h^- -\varepsilon}(Z)$. Let $\{a_i\}$ be dense in K. Consider the potential given by

$$\hat{\phi}(x) = \max_{a\in K} \left\{ \phi(a) - \tilde{|\!|} x - a \tilde{|\!|} \right\}$$

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

Lemma 6.2. *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 $\phi(a) - \tilde{\|}x - a\tilde{\|}$ must attain a minimum on K, then the transport set $\bar{\mathcal{T}}_e$ 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 $\bar{\mathcal{K}}$. The regularity of the divergence, which in general should be only a distribution, is now proved by approximation.

As in Example 4.17, we see that the potentials

$$\hat{\phi}_I(x) = \max_{j=i,\dots,I} \left\{ \phi(a_j) - \tilde{\|}x - a_j\tilde{\|} \right\}$$

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

$$d_I(x) = \sum_{i=1}^{I} d^i(x)\chi_{\Omega_i}(x) \qquad d^i(x) = \frac{x - a_i}{|x - a_i|},$$

with

$$\Omega_i = \left\{ x : \ \tilde{\|}x - a_i\tilde{\|} > \tilde{\|}x - a_j\tilde{\|},\ j \in \{1 \dots I\} \setminus i \right\}, \qquad J_I = \bigcup_i \partial\Omega_i,$$

J_I $\mathcal{H}^{n-1}$-countably rectifiable, converges p.w. $\mathcal{H}^n$-a.e. to $\hat{d}$. By d_I's membership in BV, the distribution div d_I is a Radon measure: we have thus

$$\langle \operatorname{div} d_I, \varphi \rangle = -\int \nabla\varphi \cdot d_I = \int \varphi \operatorname{div} d_I \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

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

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

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

$$(\operatorname{div} d_I)_{a.c.}(x) = \sum_i \frac{n-1}{|x - a_i|} \chi_{\Omega_i}(x) \leqslant \frac{n-1}{\operatorname{dist}(x, \cup_i a_i)}.$$

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

$$\begin{aligned}(\operatorname{div} d_I)^+(B_r(x)) - (\operatorname{div} d_I)^-(B_r(x)) &= \operatorname{div} d_I(B_r(x)) \\ &= \int_{\partial B_r(x)} d_I(y) \cdot \frac{y}{|y|} d\mathcal{H}^{n-1}(y) \geqslant -|\partial B_r(0)|.\end{aligned}$$

The last two estimates yield

$$|\operatorname{div} d_I|(B_r(x)) \leqslant |\partial B_r(0)| + \frac{2(n-1)|B_r(x)|}{\operatorname{dist}(B_r(x), \cup_i a_i)} \qquad \text{for } B_r(x) \cap K = \emptyset.$$

In particular, restrict $\operatorname{div} d_I$ on open sets O_k increasing to $\mathbb{R}^n \setminus K$. By compactness, the measures $\operatorname{div} d_I \llcorner O_k$ 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 $\varphi \in C_c^\infty(\mathbb{R}^n \setminus K)$

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

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

Lemma 6.3. *Let $\bar{\mathcal{K}}$ be the* d*-cylinder fixed above for defining* $\hat{d}$ *(Definition 6.1). Consider any couple $\mathcal{S}, c$ as in the disintegration Theorem 4.26. Then, for any* d*-sub-cylinder $\bar{\mathcal{K}}'$ of $\bar{\mathcal{K}}$, the following formulae hold: $\forall \varphi \in C_c^\infty(\mathbb{R}^n)$*

$$\partial_t c(t,y) - \Big[\big(\operatorname{div} \hat{d}\big)_{a.c.}(y + (t - d(y)\cdot y) d(y))\Big] c(t,y) = 0 \qquad \mathcal{H}^n\text{-a.e. on } \bar{\mathcal{K}}. \tag{6.1}$$

$$\int_{\bar{\mathcal{K}}'} \varphi \operatorname{div} \hat{d}_i = \int_{\bar{\mathcal{K}}'} \varphi \, (\operatorname{div} \hat{d}_i)_{a.c.} = -\int_{\bar{\mathcal{K}}'} \nabla\varphi \cdot d + \int_{\partial\bar{\mathcal{K}}'^{+} - \partial\bar{\mathcal{K}}'^{-}} \varphi d \cdot e_1. \tag{6.2}$$

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

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

Moreover, $\hat{d}$ is the vector field of directions relative to a potential $\hat{\phi}$: we can apply the disintegration Theorem 4.26, getting (4.19) for a couple $\widehat{\mathcal{S}}, \hat{c}$. Notice that on $\bar{\mathcal{K}}$, being $d = \hat{d}$, one can require $\widehat{\mathcal{S}}\llcorner_{\bar{\mathcal{K}}} \equiv \mathcal{S}$, which will lead to $\hat{c}(t,\cdot)\llcorner_{\widehat{\mathcal{S}}} \equiv c(t,\cdot)\llcorner_{\mathcal{S}}$. The local Lipschitz estimate on c (Remark 4.27), since we are integrating on a compact support, allows the integration by parts in the t variable

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

after performing this, the above equality becomes

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

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.

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

which turns out to be (6.1) on $\bar{\mathcal{K}}$. 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 $\hat{\varphi}$. 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 $\varphi = \chi_{\bar{\mathcal{K}}'}$, if Z is the relative section and $h^{\pm}$ define the height,

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

Coming back, integrating by parts again, one finds precisely (6.2). □

6.1.2 *Global divergence*

The divergence of the vector field d, generally speaking, is not a measure (see examples of Section 13.1). Nevertheless, it is not merely a distribution: it is a series of measures. Consider a covering of d-cylinders $\bar{\mathcal{K}}_i$, as in section 4.2. Repeat the construction of 6.1.1: one gets measures $\operatorname{div}\hat{d}_i$, which one can cut out of $\bar{\mathcal{K}}_i$. The finite sum of this sequence of disjoint measures converges to $\operatorname{div} d$, 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 $\{\bar{\mathcal{K}}_i\}$ one has chosen. Notwithstanding, it turns out that this is not the case. In fact, the absolutely continuous part $(\operatorname{div}\hat{d})_{\text{a.c.}}$ satisfies the following equation.

Lemma 6.4. *If one, just formally, defines on $\bar{\mathcal{T}}$ the measurable function*

$$(\operatorname{div} d)_{\text{a.c.}} := \sum_i (\operatorname{div} d_i)_{\text{a.c.}}\chi_{\bar{\mathcal{K}}_i},$$

then, for any partition into d-cylinders as in Theorem 4.26 with relative density c and sections $\mathcal{S}$, one has the relation

$$\partial_t c(t,y) - \Big[\big(\operatorname{div} d\big)_{\text{a.c.}}(y + (t - d(y)\cdot y)d(y))\Big]c(t,y) = 0 \qquad \mathcal{H}^n z\text{-a.e. on } \bar{\mathcal{T}}, \tag{6.3}$$

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

Remark 6.5. The measurable function $(\operatorname{div} d)_{a.c.}$ in general does not define a distribution, since it can fail to be locally integrable (Example 13.7, 13.8).

Proof. Since the $\bar{\mathcal{K}}_i$ are a partition of $\bar{\mathcal{T}}$, and their bases are $\mathcal{H}^n$-negligible, then the statement — which is a pointwise relation — is a direct consequence of Lemm. 6.3. □

Remark 6.6. Since c does not depend on the construction of the vector fields $\hat{d}_i$, then Equation (6.3) ensures that $(\operatorname{div} d)_{a.c.}$ is independent of the choices we made to obtain $\hat{d}_i$. Moreover, by Corollary 4.24, one has the bounds $-\frac{n-1}{b(y)\cdot d(y)-t} \leqslant \big(\operatorname{div} d\big)_{a.c.}(y+(t-d(y)\cdot y)d(y)) \leqslant \frac{n-1}{t-a(y)\cdot d(y)}$.

Lemma 6.7. *We have the equality*

$$\operatorname{div} d = \sum_i \big(\operatorname{div} \hat{d}_i\big)_{a.c.} \llcorner \bar{\mathcal{K}}_i - \mathcal{H}^{n-1} \llcorner \partial \bar{\mathcal{K}}_i^+ + \mathcal{H}^{n-1} \llcorner \partial \bar{\mathcal{K}}_i^- .$$

Therefore, $\forall \varphi \in C_c^\infty(\mathbb{R}^n \setminus \cup_x a(x) \cup b(x))$,

$$\langle \operatorname{div} \hat{d}, \varphi \rangle = \int \varphi \,(\operatorname{div} d)_{a.c.}. \tag{6.4}$$

Proof. Since $d = \hat{d}_i$ on $\bar{\mathcal{K}}_i$, Equation (6.2) can be rewritten as

$$\int_{\bar{\mathcal{K}}_i} \nabla\varphi \cdot d = -\int_{\bar{\mathcal{K}}_i} \varphi (\operatorname{div} \hat{d}_i)_{a.c.} + \int_{\partial\bar{\mathcal{K}}_i^+ - \partial\bar{\mathcal{K}}_i^-} \varphi \qquad \forall\varphi \in C_c^\infty(\mathbb{R}^n).$$

Using the partition in the proof of Lemma 4.26, it follows that the divergence of d is the sum of the above measures: $\forall \varphi \in C_c^\infty(\mathbb{R}^n)$

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

Equation (6.4) follows from the fact that one can choose a partition $\bar{\mathcal{K}}_i$ whose bases are outside of the support of φ. □

6.2 Faces of a convex function

The previous section led to a definition of a function α, on any $\mathcal{D}$-cylinder $\mathscr{C}^k = \mathscr{C}^k(\mathscr{Z}^k, C^k)$, as the Radon-Nikodym derivative in (5.73).
In the present section we find that on $\mathscr{C}^k$ the function α satisfies the system of ODEs

$$\partial_{t_\ell}\alpha\Big(t = \pi_{\langle e_1,\dots,e_k\rangle}(x), 0, x - \sum_{i=1}^k x\cdot e_i v_i(x)\Big)$$
$$= (\operatorname{div} v_\ell)_{a.c.}(x)\alpha\Big(\pi_{\langle e_1,\dots,e_k\rangle}(x), 0, x - \sum_{i=1}^k x\cdot e_i v_i(x)\Big)$$

for $\ell = 1, \dots, k$, where we assume w.l.o.g. that $0 \in C^k$, $\langle e_1, \dots, e_k \rangle$ is an axis of $\mathscr{C}$, $v_i(x)$ is the vector field

$$x \mapsto \chi_{\mathscr{C}^k}(x)(\langle \mathcal{D}(x) \rangle \cap \pi^{-1}_{\langle e_1, \dots, e_k \rangle}(e_i))$$

and $(\operatorname{div} v_i)_{a.c.}(x)$ is the density of the absolutely continuous part of the divergence of v_i, that we prove to be a measure.
This is a consequence of the Disintegration Theorem 5.23 and of the regularity estimates on α in Proposition 5.22.
Notice that even the fact that the divergence of v_i is a measure is not trivial, since the vector field is just Borel.

Heuristically, the ODEs above can be formally derived as follows.
In Chapter 5 we saw that $\mathscr{C}^k$ is the image of the product space $C^k + Z^k$, where $Z^k = \mathscr{C}^k \cap \pi^{-1}_{\langle e_1, \dots, e_k \rangle}(0)$ is a section of $\mathscr{C}^k$, with the change of variable

$$\Phi(t+z) = z + \sum_{i=1}^{k} t_i v_i(z) = \sigma^t(z) \qquad \text{for all } t = \sum_{i=1}^{k} t_i e_i \in C^k,\ z \in Z^k. \tag{6.5}$$

In Theorem 5.23 we found that the weak Jacobian of this change of variable is defined, and given by

$$|\mathfrak{J}(t+z)| = \alpha(t, 0, z).$$

From (6.5) one finds that, if v_i was smooth instead of only Borel, this Jacobian would be

$$\mathfrak{J}(t+z) = \det\left(\left[[v_j \cdot e_i]_{\substack{i=1,\dots,n \\ j=1,\dots,k}} \middle| \left[\sum_{\ell=1}^{k} t_\ell \partial_{z_j} \langle v_\ell(z) \cdot e_i \rangle + \delta_{i,j} \right]_{\substack{i=1,\dots,n \\ j=k+1,\dots,n}} \right] \right);$$

by direct computations with Cramer rule and the multilinearity of the determinant, moreover, from the last two equations above one would prove the relation

$$\partial_{t_\ell} \mathfrak{J}(t+z) = \operatorname{trace}\left(Jv_\ell(z)\, (J\Phi(t+z))^{-1} \right) \mathfrak{J}(t+z),$$

where Jg denotes the Jacobian matrix of a function g.
By the Lipschitz regularity of α w.r.t. the $\{t_i\}_{i=1}^k$ variables given in Proposition 5.22, one could then expect that

$$\partial_{t_\ell} \alpha(t, 0, z) = \left(\sum_{j=1}^{n} \partial_{x_j} (v_i(\Phi^{-1}(x)) \cdot e_j)|_{x=\Phi(t+z)} \right) \alpha(t, 0, z). \tag{6.6}$$

Notice that $\sum_j \partial_{x_j}(v_i(\Phi^{-1}(x)) \cdot e_j)|_{x=\Phi(t+z)}$ is the pointwise divergence of the vector field $v_i(\Phi^{-1}(x))$ evaluated at $x = \Phi(t+z)$. In this article, we denote it with $(\operatorname{div}(v_i \circ \Phi^{-1}))_{a.c.}$.

Finally, given a regular domain $\Omega \subset \mathbb{R}^n$, by the Green-Gauss-Stokes formula one should have

$$\int_\Omega (\operatorname{div}(v_i \circ \Phi^{-1}))_{\text{a.c.}}\, d\mathscr{L}^n(x) = \int_{\partial\Omega} v_i(\Phi^{-1}(x)) \cdot \hat{n}\, d\mathscr{H}^{n-1}(x), \tag{6.7}$$

where $\hat{n}$ is the outer normal to the boundary of Ω.

The analogue of Formulas (6.6) and (6.7) is the additional regularity we prove in this section, in a weak context, for vector fields parallel to the faces and for the current of k-faces. Actually, for simplicity of notations we will continue working with the projection of the faces on $\mathbb{R}^n$ instead of with the faces themselves. We give now the idea of the proof, in the case of one dimensional faces.

Fix the attention on a 1-dimensional $\mathcal{D}$-cylinder $\mathscr{C}$ with axis e and basis $Z = \mathscr{C} \cap \pi_{\langle e\rangle}^{-1}(0)$. Consider the distributional divergence of the vector field v giving pointwise on $\mathscr{C}$ the direction of projected faces, normalized with $v \cdot e = 1$, and vanishing elsewhere. The Disintegration Theorem 5.23 decomposes integrals on $\mathscr{C}$ to integrals first on the projected faces, with the additional density factor α, then on Z. By means of it, one then reduces the integral $\int_{\mathscr{C}} \nabla\varphi \cdot v$, defining the distributional divergence, to the following integrals on the projected faces:

$$-\int_{[h^-e,h^+e]} \nabla\varphi(x)|_{x=z+t_1 v(z)} \cdot v(z)\alpha(t,0,z)\, d\mathscr{H}^1(t) \qquad \text{where } z \text{ varies in } Z.$$

Since α is Lipschitz in t and $\nabla\varphi|_{x=\sigma^{w+t_1 e}(z)} \cdot v = \partial_{t_1}(\varphi \circ \sigma^{w+t}(z))$, by integrating by parts one arrives to

$$\int_{[h^-e,h^+e]} \varphi \circ \sigma^{w+t}(z)\partial_{t_1}\alpha(t,0,z)\, d\mathscr{H}^1(t) - [\varphi \circ \sigma^{w+t}(z)\alpha(t,0,z)]\Big|_{t=h^-e}^{t=h^+e}.$$

Applying again the disintegration theorem in the other direction, by the invertibility of α, one comes back to integrals on the $\mathcal{D}$-cylinder, where in the first addend φ is now integrated with the factor $\partial_{t_1}\alpha/\alpha$.

An argument of this kind yields an explicit representation of the distributional divergence of the truncation of a vector field v, parallel at each point x to the projected face through x, to $\mathscr{C}^k$. This divergence is a Radon measure, the absolutely continuous part is basically given by (6.6) and, as in (6.7), there is moreover a singular term representing the flux through the border of $\mathscr{C}^k$ transversal to $\mathcal{D}$, already defined as

$$\partial\mathscr{C}^k = \mathscr{C}^k \cap \pi_{\langle e_1,\dots,e_k\rangle}^{-1}(\mathrm{rb}(C^k)), \qquad \hat{n}\llcorner_{\partial\mathscr{C}^k} \text{ outer unit normal to } \pi_{\langle e_1,\dots,e_k\rangle}^{-1}(C^k). \tag{6.8}$$

As $\mathscr{C}^k$ are not regular sets, but just σ-compact, there is a loss of regularity for the divergence of v in the whole $\mathbb{R}^n$. In general, the distributional divergence will just be a series of measures.

6.2.1 *Vector fields parallel to the faces*

In the present subsection, we study the regularity of a vector field parallel, at each point, to the corresponding face through that point.

Study on model sets

As a preliminary step, fix the attention on the $\mathcal{D}$-cylinder

$$\mathscr{C}^k = \mathscr{C}^k(\mathscr{Z}^k, C^k).$$

One can assume w.l.o.g. that the axis of $\mathscr{C}^k$ is identified by vectors $\{e_1, \dots, e_k\}$ which are the first k coordinate vectors of $\mathbb{R}^n$ and that C^k is the square

$$C^k = \prod_{i=1}^{k} [-e_i, e_i].$$

Denote with Z^k the section $\mathscr{Z}^k \cap \pi^{-1}_{\langle e_1,\dots,e_k\rangle}(0)$.

Definition 6.8 (Coordinate vector fields)**.** We define on $\mathbb{R}^n$ k-*coordinate vector fields* for $\mathscr{C}^k$ as follows:

$$v_i(x) = \begin{cases} 0 & & \text{if } x \notin \mathscr{C}^k \\ v \in \langle \mathcal{D}(x) \rangle & \text{such that } \pi_{\langle e_1,\dots,e_k\rangle} v = e_i & \text{if } x \in \mathscr{C}^k. \end{cases}$$

The k-coordinate vector fields are a basis for the module on the algebra of measurable functions from $\mathbb{R}^n$ to $\mathbb{R}$ constituted by the vector fields with values in $\langle \mathcal{D}(x) \rangle$ at each point $x \in \mathscr{C}^k$, and vanishing elsewhere.

Consider the distributional divergence of v_i, denoted by $\operatorname{div} v_i$. As a consequence of the absolute continuity of the push forward with σ, and by the regularity of the density α, one gains more regularity of the divergence.

Let us fix a notation. Given any vector field $v : \mathbb{R}^n \to \mathbb{R}^n$ whose distributional divergence is a Radon measure, we will denote with $(\operatorname{div} v)_{\text{a.c.}}$ the density of the absolutely continuous part of the measure $\operatorname{div} v$.

Lemma 6.9. *The distribution* $\operatorname{div} v_i$ *is a Radon measure. Its absolutely continuous part has density*

$$(\operatorname{div} v_i)_{\text{a.c.}}(x) = \frac{\partial_{t_i} \alpha\big(t = \pi_{\langle e_1,\dots,e_k\rangle}(x), 0, x - \sum_{i=1}^k x \cdot e_i v_i(x)\big)}{\alpha\big(\pi_{\langle e_1,\dots,e_k\rangle}(x), 0, x - \sum_{i=1}^k x \cdot e_i v_i(x)\big)} \chi_{\mathscr{C}^k}(x). \tag{6.9}$$

Its singular part is $\mathscr{H}^{n-1} \llcorner \big(\mathscr{C}^k \cap \{x \cdot e_i = -1\}\big) - \mathscr{H}^{n-1} \llcorner \big(\mathscr{C}^k \cap \{x \cdot e_i = 1\}\big)$.

Proof. Consider any test function $\varphi \in C_c^\infty(\mathbb{R}^n)$ and apply the Disintegration Theorem 5.23:

$$\begin{aligned}\langle \operatorname{div} v_i, \varphi\rangle &:= -\int_{\mathscr{C}^k} \nabla\varphi(x)\cdot v_i(x)\, d\mathscr{L}^n(x)\\ &= -\int_{Z^k}\int_{C^k} \alpha(t,0,z)\,\nabla\varphi(\sigma^t(z))\cdot v_i(z)\, d\mathscr{H}^k(t)\, d\mathscr{H}^{n-k}(z),\end{aligned}$$

where we used that v_i is constant on the faces, i.e. $v_i(z) = v_i(\sigma^t(z))$. Being $\sigma^t(z) = z + \sum_{i=1}^k t_i v_i(z)$, one has

$$\nabla_x\varphi(x=\sigma^t(z))\cdot v_i(z) = \nabla_x\varphi(x=\sigma^t(z))\cdot \partial_{t_i}(\sigma^t(z)) = \partial_{t_i}(\varphi(\sigma^t(z))).$$

The inner integral is thus

$$\int_{C^k}\nabla\varphi(\sigma^t(z))\cdot v_i(z)\alpha(t,0,z)\, d\mathscr{H}^k(t) = \int_{C^k}\partial_{t_i}(\varphi(\sigma^t z))\alpha(t,0,z)\, d\mathscr{H}^k(t).$$

Since Proposition 5.22 ensures that α is Lipschitz in t, for $t \in C^k$, one can integrate by parts:

$$\begin{aligned}\int_{C^k}\partial_{t_i}(\varphi(\sigma^t(z)))\alpha(t,0,z)\, d\mathscr{H}^k(t) = &-\int_{C^k}\varphi(\sigma^t(z))\partial_{t_i}\alpha(t,0,z)\, d\mathscr{H}^k(t)\\ &+\int_{C^k\cap\{t_i=1\}}\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^{k-1}(t)\\ &-\int_{C^k\cap\{t_i=-1\}}\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^{k-1}(t).\end{aligned}$$

Substitute in the first expression. Recall moreover the definition of α in (5.73), as a Radon-Nikodym derivative of a push-forward measure, and its invertibility and Lipschitz estimates (Remark 5.21, Proposition 5.22), among with in particular the L^1 estimate on the function $\partial_{t_i}\alpha/\alpha$. Then, pushing the measure from $t = 0$ to a generic t, one comes back to the integral on the $\mathcal{D}$-cylinder

$$\begin{aligned}\langle \operatorname{div} v_i, \varphi\rangle = &\int_{Z^k}\int_{C^k}\varphi(\sigma^t(z))\partial_{t_i}\alpha(t,0,z)\, d\mathscr{H}^k(t)\, d\mathscr{H}^{n-k}(z)\\ &-\int_{Z^k}\int_{C^k\cap\{t_i=1\}}\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^{k-1}(t)\, d\mathscr{H}^{n-k}(z)\\ &+\int_{Z^k}\int_{C^k\cap\{t_i=-1\}}\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^{k-1}(t)\, d\mathscr{H}^{n-k}(z)\\ = &\int_{\mathscr{C}^k}\varphi(x)(\operatorname{div} v_i)_{\text{a.c.}}(x)\, d\mathscr{L}^n(x) - \int_{\mathscr{C}^k\cap\{x\cdot e_i=1\}}\varphi(x)\, d\mathscr{H}^{n-1}(x)\\ &+\int_{\mathscr{C}^k\cap\{x\cdot e_i=-1\}}\varphi(x)\, d\mathscr{H}^{n-1}(x).\end{aligned}$$

where $(\operatorname{div} v_i)_{\text{a.c.}}$ is the function $\frac{\partial_{t_i}\alpha}{\alpha}$ precisely written in the statement. Thus we have just proved the thesis, consisting in the last formula. □

Remark 6.10. Consider a function $\lambda \in L^1(\mathscr{C}^k;\mathbb{R})$ constant on each face, meaning that $\lambda(\sigma^t(z)) = \lambda(z)$ for $t \in C^k$ and $z \in Z^k$. One can regard this λ as a function of $\nabla f(x)$. Then the same statement of Lemma 6.9 applies to the vector field λv_i, but the divergence is clearly $\operatorname{div}(\lambda v_i) = \lambda \operatorname{div} v_i$. The proof is the same, observing that

$$
\begin{aligned}
\langle \operatorname{div}(\lambda v_i), \varphi\rangle &:= -\int_{\mathscr{C}^k} \nabla\varphi(x)\cdot\lambda(x)v_i(x)\, d\mathscr{L}^n(x)\\
&\overset{5.23}{=} -\int_{Z^k}\int_{C^k} \lambda(z)\nabla\varphi(\sigma^t(z))\cdot v_i(z)\alpha(t,0,z)\, d\mathscr{H}^k(t)\, d\mathscr{H}^{n-k}(z)\\
&= -\int_{Z^k}\int_{C^k} \lambda(z)\partial_{t_i}(\varphi(\sigma^t(z)))\alpha(t,0,z)\, d\mathscr{H}^k(t)\, d\mathscr{H}^{n-k}(z)\\
&= \int_{Z^k}\int_{C^k} \lambda(z)\varphi(\sigma^t(z))\partial_{t_i}\alpha(t,0,z)\, d\mathscr{H}^k(t)\, d\mathscr{H}^{n-k}(z) \qquad (6.10)\\
&\quad - \int_{Z^k}\int_{C^k\cap\{t_i=1\}} \lambda(z)\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^{k-1}(t)\, d\mathscr{H}^{n-k}(z)\\
&\quad + \int_{Z^k}\int_{C^k\cap\{t_i=-1\}} \lambda(z)\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^{k-1}(t)\, d\mathscr{H}^{n-k}(z)\\
&\overset{5.23}{=} \int_{\mathscr{C}^k} \varphi(x)\lambda(x)(\operatorname{div} v_i)_{\text{a.c.}}(x)\, d\mathscr{L}^n(x) - \int_{\mathscr{C}^k\cap\{x\cdot e_i=1\}} \varphi(x)\lambda(x)\, d\mathscr{H}^{n-1}(x)\\
&\quad + \int_{\mathscr{C}^k\cap\{x\cdot e_i=-1\}} \varphi(x)\lambda(x)\, d\mathscr{H}^{n-1}(x).
\end{aligned}
$$

Suitably adapting the integration by parts in the above equality (6.10) with

$$
\begin{aligned}
\int_{C^k} \lambda(\sigma^t(z))\partial_{t_i}(\varphi(\sigma^t(z)))\alpha(t,0,z)\, d\mathscr{H}^k(t) =\\
&- \int_{C^k} \lambda(\sigma^t(z))\varphi(\sigma^t(z))\partial_{t_i}\alpha(t,0,z)\, d\mathscr{H}^k(t)\\
&- \int_{C^k} \partial_{t_i}\lambda(\sigma^t(z))\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^k(t)\\
&+ \int_{C^k\cap\{t_i=1\}} \lambda(\sigma^t(z))\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^{k-1}(t)\\
&- \int_{C^k\cap\{t_i=-1\}} \lambda(\sigma^t(z))\varphi(\sigma^t(z))\alpha(t,0,z)\, d\mathscr{H}^{k-1}(t)
\end{aligned}
$$

one finds moreover that for all $\lambda \in L^1(\mathbb{R}^n;\mathbb{R})$ continuously differentiable along v_i with integrable directional derivative $\partial_{v_i}\lambda$, the following relation holds:

$$
\operatorname{div}(\lambda v_i) = \lambda \operatorname{div} v_i + \partial_{v_i}\lambda\, d\mathscr{L}^n \qquad (6.11)
$$

Notice that in (6.11) there is the addend $\lambda\mathscr{H}^{n-1} \llcorner (\mathscr{C}^k \cap \{x\cdot e_i = 1\})$, which would make no sense for a general $\lambda \in L^1(\mathbb{R}^n;\mathbb{R})$. Now we prove that the restriction to $\mathscr{C}^k \cap \{x\cdot e_i = 1\}$ of each representative of λ which is $C^1(F^k_{\nabla f(z)} \cap \mathscr{C}^k)$, for $\mathscr{H}^{n-k}$-a.e. $z \in Z^k$, identifies the same function in $L^1(\mathscr{C}^k \cap \{x\cdot e_i = 1\})$.
Indeed, any two representatives $\tilde\lambda$, $\hat\lambda$ of the L^1-class of λ can differ only on a $\mathscr{L}^n$-negligible set N. By the Disintegration Theorem 5.23, and using moreover

Fubini theorem for reducing the integral on C^k to integrals on lines parallel to e_i, one has that the intersection of N with each of the lines on the projected faces with projection on $\langle e_1, \dots, e_k\rangle$ parallel to e_i is almost always negligible:

$$\mathscr{H}^1\big(N \cap \{q + \langle v_i(q)\rangle\}\big) = 0 \qquad \text{for } q \in \mathscr{C}^k \cap \{x \cdot e_i = 0\} \setminus M, \text{ with } \mathscr{H}^{n-1}(M)=0.$$

Being continuously differentiable along v_i, one can redefine $\tilde\lambda$, $\hat\lambda$ in such a way that $N \cap \{q + \langle v_i(q)\rangle\} = \emptyset$ for all $q \in \mathscr{C}^k \cap \{x \cdot e_i = 0\} \setminus M$. As a consequence $N \cap \{x \cdot e_i = t\}$ is a subset of $\tau^{te_i}(M)$, where τ^{te_i} is the map moving along each projected face with tv_i:

$$\mathscr{C}^k \cap \{x \cdot e_i = 0\} \ni q \mapsto \tau^{te_i}(q) := q + tv_i = \sigma^{(\pi_{\langle e_1,\dots,e_k\rangle}(q)) + te_1}(q).$$

By the push forward formula (5.73), denoting $w_q := \pi_{\langle e_1,\dots,e_k\rangle}(q)$ and $z_q := \pi_{\langle e_{k+1},\dots,e_n\rangle}(q)$ for $S \subset \mathscr{C}^k \cap \{x \cdot e_1 = 0\}$

$$\mathscr{H}^{n-1} \llcorner (\tau^{te_i}(S)) = \alpha(w_q, w_q + te_i, z_q)\tau^{te_i}_\sharp(\mathscr{H}^{n-1}(q) \llcorner S)).$$

Therefore, as $\mathscr{H}^{n-1}(M) = 0$, one has that $\tilde\lambda$ and $\hat\lambda$ identify the same integrable function on each section of $\mathscr{C}^k$ perpendicular to e_i, showing that the measure $\lambda\mathscr{H}^{n-1} \llcorner (\{x \cdot e_i = 1\})$ is well defined.

Actually, the same argument as above should be used in (6.10) in order to show that $\lambda(z)$ is integrable on Z^k, so that one can separate the three integrals as we did. Indeed, being constant on each face by assumption, the restriction of λ to a section is trivially well defined as associating to a point the value of λ corresponding to the face of that point, but the integrability w.r.t. $\mathscr{H}^{n-1}$ on each slice is a consequence of the push forward estimate.

As a direct consequence of (6.11), by linearity, one gets a divergence formula for any sufficiently regular vector field which, at each point of $\mathscr{C}^k$, is parallel to the corresponding projected face of f.

Corollary 6.11. *Consider any vector field $v = \sum_{i=1}^k \lambda_i v_i$ with $\lambda_i \in L^1(\mathscr{C}^k; \mathbb{R})$ continuously differentiable along v_i, with directional derivative $\partial_{v_i}\lambda_i$ integrable on $\mathscr{C}^k$. Then the divergence of v is a Radon measure and for every $\varphi \in C^1_c(\mathbb{R}^n)$*

$$\langle \operatorname{div} v, \varphi\rangle = \int_{\mathscr{C}^k} \varphi(x)(\operatorname{div} v)_{\text{a.c.}}(x)\, d\mathscr{L}^n(x) - \int_{\partial\mathscr{C}^k} \varphi(x)\, v(x) \cdot \hat n(x)\, d\mathscr{H}^{n-1}(x),$$

where $\partial\mathscr{C}^k$, the border of $\mathscr{C}^k$ transversal to $\mathfrak{D}$, and $\hat n$, the outer unit normal, are define in Formula (6.8). Moreover, for $x \in \mathscr{C}^k$

$$(\operatorname{div} v)_{\text{a.c.}}(x) = \sum_{i=1}^k \lambda_i(x) \frac{\partial_{t_i}\alpha(t = \pi_{\langle e_1,\dots,e_k\rangle}(x), 0, x - \sum_{i=1}^k x \cdot e_i v_i(x))}{\alpha(\pi_{\langle e_1,\dots,e_k\rangle}(x), 0, x - \sum_{i=1}^k x \cdot e_i v_i(x))} + \sum_{i=1}^k \partial_{v_i}\lambda_i(x). \tag{6.12}$$

Remark 6.12. The result is essentially based on the application of the integration by parts formula when the integral on $\mathscr{C}^k$ is reduced, by the Disintegration Theorem, to integrals on C^k: this is why we assume the C^1 regularity of the λ_i, w.r.t. the directions of the k-face passing through each point of $\mathscr{C}^k$. Such regularity could be further weakened, however we do not pursue this issue here. As a consequence, one can easily extend the statement of the previous corollary to sets of the form $\mathscr{C}^k_\Omega = F^k \cap \pi^{-1}_{\langle e_1,\dots,e_n\rangle}(\overline{\Omega})$, for an open set $\Omega \subset \langle e_1,\dots,e_k\rangle$ with piecewise Lipschitz boundary, defining $\partial\mathscr{C}^k_\Omega := F^k \cap \pi^{-1}_{\langle e_1,\dots,e_n\rangle}(\mathrm{rb}(\Omega))$.

Global version

We study now the distributional divergence of an integrable vector field v on $\mathcal{T}$, as we did in Subsection 6.2.1 for such a vector field truncated on $\mathcal{D}$-cylinders.

Corollary 6.13. *Consider a vector field* $v \in L^1(\mathcal{T};\mathbb{R}^n)$ *such that* $v(x) \in \langle \mathcal{D}(x)\rangle$ *for* $x \in \mathbb{R}^n$*, where we define* $\mathcal{D}(x) = 0$ *for* $x \notin \mathcal{T}$*. Suppose moreover that the restriction to every face* E_y*, for* $y \in \mathrm{Im}\,\nabla f$*, is continuously differentiable with integrable derivatives. Then, for every* $\varphi \in C^1_c(\mathbb{R}^n)$ *one can write*

$$\begin{aligned}\langle \operatorname{div} v, \varphi\rangle = \lim_{\ell\to\infty} \sum_{i=1}^{\ell} \bigg\{ &\int_{\mathscr{C}_i} \varphi(x)(\operatorname{div}(\chi_{\mathscr{C}_i} v))_{\mathrm{a.c.}}(x)\, d\mathscr{L}^n(x) \\ &- \int_{\partial\mathscr{C}_i} \varphi(x)\, v(x)\cdot \hat{n}_i(x)\, d\mathscr{H}^{n-1}(x)\bigg\}.\end{aligned} \tag{6.13}$$

where $\{\mathscr{C}_\ell\}_{\ell\in\mathbb{N}}$ *is the countable partition of* $\mathcal{T}$ *in* $\mathcal{D}$*-cylinders given in Lemma 5.11, while* $(\operatorname{div}(\chi_{\mathscr{C}_i} v))_{\mathrm{a.c.}}$ *is the one of Corollary 6.11 and* $\partial\mathscr{C}_i$, $\hat{n}_i$ *are defined in Formula (6.8).*

Remark 6.14. By construction of the partition, each of the second integrals in the r.h.s. of (6.13) appears two times in the series, with opposite sign. Intuitively, the finite sum of these border terms is the integral on a perimeter which tends to the singular set.

Remark 6.15. Suppose that div v is a Radon measure. Then Corollary 6.13 implies that

$$\chi_{\mathscr{C}^k}(\operatorname{div} v)_{\mathrm{a.c.}} \equiv (\operatorname{div}(\chi_{\mathscr{C}^k} v))_{\mathrm{a.c.}}.$$

Proof of Corollary 6.13. The partition of $\cup_{k=1}^n E^k$ into such sets $\{\mathscr{C}_\ell\}_{\ell\in\mathbb{N}}$ is given exactly by Lemma 5.11. Moreover, Lemma 5.24 shows that the set $\mathcal{T}\setminus\cup_{k=1}^n E_k$ is Lebesgue negligible. Therefore, by dominated convergence theorem one finds that

$$\langle \operatorname{div} v, \varphi\rangle = -\int_{\mathcal{T}} v(x)\cdot\nabla\varphi(x)\, d\mathscr{L}^n(x) = -\lim_{\ell\to\infty}\sum_{i=1}^{\ell}\int_{\mathscr{C}_i} v(x)\cdot\nabla\varphi(x)\, d\mathscr{L}^n(x).$$

The addends in the r.h.s. are, by definition, the distributional divergence of the vector fields $v\chi_{\mathscr{C}_i}$ applied to φ. In particular, by Corollary 6.11, they are equal to

$$-\int_{\mathscr{C}_i} v(x)\cdot\nabla\varphi(x)\,d\mathscr{L}^n(x) = \int_{\mathscr{C}_i}\varphi(x)\big(\operatorname{div} v\big)_{\text{a.c.}}(x)\,d\mathscr{L}^n(x) + \int_{\partial\mathscr{C}_i^k}\varphi(x)\,v(x)\cdot\hat{n}_i(x)\,d\mathscr{H}^{n-1}(x),$$

proving the thesis. □

6.2.2 *The currents of* k*-faces*

In the present subsection, we change point of view. Instead of looking at vector fields constrained to the faces of f, we regard the k-dimensional faces of f as a k-dimensional current. We establish that this current is a locally flat chain, providing a sequence of normal currents converging to it in the mass norm. The border of these normal currents has the same representation one would have in a smooth setting.

We devote Appendix C to recalls on this argument, in order to fix the notations.

Divergence of the current of k*-faces on model sets*

As a preliminary study, restrict again the attention to a $\mathcal{D}$-cylinder as in Subsection 6.2.1, and keep the notation we had there.

The k-faces, restricted to $\mathscr{C}^k$, define a k-vector field

$$\xi(x) = \chi_{\mathscr{C}^k} v_1 \wedge \cdots \wedge v_k.$$

In general, this vector field does not enjoy much regularity. Nevertheless, as a consequence of the study of Chapter 5, one can find a representation of $\partial\big(\mathscr{L}^n\wedge\xi\big)$ like the one in a regular setting, (C.1). This involves the density α of the push-forward with σ which was studied before, see (5.73).

Lemma 6.16. *Consider a function λ such that it is continuously differentiable on each face and assume $\mathscr{C}^k$ bounded.*
Then, the k-dimensional current $\big(\mathscr{L}^n\wedge\lambda\xi\big)$ is normal and the following formula holds

$$\partial\big(\mathscr{L}^n\wedge\lambda\xi\big) = -\mathscr{L}^n\wedge(\operatorname{div}\lambda\xi)_{\text{a.c.}} + \big(\mathscr{H}^{n-1}\llcorner\partial\mathscr{C}^k\big)\wedge\big\langle d\hat{n},\lambda\xi\big\rangle,$$

where $\partial\mathscr{C}^k$, $\hat{n}$ are defined in (6.8), $d\hat{n}$ is the differential 1-form at each point dual to the vector field $\hat{n}$, and $(\operatorname{div}\lambda\xi)_{\text{a.c.}}$ is defined here as

$$(\operatorname{div}\lambda\xi)_{\text{a.c.}} := \sum_{i=1}^{k}(-1)^{i+1}(\operatorname{div}\lambda v_i)_{\text{a.c.}}\, v_1\wedge\cdots\wedge\hat{v}_i\wedge\cdots\wedge v_k$$

with the functions $(\operatorname{div} v_i)_{\text{a.c.}}$ of (6.9):

$$(\operatorname{div}\lambda v_i)_{\text{a.c.}}(x) = \left(\lambda(x)\frac{\partial_{t_i}\alpha\big(t=\pi_{\langle e_1,\dots,e_k\rangle}(x),0,x-\sum_{i=1}^k x\cdot e_i v_i(x)\big)}{\alpha\big(\pi_{\langle e_1,\dots,e_k\rangle}(x),0,x-\sum_{i=1}^k x\cdot e_i v_i(x)\big)} + \partial_{v_i}\lambda(x)\right)\chi_{\mathscr{C}^k}(x).$$

Proof. Actually, this is consequence of Corollary 6.11 in Subsection 6.2.1, reducing to computations in coordinates. One has to verify the equality of the two currents on a basis.

For simplicity, consider first

$$\omega = \phi\, de_2 \wedge \cdots \wedge de_k.$$

with $\phi \in C^1(\mathbb{R}^n)$. Then

$$d\omega = \partial_{x_1}\phi\, de_1 \wedge \cdots \wedge de_k + \sum_{i=k+1}^{n} \partial_{x_i}\phi\, de_i \wedge \cdots \wedge de_n,$$

$$\langle d\omega, \xi\rangle = \nabla\phi \cdot v_1 \qquad \big\langle \omega,\, (\operatorname{div}\lambda\xi)_{a.c.}\big\rangle = (\operatorname{div}\lambda v_1)_{a.c.}\phi \qquad \big\langle \omega,\, \langle d\hat{n},\, \xi\rangle\big\rangle = \phi\, \hat{n}\cdot e_1$$

and the thesis reduces exactly to Lemma 6.9, and Remark 6.10:

$$\partial\big(\mathscr{L}^n \wedge \lambda\xi\big)(\omega) := \int_{\mathscr{C}^k} \langle d\omega, \lambda\xi\rangle\, d\mathscr{L}^n \overset{6.9}{=}$$

$$-\int_{\mathscr{C}^k} \big\langle \omega,\, (\operatorname{div}\lambda\xi)_{a.c.}\big\rangle\, d\mathscr{L}^n + \int_{\partial\mathscr{C}^k} \big\langle \omega,\, \langle d\hat{n},\, \lambda\xi\rangle\big\rangle\, d\mathscr{H}^{n-1}$$
$$=: -\mathscr{L}^n \wedge (\operatorname{div}\lambda\xi)_{a.c.} + (\mathscr{H}^{n-1} \llcorner \partial\mathscr{C}^k) \wedge (\hat{n}\wedge\lambda\xi).$$

The same lemma applies with $(-1)^{i+1}v_i$ instead of v_1 if

$$\omega = \phi\, de_1 \wedge \cdots \wedge \widehat{de_i} \wedge \cdots \wedge de_k,$$

since the following formulas hold:

$$\langle d\omega, \xi\rangle = (-1)^{i+1}\nabla\phi \cdot v_i$$
$$\big\langle \omega,\, (\operatorname{div}\lambda\xi)_{a.c.}\big\rangle = (-1)^{i+1}(\operatorname{div}\lambda v_i)_{a.c.}\phi$$
$$\big\langle \omega,\, \langle d\hat{n},\, \xi\rangle\big\rangle = (-1)^{i+1}\phi\, \hat{n}\cdot e_i.$$

Let us show the equality more in general. By a direct computation, one can verify that

$$v_1 \wedge \cdots \wedge \hat{v}_i \wedge \cdots \wedge v_k$$
$$= \sum_{h=0}^{k-1} \sum_{\substack{k<i_{h+1}<\ldots \\ \cdots<i_{k-1}\leqslant n}} \sum_{\substack{\sigma\in S(1\ldots\hat{\imath}\ldots k-1) \\ \sigma(1)<\cdots<\sigma(h)}} \operatorname{sgn}\sigma\, v_{\sigma(h+1)}^{i_{h+1}} \cdots v_{\sigma(k-1)}^{i_{k-1}}\, e_{\sigma(1)\ldots\sigma(h)i_{h+1}\ldots i_{k-1}},$$

where v_i^j is the j-th component of v_i, $S(1\ldots\hat{\imath}\ldots k)$ denotes the group of permutation of the integers $\{1,\ldots,\hat{\imath},\ldots,k\}$, with i is missing, and, if $\sigma \in S(1\ldots\hat{\imath}\ldots k)$, $\operatorname{sgn}\sigma$ is 1 if the permutation is even, -1 otherwise.
On the other hand, consider now a $(k-1)$ form $\omega = \phi\, de_{i_1\ldots i_h} \wedge de_{i_{h+1}\ldots i_{k-1}}$, where $1 \leqslant i_1 < \cdots < i_h \leqslant k$, and $k < i_{h+1} < \cdots < i_{k-1} \leqslant n$. Then, again by direct computation,

$$\langle d\omega, \xi\rangle = \sum_{\substack{\sigma\in S(1\ldots k) \\ \sigma(2)=i_1,\ldots,\sigma(h+1)=i_h}} (\nabla\phi\cdot v_{\sigma(1)})\operatorname{sgn}\sigma\, v_{\sigma(h+2)}^{i_{h+1}} \cdots v_{\sigma(k)}^{i_{k-1}},$$

$$\left\langle \omega, (\operatorname{div} \lambda \xi)_{\text{a.c.}} \right\rangle = \varphi \sum_{i=1}^{k} \Bigg\{ (-1)^{i+1} (\operatorname{div} \lambda v_i)_{\text{a.c.}} \cdot$$

$$\sum_{\substack{\sigma \in S(1 \dots \hat{\imath} \dots k-1) \\ \sigma(1)=i_1, \dots, \sigma(h)=i_h}} \operatorname{sgn}\sigma\, v^{i_{h+1}}_{\sigma(h+1)} \cdots v^{i_{k-1}}_{\sigma(k-1)} \Bigg\}$$

$$= \sum_{\substack{\sigma \in S(1 \dots k) \\ \sigma(2)=i_1, \dots, \sigma(h+1)=i_h}} \left(\varphi \cdot (\operatorname{div} \lambda v_{\sigma(1)})_{\text{a.c.}} \right) \operatorname{sgn}\sigma\, v^{i_{h+1}}_{\sigma(h+2)} \cdots v^{i_{k-1}}_{\sigma(k)},$$

and finally

$$\left\langle \omega, \left\langle d\hat{n}, \xi \right\rangle \right\rangle = \sum_{i=1}^{k} (-1)^{i+1} (\hat{n} \cdot e_i) \left\langle \omega, v_1 \wedge \cdots \wedge \hat{v}_i \wedge \cdots \wedge v_k \right\rangle$$

$$= \sum_{\substack{\sigma \in S(1 \dots k) \\ \sigma(2)=i_1, \dots, \sigma(h+1)=i_h}} (\varphi\, \hat{n} \cdot v_{\sigma(1)}) \operatorname{sgn}\sigma\, v^{i_{h+1}}_{\sigma(h+2)} \cdots v^{i_{k-1}}_{\sigma(k)}$$

Therefore the thesis reduces to Corollary 6.11, being each v^i_j constant on each face. □

Divergence of the current of k*-faces in the whole space*

In the previous section, we considered a k-dimensional current $(\mathscr{L}^n \llcorner \mathscr{C}^k) \wedge \xi$ identified by the restriction to a $\mathcal{D}$-cylinder $\mathscr{C}^k$ of the k-faces of f, projected on $\mathbb{R}^n$. We established the formula analogous to (C.1) for the border of this current, which is representable by integration w.r.t. the measures $\mathscr{L}^n \llcorner \mathscr{C}^k$ and $\mathscr{H}^{n-1} \llcorner \partial \mathscr{C}^k$. In particular, when $\mathscr{C}^k$ is bounded it is a normal current.
Moreover, we have related the density of the absolutely continuous part to the function α by

$$(\operatorname{div} \xi)_{\text{a.c.}} = \sum_{i=1}^{k} \Bigg\{ (-1)^{i+1} \frac{\partial_{t_i} \alpha\big(t = \pi_{\langle e_1, \dots, e_k \rangle}(x), 0, x - \sum_{i=1}^{k} x \cdot e_i v_i(x)\big)}{\alpha\big(\pi_{\langle e_1, \dots, e_k \rangle}(x), 0, \sum_{i=1}^{k} x - x \cdot e_i v_i(x)\big)}$$

$$\chi_{\mathscr{C}^k}(x)\, v_1 \wedge \cdots \wedge \hat{v}_i \wedge \cdots \wedge v_k \Bigg\}.$$

We observe now that the partition we of $\mathbb{R}^n$ into the sets $\{F^k\}_{k=1}^n$, and the remaining set that we call now $\widetilde{F}^0$, define a $(n+1)$-uple of currents. The elements of this $(n+1)$-uple are described by the following statement, which is basically Corollary 6.13 when rephrased in this setting.

Corollary 6.17. *Let* $\{\mathscr{C}^k_\ell\}_{\ell \in \mathbb{N}}$ *be a countable partition of* E^k *in* $\mathcal{D}$*-cylinders as in Lemma 5.11 and, up to a refinement of the partition, assume moreover that the* $\mathcal{D}$*-cylinders are bounded. Consider a* k*-vector field* $\xi_k \in L^1(\mathbb{R}^n; \Lambda_k \mathbb{R}^n)$ *corresponding, at each point* $x \in E^k$*, to the* k*-plane* $\langle \mathcal{D}(x) \rangle$*, and vanishing elsewhere. Assume moreover that it is continuously differentiable if restricted to any set* $E^k_{\nabla f(x)}$*, with locally integrable derivatives, meaning more precisely that* $\xi_k \circ \sigma^{w_\ell + t}(z)$ *belongs to* $L^1_{\mathscr{H}^{n-k}(z)}(Z^k_\ell; C^1_t(C^k; \Lambda_k \mathbb{R}^n))$ *for each* ℓ*.*

Then, the k*-dimensional current* $\mathscr{L}^n \wedge \xi_k$ *is a locally flat chain, since it is the limit in the flat norm of normal currents: indeed, for* $k > 0$ *one has*

$$\partial\big(\mathscr{L}^n \wedge \xi_k\big) = \mathbf{F}\text{-}\lim_{\ell} \sum_{i=1}^{\ell} \Big\{ -\mathscr{L}^n \wedge (\operatorname{div}(\chi_{\mathscr{C}_i^k}\xi_k))_{\text{a.c.}} + \big(\mathscr{H}^{n-1} \llcorner \partial\mathscr{C}_i^k\big) \wedge \big\langle d\hat{n}_i,\, \xi_k \big\rangle \Big\},$$

where $(\operatorname{div}\chi_{\mathscr{C}_i^k}\xi_k)_{\text{a.c.}}$ *is the one of Lemma 6.16,* $\mathfrak{d}\mathscr{C}_i^k$*, the border of* $\mathscr{C}_i^k$ *transversal to* $\mathcal{D}$*, and* $\hat{n}_i$*, the outer unit normal, are defined in Formula (6.8), and* $d\hat{n}_i$ *is the dual to* $\hat{n}_i$*.*

Notice finally that the current $\mathscr{L}^n \wedge \xi_k$ is itself locally normal if restricted to the interior of E^k. However, in general E^k can have empty interior. If $\partial\big(\mathscr{L}^n \wedge \xi_k\big)$ is representable by integration, then the density of its absolutely continuous part w.r.t. $\mathscr{L}^n$, at any point $x \in \mathscr{C}_\ell^k$, is given by $\operatorname{div}(\chi_{\mathscr{C}_\ell^k}\xi_k)_{\text{a.c.}}(x)$.

Part II

Some basic problems in Optimal Mass Transportation

7 Explanation of the approach

Let (X, Ω, μ), (Y, Σ, ν) be two countably generated probability spaces, and let $(X \times Y, \Omega \otimes \Sigma)$ be the product measurable space. Using standard results on measure space isomorphisms (see for example the proof of the last theorem of [HJ]), in the following we assume that $(X, \Omega) = (Y, \Sigma) = ([0,1], \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra.

Let $\mathcal{P}([0,1]^2)$ be the set of Borel probability measures on $[0,1]^2$, and let $\Pi(\mu,\nu)$ be the subset of probability measures satisfying the marginal conditions $(P_1)_\sharp \pi = \mu$, $(P_2)_\sharp \pi = \nu$, where $P_1(x,y) = x$, $P_2(x,y) = y$ are the projection on X, Y:

$$\Pi(\mu,\nu) := \Big\{ \pi \in \mathcal{P}([0,1]^2) : (P_1)_\sharp \pi = \mu, (P_2)_\sharp \pi = \nu \Big\}.$$

For $\pi \in \Pi(\mu,\nu)$ we will denote by $\Gamma \subset [0,1]^2$ a set such that $\pi(\Gamma) = 1$: as a consequence of the inner regularity of Borel measures, it can be taken σ-compact.

For any Borel probability measure π on $[0,1]^2$, let $\Theta_\pi \subset \mathbf{P}([0,1]^2)$ be the π-completion of the Borel σ-algebra. We denote with $\Theta \subset \mathbf{P}([0,1]^2)$ the $\Pi(\mu,\nu)$-*universally measurable σ-algebra*: it is the intersection of all completed σ-algebras of the probability measures in $\Pi(\mu,\nu)$:

$$\Theta := \bigcap_{\pi \in \Pi(\mu,\nu)} \Big\{ \Theta_\pi, \pi \in \Pi(\mu,\nu) \Big\}. \tag{7.1}$$

We define the functional $\mathcal{I} : \Pi(\mu,\nu) \to \mathbb{R}$ by

$$\mathcal{I}(\pi) := \int c(x,y) \pi(dxdy), \tag{7.2}$$

where $c : [0,1]^2 \to [0,+\infty]$ is a Θ-measurable cost function. The set $\Pi^f(\mu,\nu) \subset \Pi(\mu,\nu)$ is the set of probability measures belonging to $\Pi(\mu,\nu)$ and satisfying the geometrical constraint $\mathcal{I}(\pi) < +\infty$.

The problems we are considering in the next sections are whether a given measure $\pi \in \Pi(\mu,\nu)$ satisfies one of the following properties:

- *it is extremal in $\Pi(\mu,\nu)$;*
- *it is the unique measure in $\Pi(\mu,\nu)$ concentrated on a given set $A \in \Theta$;*
- *it is minimizing the functional $\mathcal{I}(\pi)$ in $\Pi(\mu,\nu)$.*

We can restrict our analysis to the set $\Pi^f(\mu,\nu)$, by

- defining $c(x,y) = \mathbf{1}_\Gamma$ for a particular set Γ with $\pi(\Gamma) = 1$ in the first case
- defining $c(x,y) = \mathbf{1}_A$ in the second case,
- assuming that $\mathcal{I}(\pi) < +\infty$ to avoid trivialities in the third case.

In all the above cases a necessary condition can be easily obtained, namely

- π is acyclic in the first case (Definition 8.2),
- π is A-acyclic in the second case (Definition 9.2),
- π is c-cyclically monotone in the third case (Definition 10.1).

Nevertheless, there are explicit examples showing that this condition is only necessary.

The kernel is the following idea (Lemma 7.9). Let $\pi \in \Pi(\mu,\nu)$ be a transference plan.

Theorem 7.1. *Assume that there are partitions* $\{X_\alpha\}_{\alpha\in[0,1]}$, $\{Y_\beta\}_{\beta\in[0,1]}$ *such that*

1. *for all* $\pi' \in \Pi^f(\mu,\nu)$ *it holds* $\pi'(\cup_\alpha X_\alpha \times Y_\alpha) = 1$,
2. *the disintegration* $\pi = \int \pi_\alpha m(d\alpha)$ *of* π *w.r.t. the partition* $\{X_\alpha \times Y_\alpha\}_{\alpha\in[0,1]}$ *is strongly consistent,*
3. *in each equivalence class* $X_\alpha \times Y_\alpha$ *the measure* π_α *is extremal/unique/optimal in* $\Pi(\mu_\alpha,\nu_\alpha)$, *where*

$$\mu_\alpha := (P_1)_\sharp \pi_\alpha, \quad \nu_\alpha := (P_2)_\sharp \pi_\alpha.$$

Then π *is extremal/unique/optimal.*

The main tool is the Disintegration Theorem 2.7 applied to the partition $\{X_\alpha \times Y_\beta\}_{\alpha,\beta\in[0,1]}$. This partitions are constructed in order to satisfy Point (3).

Before explaining the meaning of the above conditions, we consider the following corollaries. Instead of partitions, we will equivalently speak of equivalence classes and relative equivalence relations.

Corollary 7.2 (Extremality (Theorem 8.8))**.** *Let* π *concentrated on a* σ*-compact acyclic set* Γ.

If we partition the set Γ *into axial equivalence classes (Definition 8.4), then* π *is extremal in* $\Pi(\mu,\nu)$ *if the disintegration is strongly consistent.*

We show in Theorem 8.9 that the strong consistency assumption in the above corollary is nothing more than the countable Borel limb condition of [HW].

Denote with h_X, h_Y the quotient maps w.r.t. the partitions $\{X_\alpha\}_{\alpha\in[0,1]}$, $\{Y_\beta\}_{\beta\in[0,1]}$. In Lemma 7.8 it is shown that if Conditions (1) and (2) of Theorem 7.1 are valid for π, then there exists $m \in \mathcal{P}([0,1])$ such that $(\mathbb{I},\mathbb{I})_\sharp m = (h_X,h_X)_\sharp \mu = (h_Y,h_Y)_\sharp \nu = (h_X \otimes h_Y)_\sharp \pi$.

Let now A be an analytic set and define the image set

$$A' := (h_X \otimes h_Y)(A).$$

Corollary 7.3 (Uniqueness (Page 119))**.** *Let π concentrated on a σ-compact A-acyclic set Γ.*

If we partition the set Γ into axial equivalence classes, then π is the unique measure in $\Pi(\mu, \nu)$ concentrated on A if

1. *the disintegration is strongly consistent,*
2. *A' is a subset of $\{\alpha \leqslant \beta\}$, up to measure preserving maps.*

Finally, let $c : [0,1]^2 \to [0, +\infty]$ be a coanalytic cost.

Corollary 7.4 (Optimality (Theorem 10.6))**.** *Let π concentrated on a σ-compact c-cyclically monotone set Γ and partition Γ w.r.t. the cycle equivalence relation (Definition 10.1).*

Then, π c-cyclically monotone is optimal if

1. *the disintegration is strongly consistent,*
2. *the image set $A' := (h_X \otimes h_Y)_\sharp\{c < +\infty\}$ is a set of uniqueness.*

The above result generalizes the previous known cases:

1. if μ or ν are atomic ([Pra]): clearly m must be atomic;
2. if $c(x, y) \leqslant a(x) + b(y)$ with $a \in L^1(\mu)$, $b \in L^1(\nu)$ ([RR]): m is a single δ;
3. if $c : [0,1]^2 \to \mathbb{R}$ is real valued and satisfies the following assumption ([AP])
$$\nu\left(\left\{y : \int c(x, y)\mu(dx) < +\infty\right\}\right) > 0, \quad \mu\left(\left\{x : \int c(x, y)\nu(dy) < +\infty\right\}\right) > 0 :$$
in this case m is a single δ;
4. If $\{c < +\infty\}$ is an open set O minus a $\mu \otimes \nu$ negligible set N ([BGMS]): in this case every point in $\{c < +\infty\}$ has a squared neighborhood satisfying condition (10.5) below.

In each case the equivalence classes are countably many Borel sets, so that the disintegration is strongly consistent and the acyclic set A' is a set of uniqueness (Lemma 9.9).

7.1 Interpretation

The three conditions listed in Theorem 7.1 have interesting interpretations in terms of measurability, marginal conditions and acyclic perturbations.

We first observe that the *necessary conditions* considered in all three cases can be stated as follows: *the transference plan π is unique/optimal w.r.t. the affine space generated by $\pi + \lambda_c$, where λ_c is a cyclic perturbation of π.*

Moreover, *the partitions have a natural crosswise structure w.r.t.* Γ: if $\{X_\alpha\}_\alpha$, $\{Y_\beta\}_\beta$ are the corresponding decompositions of $[0,1]$, then

$$\Gamma \cap (X_\alpha \times Y) = \Gamma \cap (X \times Y_\alpha) = \Gamma \cap (X_\alpha \times Y_\alpha). \tag{7.3}$$

This is clearly equivalent to $\Gamma \subset \cup_\alpha X_\alpha \times Y_\alpha$, so that Condition (1) is satisfied at least for π and for its cyclic perturbations.

This and consequently Condition (1) are *conditions on the geometry of the carriage* Γ, since the specific construction depends on it. In fact, fixed a procedure to partition a set Γ, it is easy to remove negligible sets obtaining different partitions: sometimes Theorem 7.1 can be satisfied or not depending on Γ, i.e. on the partition. A possible solution is to make the partition independent of Γ (Chapter 3), but maybe this decomposition does not satisfy the hypotheses of Theorem 7.1, while others do.

A consequence of the above discussion is that in the Corollary *a procedure is proposed to test a particular measure* π. Some particular cost may however imply that there is a partition valid for all transference plans: in this case the c-cyclical monotonicity becomes also sufficient, as in the known cases of Points (1)-(4) above.

Notice however that the statement is that *the necessary condition becomes sufficient if there exists a carriage* Γ *such that the corollary applies*, or more generally if there exists a partition such that Theorem 7.1 applies. When there is no such carriage, then one can modify the cost in such a way that there are transport plans satisfying the necessary condition, giving the same quotient set A' and which can be either extremal/unique/optimal or not (Proposition 11.9).

The strong consistency of the disintegration is a measure theoretic assumption: it is equivalent to the fact that the quotient space can be taken to be $([0,1],\mathcal{B})$, up to negligible sets. This is important in order to give a meaning to the optimality within the equivalence classes: otherwise the conditional probabilities π_α are useless and Condition (3) without meaning. From the geometrical point of view, we are saying that π can be represented by weighted sum of probabilities in $X_\alpha \times Y_\alpha$, and Condition (1) yields that we can decompose the problem into smaller problems in $X_\alpha \times Y_\alpha$. When the assumption is not satisfied, then one can modify the cost in order to have the same quotient measure but both c-cyclically monotone optimal and c-cyclically non optimal transport plans (Example 11.5, Proposition 11.9).

7.2 Setting and general scheme

Let $\{X_\alpha\}_{\alpha\in[0,1]}$ be a partition of X into pairwise disjoint sets, and similarly let $\{Y_\beta\}_{\beta\in[0,1]}$ be a partition of Y into pairwise disjoint sets. Let moreover $\{X_\alpha \times Y_\beta\}_{\alpha,\beta\in[0,1]}$ be the induced pairwise disjoint decomposition on $X \times Y$.

Since it is clear that the decomposition $X = \cup_\alpha X_\alpha$ with X_α pairwise disjoint induces an equivalence relation E by defining xEx' if and only if $x, x' \in X_\alpha$ for some α, we will also refer to X_α, Y_β and $X_\alpha \times Y_\beta$ as equivalence classes. We will often not distinguish an equivalence relation E on X and its graph

$$\mathrm{graph}(E) := \{(x,x') : xEx'\} \subset X \times X.$$

We will denote by $h_X : X \to [0,1]$, $h_Y : Y \to [0,1]$ the quotient maps: clearly $(h_X \otimes h_Y) : X \times Y \to [0,1]^2$ is the quotient map corresponding to the decomposition $X_\alpha \times Y_\beta$, $\alpha, \beta \in [0,1]$, of $X \times Y$.

Assumption 1. The maps h_X, h_Y are μ-measurable, ν-measurable from (X, Ω, μ), (Y, Σ, ν) to $([0,1], \mathcal{B})$, respectively, where $\mathcal{B}$ is the Borel σ-algebra.

We will consider the following disintegrations:

$$\mu = \int_0^1 \mu_\alpha m_X(d\alpha), \qquad m_X = (h_X)_\sharp \mu; \tag{7.4a}$$

$$\nu = \int_0^1 \nu_\beta m_Y(d\beta), \qquad m_Y = (h_Y)_\sharp \nu; \tag{7.4b}$$

$$\pi = \int_{[0,1]^2} \pi_{\alpha\beta} n(d\alpha d\beta), \qquad n = (h_X \otimes h_Y)_\sharp \pi. \tag{7.4c}$$

Note the fact that under the assumptions of measurability of h_X, h_Y, Theorem 2.7 implies that — up to a redefinition of μ_α, ν_α, π_α on respectively m_X, m_Y, n negligible sets — the conditional probabilities μ_α, ν_β and $\pi_{\alpha,\beta}$ satisfy

$$\mu_\alpha(X_\alpha) = \nu_\beta(Y_\beta) = \pi_{\alpha\beta}(X_\alpha \times Y_\beta) = 1 \tag{7.5}$$

for all $(\alpha, \beta) \in [0,1]^2$, i.e. they are concentrated on equivalence classes: in the following we will say that the disintegration is *strongly consistent* when the conditional probabilities are supported on the respective equivalence classes (see [Fre2], Chapter 45, Definition 452E).

The next Lemma 7.5 is valid also in the case the disintegration is not strongly consistent but just consistent, by considering the quotient measure space of Definition 2.5.

Lemma 7.5. *The measure* n *belongs to* $\Pi(m_X, m_Y)$.

Proof. This is a trivial consequence of the computation

$$n(A \times [0,1]) \overset{(7.4c)}{=} \pi(h_X^{-1}(A) \times Y) \overset{\pi \in \Pi(\mu,\nu)}{=} \mu(h_X^{-1}(A)) \overset{(7.4a)}{=} m_X(A).$$

The same computation works for $n([0,1] \times B)$. □

In the next sections, a special choice of the equivalence classes will lead to the following particular case, which under Assumption 1 is meaningful: indeed, as direct consequence of the properties of product σ-algebra (Theorem 3 in [HJ]), the set $\{\alpha = \beta\}$ belongs to the product σ-algebra $(h_X)_\sharp(\Omega) \otimes (h_Y)_\sharp(\Sigma)$ if and only if Assumption 1 holds.

Assumption 2. We assume $n = (\mathbb{I}, \mathbb{I})_\sharp m_X$.
In particular the marginals m_X and m_Y coincide: we will denote this probability measure by m.

Hence the image of $\Pi(\mu,\nu)$ under $(h_X \otimes h_Y)$ is contained in the set $\Pi(m,m)$ by Lemma 7.5. Moreover:

Lemma 7.6. *Under Assumption 2, one has $\pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha)$.*

Proof. By the marginal conditions, for any m-measurable A and Borel S

$$\int_A \mu_\alpha(S) m(d\alpha) = \mu(h_X^{-1}(A) \cap S)$$
$$\overset{\pi \in \Pi(\mu,\nu)}{=} \pi((h_X^{-1}(A) \cap S) \times [0,1]) = \int_A \pi_\alpha(S \times [0,1]) m(d\alpha).$$

Thus $(P_1)_\sharp \pi_\alpha = \mu_\alpha$ for m-a.e. α. For ν_α it is analogous. □

Under Assumption 1, a necessary and sufficient condition for Assumption 2 is the following.

Definition 7.7. We say that a set $\Gamma \subset [0,1]^2$ satisfies the *crosswise condition* w.r.t. the family $\{X_\alpha\}_{\alpha\in[0,1]}$, $\{Y_\beta\}_{\beta\in[0,1]}$, if

$$\Gamma \cap (X_\alpha \times Y) = \Gamma \cap (X \times Y_\alpha) = \Gamma \cap (X_\alpha \times Y_\alpha) \quad \forall \alpha \in [0,1]. \tag{7.6}$$

Lemma 7.8. *Assume that there exists $\Gamma \subset [0,1]^2$ such that $\pi(\Gamma) = 1$ and it satisfies the crosswise condition (7.6). Then $n = (\mathbb{I}, \mathbb{I})_\sharp m$, where $m = m_X = m_Y$.*

Conversely, if $n = (\mathbb{I}, \mathbb{I})_\sharp m$, then there exists $\Gamma \subset [0,1]^2$ such that $\pi(\Gamma) = 1$ and satisfying (7.6).

Proof. The proof follows the same line of the proof of Lemma 7.5.

The set $\Gamma' = (h_X \otimes h_Y)^{-1}(\{\alpha = \beta\})$ has full π measure if and only if $n = (h_X \otimes h_Y)_\sharp \pi = (\mathbb{I}, \mathbb{I})_\sharp m$.
Since (7.6) implies immediately $\Gamma \subset \Gamma'$, then $n = (\mathbb{I}, \mathbb{I})_\sharp m$.

Conversely, by the definition of Γ'

$$(X_\alpha \times Y) \cap \Gamma' = \Gamma' \cap (X \times Y_\alpha) = X_\alpha \times Y_\alpha.$$

This implies the (7.6) for the set Γ'. □

Along with the strong consistency of the disintegration (Assumption 1), the main assumption is the following. This assumption requires Assumption 1 and implies Assumption 2.

Assumption 3. For all $\pi \in \Pi^f(\mu,\nu)$, the image measure $n = (h_X \otimes h_Y)_\sharp \pi$ is equal to $(\mathbb{I}, \mathbb{I})_\sharp m$.

So far we do not have specified the criteria to choose the partitions X_α, Y_β. The next lemma, which is the key point of the argument, specifies it.

Lemma 7.9. *Assume that the decompositions X_α, Y_β satisfy Assumption 3 and the following:*

Assumption 4. For m-a.e $\alpha \in [0,1]$ the probability measure $\pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha)$ satisfies sufficient conditions for extremality/uniqueness/optimality.

Then $\pi \in \Pi(\mu,\nu)$ is extremal/unique/optimal.

Proof. We consider the cases separately.

Extremality. If $\pi_1, \pi_2 \in \Pi(\mu,\nu)$ are such that $\pi = (1-\lambda)\pi_1 + \lambda\pi_2$, $\lambda \in (0,1)$, then it follows from Assumption 3 that the disintegration of these measures is given by

$$\pi_1 = \int_0^1 \pi_{1,\alpha} m(d\alpha), \quad \pi_2 = \int_0^1 \pi_{2,\alpha} m(d\alpha).$$

It follows that $\pi_\alpha = (1-\lambda)\pi_{1,\alpha} + \lambda\pi_{2,\alpha}$ for m-a.e. α, so that from Assumption 1 and Assumption 4 we conclude that $\pi_\alpha = \pi_{1,\alpha} = \pi_{2,\alpha}$.

Uniqueness. The computations are similar to the previous case, only using the fact that in each class the conditional probability π_α is unique.

Optimality. For $\pi_1 \in \Pi^f(\mu,\nu)$

$$\mathcal{I}(\pi_1) = \int c(x,y)\pi_1(dxdy) \overset{(7.4c)}{=} \int_0^1 \left(\int c(x,y)\pi_{1,\alpha}(dxdy) \right) m(d\alpha).$$

From Assumption 3 it follows that $\pi_{1,\alpha}, \pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha)$, so that from Assumption 1 and Assumption 4 one has

$$\int c(x,y)\pi_{1,\alpha}(dxdy) \geqslant \int c(x,y)\pi_\alpha(dxdy) \qquad \text{for } m\text{-a.e. } \alpha.$$

The conclusion follows. □

We thus are left to perform the following steps in each of the next sections.

Procedure to verify the sufficiency of necessary conditions

1. Fix the necessary conditions under consideration.
2. Fix a measure $\pi \in \Pi^f(\mu,\nu)$ which satisfies the necessary conditions respectively for being extremal, being the unique measure concentrated on A, being optimal.
3. Construct partitions X_α, Y_β of X, Y such that:
   a) the disintegrations of μ, ν w.r.t. $X = \cup_\alpha X_\alpha$, $Y = \cup_\beta Y_\beta$ are strongly consistent. This implies that the quotient maps h_X, h_Y can be assumed to be measurable functions taking values in $([0,1], \mathcal{B})$, by Theorem 2.7;
   b) in each equivalence class $X_\alpha \times Y_\alpha$ the necessary conditions become sufficient: the measure $\pi_{\alpha\alpha}$ satisfies the sufficient conditions for extremality, uniqueness or optimality among all $\pi \in \Pi(\mu_\alpha, \nu_\alpha)$.
4. Verify that the image measure $n_{\pi'} \in \Pi(m,m)$ of all $\pi' \in \Pi^f(\mu,\nu)$ coincides with $(\mathbb{I},\mathbb{I})_\sharp m$, where $m = (h_X)_\sharp\mu = (h_Y)_\sharp\nu$.

If the above steps can be performed, then from Lemma 7.9 we deduce that π is respectively extremal, unique or optimal. In our applications, the necessary conditions reduce to a single condition on the structure of the support of π.

Remark 7.10. It is important to note that in general the decomposition depends on the particular measure π under consideration: the procedure will be used to test a particular measure π, even if in some cases it works for the whole $\Pi^f(\mu, \nu)$. In the latter case, we can test e.g. the optimality of all measures in $\Pi^f(\mu, \nu)$ using only the necessary conditions: this means that these conditions are also sufficient.

8 Extremality

The first problem we will consider is to give sufficient conditions for the extremality of transference plans in $\Pi(\mu,\nu)$. The results obtained are essentially the same as the results of [HW].

We first recall the following result ([Dou, Lin]), which we can prove by means of duality. Following the notation of Appendix B.3, we denote with $\Lambda \subset \mathcal{M}([0,1]^2)$ the set

$$\Lambda := \Big\{ \lambda \in \mathcal{M}([0,1]^2) : (P_1)_\sharp \lambda = (P_2)_\sharp \lambda = 0 \Big\}.$$

Proposition 8.1. *The transference plan $\pi \in \Pi(\mu,\nu)$ is extremal if and only if $L^1(\mu) + L^1(\nu)$ is dense in $L^1(\pi)$.*

Proof. We first prove that if $f_1 \in L^1(\mu)$, $f_2 \in L^1(\nu)$ and $(f_1 - f_2)\pi \in \Lambda$, then $f_1 - f_2 = 0$ π-a.e..

Writing

$$\pi = \int \pi_x \mu(dx) = \int \pi_y \nu(dy)$$

for the disintegration of π w.r.t. μ, ν respectively, the above condition means that

$$f_1(x) = \int f_2(y)\pi_x(dy) \quad \mu\text{-a.e. } x, \quad f_2(y) = \int f_1(x)\pi_y(dx) \quad \nu\text{-a.e. } y.$$

We then have

$$\begin{aligned} \int |f_1|\mu &= \int \left| \int f_2(y)\pi_x(dy) \right| \mu(dx) \\ &= \int |f_2|\nu + \int \left(\left| \int f_2(y)\pi_x(dy) \right| - \int |f_2(y)|\pi_x(dy) \right) \mu(dx) \leqslant \int |f_2|\nu, \end{aligned}$$

and similarly

$$\int |f_2|\nu = \int |f_1|\mu + \int \left(\left| \int f_1(x)\pi_y(dx) \right| - \int |f_1(x)|\pi_y(dx) \right) \nu(dy) \leqslant \int |f_1|\mu.$$

We thus conclude that

$$\begin{aligned} \left| \int f_2(y)\pi_x(dy) \right| &= \int |f_2(y)|\pi_x(dy) && \mu \text{ a.e. } x, \\ \left| \int f_1(x)\pi_y(dx) \right| &= \int |f_1(x)|\pi_y(dx) && \nu \text{ a.e. } y. \end{aligned}$$

i.e. π is concentrated on the set

$$\{f_1 \geqslant 0\} \times \{f_2 \geqslant 0\} \cup \{f_1 < 0\} \times \{f_2 < 0\}.$$

Since if (f_1, f_2) satisfies $(f_1 - f_2)\pi \in \Lambda$, also $[(f_1 - k) - (f_2 - k)]\pi \in \Lambda$ for all $k \in \mathbb{R}$, it follows that π is concentrated on the sets

$$\{f_1 \geqslant k\} \times \{f_2 \geqslant k\} \cup \{f_1 < k\} \times \{f_2 < k\}.$$

Hence one concludes that $f_1 - f_2 = 0$ π a.e..

$\Longleftarrow$ The previous step implies immediately that if $L^1(\mu) + L^1(\nu)$ is dense in $L^1(\pi)$, then π should be extremal: in fact, if $(f_1(x) + f_2(y))\pi \in \Pi(\mu, \nu)$, then

$$(1 - f_1(x) - f_2(y))\pi \in \Lambda,$$

so that $1 - f_1 - f_2 = 0$ and then $f_1 + f_2 = 1$ π-a.e..

$\Longrightarrow$ If instead $\overline{L^1(\mu) + L^1(\nu)} \subsetneq L^1(\pi)$, then by Hahn-Banach Theorem there exists an $L^\infty(\pi)$ function g, $|g| \leqslant 1$, such that

$$\int g(x, y)(f_1(x) + f_2(y))\pi(dxdy) = 0$$

for all $f_1 \in L^1(\mu)$, $f_2 \in L^1(\nu)$. In particular $g\pi \in \Lambda$, $g \neq 0$ on a set of positive π-measure and

$$\pi = \frac{1+g}{2}\pi + \frac{1-g}{2}\pi,$$

where the two addends in the r.h.s. above belongs to $\Pi(\mu/2, \nu/2)$. $\square$

The second result is a consequence of Proposition B.15. A cyclic perturbation λ of a measure $\pi \in \Pi(\mu, \nu)$ is specified in Definitions B.6, B.14; in particular $\pi + \lambda \in \Pi(\mu, \nu)$.

Definition 8.2 (Acyclic set and measure)**.** We say that $\Gamma \subset [0, 1]^2$ is *acyclic* if for all finite sequences $(x_i, y_i) \in \Gamma$, $i = 1, \ldots, n$, with $x_i \neq x_{i+1 \mod n}$ and $y_i \neq y_{i+1 \mod n}$ it holds

$$\Big\{(x_{i+1}, y_i), i = 1, \ldots, n, x_{n+1} = x_1\Big\} \not\subset \Gamma.$$

A measure is *acyclic* if it is concentrated on an acyclic set.

Lemma 8.3 (Th. 3 of [HW])**.** *Suppose that there is no cyclic perturbation of the measure $\pi \in \Pi(\mu, \nu)$ on $[0, 1]^2$. Then π is concentrated on an acyclic σ-compact set Γ.*

We specify now necessary and sufficient conditions for extremality:

NECESSARY CONDITION the measure π is acyclic;

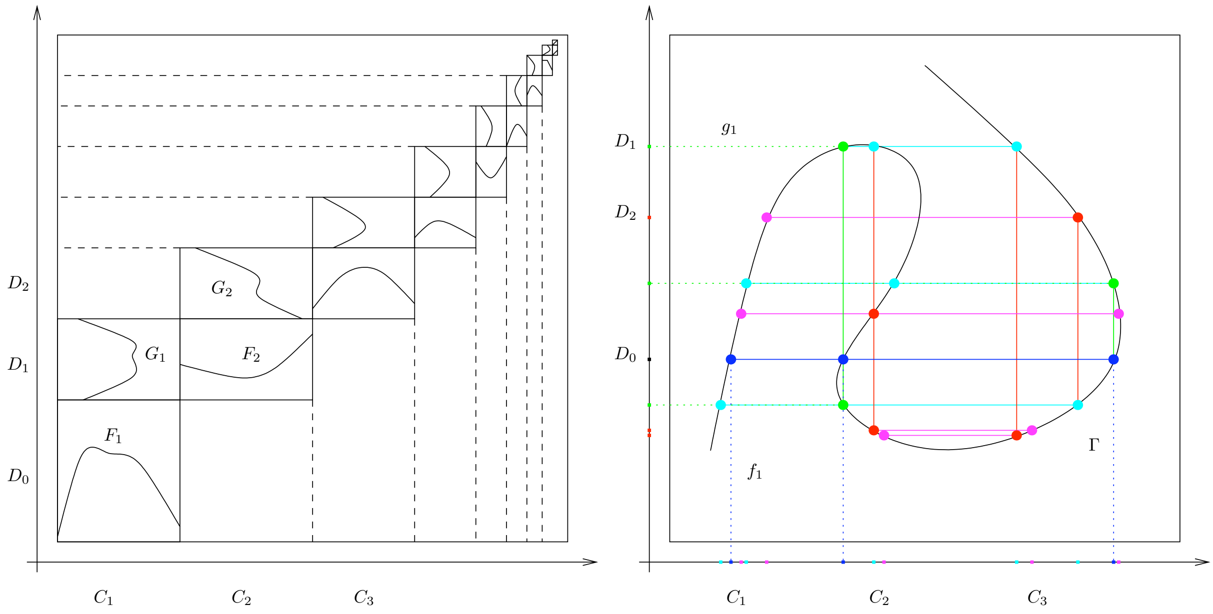


Figure 9: A limb numbering system and the axial path of a point

SUFFICIENT CONDITION the measure π is concentrated on a *Borel limb numbering system*, [HW] page 223: there are two disjoint families $\{C_k\}_{k\in\mathbb{N}_0}$, $\{D_k\}_{k\in\mathbb{N}_0}$ of Borel sets and Borel measurable functions $f_k : C_{k+1} \to D_k$, $g_k : D_{k+1} \to C_{k+1}$, $k \in \mathbb{N}_0$, such that π is concentrated on the union of the following graphs

$$F_k = \mathrm{graph}(f_k), \quad G_k = \mathrm{graph}(g_k).$$

We verify directly the second condition, [HW] Theorem 20: clearly due to the σ-additivity and inner regularity, we can always replace measurable with σ-compact sets up to a negligible set.

Proof of sufficiency of the condition. Assume first that there are only finitely many G_k, F_k, $k \leqslant N$. In this case, the uniqueness of the transference plan π follows by finite recursion, since the marginality conditions yield, setting $F_{N+1} := \emptyset$, that π must be defined by

$$\pi\llcorner_{F_k} = (\mathbb{I}, f_k)_\sharp(\mu - (P_1)_\sharp\pi\llcorner_{G_k}), \quad \pi\llcorner_{G_k} = (g_k, \mathbb{I})_\sharp(\nu - (P_2)_\sharp\pi\llcorner_{F_{k+1}}), \qquad k \in \{1, \dots, N\}. \tag{8.1}$$

For the general case, let $\pi \in \Pi(\mu, \nu)$ such that $\pi(\cup_k F_k \cup G_k) = 1$. Define the measures π_N by means of (8.1) starting at N: let

$$(\pi_N)\llcorner_{F_{N+1}} := (\mathbb{I}, f_{N+1})_\sharp\mu\llcorner_{F_{N+1}}$$

and for $k \in \{1, \dots, N\}$

$$(\pi_N)\llcorner_{F_k} := (\mathbb{I}, f_k)_\sharp(\mu - (P_1)_\sharp(\pi_N)\llcorner_{G_k}), \quad (\pi_N)\llcorner_{G_k} = (g_k, \mathbb{I})_\sharp(\nu - (P_2)_\sharp(\pi_N)\llcorner_{F_{k+1}}).$$

Since $\sum_{k>N} \mu(F_k) + \nu(G_k) \to 0$ as $N \to \infty$, it is fairly easy to see that π_N converges strongly to π.

Using the uniqueness of the limit, the uniqueness of π follows. □

The equivalence classes in order to apply Theorem 7.1 are the following.

Definition 8.4 (Axial equivalence relation)**.** We define $(x,y)E(x',y')$ if there are $(x_i,y_i) \in \Gamma$, $0 \leqslant i \leqslant I$ finite, such that

$$(x,y) = (x_0,y_0), (x',y') = (x_I,y_I) \quad \text{and} \quad (x_{i+1} - x_i)(y_{i+1} - y_i) = 0. \tag{8.2}$$

In the language of [HW], page 222, each equivalence class is an *axial path*. The next lemma is an elementary consequence of Definition 8.4.

Lemma 8.5. *The relation E of Definition 8.4 defines an equivalence relation on the acyclic set Γ. If $\Gamma = \cup_\alpha \Gamma_\alpha$ is the partition of Γ in equivalence classes, and $X_\alpha = P_1\Gamma_\alpha$, $Y_\alpha = P_2\Gamma_\alpha$ are the projections of the equivalence classes, then the crosswise condition (7.6) holds.*

By setting

$$X_0 = [0,1] \setminus P_1\Gamma, \quad Y_0 = [0,1] \setminus P_2\Gamma,$$

we have a partition of X, Y into disjoint classes.

We can thus use Theorem 2.7 to disintegrate the marginals μ, ν and every transference π plan supported on Γ. From (7.4) and Lemmata 7.6, 7.8 one has immediately the following proposition.

Proposition 8.6. *The following disintegrations w.r.t. the partitions $X = \cup_\alpha X_\alpha$, $Y = \cup_\alpha Y_\alpha$ hold:*

$$\mu = \int \mu_\alpha m(d\alpha), \quad \nu = \int \nu_\alpha m(d\alpha), \quad m = (h_X)_\sharp \mu.$$

Moreover, if π is a transference plan supported on Γ, then the disintegration of π w.r.t. the partition $\Gamma = \cup_{\alpha\in A}\Gamma_\alpha$ is given by

$$\pi = \int \pi_\alpha m(d\alpha)$$

with $\pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha)$.

The next lemma shows that in each equivalence class the sufficient condition holds.

Lemma 8.7. *Each equivalence class satisfies the Borel limb numbering condition.*

Proof. The proof is elementary: if $(x_\alpha, y_\alpha) \in \Gamma_\alpha$, then one defines recursively (Figure 9)

$$D_{0,\alpha} = \{y_\alpha\}, \quad C_{1,\alpha} = P_1(\Gamma \cap ([0,1] \times \{y_\alpha\})),$$

$$D_{k,\alpha} = P_2\big(\Gamma \cap (C_{k,\alpha} \times [0,1])\big) \setminus D_{k-1,\alpha}, \quad C_{k+1,\alpha} = \Gamma \cap ([0,1] \times D_{k,\alpha}) \setminus C_{k,\alpha}.$$

From the assumption of acyclicity, it follows immediately that each $\Gamma \cap (C_{k,\alpha} \times [0,1])$, $\Gamma \cap ([0,1] \times D_{k,\alpha})$ is the graph of a function $g_{k+1,\alpha} : C_{k+1,\alpha} \to D_{k,\alpha}$, $f_{k,\alpha} : D_{k,\alpha} \to C_{k,\alpha}$. Moreover, Γ_α is covered by the graphs $G_{k,\alpha}$, $F_{k,\alpha}$ because of the definition of the equivalence class Γ_α.

It remains to study the measurability of the functions $g_{k,\alpha}$, $f_{k,\alpha}$. First of all, if Γ is Borel, then each $\tilde{C}_{k,\alpha}$, $\tilde{D}_{k,\alpha}$ is analytic, being the projection of the analytic set $\Gamma \cap ([0,1] \times \tilde{D}_{k-1,\alpha})$, $\Gamma \cap (\tilde{C}_{k,\alpha} \times [0,1])$, respectively. Note that under the assumption of σ-compactness of Γ, then each class is actually σ-compact.

Hence each function $g_{k,\alpha}$ ($f_{k,\alpha}$) is μ_α-measurable (ν_α-measurable), so that up to a negligible set w.r.t. μ_α (ν_α) it can be taken to be Borel: clearly such a redefinition does not alter the marginal conditions. $\square$

From Lemma 7.9, it follows the following theorem.

Theorem 8.8. *If the disintegration of Proposition 8.6 is strongly consistent, then π is extremal.*

We now conclude the section showing that the existence of a Borel limb numbering systems is equivalent to the existence of an acyclic set Γ where the transference plan π is concentrated and such that the disintegration is consistent.

Theorem 8.9. *The transference plans π is concentrated on a limb numbering system Γ with Borel limbs if and only if the disintegration of π into the equivalence classes of any acyclic carriage Γ is strongly consistent.*

Proof. Assume first that π satisfies the Borel limb condition. Then from [HW], Theorem 20, it follows we can take as quotient space a Borel root set A. In particular Γ can be taken as the union of the orbits of points in A, and it is immediate to verify that the orbit of a Borel subset of A is an analytic subset of $[0,1]^2$. Hence the disintegration is consistent by the fact that $(A, \mathcal{B}(A), m)$ is a countably generated measure space.

Conversely, if the disintegration is strongly consistent, then, as a consequence of Proposition 2.9, by eventually removing a set of π-measure 0 the graph of the equivalence relation E can be taken to be Borel in the product space $\Gamma \times [0,1]$, so that there exists a measurable selection $[0,1] \ni \alpha \mapsto (x(\alpha), y(\alpha)) \in \Gamma$. Up to neglecting sets of measure 0, we can assume that $\alpha \mapsto (x(\alpha), y(\alpha))$ is Borel and the image set $\{(x(\alpha), y(\alpha)\}_{\alpha \in [0,1]}$ is σ-compact. One constructs then Borel limbs as in Lemma 8.7 from $\{(x(\alpha), y(\alpha)\}_{\alpha \in [0,1]}$. $\square$

Remark 8.10. We observe that by adding the set $G_0 = x_0 \times D_0$, where $x_0 \notin \cup_k C_k$, the disintegration is supported on a single equivalence class.

9

Uniqueness

In this section we address the question of uniqueness of transference plans concentrated on a set A.

Definition 9.1 (Set of uniqueness)**.** We say that $A \in \Theta$ is a *set of uniqueness* of $\Pi(\mu, \nu)$ if there exists a unique measure $\pi \in \Pi(\mu, \nu)$ such that $\pi(A) = 1$.

In Section 5 of [HW] (or using directly the proof of the sufficient condition, page 113) it is shown that if Γ satisfies the Borel limb condition, then Γ supports a unique transference plan.

The first lemma is a consequence of Proposition B.15.

Definition 9.2. A set $\Gamma \subset A$ is A-*acyclic* if for all finite sequences $(x_i, y_i) \in \Gamma$, $i = 1, \dots, n$, with $x_i \neq x_{i+1 \mod n}$ and $y_i \neq y_{i+1 \mod n}$ it holds

$$\Big\{(x_{i+1}, y_i), i = 1, \dots, n, x_{n+1} = x_1\Big\} \not\subset A.$$

A measure is A-*acyclic* if it is concentrated on an A-acyclic set.

Lemma 9.3. *If an analytic set A is a set of uniqueness for $\Pi(\mu, \nu)$, then the unique $\pi \in \Pi(\mu, \nu)$ is concentrated on a A-acyclic Borel set $\Gamma \subset A$.*

Necessary and sufficient condition for uniqueness are then given by:

NECESSARY CONDITION there exist a measure $\pi \in \Pi(\mu, \nu)$ and an A-acyclic Borel set $\Gamma \subset A$ such that $\pi(\Gamma) = 1$;

SUFFICIENT CONDITION A is a Borel limb numbering system (Page 113).

We will state a more general sufficient condition later at Page 119.

Let Γ as above. In particular, Γ is acyclic. We will thus use the equivalence classes of the axial equivalence relation E on Γ, Definition 8.4, assuming w.l.o.g. that $P_X\Gamma = P_Y\Gamma = [0, 1]$.

Let $h_X : X \to [0, 1]$, $h_Y : Y \to [0, 1]$ be the quotient maps. In general the image of A

$$A' := \Big\{(\alpha, \beta) : (h_X \otimes h_Y)^{-1}(\alpha, \beta) \cap A \neq \emptyset\Big\} \tag{9.1}$$

is not a subset of $\{\alpha = \beta\}$. However, for the equivalence classes in the diagonal $\{\alpha = \beta\}$, we have the following lemma.

Lemma 9.4. *For all $\alpha \in [0, 1]$,*

$$(h_X \otimes h_Y)^{-1}(\alpha, \alpha) \cap A = (h_X \otimes h_Y)^{-1}(\alpha, \alpha) \cap \Gamma.$$

Proof. The definition implies that if $x, x' \in h_X^{-1}(\alpha)$, then there exist $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I$, with $x_0 = x$, such that denoting $x_I = x'$ then (8.2) holds. A completely similar condition is valid for $y, y' \in h_Y^{-1}(\alpha)$.

Let $(\bar{x}, \bar{y}) \in (h_X^{-1}(\alpha) \times h_Y^{-1}(\alpha)) \cap (A \setminus \Gamma)$. Then there are $(x, y), (x', y') \in \Gamma$ such that $x = \bar{x}$, $y' = \bar{y}$. Consider then the axial path $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I = 2(n-1)$, connecting them inside the class α: removing by chance some points, we can assume that $(x_0, y_0) = (x, y)$, $(x_I, y_I) = (x', y')$ and

$$x_{2j} - x_{2j-1} = 0, \quad y_{2j-1} - y_{2j-2} = 0, \quad j = 1, \dots, n.$$

Hence if we add the point $(x_{I+1}, y_{I+1}) = (\bar{x}, \bar{y})$ we obtain a closed cycle, contradicting the hypotheses of acyclicity of Γ in A. $\square$

The above lemma together with Lemma 7.9 implies that non uniqueness occurs because of the following two reasons:

1. either the disintegration is not strongly consistent,
2. or the push forward of some transference plan $\pi \in \Pi(\mu, \nu)$ such that $\pi(A) = 1$ is not supported on the diagonal in the quotient space.

In the following we address the second question, and we assume that the disintegration is strongly consistent — which is equivalent to assume that the quotient maps h_X, h_Y can be taken Borel (up to a μ, ν negligible set, respectively, consequence of Proposition 2.9).

Lemma 9.5. *The set A' defined in (9.1) is analytic if A is analytic.*

Proof. Since $A' = (h_X, h_Y)(A)$, the proof is a direct consequence of the fact that Borel images of analytic sets are analytic, being the projection of a Borel set. $\square$

The next lemma is a consequence of the acyclicity of Γ in A.

Lemma 9.6. *In the quotient space, the diagonal is A'-acyclic.*

Proof. We prove the result only for 2-cycles, the proof being the same for the n-cycles.

Assume that A' has a 2-cycle, between the classes (α, α) and (α', α'). This means that there are points $(x, y) \in (h_X \otimes h_Y)^{-1}(\alpha, \alpha') \cap A$ and $(x', y') \in (h_X \otimes h_Y)^{-1}(\alpha', \alpha) \cap A$.

By definition of equivalence class, there are points $(x_i, y_i) \in (h_X \otimes h_Y)^{-1}(\alpha, \alpha)$, $i = 1, \dots, n$, and $(x'_j, y'_j) \in (h_X \otimes h_Y)^{-1}(\alpha', \alpha')$, $j = 1, \dots, n'$ forming an axial path in Γ and connecting (x, y) to (x', y') in $(h_X \otimes h_Y)^{-1}(\alpha, \alpha)$ and (x, y) to (x', y') in $(h_X \otimes h_Y)^{-1}(\alpha', \alpha')$.

The composition of the two axial paths yields a closed cycle, contradicting the assumption of acyclicity of Γ in A. $\square$

Differently from the previous section, *the consistency of the disintegration is not sufficient to deduce the uniqueness of the transference plan.*

Example 9.7 (Pratelli). Consider $\mu = \mathcal{L}^1$ and the set

$$A = \{x = y\} \cup \{y - x = \alpha \mod 1\}, \quad \Gamma = \{x = y\} \quad \text{with } \alpha \in [0,1] \setminus \mathbb{Q}.$$

In this case the quotient map is the identity, but the measure $(x, x + \alpha \mod 1)_\sharp \mathcal{L}^1$ is not concentrated on the diagonal and still belongs to $\Pi(\mathcal{L}^1, \mathcal{L}^1)$.

We now give a sufficient condition for the implication

$$n \in \Pi(m, m),\ n(A') = 1 \quad \Longrightarrow \quad n(\{\alpha = \beta\}) = 1. \tag{9.2}$$

We use the following easy observation:

Lemma 9.8. *If* $\pi = (\mathbb{I}, f)_\sharp \mu$ *and* $A = \operatorname{epi}(f)$*, then* A *is a set of uniqueness of* $\Pi(\mu, \nu)$*, where* $\nu = f_\sharp(\mu)$.

Proof. For all $a \in \mathbb{R}$, $\pi \in \Pi(\mu, \nu)$, it holds

$$\nu((a, 1]) = \mu(f^{-1}(a, 1]) = \pi\big(f^{-1}(a, 1] \times (a, 1]\big),$$

so that $\pi(f^{-1}[0, a] \times (a, 1]) = 0$ for all $\pi \in \Pi(\mu, \nu)$. Since

$$\operatorname{epi}(f) \setminus \operatorname{graph}(f) = \bigcup_{q \in \mathbb{Q}} f^{-1}[0, q] \times (q, 1],$$

π is concentrated on $\operatorname{graph}(f)$ and the result follows. □

Hence, our sufficient condition for uniqueness is the following:

Sufficient condition for uniqueness: A' is a subset of $\{\alpha \leqslant \beta\}$.

Clearly, it is enough that $(f, f)(A') \subset \{\alpha \leqslant \beta\}$, where f is an isomorphism of measurable spaces between $([0,1], \Theta_m)$ and $([0,1], \Theta_{f_\sharp m})$ — e.g. f is injective m-a.e. or a measure preserving map.

One could ask whether it is sufficient to require that A' can be completed to a Borel linear order to $[0,1]$. As the diagonal is A'-acyclic, and therefore by the Axiom of Choice it can be completed to a linear order, this would be again a measurability assumption.

An easy case is covered by the next lemma.

Lemma 9.9. *If* A' *is acyclic and* m *atomic, then* A' *is a subset of a Borel linear order on* $[0,1]$ *and a set of uniqueness.*

Proof. Let α_n be the atoms of m. The set A' can be interpreted as the graph of a relation on $\mathbb{N}$ by setting mRn if $(\alpha_m, \alpha_n) \in A'$. Denote again by A' the corresponding subset of $\mathbb{N} \times \mathbb{N}$. This establish then a relation on the atoms α_n.

We complete the relation in the following way: let A'' be a maximal set containing A' such that the diagonal is A''-acyclic. Since we are working in a countable space, A'' exists (but in general it is not unique) and it can be obtained by adding one point at most countably many times.

It is easy to verify that A'' is a partial order relation extending A': by the acyclicity, mRn implies that $(m, n) \notin A''$, and if mRn, nRo then the point (m, o)

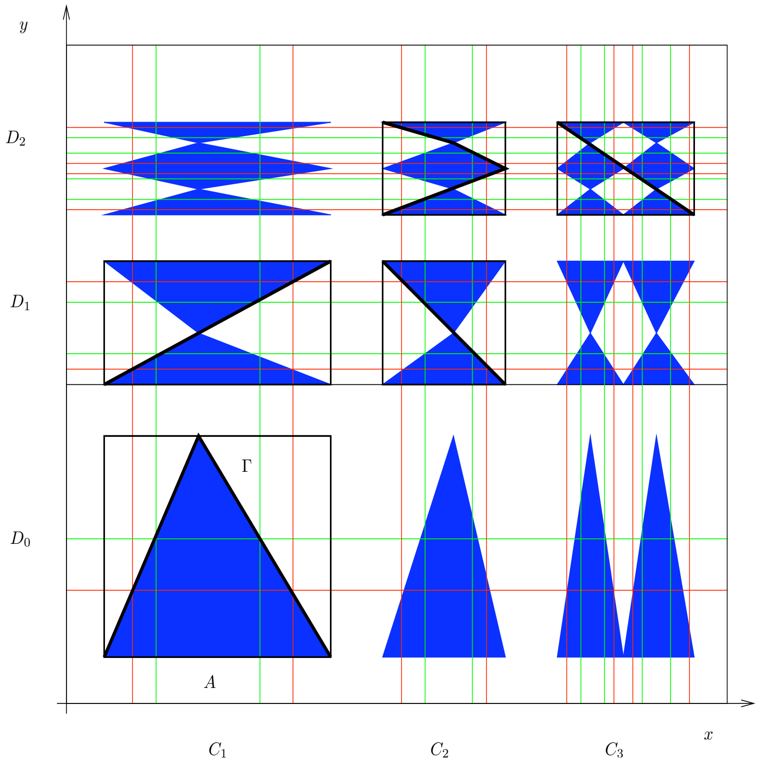


Figure 10: The set where A should be contained in order to have that A' is a subset of the epigraph of the function $\beta = \alpha$. The bold curves are the limbs of Γ, and two axial path are represented.

can be added to A'' without creating a closed cycle involving points on the diagonal.

Assume that $(n, m), (m, n) \notin A''$. Then we can add arbitrarily one of the points (m, n) or (n, m) without losing the acyclicity.

We conclude that R is a linear order relation on a countable set. One can clearly extend it to a Borel linear order $B \subset [0, 1]^2$ to $[0, 1]$.

Having an atomic measure, the map $f : x \mapsto m(B \cap \{x\} \times [0, 1])$ is an isomorphism of measurable spaces between $([0, 1], \Theta_m)$ and $([0, 1], \Theta_{f_\sharp m})$: since $(f, f)(B) \supset (f, f)(A')$ is contained in the epigraph $\{\alpha \leqslant \beta\}$ of $(\mathbb{I}, \mathbb{I})$ we conclude that $(f, f)(A')$ is a set of uniqueness, and thus also A' is a set of uniqueness. $\square$

An example of a set A for which A' is a set of uniqueness is presented in Figure 10. By setting

$$c(x, y) = \begin{cases} 1 & \Gamma \\ 0 & A \setminus \Gamma \\ +\infty & [0, 1]^2 \setminus A \end{cases}$$

the uniqueness of the transport plan in A is related to a problem of optimality.

10 Optimality and c-cyclical monotonicity

The last problem we want to address is the problem of optimality of a measure $\pi \in \Pi(\mu,\nu)$ w.r.t. the functional $\mathcal{I}$ defined in (7.2). We recall that a plan $\pi \in \Pi(\mu,\nu)$ is said to be optimal if

$$\mathcal{I}(\pi) = \int c(x,y)\pi(dxdy) = \min_{\tilde{\pi}\in\Pi(\mu,\nu)} \mathcal{I}(\tilde{\pi}).$$

In this section the function c is assumed to be a Π^1_1-function.

Definition 10.1 (Cyclical monotonicity)**.** A subset Γ of $[0,1]^2$ is c-*cyclically monotone* when for all I, $i = 1,\dots,I$, $(x_i,y_i) \in \Gamma$, $x_{I+1} := x_1$ we have

$$\sum_{i=1}^{I} \big[c(x_{i+1},y_i) - c(x_i,y_i)\big] \geqslant 0. \tag{10.1}$$

A transference plan $\pi \in \Pi(\mu,\nu)$ is c-*cyclically monotone* if there exists a π-measurable c-cyclically monotone set Γ such that $\pi(\Gamma) = 1$.

As usual, by inner regularity and by the fact that for π fixed c coincides with a Borel function up to a negligible set, the set Γ for that given measure π can be taken σ-compact and $c \llcorner \Gamma$ Borel.

We recall that a necessary condition for being optimal is that the measure is concentrated on a c-cyclically monotone set: we follow the ideas of [BGMS], which reduce to ones in [Pra] for atomic marginals (see [Kel] for the general result). A proof is provided for completeness in Lemma B.16.

Lemma 10.2. *If π is optimal, then it is c-cyclically monotone.*

Having a necessary condition which gives some structure to the problem, we have to specify a sufficient condition which should be tested in each equivalence class. We list some important remarks.

1. The optimality is implied by the fact that there exists a sequence of functions $\phi_n \in \mathcal{L}^1(\mu)$, $\psi_n \in \mathcal{L}^1(\nu)$ such that $\phi_n(x) + \psi_n(y) \leqslant c(x,y)$ and

$$\int \phi_n\mu + \int \psi_n\nu = \int (\phi_n + \psi_n)\pi \nearrow \int c\pi.$$

2. For l.s.c. costs or costs satisfying $c(x,y) \leqslant f(x) + g(y)$, $f \in L^1(\mu)$ and $g \in L^1(\nu)$-measurable, the converse of Point (1) holds.

3. Another condition is that there is an optimal pair $\phi, \psi : [0,1] \to [-\infty, +\infty)$, respectively μ-measurable and ν-measurable, such that $\phi(x) + \psi(y) \leqslant c(x,y)$ for all $(x,y) \in [0,1]^2$ and $\phi(x) + \psi(y) = c(x,y)$ π-a.e..

For completeness we prove the sufficiency of the last condition, proved also in [BS].

Lemma 10.3. *Suppose there exists Borel functions* $\phi, \psi : [0,1] \to [-\infty, +\infty)$ *and* $\Gamma \subset [0,1]^2$ *such that*

$$\phi(x) + \psi(y) < c(x,y) \qquad \forall (x,y) \in [0,1]^2 \setminus \Gamma \tag{10.2a}$$

$$\phi(x) + \psi(y) = c(x,y) \qquad \forall (x,y) \in \Gamma. \tag{10.2b}$$

If $\exists \pi \in \Pi^f(\mu, \nu)$ *such that* $\pi(\Gamma) = 1$, *then*

$$\pi \in \Pi(\mu,\nu) \text{ optimal} \quad \Longleftrightarrow \quad \pi(\Gamma) = 1.$$

It is trivial to extend the proposition to the case of $\phi : [-\infty, +\infty)$ μ-measurable and $\psi : [-\infty, +\infty)$ ν-measurable, just redefining the functions on negligible sets in order to be Borel.

Proof. Let $\bar{\pi}$ be an optimal transference plan and $\pi \in \Pi^f(\mu,\nu)$ concentrated on Γ.

Step 1. If $\lambda \in \Lambda$ and $\psi\lambda$, $(\phi + \psi)\lambda$ are Borel measures, then $\int \{\phi + \psi\}\lambda = 0$. Since $(\phi + \psi)\lambda$ is a Borel measure, one can consider the following integrals

$$\int_{[0,1]^2} \{\phi + \psi\}\lambda = \lim_M \int_{\{|\phi| < M\}} \{\phi + \psi\}\lambda.$$

Since also $\psi\lambda$ is a Borel measure

$$\begin{aligned} \lim_M \int_{\{|\phi|<M\}} \{\phi + \psi\}\lambda &= \lim_M \left\{ \int_{\{|\phi|<M\}} \phi\lambda + \int_{\{|\phi|<M\}} \psi\lambda \right\} \\ &\overset{\lambda \in \Lambda}{=} \lim_M \int_{\{|\phi|<M\}} \psi\lambda = \int \psi\lambda = \lim_M \int_{\{|\psi|<M\}} \psi\lambda = 0. \end{aligned}$$

Step 2 Let $\lambda := \bar{\pi} - \pi$.

Define $\phi_M := \phi \wedge M$ and $\psi_{-M} = \psi \vee (-M)$: it is immediate to verify that $\phi_M(x) + \psi_{-M}(y) \leqslant c(x,y)$; in particular, $\phi_M\lambda$ and $(\phi_M + \psi_{-M})\lambda$ are σ-finite Borel measures. Since $\phi_M(x) \leqslant M$, for $(x,y) \in \Gamma$ the relation $\psi(y) = c(x,y) - \phi(x)$ implies also $\phi_M(x) + \psi_{-M}(y) \geqslant 0$. As a consequence, $\phi_M + \psi_{-M}$ converges to c in $L^1(\pi)$, yielding immediately

$$\int_{[0,1]^2} c\lambda \geqslant \lim_M \int_{[0,1]^2} \{\phi_M + \psi_{-M}\}\lambda.$$

The r.h.s. vanishes by Step 1, showing the optimality of π:

$$0 \geqslant \mathcal{J}(\bar{\pi}) - \mathcal{J}(\pi) = \int_{[0,1]^2} c\lambda \geqslant \lim_M \int_{[0,1]^2} \{\phi_M + \psi_{-M}\}\lambda = 0.$$

$\square$

From the formulas

$$\phi(x,\bar{x},\bar{y}) = \inf\left\{\sum_{i=0}^{I} c(x_{i+1},y_i) - c(x_i,y_i), (x_i,y_i) \in \Gamma \text{ finite}, (x_0,y_0) = (\bar{x},\bar{y}),\ x_{I+1} = x\right\}, \tag{10.3a}$$

$$\psi(y,\bar{x},\bar{y}) = c(x,y) - \phi(x,\bar{x},\bar{y}), \tag{10.3b}$$

it is always possible to construct an optimal couple ϕ, ψ in an analytic subset of Γ containing $(\bar{x},\bar{y})$ such that $(-\phi,\psi)$ are Σ^1_1-functions (Remarks 10.15, 10.18). In Section 10.2 this idea is developed in a general framework.

Here we consider the easiest equivalence relation for which the procedure at Page 109 can be applied. This equivalence relation has been also used in [BGMS].

Definition 10.4 (Closed cycles equivalence relation)**.** We say that $(x,y)\bar{E}(x',y')$ or (x,y) *is equivalent to* (x',y') *by closed cycles* if there is a closed cycle with finite cost passing through them: there are $(x_i,y_i) \in \Gamma$ such that $(x_0,y_0) = (x,y)$ and $(x_j,y_j) = (x',y')$ for some $j \in \{0,\dots,I\}$ such that

$$\sum_{i=1}^{I} c(x_i,y_i) + c(x_{i+1},y_i) < +\infty, \qquad x_{I+1} := x_0.$$

It is easy to show that this is an equivalence relation, and it follows directly from (10.3) or the analysis of Section 10.2 that in each equivalence class there are optimal potentials ϕ, ψ.

Lemma 10.5. *The equivalence relation $\bar{E}$ satisfies the following.*

1. *Its equivalence classes are in Σ^1_1.*

2. *It satisfies the crosswise structure* (7.6)*.*

The above lemma can be seen as a straightforward consequence of Lemma 10.17 and Corollary 10.19 of Section 10.2. Since it is elementary, we give here a direct proof.

Proof. For the Point (1), just observe that for all $I \in \mathbb{N}$

$$\sum_{i=0}^{I} c(x_i,y_i) + c(x_{i+1},y_i), \quad x_{I+1} = x_0$$

is a Π^1_1-function, so that the set $Z_I(\bar{x},\bar{y})$ defined as

$$\left\{(x_1,y_1,\dots,x_I,y_I) \in \Gamma^I : \sum_{i=1}^{I} c(x_i,y_i) + c(x_{i+1},y_i) + c(\bar{x},\bar{y}) + c(x_1,\bar{y}) \in \mathbb{R}, x_{I+1} = \bar{x}\right\}$$

is in Σ^1_1.

The equivalence class of $(\bar{x}, \bar{y})$ is then given by

$$\bigcup_{I \in \mathbb{N}} \bigcup_{i=1}^{I} P_{2i-1,2i} Z_I \in \Sigma^1_1,$$

where we used the fact that Σ^1_1 is closed under projection and countable union (see Appendix B.1 or Chapter 4 of [Sri]).

The proof of Point (2) follows from the straightforward observations that $(x, y)\bar{E}(x', y)$ and $(x, y)\bar{E}(x, y')$ whenever $(x, y), (x', y), (x, y') \in \Gamma$: just consider the closed cycle with finite cost made of the two points $(x_0, y_0) := (x, y)$ and $(x_1, y_1) := (x', y)$, or $(x_1, y_1) := (x, y')$. □

Let now $\pi \in \Pi^f(\mu, \nu)$ be a c-cyclically monotone transference plan, and let Γ be a c-cyclically monotone set where π is concentrated. Let $\bar{E}$ be the equivalence class of Definition 10.4.

As in the previous section, non optimality can occur because of two reasons:

1. either the disintegration is not strongly consistent,
2. or the push forward of some measure $\pi' \in \Pi^{opt}(\mu, \nu)$ is not supported on the diagonal in the quotient space.

In the next section we give examples which show what can happen when one of the two situations above occurs. Here we conclude with two results, which yield immediately the optimality of π.

Let h_X, h_Y be the quotient maps. By redefining them on a set of measure 0, the condition of strong consistency implies that h_X, h_Y can be considered as Borel maps with values in $[0, 1]$. In particular, the set

$$A' := (h_X, h_Y)(\{c < +\infty\}) \tag{10.4}$$

is analytic. Note that

$$(h_X, h_Y)_\sharp \tilde{\pi}(A') = 1 \quad \forall \tilde{\pi} \in \Pi^f(\mu, \nu),$$

i.e. the transport plans with finite cost are concentrated on A', and moreover for the π under consideration

$$(h_X, h_Y)_\sharp \pi = (\mathbb{I}, \mathbb{I})_\sharp m,$$

where $m = (h_X)_\sharp \mu = (h_Y)_\sharp \nu$ by Lemma 10.5 and Lemma 7.8.

Theorem 10.6. *Assume that the disintegration w.r.t. the equivalence relation $\bar{E}$ is strongly consistent. If A' is a set of uniqueness in $\Pi(m, m)$, then π is optimal.*

The proof is a simple application of Lemma 7.9.

The next corollary is a direct consequence of Lemma 9.9.

Corollary 10.7. *If $m = (h_X)_\sharp \mu$ is purely atomic, then the c-cyclical monotone measure π is optimal.*

We now give a simple condition which implies that the image measure m is purely atomic.

Proposition 10.8. *Assume that c satisfies the following assumption: for π-a.e. (x, y) there exist Borel sets $A_{(x,y)} \subset X$, $B_{(x,y)} \subset Y$ such that*

$$\pi(A_{(x,y)} \times B_{(x,y)}) > 0, \quad (x, y) \in A_{(x,y)} \times B_{(x,y)},$$

and

$$\mu \otimes \nu\big((A_{(x,y)} \times B_{(x,y)}) \cap \{c = +\infty\}\big) = 0. \tag{10.5}$$

Then the image measure is purely atomic.

Proof. First of all, we can assume that the condition holds for all $(x, y) \in \Gamma$, where Γ is a c-cyclically monotone set such that $\pi(\Gamma) = 1$.

For all $(x, y) \in \Gamma$, this assumption and Fubini theorem imply that there is $\bar{x} \in A_{(x,y)}$ such that

$$\bar{B} := P_2((A_{(x,y)} \times B_{(x,y)} \cap \{c < +\infty\})_{\bar{x}}),$$

where $(A_{(x,y)} \times B_{(x,y)})_{\bar{x}} := (A_{(x,y)} \times B_{(x,y)}) \cap (\{\bar{x}\} \times [0, 1]))$, has full ν-measure in $B_{(x,y)}$, and then there are $\bar{y}_1, \bar{y}_2 \in \bar{B}$ such that

$$\bar{A} := P_2((A_{(x,y)} \times B_{(x,y)} \cap \{c < +\infty\})_{\bar{y}_1}) \cap P_2((A_{(x,y)} \times B_{(x,y)} \cap \{c < +\infty\})_{\bar{y}_2})$$

has full μ-measure in $A_{(x,y)}$. The functions ϕ, ψ given by formula (10.3) provide then potentials on the sets $\bar{A} \times \bar{B}$.

From the cross structure of the equivalence relation $\bar{E}$, it follows that m-a.e. equivalence class has positive measure, so that m is purely atomic. □

Remark 10.9. Let Γ be a c-cyclically monotone set where π is concentrated. The proof shows actually that in each equivalence class

$$\phi(x) + \psi(y) = c(x, y)$$

up to a cross-negligible set. This is clearly a stronger condition than $c_{\llcorner\Gamma} < +\infty$ π-a.e..

Remark 10.10. From the definition of the optimal couple $(\phi(\cdot, \bar{x}, \bar{y}), \psi(\cdot, \bar{x}, \bar{y}))$, we can define the following relation on $P_1(\Gamma)$.

Definition 10.11. We say that $x \geqslant_c x'$ if $\phi(x', x, y) < +\infty$: equivalently there are points $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I$ such that $x_0 = x$, $x_{I+1} = x'$ and $\sum_i c(x_{i+1}, x_i) + c(x_i, y_i) < +\infty$.

The result of this section can be rephrased as the fact that $\geqslant_c$ can be completed into a Borel relation R (up to cross negligible sets) such that

1. xRx;
2. for all x, x', at least xRx' or $x'Rx$.
3. xRx' and $x'Rx$ implies that they belong to a closed cycle with finite cost.

By considering the map

$$L_X(x) := \mu(\{y : xRy\})$$

it follows by Fubini theorem that $L_X : [0,1] \to [0,1]$ in μ-measurable and Condition (3) concludes that it is exactly our quotient map. Moreover, considering the twin map

$$L_Y(y) := L_X(x) \quad \text{for any } (x,y) \in \Gamma,$$

it is fairly easy to check that $A' \subset \{\alpha \leqslant \beta\}$.

10.1 Extension of the construction

The approach we are proposing can be generalized as follows.

Assumption 5. Assume that for any $(\bar{x}, \bar{y}) \in \Gamma$ there exist universally measurable subsets $A_{(\bar{x},\bar{y})}$,$B_{(\bar{x},\bar{y})}$ of $[0,1]$ and universally measurable functions $\phi_{(\bar{x},\bar{y})}$, $\psi_{(\bar{x},\bar{y})}$ satisfying

$$\phi_{(\bar{x},\bar{y})}(x) + \psi_{(\bar{x},\bar{y})}(y) < c(x,y) \qquad \forall (x,y) \in A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})} \setminus \Gamma \tag{10.6a}$$
$$\phi_{(\bar{x},\bar{y})}(x) + \psi_{(\bar{x},\bar{y})}(y) = c(x,y) \qquad \forall (x,y) \in A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})} \cap \Gamma. \tag{10.6b}$$

We can define the relation R

$$(x,y)R(x',y') \qquad \Longleftrightarrow \qquad (x',y') \in A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})}.$$

Assume that there exist partitions $\{X_\alpha\}_\alpha$, $\{Y_\alpha\}_\alpha$ of $[0,1]$ such that each $X_\alpha \times Y_{(x_\alpha,y_\alpha)} \subset A_{(x_\alpha,y_\alpha)} \times B_{(x_\alpha,y_\alpha)}$ for some $(x_\alpha, y_\alpha) \in \Gamma$. Then optimality holds if the equivalence relation induced by $\{X_\alpha \times Y_\beta\}_{\alpha,\beta}$ satisfies Assumptions 1, 2, 3, i.e. if the disintegrations w.r.t. $\{X_\alpha \times Y_\beta\}_{\alpha,\beta}$ is strongly consistent, $\pi(\cup_\alpha X_\alpha \times Y_\alpha) = 1$ and A' of (10.4) is a set of uniqueness.

A method for constructing a relation R satisfying Assumption 5 and the crosswise condition w.r.t. Γ (Definition 7.7) is exploited in Appendix 10.2.

A special case is covered by the following theorem.

We say that the function $\phi_{\bar{x},\bar{y}} : [0,1] \to \mathbb{R}$ is c-*cyclically monotone* if for all $x, x' \in A_{(\bar{x},\bar{y})}$ and for all $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I$, $x_0 = x$, $x_{I+1} = x'$, it holds

$$\phi_{\bar{x},\bar{y}}(x') \leqslant \phi_{\bar{x},\bar{y}}(x) + \sum_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i).$$

Theorem 10.12. *Let* $\{A_{(\bar{x},\bar{y})}, B_{(\bar{x},\bar{y})}\}_{(\bar{x},\bar{y}) \in \Gamma}$ *satisfying Assumption 5 and s.t.* $\pi(A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})}) > 0$ *and* $\phi_{(\bar{x},\bar{y})}$ *is* c*-cyclically monotone. Then* π *is optimal.*

Proof. We remind that, since c is π-measurable, we can assume w.l.o.g. $c\llcorner_\Gamma$ Borel and Γ σ-compact.
We extend the potentials $\phi_{(\bar{x},\bar{y})}$, $\psi_{(\bar{x},\bar{y})}$ which by assumption exist on $A_{(\bar{x},\bar{y})}$, $B_{(\bar{x},\bar{y})}$ to Borel sets $X_{(\bar{x},\bar{y})}$, $Y_{(\bar{x},\bar{y})}$ satisfying the crosswise structure

$$\Gamma \cap (X_{(\bar{x},\bar{y})} \times Y) = \Gamma \cap (X \times Y_{(\bar{x},\bar{y})}) = \Gamma \cap (X_{(\bar{x},\bar{y})} \times Y_{(\bar{x},\bar{y})}).$$

We derive then a partition $\{X_i \times Y_i\}_{i\in\mathbb{N}}$ of Γ, up to a negligible set, and apply Theorem 10.6.

Step 0. By the inner regularity of measures, one can assume w.l.o.g. each $A_{(\bar{x},\bar{y})}$, $B_{(\bar{x},\bar{y})}$ to be σ-compact and that $\phi_{(\bar{x},\bar{y})}$, $\psi_{(\bar{x},\bar{y})}$ satisfying (10.6) to be Borel. One can as well require w.l.o.g. that $A_{(\bar{x},\bar{y})} = P_1(\Gamma \cap (A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})}))$, $B_{(\bar{x},\bar{y})} = P_2(\Gamma \cap (A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})}))$.

Step 1. Fix a point $(\bar{x},\bar{y}) \in \Gamma$, set for simplicity $A = A_{(\bar{x},\bar{y})}$, $B = B_{(\bar{x},\bar{y})}$, $\phi = \phi_{(\bar{x},\bar{y})}$, $\psi = \psi_{(\bar{x},\bar{y})}$.
Define the σ-compact set $B' := P_2(\Gamma \cap P_1^{-1}A) = \Gamma(A) \supset B$ and the Borel function

$$\psi'(y) = \inf_{x\in A\times\{y\}} \{c(x,y) - \phi(x)\} \overset{\text{Lemma 10.17}}{=} \begin{cases} \psi(y) & y \in B \\ c(x,y) - \phi(x) & (x,y) \in \Gamma \cap A \times B' \end{cases}.$$

The couple ϕ, ψ' is an extension of ϕ, ψ satisfying (10.6) on A, B' and s.t. $\Gamma \cap A \times (Y \setminus B') = \emptyset$. Repeating the procedure for $A' := P_1(\Gamma \cap P_2^{-1}B') = \Gamma^{-1}(A) \supset A$ with

$$\phi'(x) = \inf_{x\in\{x\}\times B'} \{c(x,y) - \psi(y)\} \overset{\text{Lemma 10.17}}{=} \begin{cases} \phi(y) & y \in A \\ c(x,y) - \psi(y) & (x,y) \in \Gamma \cap A' \times B' \end{cases}$$

and iterating it at most countably many times we can extend ϕ, ψ preserving (10.6) on

$$X = \cup_n (\Gamma^{-1} \circ \Gamma)(A) \qquad Y = \Gamma(X).$$

Step 3. Let $\{X_{(\bar{x},\bar{y})}, Y_{(\bar{x},\bar{y})}\}_{(\bar{x},\bar{y})\in\Gamma}$ the covering of Γ constructed in the previous steps. Since $\pi(X_{(\bar{x},\bar{y})} \times X_{(\bar{x},\bar{y})}) > 0$ and $X_{(\bar{x},\bar{y})} = \Gamma(X_{(\bar{x},\bar{y})})$, then one can extract from a refinement countable partitions $\{X_i\}_i$, $\{Y_i\}_i$ of $P_1(\Gamma)$, $P_2(\Gamma)$ such that

$$\Gamma \cap (X_i \times Y) = \Gamma \cap (X \times Y_i) = \Gamma \cap (X_i \times Y_i) \quad \forall i \in \mathbb{N}.$$

The thesis follows then from Theorem 10.6, by Corollary 10.7. $\square$

The case of Point (3) corresponds to a single global class.

10.2 The c-cyclically monotone relation

Let c be a $\Pi_1^1([0,1]^2; [0,+\infty])$-function and let Γ be a c-cyclically monotone σ-compact set such that $c\llcorner_\Gamma$ is Borel and $c : \Gamma \to \mathbb{R}^+$ (we use Lemma 10.5 and inner regularity). In the following this will be the set where a transference plan is concentrated.

The next definition is not the standard one, but it is useful for our construction.

Definition 10.13 (Cyclically Monotone Envelope)**.** For a given function $f : P_1(\Gamma) \to (-\infty, +\infty]$ define as the c-*cyclically monotone envelope* of f the function

$$\phi(x) = \begin{cases} \inf\limits_{x_{I+1}=x, I \in \mathbb{N}} \left\{ \sum\limits_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i) + f(x_0), (x_i, y_i) \in \Gamma \right\} & \text{if } < +\infty \\ -\infty & \text{otherwise} \end{cases} \tag{10.7}$$

Similarly, for a given function $g : P_2(\Gamma) \to (-\infty, +\infty]$ define the c_{-1}-*cyclically monotone envelope* of g the function

$$\psi(y) = \begin{cases} \inf\limits_{(x_i,y_i) \in \Gamma, y_{I+1}=y, I \in \mathbb{N}} \left\{ \sum\limits_{i=0}^{I} c(x_i, y_{i+1}) - c(x_i, y_i) + g(y_0) \right\} & \text{if } < +\infty \\ -\infty & \text{otherwise} \end{cases} \tag{10.8}$$

In the following we will denote them by

$$\mathtt{C}(f) \qquad \text{and} \qquad \mathtt{C}^{-1}(g).$$

Moreover, we will often call the first case of formulas (10.7), (10.8) as the inf-formula.

Lemma 10.14. *If* f, g *belong to the* Δ^1_n*-class with* $n \geqslant 2$*, then the functions* $\phi, \psi : [0, 1] \to [-\infty, +\infty)$ *belong to the* Δ^1_{n+1}*-class. Moreover* $\phi(x) \leqslant f(x)$*,* $\psi(y) \leqslant g(y)$ *for* $x \in P_1(\Gamma)$*,* $y \in P_2(\Gamma)$*.*

Proof. The second part of the lemma holds trivially, because of the particular path $(x_i, y_i) = (x, y) \in \Gamma$ for all i.

Consider thus the function

$$\phi_I(x_0, y_0, \ldots, x_I, y_I, x) = \sum_{i=0}^{I} c(x_{i+1}, x_i) - c(x_i, y_i) + f(x_0), \quad (x_i, y_i) \in \Gamma,\ x_{I+1} = x.$$

Being the sum of the Π^1_1 functions $c(x_{i+1}, x_i) - c(x_i, y_i)$ ($c\llcorner\Gamma$ is Borel) with the Δ^1_n-function f, $\phi_I(x_0, y_0, \ldots, x_I, y_I, x)$ is Δ^1_n with $n \geqslant 2$.

If $g(x, y)$ is a Δ^1_n-function, then $\tilde{g}(x) = \inf_y g(x, y)$ satisfies

$$\tilde{g}^{-1}(-\infty, s) = P_1(g^{-1}(-\infty, s)) \in \Sigma^1_n,$$

so that $\tilde{g}$ is in the Π^1_n-class.

It follows that

$$\phi_I(x) = \inf \left\{ \sum_{i=0}^{I} c(x_{i+1}, x_i) - c(x_i, y_i) + f(x_0), (x_i, y_i) \in \Gamma, x_{I+1} = x \right\}$$

is Π^1_n, and finally $\inf_I \phi_I(x)$ is also Π^1_n. We conclude the proof by just observing that the set $\{x : \inf_I \phi_I(x) = +\infty\}$ is in Π^1_n, being the countable intersection of the Π^1_n-sets $\{x : \inf_I \phi_I(x) > k\}$. Hence $\{x : \inf_I \phi_I(x) < +\infty\} \in \Sigma^1_n$, so that the conclusion follows from the fact that $\Delta^1_{n+1} \supset \Sigma^1_n \cup \Pi^1_n$ and it is a σ-algebra. □

Remark 10.15. In the case $n = 1$ the same proof shows that ϕ, ψ are $\mathcal{A}$-functions.

Definition 10.16. A function $f : [0,1] \to [-\infty,+\infty]$ is c-*cyclically monotone* if for all $x, x' \in [0,1]$ such that $f(x) > -\infty$ and for all $(x_i, y_i) \in \Gamma$, $i = 0,\dots,I$, $x_0 = x$, $x_{I+1} = x'$, it holds

$$f(x') \leqslant f(x) + \sum_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i). \tag{10.9}$$

Similarly, a function $g : [0,1] \to [-\infty,+\infty]$ is c_{-1}-*cyclically monotone* if for all $y, y' \in [0,1]$ such that $g(y) > -\infty$ and for all $(x_i, y_i) \in \Gamma$, $i = 0,\dots,I$, $y_0 = y$, $y_{I+1} = y'$, it holds

$$g(y') \leqslant g(y) + \sum_{i=0}^{I} c(x_i, y_{i+1}) - c(x_i, y_i). \tag{10.10}$$

The following are well known results: we give the proof for completeness. We recall that for any function $h : [0,1] \mapsto [-\infty,+\infty]$ the set F_h is the set where h is finite:

$$F_h = h^{-1}(\mathbb{R}) = \big\{x \in [0,1] : h(x) \in \mathbb{R}\big\}. \tag{10.11}$$

Lemma 10.17. *The following holds:*

1. *The function* ϕ *(*ψ*) defined in* (10.7) *(in* (10.8)*) is* c*-cyclically monotone (*c_{-1}*-cyclically monotone).*

2. *If* f *is* c*-cyclically monotone (*g *is* c_{-1}*-monotone), then* $\phi(x) = f(x)$ *on* $F_f \cap \pi_1(\Gamma)$ *(*$\psi(x) = g(x)$ *on* $F_g \cap \pi_2(\Gamma)$*).*

3. *If we define the function*

$$\begin{gathered} g'(y) = \begin{cases} c(x,y) - \phi(x) & (x,y) \in (F_\phi \times [0,1]) \cap \Gamma \\ +\infty & \textit{otherwise} \end{cases} \\ \left(f'(x) = \begin{cases} c(x,y) - \psi(y) & (x,y) \in ([0,1] \times F_\psi) \cap \Gamma \\ +\infty & \textit{otherwise} \end{cases} \right) \end{gathered} \tag{10.12}$$

   then g' *is* c_{-1}*-monotone (*f' *is* c*-cyclically monotone) and belongs to the* Δ^1_n *pointclass if* f *is in the* Δ^1_n *class (belongs to the* Δ^1_n *pointclass if* g *is in the* Δ^1_n *class).*

A part of the statement is that $c(x,y) - \phi(x)$ does not depend on x for fixed y in $(F_\phi \times [0,1]) \cap F_c$ ($c(x,y) - \psi(y)$ does not depend on y for fixed x in $([0,1] \times F_\phi) \cap F_c$).

Remark 10.18. If ϕ, ψ are $\mathcal{A}$-functions, it is fairly easy to see that g', f' are $\mathcal{A}$-functions.

Proof. The proof will be given on for ϕ, the analysis for ψ being completely similar.

Point (1). The first part follows by the definition: for any axial path as in Definition 10.16 we have

$$\begin{aligned}
&\phi(x)+\sum_{i=0}^{I} c(x_{i+1},y_i)-c(x_i,y_i)\\
&= \inf\left\{\sum_{i=0}^{I'} c(x_{i+1},x_i)-c(x_i,y_i)+f(x_0),(x_i,y_i)\in\Gamma,x_{n+1}=x,I'\in\mathbb{N}\right\}\\
&\qquad\qquad+\sum_{i=0}^{I} c(x_{i+1},y_i)-c(x_i,y_i)\\
&\geqslant \inf\left\{\sum_{i=0}^{I'+I} c(x_{i+1},x_i)-c(x_i,y_i)+f(x_0),(x_i,y_i)\in\Gamma,x_{I'+I+1}=x',I'\in\mathbb{N}\right\}\\
&\geqslant\phi(x').
\end{aligned}$$

Notice that we have used that $\phi(x)>-\infty$ to assure that its value is given by the inf-formula.

Point (2). The second point follows by the definition of c-monotonicity: first of all, if $x\in F_f\cap P_1(\Gamma)$, then the value of ϕ is computed by the inf-formula in (10.7). Then we have from the c-monotonicity of Γ

$$f(x')\leqslant\sum_{i=0}^{I} c(x_{i+1},y_i)-c(x_i,y_i)+f(x_0),\quad x_0\in F_f,(x_i,y_i)\in\Gamma,x_{I+1}=x'.$$

Hence we obtain $\phi(x)\geqslant f(x)$, and using Lemma 10.14 we conclude the proof of the second point.

Point (3). Assume that for y fixed there are $x,x'\in F_\phi$ such that $(x,y)\in\Gamma$ and

$$c(x,y)-\phi(x)\geqslant c(x',y)-\phi(x')+\epsilon.$$

Then, since $x,x'\in F_\phi$, there are points $(x_i,y_i)\in\Gamma$, $i=0,\dots,I$, $x_{I+1}=x$ such that

$$\sum_{i=0}^{I} c(x_{i+1},x_i)-c(x_i,y_i)+f(x_0)<\phi(x)+\frac{\epsilon}{2}.$$

Add then the point $(x_{I+1},y_{I+1})=(x,y)\in\Gamma$ to the previous path: the definition of ϕ implies then for $x_{I+2}=x'$

$$\begin{aligned}
\phi(x')&\leqslant\sum_{i=0}^{I+1} c(x_{i+1},x_i)-c(x_i,y_i)+f(x_0)\\
&=c(x',y_{I+1})-c(x_{I+1},y_{I+1})+\sum_{i=0}^{I} c(x_{i+1},y_i)-c(x_i,y_i)+f(x_0)\\
&<c(x',y)-c(x,y)+\phi(x)+\frac{\epsilon}{2},
\end{aligned}$$

yielding a contradiction. This shows that the definition of g makes sense.

The proof of the c-monotonicity is similar: assume that there exist points $(x_i, y_i) \in \Gamma$, $i = 1, \dots, I$, such that $g'(y_0) > -\infty$ and

$$g'(y') > g'(y) + \sum_{i=0}^{I} c(x_i, y_{i+1}) - c(x_i, y_i), \quad y_0 = y,\ y_{I+1} = y'$$

Using the fact that $g'(y), g'(y') > -\infty$, it follows that there exists $(x, y), (x', y') \in (F_\phi \times [0,1]) \cap \Gamma$ such that $g'(y) = c(x, y) - \phi(x)$, $g'(y') = c(x', y') - \phi(x')$ so that for $(x_{I+1}, y_{I+1}) = (x', y')$, $(x_0, y_0) = (x, y)$

$$\begin{aligned} g'(y') &> g'(y) + \sum_{i=0}^{I} c(x_i, y_{i+1}) - c(x_i, y_i) \\ &= c(x, y) - \phi(x) + c(x', y') + \sum_{i=1}^{I+1} c(x_{i-1}, y_i) - c(x_i, y_i) - c(x_0, y_0) \\ &\geqslant c(x', y') - \phi(x) - \phi(x') + \phi(x) \\ &= c(x', y') - \phi(x') = g'(y'), \end{aligned}$$

yielding a contradiction. We have used the c-monotonicity of ϕ.

Finally, since $c_{\llcorner\Gamma}$ is Borel, then it follows immediately that g' is in the Δ^1_n-class. □

For fixed $(\bar{x}, \bar{y}) \in \Gamma$, we can thus define recursively for $i \in \mathbb{N}_0$ the following sequence of functions ψ_{2i}, ϕ_{2i+1}.

1. Set $\psi_0(y; \bar{x}, \bar{y}) = -\mathbb{1}_{\bar{y}}(y)$.
2. Assume that $\psi_{2i}(\bar{x}, \bar{y})$ is given. For $i \in \mathbb{N}_0$, define then the function $\phi_{2i+1}(x; \bar{x}, \bar{y})$ as
$$\phi_{2i+1}(x; \bar{x}, \bar{y}) = \mathtt{C}\big((\psi_{2i})'(\bar{x}, \bar{y})\big), \tag{10.13}$$
where $(\psi_{2i}(\bar{x}, \bar{y}))'$ is defined in (10.12).
3. Similarly, if $\phi_{2i+1}(\bar{x}, \bar{y})$ is given, define
$$\psi_{2i+2}(y; \bar{x}, \bar{y}) = \mathtt{C}_{-1}\big(\phi_{2i+1}(\bar{x}, \bar{y})\big). \tag{10.14}$$

Note that ϕ_{2i+1} is a Δ^1_{2i+2}-function, ψ_{2i+2} is a Δ^1_{2i+3}-function for $i \in \mathbb{N}_0$ (Lemma 10.17), so that the sets

$$A_{2i+1}(\bar{x}, \bar{y}) = F_{\phi_{2i+1}(\bar{x}, \bar{y})}, \quad B_{2i+2}(\bar{x}, \bar{y}) = F_{\psi_{2i+2}(\bar{x}, \bar{y})}, \quad i \in \mathbb{N}_0, \tag{10.15}$$

are in Δ^1_{2i+2}, Δ^1_{2i+3}, respectively.

From Lemma 10.17 it follows the next corollary.

Corollary 10.19. *If $\phi_{2i+1}(x, \bar{x}, \bar{y})$, $\psi_{2i}(y, \bar{x}, \bar{y})$ are constructed by (10.13), (10.14) and $A_{2i+1}(\bar{x}, \bar{y})$, $B_{2i}(\bar{x}, \bar{y})$ are defined by (10.15), then the following holds:*

1. $A_{2i+1} \subset A_{2j+1}$, $B_{2i} \subset B_{2j}$ *if* $i \leqslant j$, *and*

$$\phi_{2j+1}(\bar{x},\bar{y})\llcorner_{A_{2i+1}(\bar{x},\bar{y})} = \phi_{2i+1}(\bar{x},\bar{y}), \quad \psi_{2j}(\bar{x},\bar{y})\llcorner_{A_{2i}(\bar{x},\bar{y})} = \psi_{2i}(\bar{x},\bar{y}).$$

2. $A_1(\bar{x},\bar{y}) \supseteq P_1(\Gamma \cap ([0,1] \times \{\bar{y}\}))$ *and in general*

$$A_{2i+1}(\bar{x},\bar{y}) \supseteq P_1\Big(\big([0,1] \times B_{2i}(\bar{x},\bar{y})\big) \cap \Gamma\Big), \quad B_{2i+2}(\bar{x},\bar{y}) \supseteq P_2\Big(\big(A_{2i+1}(\bar{x},\bar{y}) \times [0,1]\big) \cap \Gamma\Big).$$

3. *On the set* $(A_{2i+1}(\bar{x},\bar{y}) \times A_{2j}(\bar{x},\bar{y})) \cap \Gamma$ *it holds*

$$\phi_{2i+1}(x,\bar{x},\bar{y}) + \psi_{2j}(x,\bar{x},\bar{y}) = c(x,y).$$

Proof. Point (1). Point (3) of Lemma 10.17 implies that at each step we are applying formula (10.7) to the c-cyclically monotone function $c(x,y) - \psi_{2i}(y)$ or the c_{-1}-cyclically monotone $c(x,y) - \phi_{2i+1}(y)$. From Point (2) of the same lemma we deduce Point (1).

Point (2). The second point is again a consequence of the c-cyclically monotonicity or c_{-1}-cyclically monotonicity of the functions $c(x,y) - \psi_{2i}(y)$, $c(x,y) - \phi_{2i+1}(y)$ on the set $([0,1] \times B_{2i}(\bar{x},\bar{y})) \cap \Gamma$, $(A_{2i+1}(\bar{x},\bar{y}) \times [0,1]) \cap \Gamma$, respectively.

Point (3). The last point follows from Point (2) by Lemma 10.17. $\square$

For all $(x,y) \in \Gamma$, define the set $\Gamma_{(x,y)}$ as

$$\Gamma_{(x,y)} := \Gamma \cap \bigg(\bigcup_i A_{2i+1}(x,y) \times B_{2i}(x,y)\bigg). \tag{10.16}$$

Observe that under (PD) $\Gamma_{(x,y)}$ is measurable for all Borel measures (Section B.1).

We then define the following relations in $[0,1]^2$.

Definition 10.20 (c-cyclically monotone relation)**.** We say that $(x,y)R(x',y')$ if $(x',y') \in \Gamma_{(x,y)}$. We call this relation R the c-*cyclically monotone relation*.

Clearly $\bar{E} \subset R$, where $\bar{E}$ is given in Definition 10.4: actually the equivalence class of $(\bar{x},\bar{y})$ w.r.t. $\bar{E}$ is already contained in $(A_1 \times [0,1]) \cap \Gamma$.

Remark 10.21. The following are easy observations.

1. If $(x,y)R(x',y')$, then from Point (2) of Corollary 10.19 also $(x,y)R(\Gamma \cap (\{x'\} \times Y))$ and $(x,y)R(\Gamma \cap (X \times \{y'\}))$: this means that Γ satisfies the crosswise condition w.r.t. R (Definition 7.7). In particular to characterize R it is enough to define the projected relations

$$xR_1x' \;\Leftrightarrow\; x' \in \bigcup_{i \in \mathbb{N}_0} A_{2i+1}(x,y), \quad yR_2y' \;\Leftrightarrow\; y' \in \bigcup_{i \in \mathbb{N}_0} B_{2i}(x,y).$$

2. The relation R is nor transitive neither symmetric, as the following example shows (see Figure 11).

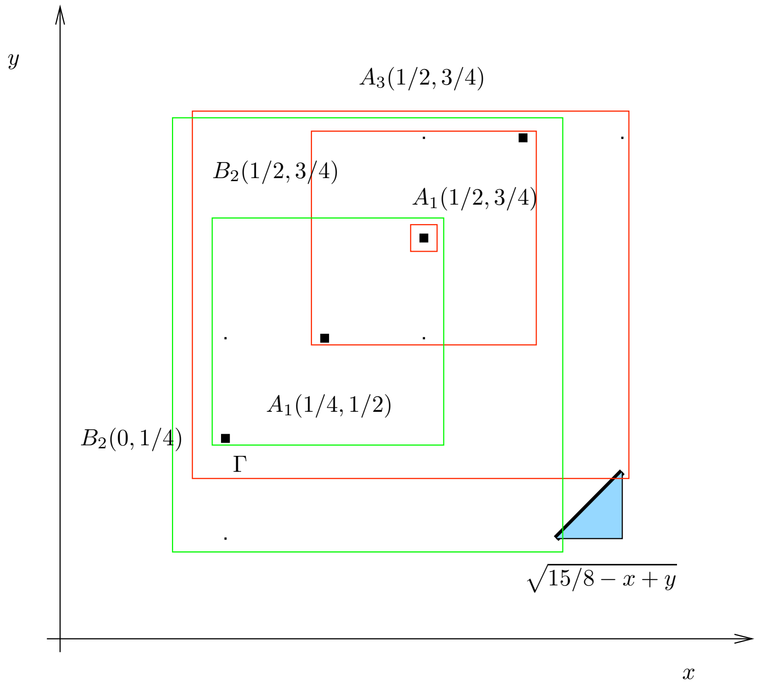


Figure 11: The cost of Point (2) of Remark 10.21: the c-cyclically monotone relation is not an equivalence relation.

Consider the cost

$$c(x,y) = \begin{cases} 0 & (x,y) \in A \\ \sqrt{15/8 - x + y} & 7/8 \leqslant y + 7/8 \leqslant x \leqslant 1 \\ +\infty & \text{otherwise} \end{cases}$$

where

$$A = \Big\{(0,0), (0,1/4), (0,1/2), (1/4,1/2), (1/2,1/2), (1/2,3/4), (1/2,1), (3/4,1), (1,1)\Big\}.$$

Let Γ be the set

$$\Gamma = \big\{(0,1/4), (1/4,1/2), (1/2,3/4), (3/4,1)\big\} \cup \big\{(x, x - 7/8), x \in [7/8,1]\big\}.$$

It is easy to see that

$$\begin{aligned} \Gamma_{(1/4,1/2)} &= \Gamma \cap (\{0, 1/4, 1/2, 3/4, 7/8\} \times [0,1]) \\ &\neq \Gamma \cap ([0,1] \times \{1/8, 1/4, 1/2, 3/4, 1\}) = \Gamma_{(1/2,3/4)}. \end{aligned}$$

3. Another possible definition can for example be the following symmetric relation on Γ.

Definition 10.22. We say that $(x', y') R (x'', y'')$ if there exists Borel functions $\phi, \psi : [0,1] \mapsto \mathbb{R} \cup \{-\infty\}$ such that

$$\phi(x') + \psi(y') = c(x', y'), \quad \phi(x'') + \psi(y'') = c(x'', y''), \quad \phi(x) + \psi(y) \leqslant c(x,y)\ \forall (x,y).$$

(10.17)

However, the following points are in order.

a) The relation R depends deeply on the choice of Γ. Since we will use R to generate a disjoint partition of $[0,1]^2$, the disintegration of the measure π will depend on Γ, even if in Chapter 3 it is shown a way of making it in some sense independent.
b) We observe that even if $R_x = \{y : yRx\} = [0,1]$ for some x, this does not mean that the measure is optimal. As an example, consider (Figure 12)

$$c(x,y) = \begin{cases} 1 & 0 < x = y < 1 \\ 0 & 1 > x = y - \alpha \quad \text{mod } 1 \\ 0 & y = 0 \\ +\infty & \text{otherwise} \end{cases}$$

with $\alpha \in [0,1] \setminus \mathbb{Q}$, and the transport problem $\mu = \delta_1 + \mathcal{L}^1$, $\nu = \delta_0 + \mathcal{L}^1$. The transference plan $\pi = \delta_{(1,0)} + (\mathbb{I},\mathbb{I})_\sharp \mathcal{L}^1$ is clearly not optimal, but since the set

$$\Gamma = \big\{(x,x), x \in [0,1)\big\} \cup \{(1,0)\}$$

has not closed cycles, it follows that it is c-cyclically monotone and moreover $R_x = [0,1]$.

The main use of the c-cyclically monotone relation R is that any crosswise equivalence relation whose graph is contained in R and such that the disintegration is strongly consistent can be use to apply Theorem 10.6: the relation $\bar{E}$ of Definition 10.4 is a possible choice. Note that the strong consistency of the disintegration allows to replace the universally measurable equivalence classes with Borel one, up to a π-negligible set.

Remark 10.23. Under Cantor Hypothesis we can give a procedure to construct an equivalence relation $E' \subset R$ maximal w.r.t. inclusion: if R_α, $\alpha \in \omega_1$, is an ordering of the partition $R_{(\bar{x},\bar{y})} = \{(x,y) : (\bar{x},\bar{y}) R (x,y)\}$, one then defines the partition

$$E_\alpha = \Gamma \cap \left[\left(P_1 R_\alpha \setminus \bigcup_{\beta<\alpha} P_1 R_\beta \right) \times [0,1] \right].$$

Being R_β universally measurable and $\sharp\{\beta < \alpha\} = \omega_0$, we have that each E_α is universally measurable. Moreover it is a partition, and from the definition of R it follows that in each class there are optimal ϕ, ψ. Finally it is clearly maximal w.r.t. graph inclusion among all equivalence relations containing $\bar{E}$ and contained in R.

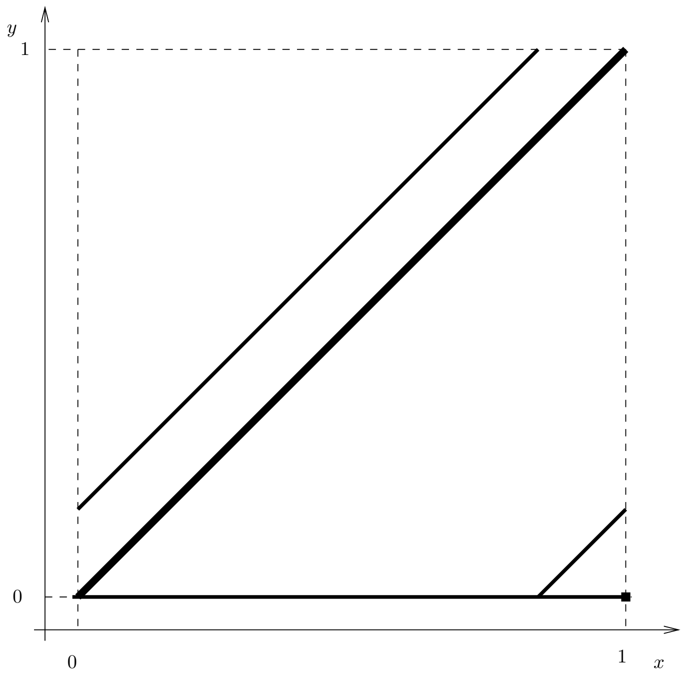


Figure 12: The cost function considered in Point (3b) of Remark 10.21. Non optimality even with $R_x = \{y : yRx\} = [0, 1]$ for some x.

11 Examples

In this section we study the dependence of our construction w.r.t. the choice of Γ, and the necessity of the assumptions in Theorem 10.6.

11.1 Dependence w.r.t. the set Γ

We consider the situation where the assumptions of Theorem 10.6 do not hold, so that either we do not have the strong consistency of the disintegration, or the set A' is not a set of uniqueness. Keeping fixed μ, ν, c and the plan $\pi \in \Pi(\mu,\nu)$, varying Γ, the following cases are possible:

1. Strong consistency of the disintegration is not satisfied for any choice of Γ, and the plan we are testing can either be optimal or not (Example 11.1, Example 11.2).
2. Strong consistency can be satisfied or not, depending on Γ, and, when it is, the quotient problem can be both well posed (A' is a set of uniqueness) or not (Example 11.3, Example 11.2). We are testing an optimal plan.
3. Strong consistency is always satisfied, but the image measure m is not atomic (Example 11.4). The plan we are testing can either be optimal or not.

In Figure 13, for each example we draw the pictures of the set in $[0,1]^2$ where c is finite.

Example 11.1. Consider $\mu = \nu = \mathcal{L}^1$ with the cost given by

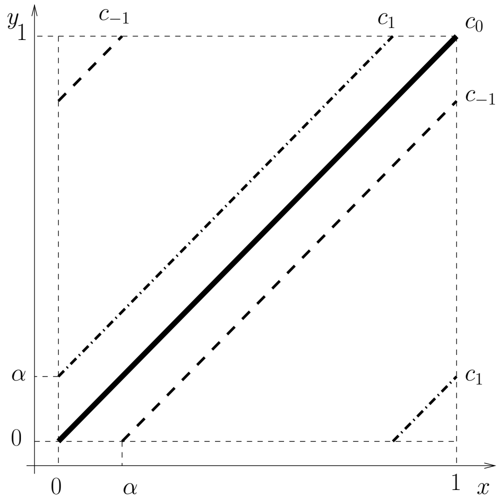
$$c(x,y) = \begin{cases} c_0 & y - x = 0 \\ c_1 & y - x = \alpha \ (\mathrm{mod}\ 1) \\ c_{-1} & y - x = -\alpha \ (\mathrm{mod}\ 1) \\ +\infty & \text{otherwise} \end{cases} \qquad \text{with } \alpha \in [0,1] \setminus \mathbb{Q} \text{ and } c_1 + c_{-1} \geqslant 2c_0.$$

The extremal points in $\Pi(\mu,\nu)$ are, for $i \in \{0,1,-1\}$,

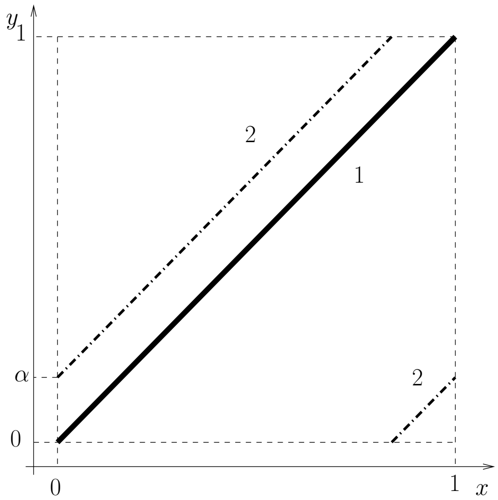
$$\pi_i = \big(\mathrm{Id}, \mathrm{Id} + i\alpha(\mathrm{mod}\ 1)\big)_\sharp \mathcal{L}^1 \quad \Longrightarrow \quad \int c(x,y)\, d\pi_i = c_i,$$

the optimal one will be the one corresponding to the lowest c_i.

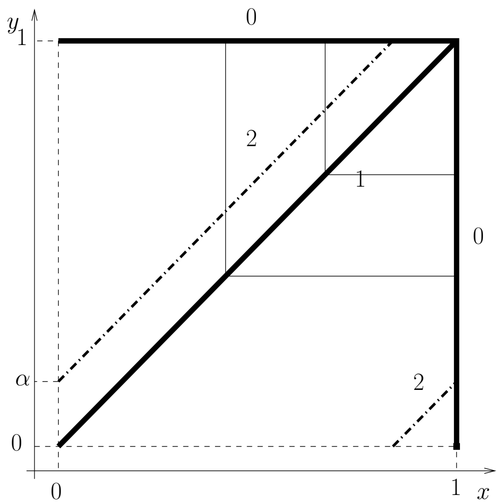
Fix the attention on π_0, which is c-cyclically monotone when $c_1 + c_{-1} \geqslant 2c_0$. Take as Γ the diagonal $\{x = y\}$: the equivalence classes are given by $\{x + n\alpha \mod 1\}$,



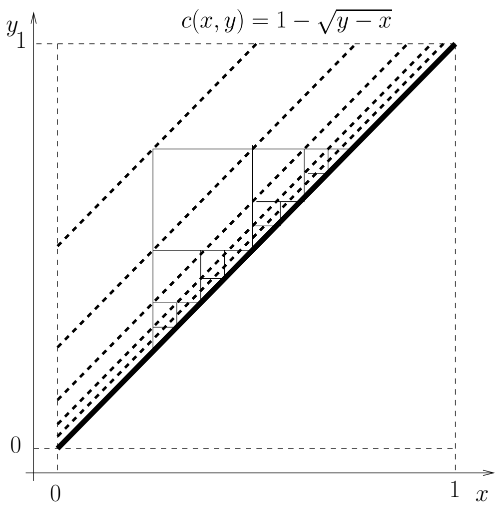
(a) When $\alpha \notin \mathbb{Q}$, the cycle decomposition of the plan $\pi = (\mathbb{I}, \mathbb{I})_\sharp \mathcal{L}^1$ gives always a non-measurable disintegration (Ex. 11.1).



(b) Disintegration sometimes badly supported, sometimes well, but with no answer (Ex. 11.2).



(c) Disintegration either measurable or not, quotient problem either well posed or not (Ex.: 11.3).



(d) A set of uniqueness with no optimal pair and well posed quotient problem (Ex.: 11.4).

Figure 13: . In the picture you find, in bold, the set where c is finite. We analyze different choices of Γ.

the quotient is a Vitali set, and thus the unique consistent disintegration is the trivial disintegration

$$\mathcal{L}^1 = \int \mathcal{L}^1 m.$$

Moreover, one can verify that there is no choice of Γ for which the disintegration is strongly consistent. When $c_1 < c_0$, we have a c-cyclically monotone transference plan π for which the decomposition gives a disintegration which is not strongly consistent and π is not optimal. When $c_1, c_2 > c_0$, we have an optimal c-cyclically monotone transference plan π for which the disintegration consistent with the decomposition in cycles is not strongly consistent.

Example 11.2. Consider an example given in [AP], page 135: $\mu = \nu = \mathcal{L}^1$ with the cost given by

$$c(x,y) = \begin{cases} 1 & y - x = 0 \\ 2 & y - x = \alpha \pmod 1 \\ +\infty & \text{otherwise} \end{cases} \qquad \text{with } \alpha \in [0,1].$$

The extremal plans in $\Pi(\mu,\nu)$ with finite costs are, for $i \in \{0,1\}$,

$$\pi_i = \big(\mathrm{Id}, \mathrm{Id} + i\alpha(\mathrm{mod}\ 1)\big)_\sharp \mathcal{L}^1 \quad \Longrightarrow \quad \int c(x,y) d\pi_i = 1 + i;$$

both are c-cyclically monotone, and the optimal one is π_0. Take $\Gamma = \{x = y\}$: then there is no cycle of finite cost, therefore the cycle decomposition gives classes consisting in singletons, the quotient space is the original one, $m = \mathcal{L}^1$, $\pi_\alpha = \delta_{\{(x,y)\}}$, where α is the class of (x,y). This means that the measurability condition is satisfied, but the quotient problem (which here is essentially the original one) has not uniqueness. Take instead $\Gamma = \{c < \infty\}$: now we have cycles, all with zero cost, obtained by going on and coming back along the same way; consider for example the cycle

$$(w_1, w_1) = (0,0) \; \to \; (w_2, w_2) = (0, \alpha) \; \to \; (w_3, w_3) = (\alpha, \alpha) \; \to \; (w_4, w_4) = (0,0).$$

The situation is similar to Example 11.1 and, as it was there, the disintegration is not strongly consistent. Thus we have that, depending on Γ, strong consistency can be satisfied or not, and when it is, the quotient problem has not uniqueness. This behavior holds when testing either π_0 or π_1, thus it does not depend on the optimality of the plan we are testing.

Example 11.3. Consider the same setting as in Example 11.2, but put the cost to be finite, say zero, also on the lines $\{x = 1\}$ and $\{y = 1\}$. Now, considering $\pi = (\mathbb{I}, \mathbb{I})_\sharp \mathcal{L}^1$,

- with Γ containing $(1,1)$, all the points are connected by a cycle of finite cost, we have just one class and optimality follows by c-monotonicity;

- with $\Gamma = \{(x,x) : x \in [0,1)\}$ the classes are made of single points, the disintegrations is trivially measurable, the quotient problem is essentially the original one and we are in the non-uniqueness case;
- when you consider instead $\Gamma = \{(x,x) : x \in [0,1)\} \cup \{(x, x+\alpha) : x \in [0,1] \setminus \{1-\alpha\}\}$ again the quotient space is a Vitali set, the strong consistency of the disintegration is lost.

Depending on the choice of Γ, we can have or not strong consistency; moreover, when we have strong consistency, the quotient space can have uniqueness or non-uniqueness. Notice that since there exists Γ for which Theorem 10.6 holds, π must optimal: the first argument does not hold for $(\mathbb{I}, \mathbb{I}+\alpha)_\sharp \mathcal{L}^1$, since it is not c-cyclically monotone.

Example 11.4 (A set of uniqueness with nonexistence of ϕ, ψ). Consider $\mu = \nu$ with the cost given by

$$c(x,y) = \begin{cases} 1 & y = x \\ 1 - \sqrt{y-x} & y - x = 2^{-n} \\ +\infty & \text{otherwise} \end{cases} \qquad \text{with } n \in \mathbb{N}.$$

Unless μ is purely atomic with a finite number of atoms, there is no optimal potential. However, applying the procedure one can deduce optimality: $\{c < \infty\}$ is acyclic and therefore the cycle decomposition consists in singletons, the quotient spaces are the original ones, and therefore A' of (10.4) is a set of uniqueness, and $\pi_{\{(x,x)\}} = \delta_{\{(x,x)\}}$, $m = \mu$.

Example 11.5. The final example shows that in the case of non strong consistency, then we can construct a cost $\tilde{c}$ such that the image measure m is the same but there are non optimal transference plans. We just sketch the main steps.

Let h_X, h_Y be the quotient maps for the equivalence relation of Definition 2.5.

Step 1. The conditional probabilities μ_α, ν_β cannot be purely atomic for m-a.e. α, β. By the regularity of the disintegration one can in fact show ([BC2]) that there exists a Borel set B such that $B \cap h_X^{-1}(\alpha)$ is countable and the atomic part of μ_α is concentrated on B. Hence if μ_α is purely atomic we can reduce to the case where $h_X^{-1}(\alpha)$ is countable for all α.

Assume by contradiction that each equivalence class has countably many counterimages. We can use Lusin Theorem (Theorem 5.10.3 in [Sri]) to find a countable family of Borel maps $h'_n : [0,1] \supset B_n \to [0,1]$, $B_n \in \mathcal{B}([0,1])$, $n \in \mathbb{N}$, such that $h_X \circ h'_n = \mathbb{I}\llcorner_{B_n}$ and

$$\operatorname{graph}(h_X) = \bigcup_n \operatorname{graph}(h'_n), \quad \operatorname{graph}(h'_n) \bigcap \operatorname{graph}(h'_m) = \emptyset.$$

Define the analytic $\bar{E}$-saturated sets

$$Z_n = P_1\big(\bar{E} \cap [0,1] \times h'_n(B_n)\big) \setminus \bigcup_{i=1}^{n-1} Z_i.$$

By construction, $h'_n(B_n) \cap Z_n$ is an analytic section of Z_n, so that Proposition 2.9 implies that the disintegration is strongly consistent.

Step 2. We restrict to the case where μ_α, ν_α have no atoms.

The previous step shows that there is a set of positive m-measure for which the conditional probability μ_α is not purely atomic. Let $\mu_{\alpha,c}$ be the continuous part of μ_α: in [BC2] it is shown that

$$\int \mu_{\alpha,c} m(d\alpha) = \mu \llcorner B$$

for some Borel set B, so that we can assume C compact and restrict the transport to $C \times [0,1]$.

Repeating the procedure for Y, there exists D compact such that for the transport problem in $C \times D$ the conditional probabilities μ_α, ν_α are continuous.

Step 3. We redefine the cost in the set $C \times D$ is order to have the same equivalence classes for h_X, h_Y but for which there are non optimal cyclically monotone costs.

Define the map

$$H_X(\alpha, x) = \mu_\alpha((0,x)), \quad H_Y(\beta, y) = \nu_\beta((0,y)).$$

By measurability of $\mu_\alpha(B)$, $\nu_\beta(B)$ for all Borel sets B, by restricting C, D we can assume that H is continuous in α and x. If $\bar{c}$ is the cost of Example 11.2, then define

$$\tilde{c}(x,y) = \begin{cases} \bar{c}(H_X(\alpha,x), H_Y(\alpha,y)) & (x,y) \in (h_X \otimes h_Y)^{-1}(\alpha,\alpha) \\ c & \text{otherwise} \end{cases}$$

With the notation of Example 11.2, for any pseudoinverse $H_X^{-1}(\alpha)$, $H_y^{-1}(\alpha)$ it is fairly easy to verify that

$$\pi = \int (H_X^{-1}(\alpha), H_Y^{-1}(\alpha))_\sharp \pi_1 m(d\alpha)$$

is a $\tilde{c}$-cyclically monotone transference plan which is not optimal: the optimal is

$$\pi' = \int (H_X^{-1}(\alpha), H_Y^{-1}(\alpha))_\sharp \pi_0 m(d\alpha).$$

11.2 Analysis of the transport problem in the quotient space

In this section we consider some examples related to the study of the quotient transport problem. The examples are as follows.

1. The regularity properties of the original cost (e.g. l.s.c.) are in general not preserved (Example 11.6).

2. In general, there is no way to construct a quotient cost c independently of the transference plan π and different from $\mathbb{I}_{A'}$ (Example 11.7).

3. The set $\Pi^f(m,m)$ strictly contains the set $(h_X \otimes h_Y)_\sharp \Pi^f(\mu,\nu)$ (Examples 11.7, 11.8).

4. If the uniqueness assumption of Theorem 10.6 does not hold, then we can construct a cost c' which gives the same equivalence classes and quotient transport problem and such that the original π is c'-cyclically monotone but not optimal for c' (Proposition 11.9).

Example 11.6 (Fig. 14). Consider the cost

$$c(x,y) = \begin{cases} 0 & y = x, x \in [0,1/2] \\ 1 & y = x + 1/2 \quad \mathrm{mod}\ 1 \\ 1 & y = x, x \in (1/2,1] \end{cases}$$

and the measures

$$\mu = \nu = \sum_{i=1}^{+\infty} 2^{-i-1}\delta\left(x - \frac{1}{2} + 2^{-i}\right) + \frac{1}{2}\delta(x - 3/2), \qquad \pi = (x,x)\sharp\mu.$$

The quotient cost c should satisfy

$$c_\pi(\alpha,\alpha) = \int c(x,y)d\pi_\alpha, \qquad \pi_\alpha \in \Pi(\mu_\alpha,\nu_\alpha), \tag{11.1}$$

so that one obtains

$$c(\alpha,\beta) = \begin{cases} 0 & \beta = \alpha = 1/2 - 2^{-i}, i \in \mathbb{N} \\ 1 & \beta = \alpha = 1/2 \end{cases}$$

Clearly this cost is not l.s.c., and there is no way to make it l.s.c.. This example shows that we cannot preserve regularity properties for the quotient cost c.

Example 11.7 (Fig. 15). Let $r \in [0,1/4] \setminus \mathbb{Q}$ and consider the cost

$$c(x,y) = \begin{cases} 1 & x = y \\ 1+d & y = x + 1/2, x \in [0,1/2] \\ 1+d & y = x - 1/2, x \in [1/2,1] \\ 0 & y = x + r, x \in [0,1/2-r] \\ e & y = x + r, x \in [1/2, 1-r] \\ 0 & y = x - 1/2 + r, x \in [1-r,1] \\ +\infty & \text{otherwise} \end{cases} \qquad d, e \geqslant 0.$$

The settings are

$$\mu = \nu = \mathcal{L}^1, \qquad \Gamma = \{y = x\}.$$

The equivalence relation is $(x,x) \simeq (x+1/2, x+1/2)$: for simplicity we consider the quotient space as $[0,1/2)$.

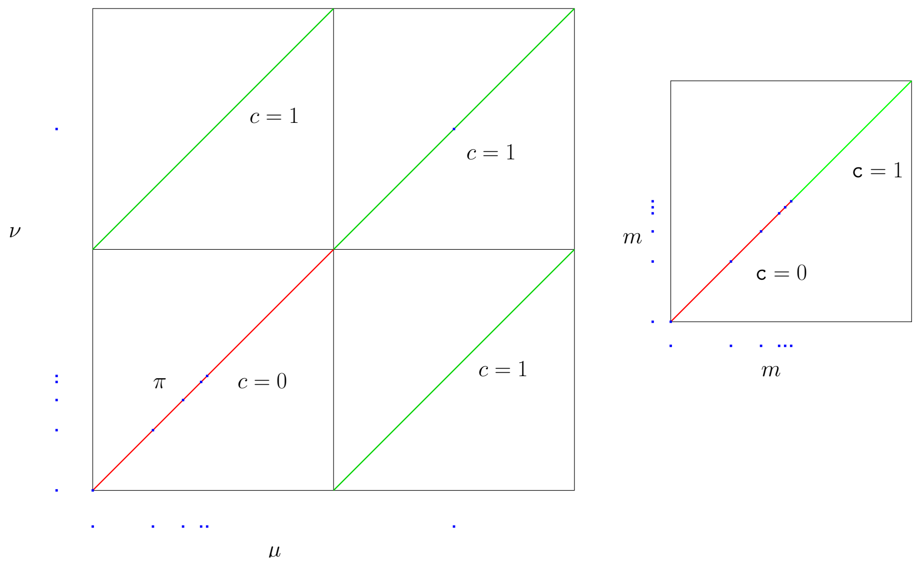


Figure 14: Example 11.6. Outside the segments the cost is defined as $+\infty$. The quotient cost in general is not regular.

In the quotient space, the cost c_π is finite only on $y = x$ and $y = x + r \mod 1/2$. However we have several linear independent plans on $y = x + r$.

The easiest to consider is $\pi_0 = (x, f_0(x))_\sharp \mathcal{L}^1$, where

$$f_0(x) = \begin{cases} x & x \in [0, \frac{1}{2}] \\ x + r & x \in (\frac{1}{2}, 1 - r] \\ x - \frac{1}{2} + r & x \in (1 - r, 1] \end{cases}$$

for which by formula (11.1) we obtain a quotient cost of

$$c_0 = \begin{cases} 1 & \beta = \alpha \\ e & \beta = \alpha + r, \alpha \in [0, \frac{1}{2} - r] \\ 0 & \beta = \alpha - \frac{1}{2} + r, \alpha \in [\frac{1}{2} - r, \frac{1}{2}] \end{cases} \tag{11.2}$$

Another cost is obtained by $\pi_1 = (x, f_1(x))_\sharp \mathcal{L}^1$, where

$$f_1(x) = \begin{cases} x + r & x \in [0, \frac{1}{2} - r] \\ x + \frac{1}{2} & x \in (\frac{1}{2} - r, \frac{1}{2}] \\ x - \frac{1}{2} & x \in (\frac{1}{2}, \frac{1}{2} + r] \\ x & x \in (\frac{1}{2} + r, 1 - r] \\ x - \frac{1}{2} + r & x \in (1 - r, 1] \end{cases}$$

In this case the cost is

$$c_1 = \begin{cases} 1 + d & \beta = \alpha, \alpha \in [0, r) \cup (1/2 - r, 1/2] \\ 1 & \beta = \alpha, \alpha \in [r, 1/2 - r] \\ 0 & \beta = \alpha + r \mod 1/2 \end{cases} \tag{11.3}$$

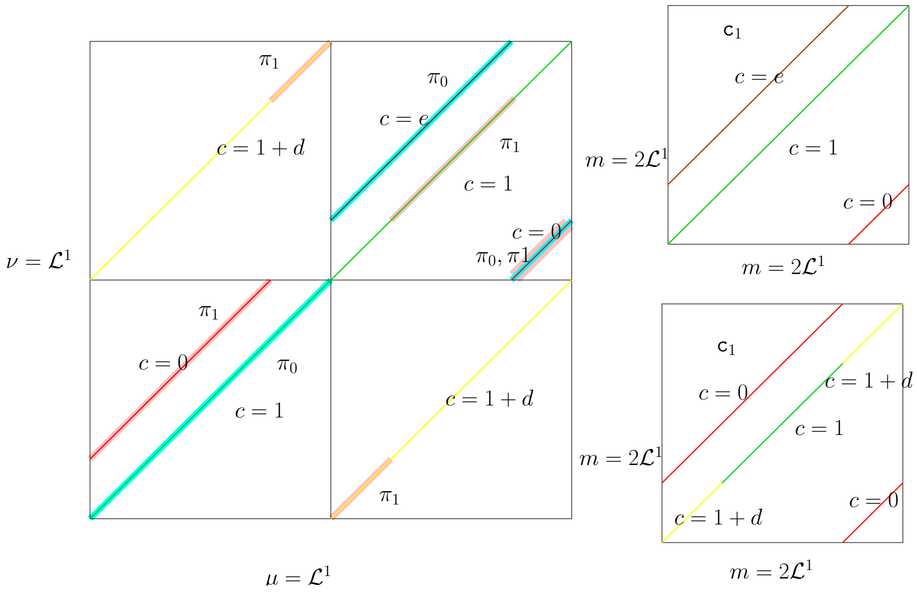


Figure 15: Example 11.7. Outside the segments the cost is $+\infty$, while for the two different transport the costs are given by (11.2), (11.3). There is no universal cost on the quotient space.

Since it is impossible to have a transference plan π in the original coordinates such that

$$c_\pi = \begin{cases} 1 & \beta = \alpha \\ 0 & \beta = \alpha + r \bmod 1/2 \end{cases}$$

then it follows that there is no clear way to associate the cost c in the quotient space independently of the transport plan π, even requiring the weak condition that for π c-cyclically monotone and $\pi' \in \Pi(\mu, \nu)$

$$\int cm = \int c\pi, \quad \int cn \leqslant \int c\pi', \; n = (h_X \otimes h_Y)_\sharp \pi'.$$

We note that there is no transference plan whose image is concentrated only on $\beta = \alpha + r \bmod 1$, so that in general the image of $\Pi^f(\mu, \nu)$ under the map (h_X, h_Y) is a strict subset of $\Pi^f(m, m)$.

Example 11.8 (Fig. 16). We consider the cost for $r \in [\frac{1}{4}, \frac{1}{2}] \setminus \mathbb{Q}$

$$c(x, y) = \begin{cases} 1 & y = x \\ 1 + d & y = \frac{x}{2} + \frac{1}{2} \\ 1 + d & y = 2x - 1 \\ e & y = x + r, x \in [0, \frac{1}{2} - r] \\ f & y = x - 2^{-i}(\frac{1}{2} - r), x \in (1 - 2^{-i}) + 2^{-i}[\frac{1}{2} - 2^{-i+1}r, \frac{1}{2} - 2^{-i}r), i \in \mathbb{N} \\ +\infty & \text{otherwise} \end{cases}.$$

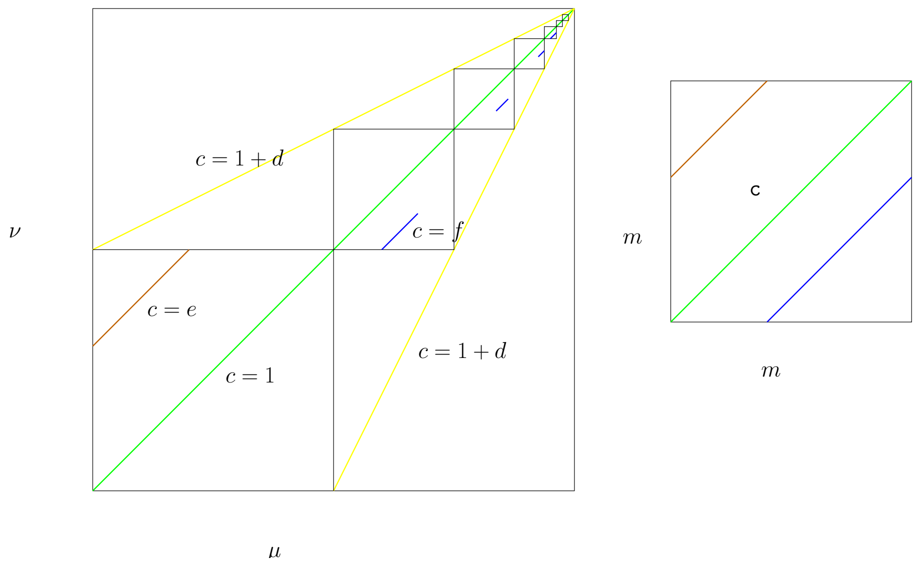


Figure 16: Example 11.8. $\Pi^f(m, m)$ is strictly contained in $(h_X \otimes h_Y)_\sharp \Pi(\mu, \nu)$.

We consider the measures

$$\mu = \nu = \frac{3}{2} \sum_{i=0}^{+\infty} 2^{-i} \mathcal{L}^1 \llcorner_{[1-2^{-i}, 1-2^{-i-1})}.$$

Since the measure of the segment $[1 - 2^{-i}, 1 - 2^{-i-1})$ is 2^{-2i-1}, all measures π with finite cost in $\Pi(\mu, \nu)$ are concentrated on the segments

$$\{y = x, x \in [0, 1]\} \cup \{y = x/2 + 1/2, x \in [0, 1]\} \cup \{y = 2x - 1, x \in [1/2, 1]\}.$$

This can be seen in the quotient space, because

$$m = \mathcal{L}^1,$$

and every measure $\tilde{m} \in \Pi^f(m, m)$ is of the form $\tilde{m} = a_1(x, x)_\sharp \mathcal{L}^1 + a_2(x, x + r \bmod 1)_\sharp \mathcal{L}^1$, $a_1, a_2 \geqslant 0$ and $a_1 + a_2 = 1$. But clearly this cannot be any image of a measure with finite cost in $\Pi^f(\mu, \nu)$.

The next proposition shows that if A' is not a set of uniqueness, then the problem of optimality cannot be decided by just using c-monotonicity.

Proposition 11.9. *If there exists a transference plan $\tilde{m} \in \Pi^f(m, m)$ different from $(\mathbb{I}, \mathbb{I})_\sharp m$, then there exists a cost $\hat{c}(x, y)$ for which the following holds:*

1. *the set Γ is $\hat{c}$-cyclically monotone;*
2. *there are two measures π_0, π_1 in $\Pi(\mu, \nu)$ such that*

$$\int \hat{c}(x, y) d\tilde{\pi} < \int \hat{c}(x, y) d\pi < +\infty.$$

A variation of the following proof (using Lusin Theorem and inner regularity) allows to construct a cost which is also l.s.c..

Proof. Let $m, m' \in \Pi^f(m, m)$, $m \neq m'$, and consider a Borel cost c such that

$$c([0,1]^2 \setminus A') = +\infty, \qquad \int cm' < \int cm < +\infty.$$

It is fairly easy to construct such a cost.

Define now

$$\pi = \int \mu_\alpha \times \nu_\beta \, m(d\alpha d\beta), \quad \pi' = \int \mu_\alpha \times \nu_\beta \, m'(d\alpha d\beta), \quad \hat{c} = c(h_X(x), h_Y(y)).$$

It follows that

$$\int \hat{c}\pi' = \int cm' < \int cm = \int \hat{c}\pi < +\infty.$$

Moreover, since A' is acyclic, the equivalence classes w.r.t. the equivalence relation $\bar{E}$ do not change. $\square$

12 Existence of an optimal potential

In this section we consider the following problem.
Let $\pi \in \mathcal{P}([0,1]^2)$ be concentrated on a c-cyclically monotone set Γ. Assume that there exist partitions $\{X_\alpha\}_\alpha$, $\{Y_\beta\}_\beta$ of $[0,1]$ into Borel sets such that

- $\Gamma \subset \cup_\alpha X_\alpha \times Y_\alpha$ — i.e. Γ satisfies the crosswise condition of Definition 7.7 w.r.t. the partition;
- in each set $X_\alpha \times Y_\alpha$ there exists an optimal couple $\phi_\alpha \in \mathcal{B}(X_\alpha, \mathbb{R})$, $\psi_\alpha \in \mathcal{B}(Y_\alpha; \mathbb{R})$:

$$\phi_\alpha + \psi_\alpha \leqslant c \text{ on } X_\alpha \times Y_\alpha \qquad \phi_\alpha + \psi_\alpha = c \text{ on } \Gamma \cap X_\alpha \times Y_\alpha.$$

Is it possible to find a Borel couple of functions ϕ, ψ s.t.

$$\phi + \psi \leqslant c \text{ on } \cup_\alpha X_\alpha \times Y_\alpha \qquad \phi + \psi = c \ \pi\text{-a.e.?}$$

We show that this is the case under Assumption 1, i.e. if the disintegration of π w.r.t the partition $\{X_\alpha \times Y_\alpha\}$ is strongly consistent. If $\{c < +\infty\} \subset \cup_\alpha X_\alpha \times Y_\alpha$ this provides clearly an optimal couple.

The approach is to show that the set

$$\left\{ (\alpha, \tilde{\phi}, \tilde{\psi}) : \ \tilde{\phi}, \tilde{\psi} \text{ optimal couple in } X_\alpha \times Y_\alpha \right\}$$

is an analytic subset of a suitable Polish space, that we are first going to define. We apply then a selection theorem in order to construct an optimal couple.

In order to structure the ambient space with a Polish topology, we need some preliminary lemmata.

Lemma 12.1. *For every nonnegative function $\bar{\varphi} \in C^0([0,1])$ the map*

$$G_{\bar{\varphi}} \quad : \quad \mathcal{M}([0,1]) \quad \ni \quad \mu \quad \mapsto \quad \int \bar{\varphi}\mu^+ \quad \in \quad \mathbb{R}$$

is convex l.s.c. is w.r.t. weak-topology.*

Proof. Since for every $\mu \in \mathcal{M}([0,1])$

$$\sup \left\{ \int \varphi\mu : \ 0 \leqslant \varphi \leqslant \bar{\varphi} \right\} = \int \bar{\varphi}\mu^+,$$

then $G_{\bar{\varphi}}$ is the supremum of bounded linear functionals, proving the thesis. $\square$

Corollary 12.2. *The map*

$$\mathcal{M}([0,1]) \quad \ni \quad \mu \quad \mapsto \quad \mu^+ \quad \in \quad \mathcal{M}^+([0,1])$$

is Borel w.r.t. weak-topology. For every nonnegative measure ξ, the sublevel set $\{\mu : \mu^+ \leqslant \xi\}$ is closed and convex: in fact $\mu \mapsto \mu^+$ is order convex, meaning that*

$$(\lambda\mu + (1-\lambda)\nu)^+ \leqslant \lambda\mu^+ + (1-\lambda)\nu^+.$$

Proof. It is enough to observe that any function $f : \mathcal{M}([0,1]) \to \mathcal{M}([0,1])$ is Borel if and only if the function $\mu \mapsto \int \varphi f(\mu)$ is Borel for every nonnegative $\varphi \in C^0([0,1])$: the Borel measurability then follows by Lemma 12.1. As well, f is order convex if and only if $\mu \mapsto \int \varphi f(\mu)$ is convex $\forall \varphi \in C^0([0,1];\mathbb{R}^+)$. □

Corollary 12.3. *The function*

$$\mathcal{M}([0,1]) \times \mathcal{M}([0,1]) \quad \ni \quad (\mu_1, \mu_2) \quad \mapsto \quad \mu_1 \wedge \mu_2 \quad \in \quad \mathcal{M}([0,1])$$

is Borel w.r.t. weak-topology.*

Proof. The thesis follows by the relation $\mu_1 \wedge \mu_2 = \mu_1 - [\mu_1 - \mu_2]^+$ and Corollary 12.2. □

Lemma 12.4. *The function*

$$\mathcal{M}^+([0,1])^3 \times C^0([0,1];\mathbb{R}^+) \ni (\mu_1, \mu_2, \mu_3, \phi) \;\mapsto\; \int \phi \frac{d\mu_2}{d\mu_1} \frac{d\mu_3}{d\mu_1} \mu_1 \in [0, +\infty]$$

is Borel w.r.t. weak-topology.*

Proof. Let $\{h_{j,I}\}_{j=1}^{2^I}$ be a partition of $[0,1]$ into continuous functions such that

$$0 \leqslant h_{j,I} \leqslant 1, \; \sum_{j=1}^{2^I} h_{j,I} = 1, \; \operatorname{supp} h_{j,I} \subset \left[(j-1)2^{-I} - 2^{-I-2}, j2^{-I} + 2^{-I-2}\right].$$

Define the l.s.c. and continuous functions, respectively,

$$\mathbb{R}^+ \ni x \mapsto x^{-1^*} := \begin{cases} 0 & x = 0 \\ 1/x & x > 0 \end{cases}$$

$$\mathcal{M}([0,1]) \times C^0([0,1]) \ni \mu \mapsto \left(\int h_{j,I} \phi \mu \right)_{j=1}^{2^I} \in \mathbb{R}^{2^I}$$

If $\phi \in C^0([0,1])$ and $\mu_2 = (d\mu_2/d\mu_1)\mu_1$, then

$$g_I(\mu_1, \mu_2) := \sum_{j=1}^{I} h_{j,I}(x) \left(\int h_{j,I} \mu_1 \right)^{-1^*} \left(\int h_{j,I} \mu_2 \right) \;\to\; \frac{d\mu_2}{d\mu_1}$$

in $L^1(\mu)$, so that for $0 \leqslant \mu_2 \leqslant k\mu_1$ and $0 \leqslant \mu_3 \leqslant k\mu_1$ it follows

$$\int \phi \frac{d\mu_2}{d\mu_1}\frac{d\mu_3}{\mu_1}\mu_1 = \lim_{I\to+\infty}\int \phi g_I(\mu_1,\mu_2)g_I(\mu_1,\mu_3)\mu_1.$$

We finally reduce to the case $\mu_2 \leqslant k\mu_1$ and $\mu_3 \leqslant k\mu_1$: indeed

$$\int \phi \frac{d\mu_2}{d\mu_1}\frac{d\mu_3}{\mu_1}\mu_1 = \lim_{k\to\infty}\lim_{I\to+\infty}\int \phi g_I(\mu_1,\mu_2\wedge(k\mu_1))g_I(\mu_1,\mu_3\wedge(k\mu_1))\mu_1$$

and by Corollary 12.3 the map

$$(\mu_1,\mu_2,\mu_3)\mapsto\Big(\mu_1,(k\mu_1)\wedge\mu_2,(k\mu_1)\wedge\mu_3\Big)$$

is Borel. By composition of the above Borel maps, the statement of the lemma is proved. □

Lemma 12.5. *The function*

$$\begin{array}{rcccc} H_M & : & \mathcal{M}^+([0,1])^2\times\mathcal{M}([0,1])\times\mathcal{M}^+([0,1]^2) & \to & (-\infty,+\infty] \\ & & (\mu,\nu,\eta,\xi,\pi) & \mapsto & H_M := \int\big(\frac{d(\eta+M\mu)^+}{d\mu}\big)\big(\frac{d(P_1)_\sharp\pi}{d\mu}\big)\mu \\ & & & & +\int\big(\frac{d(\xi+M\nu)^+}{d\nu}\big)\big(\frac{d(P_2)_\sharp\pi}{d\nu}\big)\nu \end{array}$$

is Borel w.r.t. weak-topology for all $k\in\mathbb{R}^+$.*

Proof. It follows immediately from Corollary 12.2 and Lemma 12.4. □

Lemma 12.6. *The subset of sequences in $\mathbb{R}^{\mathbb{N}}$ converging to zero is analytic w.r.t. the product topology.*

Proof. The family on nondecreasing sequences m_n is a closed subset of $\mathbb{N}^{\mathbb{N}}$, with the product topology. The sequences of $\mathbb{R}^{\mathbb{N}}$ converging to zero are then the projection of the closed subset of $\mathbb{R}^{\mathbb{N}}\times\mathbb{N}^{\mathbb{N}}$

$$C := \{(\{f_\ell\}_{\ell\in\mathbb{N}},\{m_n\}_{n\in\mathbb{N}}):\ |f_i|\leqslant 2^{-n}\ \forall i\geqslant m_n\}.$$

We now show that C is closed.
Consider sequences $\{f_{\ell,k}\}_\ell$, $\{m_{n,k}\}_n$ converging pointwise to $\{f_\ell\}_\ell$, $\{m_n\}_n$, with $(\{f_{\ell,k}\}_\ell,\{m_{n,k}\}_n)\in C$. Then for each $n\in\mathbb{N}$ exists $k(n)$ such that the sequence $\{m_{n,k}\}_k$ is constantly m_n for $k>k(n)$. As a consequence, for all $k>k(n)$ one has $|f_{i,k}|\leqslant 2^{-n}$ for $i\geqslant m_n$. Since $\{f_{i,k}\}_i$ converges pointwise, it follows that $|f_i|\leqslant 2^{-n}$ for $i\geqslant m_n$. Hence $(\{f_\ell\}_\ell,\{m_n\}_n)\in C$. □

Given a subset J of $\mathbb{R}\cup\{\pm\infty\}$, we denote by $L(\mu;J)$ the μ-measurable maps from $[0,1]$ to J. If not differently stated, μ-measurable functions are equivalence classes of functions which coincide μ-a.e..

Proposition 12.7. *There exists a Polish topology on linear space*

$$L = \Big\{ (\mu, \varphi) : \ \mu \in \mathcal{P}([0,1]), \ \varphi \in L(\mu; \mathbb{R} \cup \{\pm\infty\}) \Big\}$$

such that the map

$$\begin{array}{rccc} I : & L & \to & \mathcal{P}([0,1]) \times \prod_{M=1}^{\infty} \mathcal{M}([0,1]) \\ & (\mu,\varphi) & \mapsto & \big(\mu, \big\{ (\varphi \wedge M) \vee (-M) \mu \big\}_{M \in \mathbb{N}} \big) \end{array}$$

is continuous.

Proof. We inject L in $\mathcal{P}([0,1]) \times \prod_{M=1}^{\infty} \mathcal{M}([0,1])$ by the map I. The image of L is the set

$$\mathrm{Im}(I) = \Big\{ (\mu, \eta_M) : \ \eta_N = (\eta_M \wedge N\mu) \vee (-N\mu) \quad \text{for } M > N \Big\}. \tag{12.1}$$

Notice that the compatibility condition $\eta_N = (\eta_M \wedge N\mu) \vee (-N\mu)$ implies that the Radon-Nikodym derivative $\varphi_N := \frac{d\eta_N}{d\mu}$ converges μ-a.e. to a uniquely identified $\varphi \in L(\mu; \mathbb{R} \cup \{\pm\infty\})$.

We observe that by Corollary 12.3 the function

$$(\mu, \eta) \mapsto F_N(\mu, \eta) := -(-(\eta \wedge N\mu) \wedge N\mu)$$

is Borel and $\mathrm{Im}(I)$ is the intersection of the following countably many graphs

$$\mathrm{Im}(I) = \bigcap_{N<M} \Big\{ (\mu, \{\eta_Q\}_Q) : F_N(\mu, \eta_M) = \eta_N \Big\}.$$

Being a Borel subset of a Polish space, by Theorem 3.2.4 of [Sri] there is a finer Polish topology on $\mathcal{P}([0,1]) \times \prod_{M=1}^{\infty} \mathcal{M}([0,1])$ such that $\mathrm{Im}(I)$ itself is Polish, and this Polish topology can be pulled back to L by the injective map I. The continuity of I, also w.r.t. the product weak* topology on the image space, is then immediate. □

Lemma 12.8. *The subset* $L^f := \{(\mu, \varphi) : \ \varphi \in L(\mu; \mathbb{R})\}$ *of* L *is analytic.*

Proof. If $\varphi \in L(\mu; \mathbb{R} \cup \{\pm\infty\})$, the condition $\varphi \in L(\mu; \mathbb{R})$ is clearly equivalent to $\lim_M \|\mu\|(|\varphi| > M) = 0$.
Since the injection I is continuous and $I(L)$ is Borel by Proposition 12.7, it is enough to prove that

$$I(L^f) = \Big\{ (\mu, \{\xi_M\}_M) : \ \lim_M \|\xi_{M+1} - \xi_M\| = 0 \Big\}$$

is analytic. By Lemma 12.6 this follows by the l.s.c. of the map

$$\begin{array}{ccc} \mathcal{P}([0,1]) \times \prod_{M=1}^{\infty} \mathcal{M}([0,1]) & \to & \mathbb{R}^{\mathbb{N}} \\ (\mu, \{\xi_M\}_M) & \mapsto & \{\|\xi_{M+1} - \xi_M\|\}_M \end{array}.$$

□

Theorem 12.9. *Let* c *be l.s.c.. Assume that the disintegration of* π *w.r.t. a partition* $\{X_\alpha \times Y_\alpha\}_\alpha$ *is strongly consistent and that there exist optimal couples* $\phi_\alpha \in \mathcal{B}(X_\alpha, \mathbb{R})$, $\psi_\alpha \in \mathcal{B}(Y_\alpha; \mathbb{R})$:

$$\phi_\alpha + \psi_\alpha \leqslant c \text{ on } X_\alpha \times Y_\alpha \qquad \phi_\alpha + \psi_\alpha = c \text{ on } \Gamma \cap X_\alpha \times Y_\alpha.$$

Then there exist Borel optimal potentials on $\cup_\alpha X_\alpha \times Y_\alpha$.

Proof. We prove the theorem by means of Von Neumann's selection principle.

Step 1. Consider the Polish space

$$Z := L \times L \times \mathcal{P}([0,1]^2).$$

We first prove the analyticity of the subset A of Z made of those

$$((\mu,\varphi),(\nu,\psi),\pi) \in L^f \times L^f \times \mathcal{P}([0,1]^2)$$

satisfying the relations

1. $(P_1)_\sharp \pi = \mu$, $(P_2)_\sharp \pi = \nu$;
2. $\phi + \psi \leqslant c$ out of cross-negligible sets w.r.t. the measures μ, ν;
3. $\phi + \psi = c$ π-a.e..

Since Σ^1_1 is closed under countable intersections, it suffices to show that each of the conditions above defines an analytic set.

Constraint (1) defines a closed set, by the continuity of the immersion I in Proposition 12.7 and because $\{(\mu,\nu,\pi) : \pi \in \Pi(\mu,\nu)\}$ is compact in $\mathcal{P}([0,1]) \times \mathcal{P}([0,1]) \times \mathcal{P}([0,1]^2)$.

Setting $\phi_M = ((\phi \wedge M) \vee (-M))$, $\psi_M = ((\psi \wedge M) \vee (-M))$ for $M \in \mathbb{N}$, Condition (2) is equivalent to

$$\int \phi_M (P_1)_\sharp \pi + \int \psi_M (P_2)_\sharp \pi \leqslant \int c\pi \qquad \forall \pi \in \Pi^{\leqslant}(\mu,\nu), \forall M \in \mathbb{N}. \tag{12.2}$$

Indeed, suppose that Condition (2) is not satisfied, i.e. the set $\{(x,y) : \phi(x) + \psi(y) > c(x,y)\}$ is not cross-negligible. Then, since ϕ_M, ψ_M converge to ϕ, ψ, the set $\{(x,y) : \phi_M(x) + \psi_M(y) > c(x,y)\}$ can't be cross-negligible. By the duality Theorem B.2 there exists a non-zero $\pi \in \Pi^{\leqslant}(\mu,\nu)$ concentrated on $\{(x,y) : \phi_M(x) + \psi_M(y) > c(x,y)\}$ and therefore (12.2) does not hold. The converse is immediate, as $\phi_M + \psi_M \leqslant c$.

We consider the Borel set (Lemma 12.5)

$$C_{n,M} := \left\{ (\mu,\nu,\xi,\eta,\pi) : H_M(\mu,\nu,\xi,\eta,\pi) - \int (c + 2M)\pi \geqslant 2^{-n}, \pi \in \Pi^{\leqslant}(\mu,\nu) \right\}. \tag{12.3}$$

Since for $\pi \in \Pi^{\leqslant}(\mu,\nu)$ one has

$$H_M(\mu,\nu,\xi,\eta,\pi) = \int \frac{d(\xi + M\mu)^+}{d\mu} (P_1)_\sharp \pi + \int \frac{d(\eta + M\nu)^+}{d\nu} (P_2)_\sharp \pi$$

then for fixed (μ, ν, η, ξ) the function

$$\pi \mapsto \begin{cases} H_M(\mu, \nu, \xi, \eta, \pi) - \int (c + 2M)\pi & \pi \in \Pi^{\leqslant}(\mu, \nu) \\ +\infty & \text{otherwise} \end{cases}$$

is u.s.c. for l.s.c. cost c. In particular the section

$$C_{n,M} \cap \{(\mu, \nu, \xi, \eta)\} \times \mathcal{P}([0,1]^2)$$

is closed, hence compact. By Novikov Theorem (Theorem 4.7.11 of [Sri]), it follows that

$$P_{1234}(C_{n,M}) = \left\{ (\mu, \nu, \eta, \xi) : \exists \pi \in \Pi^{\leqslant}(\mu, \nu), H_M(\mu, \nu, \eta, \xi, \pi) - \int (c + 2M)\pi \geqslant 2^{-n} \right\}$$

is Borel. Finally, the set

$$\begin{aligned} D_M &:= \bigcup_{n \in \mathbb{N}} P_{1234}(C_{n,M}) \\ &= \left\{ (\mu, \nu, \eta, \xi) : \exists \pi \in \Pi^{\leqslant}(\mu, \nu), H_M(\mu, \nu, \eta, \xi, \pi) - \int (c + 2M)\pi > 0 \right\} \end{aligned}$$

is Borel.

Condition (12.2) thus can be rewritten as

$$\left\{ (\mu, \nu, \{\xi_M\}_M, \{\eta_M\}) \in \mathcal{P}([0,1])^2 \times \left(\prod_{M=1}^{\infty} \mathcal{M}([0,1]) \right)^2 : (\mu, \nu, \xi_M, \eta_M) \notin D_M \right\},$$

and the above discussion implies that this is a Borel set.

We prove finally that Condition (3) identifies an analytic set. Consider the map

$$\begin{aligned} \left(\prod_{M=1}^{\infty} \mathcal{M}([0,1]) \times \mathcal{M}([0,1]) \right) \times \mathcal{P}([0,1]^2) &\to \mathbb{R}^{\mathbb{N}} \\ (\{\xi_M, \eta_M\}_M, \pi) &\mapsto \left\{ \int \xi_M + \int \eta_M - \int c\pi \right\}_M . \end{aligned}$$

This function is clearly Borel. Moreover, by Lemma 12.6 the family of sequences converging to 0 is an analytic subset of $\mathbb{N}^{\mathbb{N}}$, and therefore his counterimage is analytic. The thesis follows again by the continuity of the immersion I of Proposition 12.7.

Step 2. Since the set A of Step 1 is analytic, and the map $[0,1] \ni \alpha \mapsto \pi_\alpha \in \mathcal{P}([0,1]^2)$ can be assumed to be Borel, then the set $B = [0,1] \times A \cap \{(\alpha, (\mu, \varphi), (\nu, \psi), \pi : \pi = \pi_\alpha)\}$ is analytic.

Step 3. By Von Neumann's selection principle applied to B, there exists an analytic map

$$[0,1] \ni \alpha \mapsto \big((\mu_\alpha, \phi_\alpha), (\nu_\alpha, \psi_\alpha)\big) \in L \times L .$$

Hence, by the immersion of I of Proposition 12.7 we can define the sequence of measures

$$\xi_M := \int \xi_{M,\alpha} m(d\alpha) \qquad \eta_M := \int \eta_{M,\alpha} m(d\alpha).$$

It is not difficult to show that $\left(\mu, \{\xi_M\}_{M \in \mathbb{N}}\right)$ and $\left(\nu, \{\eta_M\}_{M \in \mathbb{N}}\right)$ belong to the image (12.1) of I: by the formula

$$\xi_M = \frac{d\xi_M}{d\mu}\mu = \int \left(\frac{d\xi_M}{d\mu}\mu_\alpha\right) m(d\alpha) = \int \xi_{M,\alpha} m(d\alpha)$$

it follows that $(\mu, \{\xi_M\}_M) \in L^f$ and satisfies the compatibility condition. Therefore taking the counterimage with I one can define functions $\phi \in L(\mu)$, $\psi \in L(\nu)$ which are global potentials. $\square$

Remark 12.10. Theorem 12.9 does not provide an optimal couple for a generic equivalence relation different from the axial one, and in particular it does not apply for the cycle equivalence relation (see Example 11.4). This holds if e.g. $\{c < \infty\} \subset \cup_\alpha X_\alpha \times Y_\alpha$.

Remark 12.11. Even if every two points are connected by an axial path and there exist Borel potentials, in general there is no point $(\bar{x}, \bar{y})$ such that the extensions of Corollary 10.19 define Borel potentials $\tilde{\phi}$, $\tilde{\psi}$.

Remark 12.12. In the proof one can observe that we can replace the cost c with any other cost c', just requiring that for m-a.e. α it holds $\phi_\alpha + \psi_\alpha \leqslant c'$. In particular, we can take a cost whose graph is σ-compact in each equivalence class and prove that the sets $C_{n,M}$ of (12.3) are σ-compact.

This shows how Theorem 12.9 can be extended to π-measurable costs.

13 Sudakov theorem with strictly convex, possibly asymmetric norms

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 one dimensional theory is well established: we state the known result in Theorem 13.1, 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

$$\phi(x) := \inf_{\substack{n,\, x_n := x, \\ i=1,\dots,n-1, \\ (x_i, y_i) \in \Gamma}} \left\{ \sum_{i=0}^{n-1} \left[\tilde{|\!|} y_i - x_{i+1} \tilde{|\!|} - \tilde{|\!|} y_i - x_i \tilde{|\!|} \right] \right\}, \qquad x \in \mathbb{R}^n,\ (x_0, y_0) \in \Gamma \text{ fixed}, \tag{13.1}$$

is 1-Lipschitz w.r.t. $\tilde{|\!|} \cdot \tilde{|\!|}$ and satisfies $\phi(x) - \phi(y) = \tilde{|\!|} y - x \tilde{|\!|}$ for every $(x, y) \in \Gamma$. For completeness we sketch the proof in Remark 13.6.

In the present section we fix the potential ϕ as in 13.1 and we consider the related transport set $\bar{\mathcal{T}}_e$. The transport rays of $\bar{\mathcal{T}}_e$ 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 $y \in \bar{\mathcal{P}}(x)$, being $\phi(x) - \phi(y) = \tilde{|\!|} y - x \tilde{|\!|}$: 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 $\bar{\mathcal{P}}(x)$ 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 4 yields the following information. Up to removing an $\mathscr{L}^n$-negligible set from $\bar{\mathcal{T}}_e$, the relation which defines the transport rays, precisely

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

is an equivalence relation. Equivalently, we are saying that the transport rays provide a partition of $\bar{\mathcal{T}}$ up to a μ-negligible set, into segments since the norm is strictly convex. Moreover, if one disintegrates the Lebesgue measure on $\bar{\mathcal{T}}_e$ w.r.t. this partition — identifying points on a same transport ray — the conditional measures are absolutely continuous by Theorem 4.26.

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 4.26 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

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

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. $c(x, y) = \tilde{\|} y - x \tilde{\|}$. 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 one dimensional transport problems, *and by the study of Section 4 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 fundamental, 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 $\bar{\mathcal{T}}_e \setminus \bar{\mathcal{T}}$ 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 $\bar{\mathcal{Z}}$, one can determine the portion of mass $\nu(\bar{\mathcal{Z}} \setminus \bar{\mathcal{T}})$ on the terminal points of $\bar{\mathcal{Z}}$ which comes from the rays in $\bar{\mathcal{Z}}$ as the mass initially present in $\bar{\mathcal{Z}}$ which has not been carried to $\bar{\mathcal{Z}} \cap \bar{\mathcal{T}}$; it is just the difference

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

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

Theorem 13.1 (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*

$$t(x) = \sup\left\{y \in \mathbb{R} : \ F(y) \leqslant G(x)\right\} \tag{13.2}$$

  (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 $c(x, y) = \tilde{\|}y - x\tilde{\|}$, and we assume that the optimal cost is finite.
- We fix the potential ϕ in 13.1 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 4.2, so that Theorem 4.26 holds.
- $p : \bar{\mathcal{T}} \to \mathcal{S}$ 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 $\mu = \mu \llcorner \bar{\mathcal{T}}$ w.r.t. the partition in transport rays.
- Let $\nu \llcorner \bar{\mathcal{T}} = \int_S \tilde{\nu}_z \tilde{m}(z)$ be the disintegration of $\nu \llcorner \bar{\mathcal{T}}$ 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 13.2. *Define on* $\bar{\mathcal{T}}$ *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*

$$T : \ z \mapsto \begin{cases} z & \text{if } z \notin \bar{\mathcal{T}} \\ x + td(x) & \text{where } t = \sup\{s : \ F(x + sd(x)) \leqslant G(z)\}, \text{ if } z \in \bar{\mathcal{T}} \end{cases}. \tag{13.3}$$

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

Proof. The proof follows the one 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 4.26

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

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

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

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 one dimensional theory, Theorem 13.1, an optimal transport map from $(\bar{\mathcal{R}}(y), \mu_y)$ to $(\bar{\mathcal{R}}(y), \nu_y)$ is given by the restriction of T in 13.3 to $\bar{\mathcal{R}}(y)$.

Step 3: Measurability of T. The map T in 13.3 is Borel, not only on the rays, but in the whole $\bar{\mathcal{T}}$. To see it, consider the countable partition of $\bar{\mathcal{T}}$ into σ-compact sets $\bar{\mathcal{Z}}(Z_k)$, by Lemma 4.11. In particular, a subset C of $\bar{\mathcal{T}}$ is Borel if and only if its intersections with the $\bar{\mathcal{Z}}(Z_k)$ are Borel. Moreover, composing T with the Borel change of variable given in Remark 4.15, from the sheaf set $\bar{\mathcal{Z}}$ we can reduce to $(0,1)e_1 + Z$, $d(x) = e_1$ and the map T takes the form $T(y) = y + (T\cdot e_1 - y\cdot e_1)e_1$. One, then, has just to prove that the map $T\cdot e_1$ is Borel: this map is monotone in the first variable, and Borel in the second; in particular, it is Borel on $(0,1)\cdot Z$.

Step 4: Disintegration of optimal transport plans. Consider any transport plan $\pi \in \Pi(\mu,\nu)$. By construction of ϕ with 13.1, if (x,y) belong to the support of π then $\phi(x) - \phi(y) = |\!|\!|y-x|\!|\!|$: the support of π is then contained in $\cup_{y\in S}\bar{\mathcal{R}}(y)\times\bar{\mathcal{R}}(y) \cup \{x=y\}$.

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

As a consequence, one can disintegrate π w.r.t. $\{(\!|a(y), b(y)|\!) \times \bar{\mathcal{R}}(y)\}_{y\in\mathcal{S}}$ by Theorem 2.7.

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

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

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

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

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

$$\int_{\bar{\mathcal{R}}(z)\times\bar{\mathcal{R}}(z)} \tilde{\|}x - y\tilde{\|} \, d\pi_z(x, y) \geqslant \int_{\bar{\mathcal{R}}(z)} \tilde{\|}x - T(x)\tilde{\|} \, d\mu_z(x), \tag{13.4}$$

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

$$\begin{aligned}
\int \tilde{\|}x - y\tilde{\|} \, d\pi' \geqslant \int \tilde{\|}x - y\tilde{\|} \, d\pi &= \int_{\mathcal{S}} \left\{ \int_{\bar{\mathcal{R}}(z)\times\bar{\mathcal{R}}(z)} \tilde{\|}x - y\tilde{\|} \, d\pi_z(x, y) \right\} dm(z) \\
&\overset{13.4}{\geqslant} \int_{\mathcal{S}} \left\{ \int_{\bar{\mathcal{R}}(z)\times\bar{\mathcal{R}}(z)} x - T(x) \, d\mu_z(x) \right\} dm(z) = \int \tilde{\|}x - T(x)\tilde{\|} \, d\mu
\end{aligned}$$

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

$$T = \mathbb{I}_{\mathbb{R}^n \setminus \bar{\mathcal{T}}} + \sum_{y\in S} T_z \chi_{\llcorner(\!|a(z), b(z)|\!)},$$

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

Definition 13.3. We call t^{-1} the surjective multivalued function, monotone along each ray, whose graph contains the transpose of the graph of T. Let $\tilde{t}^{-1}$ be the single valued function whose graph is contained in the graph of t^{-1} and which is left continuous (and monotone nondeacreasing) on secondary transport rays.
Then $\nu((\!|a(x), x|\!)) = \mu((\!|a(x), \tilde{t}^{-1}(x)|\!))$ and $\nu((\!|a(x), x]\!]) = \mu((\!|a(x), t^{-1}(x)|\!))$, where $a(x)$ denotes formally the first endpoint of $r(x)$.

THE TRANSPORT DENSITY. As a further application of the disintegration theorem, we write the expression of the transport density relative to optimal secondary transport plans in terms of the conditional measures μ_z, ν_z, $z \in Q$ of μ, ν, for the ray equivalence relation. In particular, one can see its absolute continuity. As in the smooth setting the density function w.r.t. $\mathscr{L}^n \llcorner \bar{\mathcal{T}}$ vanishes approaching initial points along secondary transport rays. As known, the same property does not hold for the terminal points — see Example 13.5 below taken from [FM2].

We omit the verification, since it is quite standard (see e.g. Section 8 in [BG]).

Let f, the Radon-Nycodim derivative of μ w.r.t. $\mathscr{L}^n$, and γ, introduced in the disintegration, be Borel functions from $\mathbb{R}^n$ to $\mathbb{R}$ such that

$$\mu \llcorner \bar{\mathcal{T}} = \int_Q \mu_z \, d\mathcal{H}^{n-1}(z) = \int_Q (f\gamma\mathcal{H}^1 \llcorner r_z) \, d\mathcal{H}^{n-1}(z) \qquad \nu \llcorner \bar{\mathcal{T}} = \int_Q \nu_z \, d\mathcal{H}^{n-1}(z).$$

Let $z : \bar{\mathcal{T}}_e \to Q$ be the Borel multivalued quotient projection. Set $d = 0$ where $\bar{\mathcal{D}}$ is multivalued.

Lemma 13.4. *A particular solution $\rho \in \mathcal{M}^+_{loc}(\mathbb{R}^n)$ to the transport equation*

$$\operatorname{div}(d\rho) = \mu - \nu$$

is given by

$$\rho(x) = \frac{(\mu_{z(x)} - \nu_{z(x)})((\!|a(x), x|\!))}{\gamma(x)} \mathscr{L}^n(x) \llcorner \bar{\mathcal{T}} = \left(\frac{\chi_{\bar{\mathcal{T}}}(x)}{\gamma(x)} \int_{(\!|\bar{t}^{-1}(x), x|\!)} f\gamma \, d\mathcal{H}^1 \right) \mathscr{L}^n(x). \tag{13.5}$$

Example 13.5 (Taken from [FM2]). Consider in $\mathbb{R}^2$ the measures $\mu = 2\mathscr{L}^2 \llcorner \mathbf{B}_1$ and $\nu = \frac{1}{2|x|^{3/2}} \mathscr{L}^2 \mathcal{B}_1$, where $|\cdot|$ here denotes the Euclidean norm. A Kantorovich potential is provided by $|x|$. The transport density is $\rho = (|x|^{-\frac{1}{2}} - |x|) \mathscr{L}^2 \llcorner \mathbf{B}_1$. While vanishing towards $\partial \mathbf{B}_1$, the density of ρ blows up towards the origin. Concentrating ν at the origin, the density would be instead $\rho = -|x|^2 \llcorner \mathbf{B}_1$.

We sketch finally in the following remark the proof of the standard claim in the introduction. One could see that the potential defined by 13.1 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 13.6. 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

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

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 13.1,

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

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 $\phi(y)$ is estimated from above by

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

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

13.1 Remarks on the Decomposition in Transport Rays

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 $\bar{\mathcal{T}}$, which extends it and has analogous properties; however in dimension $n \geqslant 2$ even this extension does not fill the space, for any Kantorovich potential of the transport problem — as shown in Example 13.8.

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

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

This means that $\pm\nabla\phi \in D$. Consider now a point where, moreover, there is an outgoing ray. As an immediate consequence of 4.2, 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

$$d \in \bar{\mathcal{D}}(x) \quad \text{satisfies} \quad \frac{d(x)}{\tilde{\|}d(x)\tilde{\|}} \in \delta D(-\partial\phi(x)) \qquad (d(x) \in \partial\phi(x) \text{ if } D \text{ is a ball}). \tag{13.6}$$

Equation 13.6 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

$$d(x) = \delta D(-\nabla\phi(x)), \qquad \text{where } -\nabla\phi(x) \in \partial D. \tag{13.7}$$

This, generally, extends the vector field we analyzed (see Example 13.7), 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 13.8 we find that the gradient of ϕ can vanish on sets with $\mathscr{L}^n$-postive measure.

Example 13.7 (Transport rays do not fill continuously the line). Consider in $[0,1]$ the following transport problem, with $c(x,y) = |y-x|$ (Figure 17).

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

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

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

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

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 13.7.

Notice that the divergence of the vector field of ray directions is not a locally finite measure.

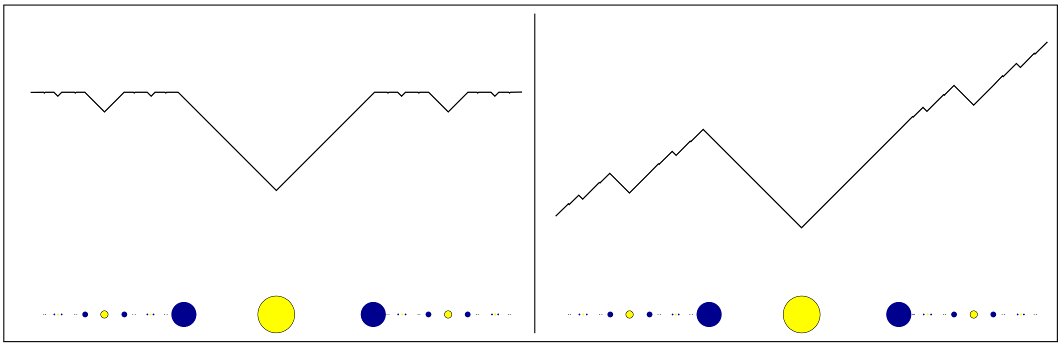


Figure 17: Example 13.7. The pictures show the graphs of the Kantorovich potentials ϕ (on the left) and $\tilde{\phi}$ (on the right), for the same transport problem — source masses are the blue ones, destinations the yellow ones. With $\tilde{\phi}$ the vector field of ray directions is defined $\mathscr{L}^1$-a.e., while with ϕ this is not the case, since the gradient vanishes in a $\mathscr{L}^1$-positive measure set.

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

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

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 13.1. 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 13.8 ($\bar{\mathfrak{T}}$ does not fill the space). Consider in the unit square $X = Y = [0,1]^2$ the following transport problem (see Figure 18).

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

$$\ell_i = \frac{\lambda^{\frac{1}{2^{i+1}}} \ell_{i-1}}{2} = \lambda^{\sum_{j=2}^{i+1} 2^{-j}} 2^{-i-1}, \qquad a_i = \ell_{i-1} - 2\ell_i,$$

$$n_i \text{ maximum in } 2\mathbb{N} \text{ s.t. } r_i := \frac{\ell_i + a_i}{n_i} < a_i.$$

Define moreover the sequence of centers, for $i \in \mathbb{N}$,

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

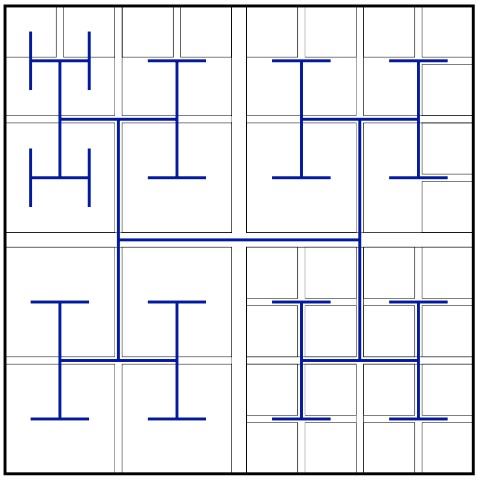


Figure 18: Example 13.8. In this case, however one chooses the potential, the vector field of direction is not defined on the whole square. In fact, the potential must be constant in points, belonging to the blue skeleton we begin to draw, dense in a $\mathscr{L}^2$-positive measure set.

and, finally, the intermediate points

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

Then, the marginal measures be given by

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

One can immediately verify that the transport plan

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

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

Let ϕ be any 1-Lipschitz function whose c-subdifferential contains the support of π. Since $\phi(y) - \phi(x) = \tilde{\|} y - x \tilde{\|}$ 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 13.6, that K is in the complementary of $\bar{\mathcal{T}}$. 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 13.7 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. $\mathscr{L}^2$.

Part III

Appendix

A

An Isomorphism Theorem

We describe here a standard measurable space isomorphism among countably generated measure spaces to transform the mass transport problem from (X, Ω), (Y, Σ) to Borel probability spaces on $[0, 1]$. We refer to the last theorem in [HJ].

The notations follow the ones of Chapter 2. We recall that a probability space (X, Ω, μ) is essentially countably generated by the family $\{B_m\}_{m \in \mathbb{N}}$ if $B_m \in \Omega$, $m \in \mathbb{N}$, and $\forall S \in \Omega$ there exists A in the σ-algebra generated by B_m such that $\mu(S \triangle A) = 0$.

Lemma A.1. *Consider a probability space (X, Ω, μ) essentially countably generated by the family $\{B_m\}_{m \in \mathbb{N}}$. If $\mathcal{B}$ is the Borel σ algebra of $[0, 1]$, then the map $([0, 1], \Omega) \to ([0, 1], \mathcal{B})$*

$$f(x) = \sum_{m \in \mathbb{N}} 10^{-m} \chi_{B_m}(x)$$

realizes an isomorphism between the measure algebras of (X, Ω, μ) and of the Borel probability space $(X, \mathcal{B}, f_\sharp \mu)$.

Proof. Since the measure algebra of a space is isomorphic to the one of its completion, we directly assume that Ω is generated by $\{B_m\}_{m \in \mathbb{N}}$.
By the measurability of f one has that $f_\sharp \Omega \supset \mathcal{B}$ by definition.
Moreover, if we identify those points of X which are not separated by $\{B_m\}_{m \in \mathbb{N}}$ then f becomes injective. This implies that for $B \in \Omega$ then $f^{-1}(f(B)) = B$ (every map $g : ([0, 1], f_\sharp \Omega) \to (X, \Omega)$ such that $g(z) \in f^{-1}(z)$ is measurable). Thus $f_\sharp \Omega$ is generated by $f(B_m)$, $m \in \mathbb{N}$, and f immediately induces an isomorphism of measure algebras. $\square$

Since f is injective except for collapsing the atoms, and since every measurable function and measure must be constant on atoms, while measures do not 'break' them, this map is suitable to translate the transport problems we considered in Chapter 7-10 to

1. $X = Y = [0, 1]$;
2. $\mu, \nu \in \mathcal{P}([0, 1], \mathcal{B})$.

If for example $c : X \times Y \to [0, +\infty]$ is a $\Pi(\mu, \nu)$-universally measurable cost, then

$$\hat{c} = c \circ (f^1 \otimes f^2).$$

will be the $\Pi(f^1_\sharp \mu, f^2_\sharp \nu)$-universally measurable cost, where f^1, f^2 are given by Lemma A.1.

This is possible since the formulation of our problems are invariant for measure space isomorphism, if the necessary and sufficient conditions are — and it takes a while to see that this is the case. Indeed, $\Pi(\mu,\nu)$ is isomorphic to $\Pi(f^1_\sharp\mu, f^2_\sharp\nu)$, $\Pi^f(\mu,\nu)$ is isomorphic to $\Pi^f(f^1_\sharp\mu, f^2_\sharp\nu)$, measurable c-monotone sets are mapped to measurable c'-monotone sets, measurable acyclic sets to measurable acyclic sets, measurable A-acyclic sets to measurable A-acyclic sets and so on.

We have similarly the following particular result of a more general classification of Polish spaces, that we state for the sake of completeness. The proof is straightforward.

Let Z be a Polish space, and $\{B_m\}_{m\in\mathbb{N}}$ a family of open balls which is a basis for the topology. Define $f : Z \to [0,1]$ as in Lemma A.1.

Proposition A.2. *The map* f *is one to one and lower semicontinuous. Moreover,* f^{-1} *is continuous on* $f(z)$, $f_\sharp(\mathscr{B}(Z)) \supset \mathscr{B}([0,1])$ *and* ${f^{-1}}_\sharp(\mathscr{B}\llcorner_{f(Z)}) \subset \mathscr{B}(Z)$.

Given a lower semicontinuous cost function c, in this case one can define similarly to above the cost $\tilde{c}$ in $[0,1]^2$ as the following lower semicontinuous envelope

$$\tilde{c} = \text{l.s.c. env.}\left(\begin{cases} c(f^1(s), f^2(t)) & \text{for } (s,t) \in f(X \times Y) \\ +\infty & \text{otherwise} \end{cases}\right).$$

One easily verifies that $\tilde{c}$ coincides with c on $f(X \times Y)$. Therefore also in this case it is not restrictive to work in $[0,1]^2$ preserving the lower semicontinuity of the cost.

B

Perturbation by cycles

For the particular applications we are considering, the geometrical constraints allow only to perturb a given measure $\pi \in \Pi(\mu, \nu)$ by means of bounded measures $\lambda \in \mathcal{M}([0,1]^2)$ with 0 marginals, and such that $\pi + \lambda \geqslant 0$. The simplest way of doing this perturbation is to consider closed cycles in $[0,1]^2$: we will call this types of perturbation *perturbation by cycles* (a more precise definition is given below).

The problem of checking whether a measure μ can be perturbed by cycles has been considered in several different contexts, see for example [AP, BGMS, HW]. Here we would like to construct effectively a perturbation, which will be (by definition) a perturbation by cycles.

Since we are using a duality result valid only for analytic costs, in the following we will restrict to a coanalytic cost c. We first recall some useful results on analytic subsets of Polish spaces (in our case $[0,1]$), and the main results of [Kel].

B.1 Borel, analytic and universally measurable sets

Our main reference is [Sri].

The *projective class* $\Sigma^1_1(X)$ is the family of subsets A of the Polish space X for which there exists Y Polish and $B \in \mathcal{B}(X \times Y)$ such that $A = P_1(B)$. The *coprojective class* $\Pi^1_1(X)$ is the complement in X of the class $\Sigma^1_1(X)$. The σ-algebra generated by Σ^1_1 is denoted by $\mathcal{A}$.

The *projective class* $\Sigma^1_{n+1}(X)$ is the family of subsets A of the Polish space X for which there exists Y Polish and $B \in \Pi^1_n(X \times Y)$ such that $A = P_1(B)$. The *coprojective class* $\Pi^1_{n+1}(X)$ is the complement in X of the class $\Sigma^1_{n+1}(X)$.

If Σ^1_n, Π^1_n are the projective, coprojective pointclasses, then the following holds (Chapter 4 of [Sri]):

1. Σ^1_n, Π^1_n are closed under countable unions, intersections (in particular they are monotone classes);

2. Σ^1_n is closed w.r.t. projections, Π^1_n is closed w.r.t. coprojections;

3. the *ambiguous class* $\Delta^1_n := \Sigma^1_n \cap \Pi^1_n$ is a σ-algebra and $\Sigma^1_n \cup \Pi^1_n \subset \Delta^1_{n+1}$.

We recall that a subset of X Polish is *universally measurable* set if it belongs to all completed σ-algebras of all Borel measures on X: it can be proved that every set in $\mathcal{A}$ is universally measurable.

Under the axiom of Projective Determinacy (PD) all projective sets are universally measurable, and PD is undecidable in ZFC ([MS, Mos]). *In the rest of the* *present Appendix we choose to assume* (PD). One could avoid this assumption by recovering independently the measurability of the functions we are going to define by countable limit procedures (see for example [Car1]), but since our aim is to describe a construction is not sufficiently motivated here.

In the following we will then use the fact that Borel counterimages of universally measurable sets are universally measurable.

B.2 General duality results

All the results recalled in this sections are contained in [Kel].

Let $A \subset [0,1]^d$ be a subset of $[0,1]^d$, and consider Borel probabilities $\mu_i \in \mathcal{P}([0,1])$, $i = 1, \dots, d$. We want to know if there is a measure π such that $\pi^*(A) > 0$ and its marginals are bounded by the measure μ_i: $(P_i)_\sharp \pi \leqslant \mu_i$. We recall that

$$\pi^*(A) := \inf \left\{ \pi(A') : A' \in \mathcal{B}([0,1]^d), A \subset A' \right\} \tag{B.1}$$

is the outer π measure. For simplicity, we will denote the i-th measure space in the product with X_i.

Definition B.1. A set $A \subset [0,1]^d$ is *cross-negligible w.r.t. the measures* μ_i, $i = 1, \dots, d$, if there are negligible sets N_i, $i = 1, \dots, d$, such that $A \subset \cup_i P_i^{-1}(N_i)$.

Given $A_1, A_2 \in [0,1]^d$, we define

$$\mathrm{dist}(A_1, A_2) := \inf \left\{ \sum_{i=1}^d \int h_i \mu_i : \chi_{A_1 \triangle A_2}(x) \leqslant \sum_{i=1}^d h_i(x_i),\ h_i \in L^\infty(\mu_i) \right\}. \tag{B.2}$$

We say that $A_1, A_2 \subset [0,1]^d$ are equivalent and we write $A_1 \sim_{\mathrm{dist}} A_2$ if $A_1 \triangle A_2$ is cross negligible, i.e. $\mathrm{dist}(A_1, A_2) = 0$.

This definition is the same as the L-shaped sets defined in [BGMS]. Clearly the cross-negligible sets can be taken to be G_δ-sets. The fact that $\sim_{\mathrm{dist}}$ is an equivalence relation and that $(\mathbf{P}(X)/\sim_{\mathrm{dist}}, \mathrm{dist})$ is a metric space is proved in [Kel], Proposition 1.15. Following again [Kel], given $\mathcal{A} \subset \mathbf{P}(X)$, we denote $\bar{\mathsf{A}}$ the closure of $\mathcal{A}$ w.r.t. the distance dist.

The next theorem collects some of the main results of [Kel]. This results are duality results, which compare the supremum of a linear function in the convex set

$$\Pi(\mu_1, \dots, \mu_d) := \left\{ \pi \in \mathcal{P}([0,1]^d) : (P_i)_\sharp \pi = \mu_i, i = 1, \dots, d \right\}$$

with the infimum of a convex function in a predual space.

Theorem B.2. *If* $A \in \overline{\Sigma_1^1(\mathbb{R}^d)}$, *then the following duality holds*

$$\sup \left\{ \pi(A) : \pi \in \Pi(\mu_1, \dots, \mu_d) \right\} = \min \left\{ \sum_{i=1}^d \int h_i \mu_i : \sum_{i=1}^d h_i(x_i) \geqslant \chi_A(X), 0 \leqslant h_i \leqslant 1 \right\}.$$

(B.3)

Moreover, if A is in closure w.r.t. d of the family of closed sets, then the max on the l.h.s. is reached. In particular the maximum is reached when A is in the class of countable intersections of elements of the product algebra.

Proof. The fact that the duality (B.3) holds with the infimum in the r.h.s. is a consequence of [Kel], Theorem 2.14. In our settings the analytic sets contain all the Borel sets, so that in particular the duality holds for Borel sets.

The fact that the minimum is reached is a consequence of [Kel], Theorem 2.21.

Finally, the last assertion follows from [Kel], Theorem 2.19, and the subsequent remarks. □

A fairly easy corollary is that if the supremum of (B.3) is equal to 0, then A is cross negligible.

Remark B.3. Note that since we are considering a maximum problem for a positive linear functional, then the problem is equivalent when considered in the larger space

$$\Pi^{\leqslant}(\mu_1,\dots,\mu_d) := \Big\{0 \leqslant \pi \in \mathcal{M}([0,1]^d) : (P_i)_\sharp \pi \leqslant \mu_i, i = 1,\dots,d\Big\}.$$

B.3 Decomposition of measures with 0 marginals

In this section we decompose a measure with 0 marginals into its essentially cyclic part and acyclic part. The decomposition is not unique, even if we can determine if a perturbation is essentially cyclic or acyclic.

Let Λ be the convex closed set of Borel measures on $[0,1]^2$ with 0 marginals:

$$\Lambda := \Big\{\lambda \in \mathcal{M}([0,1]^d) : (P_i)_\sharp \lambda = 0,\ i = 1,\dots,d\Big\}. \tag{B.4}$$

In the following we restrict to $d = 2$, in view of applications to the transport problem in $[0,1]^2$.

Definition B.4. We define the following sets.

The *configuration set*

$$C_n := \Big\{w \in [0,1]^{2n} : P_{2i-1}w \neq (P_{2i+1 \mod 2n})w,\ P_{2i}w \neq (P_{2i+2 \mod 2n})w, i = 1,\dots,n\Big\}.$$

The *phase set*

$$D_n := \Big\{z \in [0,1]^{4n} : (P_{4i-1}, P_{4i})z = (P_{4i+1 \mod 4n}, P_{4i-2 \mod 4n})z, i = 1,\dots,n\Big\}.$$

The the set of finite cycles, of arbitrary length, D_∞

$$D_\infty := \Big\{z \in [0,1]^{2\mathbb{N}} : (P_{4i-1}, P_{4i})z = (P_{4i+1}, P_{4i-2})z, \exists k : P_{4kj+i}z = P_i z, 1 \leqslant i \leqslant k, j \in \mathbb{N}\Big\}.$$

(B.5)

The *projection operator*

$$q : [0,1]^{4n} \to [0,1]^{2n}, \quad (P_{2i-1}, P_{2i})q(z) = (P_{4i-3}, P_{4i-2})z, \ i = 1, \dots, n.$$

The *reduced phase set*

$$\tilde{D}_n := q^{-1}(C_n) \cap D_n.$$

The *narrow configuration set* and *narrow phase space*

$$\hat{C}_n := \Big\{ w \in [0,1]^{2n} : \ (P_{2i-1}, P_{2i})w \neq (P_{2j-1}, P_{2k})w, i \neq j, k \Big\}, \quad \hat{D}_n := q^{-1}(\hat{C}_n) \cap D_n. \tag{B.6}$$

Remark B.5. The following remarks are straightforward.

1. The set C_n is open not connected in $[0,1]^{2n}$, and its connected components are given by the family of sets

$$C_{n,I} := \Big\{ w \in [0,1]^{2n} : \ P_{2i-1}w \gtrless (P_{2i+1 \mod 2n})w, \ P_{2i}w \gtrless (P_{2i+2 \mod 2n})w, i = 1, \dots, n \Big\}$$

   for the 4 possible choices of the inequalities and of $i \in \{1, \dots, n\}$.

2. The set D_n is compact connected, and the set $\tilde{D}_n$ can be written as

$$\tilde{D}_n := \Big\{ z \in [0,1]^{4n} : \ (P_{4i-1}, P_{4i})z = (P_{4i+1 \mod 4n}, P_{4i-2 \mod 4n})z, \\ P_{4i-3}z \neq (P_{4i+1 \mod 4n})z, \ P_{4i-2}z \neq (P_{4i+2 \mod 4n})z, i = 1, \dots, n \Big\}$$

3. Both sets C_n and D_n are invariant for the cyclical permutation of coordinates T defined by $(P_{i+2 \mod n})(Tw) = P_i w$, $i = 1, \dots, 2n$ in $[0,1]^{2n}$ and by $q^{-1}Tq$ on D_n.

4. The narrow phase set is made by cycles of length exactly n.

We give now the following definitions.

Definition B.6. A measure λ is n-*cyclic*, or a n-*cycle*, if there exists $m \in \mathcal{M}^+(C_n)$ such that

$$\lambda^+ = \frac{1}{n} \int_{C_n} \sum_{i=1}^{n} \delta_{P_{(2i-1,2i)}w} m(dw), \quad \lambda^- = \frac{1}{n} \int_{C_n} \sum_{i=1}^{n} \delta_{P_{(2i+1,2i \mod 2n)}w} m(dw). \tag{B.7}$$

A n-cyclic measure λ is a *simple* n-*cycle* if m is supported on a set $q(Q)$ with

$$Q = \Big\{ z \in D_n : (P_{4i-3}, P_{4i-2})z \in (x_i, y_i) + [-\epsilon, \epsilon]^2, \min_{i,j} \big\{ |x_i - x_j|, |y_i - y_j| \big\} \geqslant 2\epsilon \Big\}.$$

A measure λ is *cyclic* if there exist $m_n \in \mathcal{M}^+(C_n)$, $n \in \mathbb{N}$, such that $\sum_n m_n(C_n) < \infty$ and

$$\lambda^+ = \sum_n \frac{1}{n} \int_{C_n} \sum_{i=1}^n \delta_{P_{(2i-1,2i)} w} m_n(dw), \quad \lambda^- = \sum_n \frac{1}{n} \int_{C_n} \sum_{i=1}^n \delta_{P_{(2i+1 \mod 2n, 2i)} w} m_n(dw). \tag{B.8}$$

From the definition of simple n-cycles it follows that there are disjoint $2n$ sets $(x_i, y_i) + [-\epsilon, \epsilon]^2$, $(x_{i+1 \mod n}, y_i) + [-\epsilon, \epsilon]^2$, $i = 1, \dots, n$, such that

$$\lambda^+ \left(\bigcup_{i=1}^n (x_i, y_i) + [-\epsilon, \epsilon]^2 \right) + \lambda^- \left(\bigcup_{i=1}^n (x_{i+1 \mod n}, y_i) + [-\epsilon, \epsilon]^2 \right) = |\lambda|.$$

The next lemma is a simple consequence of the separability of $[0,1]^{4n}$ and the fact that $\hat{C}_n$ is open.

Lemma B.7. *Each n-cyclic measure λ of the form*

$$\lambda^+ = \frac{1}{n} \int_{\hat{C}_n} \sum_{i=1}^n \delta_{P_{(2i-1,2i)} w} m(dw), \quad \lambda^- = \frac{1}{n} \int_{\hat{C}_n} \sum_{i=1}^n \delta_{P_{(2i+1, 2i \mod 2n)} w} m(dw)$$

can be written as the sum of simple n-cycles λ_i so that

$$\lambda^+ = \sum_i \lambda_i^+, \quad \lambda^- = \sum_i \lambda_i^-.$$

B.3.1 n-*cyclic components of a measure*

Consider the Jordan decomposition of $\lambda \in \Lambda$,

$$\lambda = \lambda^+ - \lambda^- \qquad \lambda^+ \perp \lambda^-, \ \lambda^+, \lambda^- \geqslant 0,$$

and the Borel sets A^+, A^- of the Hahn decomposition:

$$A^+ \cap A^- = \emptyset, \ A^+ \cup A^- = [0,1]^2, \qquad \lambda^+ = \lambda \llcorner_{A^+}, \ \lambda^- = \lambda \llcorner_{A^-}.$$

Define then

$$\mu_{2i-1} := \lambda^+, \quad \mu_{2i} := \lambda^- \tag{B.9}$$

with $i = 1, \dots, n$.

From Theorem B.2 and the fact that D_n is compact, the following proposition follows.

Proposition B.8. *Let μ_i as in (B.9). There exists a solution to the marginal problem, for $n \in \mathbb{N}$,*

$$\max \left\{ \pi(\tilde{D}_n) : \pi \in \Pi(\mu_1, \dots, \mu_{2n}) \right\} = \min \left\{ \sum_{i=1}^{2n} \int_{[0,1]^2} h_i \mu_i : \sum_{i=1}^{2n} h_i((P_{2i-1}, P_{2i})z) \geqslant \chi_{\tilde{D}_n}(z) \right\}. \tag{B.10}$$

Proof. It is enough to prove that D_n is in the equivalence class of $\tilde{D}_n$ w.r.t. $\sim_{\mathrm{dist}}$: from this it follows that for every measure in $\Pi(\mu_i)$ one has $\pi(D_n) = \pi(\tilde{D}_n)$, and then one can apply Theroem B.2.

Step 1. By definition

$$D_n \setminus \tilde{D}_n \subset \bigcup_{i=1}^{n} \Big\{ z : P_{4i-3}z = (P_{4i+1 \mod 4n}) \text{ or } P_{4i-2}z = (P_{4i+2 \mod 4n})z \Big\},$$

so that if $z \in D_n \setminus \tilde{D}_n$ for at least one i

$$(P_{4i-3}, P_{4i-2})z = (P_{4i-1}, P_{4i})z \quad \text{or} \quad (P_{4i-1}, P_{4i})z = (P_{4i+1 \mod 4n}, P_{4i+2 \mod 4n})z. \tag{B.11}$$

Step 2. Consider the functions, for $i = 1, \dots, n$,

$$f_{2i-1} = \chi_{[0,1]^2 \setminus A^+}, \quad f_{2i} = \chi_{[0,1]^2 \setminus A^-}.$$

Since $f_{2i-1} + f_{2i} \geqslant 1$, it follows from (B.11) that

$$\sum_{i=1}^{n} f_{2i-1}\big((P_{4i-3}, P_{4i-2})z\big) + f_{2i}\big((P_{4i-1}, P_{4i})z\big) \geqslant \chi_{D_n \setminus \tilde{D}_n}.$$

Step 3. Since $\lambda^+(A^-) = \lambda^-(A^+) = 0$, then

$$\sum_{i=1}^{n} \int_{[0,1]^2} f_{2i-1}\mu_{2i-1} + \int_{[0,1]^2} f_{2i}\mu_{2i} = \sum_{i=1}^{n} \lambda^+(A^-) + \lambda^-(A^+) = 0.$$

Hence $\mathrm{dist}(D_n, \tilde{D}_n) = 0$. □

We now define the n-cyclic components of λ.

Definition B.9. Let π be a maximizer for (B.10) and define the measure

$$\lambda_n := \frac{1}{n}\sum_{i=1}^{n} (P_{4i-3}, P_{4i-2})_\sharp \pi \llcorner_{\tilde{D}_n} - \frac{1}{n}\sum_{i=1}^{n} (P_{4i-1}, P_{4i})_\sharp \pi \llcorner_{\tilde{D}_n}.$$

We say that λ_n is the (or better a) n-*cyclic component* of λ.

Remark B.10. The following are easy remarks.

1. $0 \leqslant \lambda_n^+ \leqslant \lambda^+$ and $0 \leqslant \lambda_n^- \leqslant \lambda^-$: in fact, by construction

$$0 \leqslant (P_{4i-3}, P_{4i-2})_\sharp \pi \llcorner_{\tilde{D}_n} \leqslant \lambda^+, \quad 0 \leqslant (P_{4i-1}, P_{4i})_\sharp \pi \llcorner_{\tilde{D}_n} \leqslant \lambda^+. \tag{B.12}$$

Moreover, by the definition of D_n, it follows that

$$\big|(P_{4i-3}, P_{4i-2})_\sharp \pi \llcorner_{\tilde{D}_n}\big| = \big|(P_{4i-1}, P_{4i})_\sharp \pi \llcorner_{\tilde{D}_n}\big| = \pi(D_n),$$

so that $|\lambda_n| = 2\pi(D_n)$.

2. If π is a maximum, also the symmetrized measure

$$\tilde{\pi} := \frac{1}{n} \sum_{i=0}^{n-1} \Big(\underbrace{T \circ \cdots \circ T}_{i-\text{times}} \Big)_\sharp \pi$$

is still a maximum. For this measure $\tilde{\pi}$ it follows that

$$\lambda_n = (P_{4i-3}, P_{4i-2})_\sharp \tilde{\pi} \llcorner_{\tilde{D}_n} - (P_{4i-1}, P_{4i})_\sharp \tilde{\pi} \llcorner_{\tilde{D}_n} \tag{B.13}$$

for all $i = 1, \dots, n$. In particular, if we consider again the problem (B.10) with $\lambda_n^\pm$ as marginals in (B.9), then $\tilde{\pi}$ is still a maximum. However, there are maxima which are not symmetric, and for which the projection on a single component does not exhibit a cyclic structure, as in Example B.17.

3. The n-cyclic part of λ is a n-cyclic measure, as one can see by the trivial disintegration

$$\pi = \int_{C_n} \delta_{q^{-1}(w)} m(w), \qquad m(w) := (q_\sharp \pi)(w).$$

Conversely, if λ is n-cyclic, then $\pi \llcorner_{D_n} = (q^{-1})_\sharp m$ is a maximum for the problem (B.10).

Note that the condition

$$\lambda = \frac{1}{n} \int_{C_n} \sum_{i=1}^{n} \Big(\delta_{P_{(2i-1,2i)} w} - \delta_{P_{(2i+1 \mod 2n, 2i)} w} \Big) m(dw)$$

is not sufficient, because of cancellation, as it can be easily seen by the measure

$$\lambda = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

4. If $\lambda_n = 0$, it follows from the duality stated in Theorem B.2 that D_n is cross negligible, so that there exists Borel sets N_i, $i = 1, \dots, n$ such that

$$\lambda^+(N_{2i-1}) = \lambda^-(N_{2i}) = 0 \quad \text{and} \quad D_n \subset \bigcup_{i=1}^{2n} (P_i)^{-1}(N_i).$$

Hence the sets

$$N^+ = \bigcup_{i=1}^{n} N_{2i-1}, \quad N^- = \bigcup_{i=1}^{n} N_{2i}$$

still satisfy $\lambda^+(N^+) = \lambda^-(N^-) = 0$ and

$$D_n \cap \bigcap_{i=1}^{n} (P_{2i-1})^{-1}(N^+)^c \cap (P_{2i})^{-1}(N^-)^c = \emptyset.$$

We thus conclude that if $\lambda_n = 0$ there exist Borel sets A^+, A^- such that λ^+ is concentrated in A^+, λ^- is concentrated in A^- and there is no n-cycle $\{(x_i, y_i), i = 1, \dots, n\}$ such that $(x_i, y_i) \in A^+$ and $(x_{i+1 \mod n}, y_i) \in A^-$ for all $i = 1, \dots, n$.

Define the measure $\lambda_{\not n} := \lambda - \lambda_n$.

Lemma B.11. *The n-cyclic component of $\lambda_{\not n}$ is zero. Equivalently, $\lambda_{\not n}$ satisfies*

$$\max\left\{\pi(\tilde{D}_n), \pi \in \Pi(\mu_1, \dots, \mu_{2n})\right\} = 0 \tag{B.14}$$

for the marginal problem

$$\mu_i = \begin{cases} \lambda^+_{\not n} & i \text{ odd} \\ \lambda^-_{\not n} & i \text{ even} \end{cases}.$$

Proof. If in (B.14) we have a positive maximum π', then we can assume this maximum to be symmetric, so that (B.13) holds. Let π be a symmetric positive maximum of the original (B.10): by construction we have that

$$0 \leqslant \lambda^+_n = (P_{(1,2)})_\sharp \pi \llcorner D_n \leqslant \lambda^+, \quad 0 \leqslant \lambda^-_n = (P_{(3,4)})_\sharp \pi \llcorner D_n \leqslant \lambda^-$$

$$0 \leqslant (P_{(1,2)})_\sharp \pi' \leqslant \lambda^+ - \lambda^+_n, \quad 0 \leqslant (P_{(3,4)})_\sharp \pi' \leqslant \lambda^- - \lambda^-_n,$$

so that

$$0 \leqslant \lambda^+_n + (P_{(1,2)})_\sharp \pi' = (P_{(1,2)})_\sharp (\pi + \pi') \leqslant \lambda^+, \quad 0 \leqslant \lambda^-_n + (P_{(3,4)})_\sharp \pi' = (P_{(3,4)})_\sharp (\pi + \pi') \leqslant \lambda^-,$$

and $(\pi + \pi')(D_n) > \pi(D_n)$, contradicting the maximality of π. $\square$

A measure can be decomposed into a *cyclic and an acyclic part* by removing n-cyclic components for all $n \in \mathbb{N}$ (see Remark B.13). However, when removing a n-cyclic component the m-cyclic components are affected, for $m \neq n$. More clearly, the following observations are in order.

For all $n, k \in \mathbb{N}$ one has

$$\max\left\{\pi(\tilde{D}_n), \pi \in \Pi(\mu_1, \dots, \mu_{2n})\right\} \leqslant \max\left\{\pi(\tilde{D}_{kn}), \pi \in \Pi(\mu_1, \dots, \mu_{2kn})\right\},$$

because if π_1 is a measure in $\Pi(\mu_1, \dots, \mu_{2n})$, then the measure

$$\pi_2 = \big(\underbrace{\mathbb{I}_n, \dots, \mathbb{I}_n}_{k-\text{times}}\big)_\sharp \pi_1$$

belongs to $\Pi(\mu_1, \dots, \mu_{2kn})$ and $\pi_2(\tilde{D}_{kn}) = \pi_2(D_{kn}) = \pi_1(D_n) = \pi_1(\tilde{D}_n)$.

However, in general

$$\max\left\{\pi(\tilde{D}_n), \pi \in \Pi(\mu_1, \dots, \mu_{2n})\right\} + \max\left\{\pi(\tilde{D}_n), \pi \in \Pi(\nu_1, \dots, \nu_{2kn})\right\}$$
$$< \max\left\{\pi(\tilde{D}_{kn}), \pi \in \Pi(\mu_1, \dots, \mu_{2kn})\right\},$$

where we define

$$\nu_i = \begin{cases} \lambda^+_{\not 1} & i \text{ odd} \\ \lambda^-_{\not 1} & i \text{ even} \end{cases} \qquad \mu_i = \begin{cases} \lambda^+ & i \text{ odd} \\ \lambda^- & i \text{ even} \end{cases}.$$

This can be seen in Example B.17, by taking $n = 2$ and $k = 4$: in fact for any choice of the maximal solution for $n = 2$ the remaining measure $\lambda - \lambda_{\not 2}$ does not contain any cycle of length 8, while λ itself is a cycle of length 8. It follows

$$\max\Big\{\pi(\tilde{D}_n) : \pi \in \Pi(\mu_1, \dots, \mu_{2n})\Big\} + \max\Big\{\pi(\tilde{D}_n) : \pi \in \Pi(\nu_1, \dots, \nu_{2kn})\Big\} = 2 \\ < 8 = \max\Big\{\pi(\tilde{D}_{kn}) : \pi \in \Pi(\mu_1, \dots, \mu_{2kn})\Big\}.$$

An even more interesting example is provided in Example B.18, where it is shown that a measure can be decomposed into a cyclic and an acyclic part in different ways, and the mass of each part depends on the decomposition one chooses.

B.3.2 *Cyclic and essentially cyclic measures*

Given a sequence of marginals μ_i, let

$$\Pi_\infty(\{\mu_i\}_i) = \Big\{\pi \in \mathcal{P}([0,1]^{2\mathbb{N}}) : (P_i)_\sharp \pi = \mu_i, i \in \mathbb{N}\Big\},$$

Consider following problem in $[0,1]^{2\mathbb{N}}$:

$$\sup\Big\{\pi(D_\infty), (P_{2i-1})_\sharp \pi = \lambda^+, (P_{2i})_\sharp \pi = \lambda^-, i \in \mathbb{N}\Big\}. \tag{B.15}$$

Definition B.12. We say that a measure $\lambda \in \Lambda$ is *essentially cyclic* if

$$\sup\Big\{\pi(D_\infty), (P_{2i-1})_\sharp \pi = \lambda^+, (P_{2i})_\sharp \pi = \lambda^-, i \in \mathbb{N}\Big\} = \lambda^+([0,1]^2) = \lambda^-([0,1]^2).$$

It is clear that if λ is cyclic, then the maximum exists, and viceversa (Remark B.10, Point (3), observing that $D_n \hookrightarrow D_\infty$). If λ is acyclic, then the supremum is equal to 0. Since D_∞ is not closed in $[0,1]^{2\mathbb{N}}$, we cannot state that such a maximum exists.

Remark B.13. We now construct a special decomposition, whose cyclic part however is not necessary maximal.

Define recursively the marginal problem in D_n by

$$\mu_{2n-1} := \lambda^+ - \sum_{i=2}^{n-1} \lambda_i^+ \quad \mu_{2n} := \lambda^- - \sum_{i=2}^{n-1} \lambda_i^-, \tag{B.16}$$

where λ_i is given at the i-th step and λ_n is obtained by

$$\lambda_n := \frac{1}{n}\sum_{i=1}^{n} (P_{4i-3}, P_{4i-2})_\sharp \pi \llcorner_{\tilde{D}_n} - \frac{1}{n}\sum_{i=1}^{n} (P_{4i-1}, P_{4i})_\sharp \pi \llcorner_{\tilde{D}_n}$$

solving the problem

$$\max\Big\{\pi(\hat{D}_n), \pi \in \Pi(\mu_1, \dots, \mu_{2n})\Big\} = \min\bigg\{\sum_{i=1}^{2n} \int_{[0,1]^2} h_i(x)\mu_i,\ \sum_{i=1}^{2n} h_i((P_{2i-1}, P_{2i})z) \geqslant \chi_{\hat{D}_n}\bigg\}. \tag{B.17}$$

Let $\{\pi_n\}_{n\in\mathbb{N}}$ be the sequence of maxima for (B.17). There is a canonical way to embed π_n in $\Pi_\infty(\{\mu_i\}_i)$, with μ_i given by (B.16) for $i \in \mathbb{N}$ (here we assume that $\|\lambda\| = 2$). In fact, it is enough to take

$$T_n : [0,1]^{4n} \to [0,1]^{2\mathbb{N}}, \quad z \mapsto T_n(z) = (z, z, z, \dots), \quad \tilde{\pi}_n = (T_n)_\sharp \pi_n.$$

Hence the measure $\tilde{\pi} = \sum_n \tilde{\pi}_n$ belongs to π_∞, the series being strongly converging, and since every map T_n takes values in D_∞, the measure $\tilde{\pi}$ satisfies

$$\tilde{\pi}(D_\infty) = \sum_n \pi_n(D_n).$$

Example B.18 implies that in general $\tilde{\pi}$ it is not a supremum.

Similarly, the measures $\sum_i^n \lambda_i^+$, $\sum_i^n \lambda_i^-$ are strongly convergent to measures λ_c^+, λ_c^-.

The sets D_n are cross negligible for the marginals

$$\mu_i = \begin{cases} \lambda_a^+ = \lambda^+ - \lambda_c^+ & i \text{ odd} \\ \lambda_a^- = \lambda^- - \lambda_c^- & i \text{ even} \end{cases}$$

This follows easily from (B.14) and the fact that the series of λ_n is converging.

Hence, from Point (4) of Remark B.10, one concludes that λ_a^+, λ_a^- are supported on two disjoint sets A_a^+, A_a^-, respectively, so that there are no closed cycles $\{(x_i, y_i), i = 1, \dots, n\}$, $n \in \mathbb{N}$, such that $(x_i, y_i) \in A_a^+$ and $(x_{i+1}, y_i) \in A_a^-$ for all $i = 1, \dots, n$ and $(x_{n+1}, y_{n+1}) = (x_1, y_1)$.

B.3.3 *Perturbation of measures*

For a measure $\pi \in \mathcal{P}([0,1]^2)$ and an analytic set $A \subset [0,1]^2$ such that $\pi(A) = 1$, we give the following definition.

Definition B.14. A *cyclic perturbation on* A *of the measure* π is a cyclic, nonzero measure λ concentrated on A and such that $\lambda^- \leqslant \pi$. When not specified, $A = [0,1]^2$.

Proposition B.15. *If there is no cyclic perturbation of* π *on* A*, then there is* Γ *with* $\pi(\Gamma) = 1$ *such that for all finite sequences* $(x_i, y_i) \in \Gamma$, $i = 1, \dots, n$, *with* $x_i \neq x_{i+1 \mod n}$ *and* $y_i \neq y_{i+1 \mod n}$ *it holds*

$$\Big\{(x_{i+1}, y_i), i = 1, \dots, n, x_{n+1} = x_1\Big\} \not\subset A.$$

Proof. Define the set on n-cycles in A as

$$C_{n,A} := q(D_n \cap \prod^{2n} A).$$

The fact that there is no cyclic perturbation means that for all $n \in \mathbb{N}$

$$\sup\left\{m \in \Pi(\pi, \dots, \pi) : \ m(q(D_n \cap \Pi^{2n} A))\right\} = 0.$$

Then $q(D_n \cap \Pi^{2n} A)$ is cross-negligible by Theorem B.2: there exist π-negligible sets N_i such that

$$\prod^{n}(\Gamma \setminus (\cup_i N_i)) \cap q(D_n \cap \Pi^{2n} A) = \emptyset.$$

The set $\Gamma \setminus (\cup_i N_i)$ satisfies the statement of the proposition. □

Proposition B.16. *If there is no cyclic perturbation λ of π such that $\mathcal{I}(\pi + \lambda) < \mathcal{I}(\pi)$, then there is Γ with $\pi(\Gamma) = 1$ such that for all finite sequences $(x_i, y_i) \in \Gamma$, $i = 1, \dots, I$, $x_{I+1} := x_1$ it holds*

$$\sum_{i=1}^{I} \left[c(x_{i+1}, y_i) - c(x_i, y_i)\right] \geqslant 0. \tag{B.18}$$

Proof. Let Γ a σ-compact carriage of π such that $c \llcorner_\Gamma$ is Borel. The set

$$Z_n = \left\{(x_1, y_1, \dots, x_n, y_n) \in \Gamma^n : \sum_{i=1}^{n} \left[c(x_{i+1}, y_i) - c(x_i, y_i)\right] < 0\right\} \cap C_n$$

is analytic: in fact, being the sum of a Borel function and an Π^1_1-function, the function

$$\sum_{i=1}^{n} \left[c(x_{i+1}, y_i) - c(x_i, y_i)\right]$$

is a Π^1_1-function.

The fact that there is no cyclic perturbation λ of π which lowers the cost $\mathcal{I}$ means that for all n

$$\sup\left\{m \in \Pi(\pi, \dots, \pi) : \ m(Z_n)\right\} = 0,$$

otherwise the projected measure λ satisfies

$$\int c\lambda = \int_{Z_n} \frac{1}{n} \sum_{i=1}^{n} \left[c(x_{i+1}, y_i) - c(x_i, y_i)\right] m(dx_1 dy_1 \dots dx_n dy_n) < 0$$

contradicting optimality of π, as $\pi + \lambda$ would be a transference plan with lower cost.

Theorem B.2 implies that there are π-negligible sets $N_{n,i} \subset [0,1]^2$, $i = 1, \dots, n$, such that

$$Z_n \subset \bigcup_{i=1}^{n} (P_{2i-1,2i})^{-1}(N_{n,i}).$$

The set $\Gamma \setminus \cup_{i=1}^{n} N_{n,i}$ satisfies then (B.18) for cycles of length I at most n. The c-cyclically monotone set Γ proving the lemma is finally $\Gamma \setminus \cup_n \cup_{i=1}^{n} N_{n,i}$. □

B.3.4 *Examples*

We give now some examples.

Example B.17. Here we show that there are maxima of the problem (B.10) which are not symmetric, and for which the projection on a single component does not exhibit any cyclic structure. Consider the following example (since the measures are atomic, we use a matrix notation):

$$\lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & -1 & 0 \\ -1 & 1 & 0 & -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

It is easy to verify that the maximum in the problem (B.10) with $n = 2$ is 2, by just considering the functions

$$h_1 = h_3 = 1 - \chi_{\mathrm{supp}\lambda^+}, \quad h_2 = h_4 = 1 - \chi_{\mathrm{supp}\lambda^-} + \delta_{\{3,2\}},$$

and that a maximizer is the measure:

$$\bar{\pi} \llcorner_{\tilde{D}_2} = \delta_{(\{3,1\},\{3,2\},\{4,2\},\{4,1\})} + \delta_{(\{2,2\},\{2,5\},\{3,5\},\{3,2\})}.$$

It follows that $\lambda_2 \neq (P_1, P_2)_\sharp \bar{\pi} \llcorner_{\tilde{D}_2} - (P_3, P_4)_\sharp \bar{\pi}$: indeed

$$(P_1, P_2)_\sharp \bar{\pi} \llcorner_{\tilde{D}_2} - (P_3, P_4)_\sharp \bar{\pi} \llcorner_{\tilde{D}_2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Conversely the symmetrized measure yields

$$\lambda_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & -1/2 & 0 & 0 \\ 1/2 & -1 & 0 & 0 & 1/2 & 0 & 0 \\ -1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

This example proves also that we do not have uniqueness, by just observing that

$$\begin{aligned} &\big\{\pi : \pi(D_2) = 2, \pi \in \Pi(\lambda^+, \lambda^-, \lambda^+, \lambda^-)\big\} \\ &\qquad = \Big\{\alpha_1 \delta_{(\{3,1\},\{3,2\},\{4,2\},\{4,1\})} + \alpha_2 \delta_{(\{4,2\},\{4,1\},\{3,1\},\{3,2\})} \\ &\qquad\qquad + \alpha_3 \delta_{(\{2,2\},\{2,5\},\{3,5\},\{3,2\})} + \alpha_4 \delta_{(\{3,5\},\{3,2\},\{2,2\},\{2,5\})}, \alpha_i \geqslant 0, \sum_{i=1}^4 \alpha_i = 1\Big\}. \end{aligned}$$

Hence the symmetrized set of π and the projected set are

$$\Big\{\alpha_1\Big(\delta_{(\{3,1\},\{3,2\},\{4,2\},\{4,1\})}+\delta_{(\{4,2\},\{4,1\},\{3,1\},\{3,2\})}\Big)\Big)$$
$$+\alpha_2\Big(\delta_{(\{2,2\},\{2,5\},\{3,5\},\{3,2\})}+\delta_{(\{3,5\},\{3,2\},\{2,2\},\{2,5\})}\Big),\alpha_i\geqslant 0,\sum_{i=1}^{2}\alpha_i=\frac{1}{2}\Big\},$$

$$\left\{\alpha_1\begin{bmatrix}0&0&0&0&0&0&0\\0&1&0&0&-1&0&0\\0&-1&0&0&1&0&0\\0&0&0&0&0&0&0\\0&0&0&0&0&0&0\\0&0&0&0&0&0&0\end{bmatrix}+\alpha_2\begin{bmatrix}0&0&0&0&0&0&0\\0&0&0&0&0&0&0\\1&-1&0&0&0&0&0\\-1&1&0&0&0&0&0\\0&0&0&0&0&0&0\\0&0&0&0&0&0&0\end{bmatrix},\alpha_i\geqslant 0,\sum_{i=1}^{2}\alpha_i=1\right\}.$$

Example B.18. Here we decompose the measure

$$\lambda:=\begin{bmatrix}0&0&0&0&-m_1&m_1\\-m_1&0&0&m_1&m_1&-m_1\\0&0&0&-m_2&m_1&0\\m_1&-m_1&0&0&0&0\\0&m_1&-m_1&0&0&0\\0&0&m_1&0&-m_1&0\end{bmatrix}$$

into an essentially cyclic and an acyclic part in two different ways, and the two acyclic part will not even have the same mass. Let

$$m_1:=(\mathbb{I},\mathbb{I})_\sharp\mathcal{L}^1\llcorner_{[0,a]},\quad m_2:=(\mathbb{I}+\alpha\mod a,\mathbb{I})_\sharp\mathcal{L}^1\llcorner_{[0,a]},$$

with $\alpha\in\mathbb{R}\setminus\mathbb{Q}$. Depending on $n=2$ or $n=4$ we obtain the following two decompositions:

$$\lambda=\begin{bmatrix}0&0&0&0&-m_1&m_1\\0&0&0&0&m_1&-m_1\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix}+\begin{bmatrix}0&0&0&0&0&0\\-m_1&0&0&m_1&0&0\\0&0&0&-m_2&m_1&0\\m_1&-m_1&0&0&0&0\\0&m_1&-m_1&0&0&0\\0&0&m_1&0&-m_1&0\end{bmatrix},$$

$$\lambda=\begin{bmatrix}0&0&0&0&0&0\\-m_1&0&0&0&m_1&0\\0&0&0&0&0&0\\m_1&-m_1&0&0&0&0\\0&m_1&-m_1&0&0&0\\0&0&m_1&0&-m_1&0\end{bmatrix}+\begin{bmatrix}0&0&0&0&-m_1&m_1\\0&0&0&m_1&0&-m_1\\0&0&0&-m_2&m_1&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix}.$$

The first measure is cyclic and the second is acyclic, because of m_2.

C

Tensors and currents

Our main references are Chapter 4 of [Mor] and Sections 1.5.1, 4.1 of [Fed].

Let $\{e_1, \dots, e_n\}$ be a basis of $\mathbb{R}^n$. The **wedge product** between vectors is multilinear and alternating: for $m \in \mathbb{N}$, $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, $0 < i \leqslant m$, $u_0, \dots, u_m \in \mathbb{R}^n$

$$\Big(\sum_{i=1}^{n} \lambda_i e_i\Big) \wedge u_1 \wedge \cdots \wedge u_m = \sum_{i=1}^{n} \lambda_i (e_i \wedge u_1 \wedge \cdots \wedge u_m)$$
$$u_0 \wedge \cdots \wedge u_i \wedge \cdots \wedge u_m = (-1)^i u_i \wedge u_0 \wedge \cdots \wedge \widehat{u_i} \wedge \cdots \wedge u_m,$$

where the element under the hat is missing. The space of all linear combinations of

$$\big\{e_{i_1 \dots i_m} := e_{i_1} \wedge \cdots \wedge e_{i_m} : \ i_1 < \cdots < i_m \text{ in } \{1, \dots, n\}\big\}$$

is the space of m-**vectors**, denoted by $\Lambda_m \mathbb{R}^n$. The space $\Lambda_0 \mathbb{R}$ is just $\mathbb{R}$. $\Lambda_m \mathbb{R}^n$ has the **inner product** given by

$$e_{i_1 \dots i_m} \cdot e_{j_1 \dots j_m} = \prod_{k=1}^{m} \delta_{i_k j_k} \qquad \text{where } \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

The induced **norm** is denoted by $|\cdot|$. An m-**vector field** is a map $\xi : \mathbb{R}^n \to \Lambda_m \mathbb{R}^n$.

The dual Hilbert space to $\Lambda_m \mathbb{R}^n$, denoted by $\Lambda^m \mathbb{R}^n$, is the space of m-**covectors**. The element dual to $e_{i_1 \dots i_m}$ is denoted by $de_{i_1 \dots i_m}$. A **differential** m-**form** is a map $\omega : \mathbb{R}^n \to \Lambda^m \mathbb{R}^n$.

We denote with $\langle \cdot, \cdot \rangle$ the duality pairing between m-vectors and m-covectors. Moreover, the same symbol denotes in this paper the bilinear **pairing**, which is a map $\Lambda^p \mathbb{R}^n \times \Lambda_q \mathbb{R}^n \to \Lambda^{p-q} \mathbb{R}^n$ for $p > q$ and $\Lambda^p \mathbb{R}^n \times \Lambda_q \mathbb{R}^n \to \Lambda_{q-p} \mathbb{R}^n$ for $q > p$ whose non-vanishing images on a basis are

$$de_{i_1 \dots i_\ell} = \big\langle de_{i_1 \dots i_\ell} \wedge de_{i_{\ell+1} \dots i_{\ell+m}}, e_{i_{\ell+1} \dots i_{\ell+m}} \big\rangle \qquad \text{if } p = \ell + m > m = q$$
$$e_{i_{\ell+1} \dots i_{\ell+m}} = \big\langle de_{i_1 \dots i_\ell}, e_{i_1 \dots i_\ell} \wedge e_{i_{\ell+1} \dots i_{\ell+m}} \big\rangle \qquad \text{if } p = \ell < \ell + m = q.$$

Consider any differential m-form

$$\omega = \sum_{i_1 \dots i_m} \omega_{i_1 \dots i_m} \, de_{i_1 \dots i_m}$$

which is differentiable. The **exterior derivative** $d\omega$ of ω is the differential $(m+1)$-form

$$d\omega = \sum_{i_1 \dots i_m} \sum_{j=1}^{n} \frac{\partial \omega_{i_1 \dots i_m}}{\partial x_j} \, de_j \wedge de_{i_1 \dots i_m}.$$

If $\omega \in C^i(\mathbb{R}^n; \Lambda^m \mathbb{R}^n)$, the i-th exterior derivative is denoted with $d^i\omega$.
Consider any m-vector field

$$\xi = \sum \xi_{i_1\dots i_m} e_{i_1\dots i_m}$$

which is differentiable. The **pointwise divergence** $(\operatorname{div}\xi)_{\text{a.c.}}$ of ξ is the $(m-1)$-vector field

$$(\operatorname{div}\xi)_{\text{a.c.}} = \sum_{i_1\dots i_m} \sum_{j=1}^{n} \frac{\partial \xi_{i_1\dots i_m}}{\partial x_j} \big\langle de_j,\, e_{i_1\dots i_m}\big\rangle.$$

Consider the space $\mathscr{D}^m$ of C^∞-differential m-form with compact support. The topology is generated by the seminorms

$$\nu_K^i(\phi) = \sup_{x\in K,\, 0\leqslant j\leqslant i} |d^j\phi(x)| \qquad \text{with } K \text{ compact subset of } \mathbb{R}^n,\ i \in \mathbb{N}.$$

The dual space to $\mathscr{D}^m$, endowed with the weak topology, is called the space of **m-dimensional currents** and it is denoted by $\mathscr{D}_m$. The support of a current $T \in \mathscr{D}_m$ is the smallest close set $K \subset \mathbb{R}^n$ such that $T(\omega) = 0$ whenever $\omega \in \mathscr{D}^m$ vanishes out of K. The **mass** of a current $T \in \mathscr{D}_m$ is defined as

$$\mathbf{M}(T) = \sup\left\{T(\omega):\ \omega \in \mathscr{D}^m,\ \sup_{x\in\mathbb{R}^n} |\omega(x)| \leqslant 1\right\}.$$

The **flat norm** of a current $T \in \mathscr{D}_m$ is defined as

$$\mathbf{F}(T) = \sup\left\{T(\omega):\ \omega \in \mathscr{D}^m,\ \sup_{x\in\mathbb{R}^n} |\omega(x)| \leqslant 1,\ \sup_{x\in\mathbb{R}^n} |d\omega(x)| \leqslant 1\right\}.$$

An m-dimensional current $T \in \mathscr{D}_m$ is **representable by integration**, and we denote it by $T = \mu \wedge \xi$, if there exists a Radon measure μ over $\mathbb{R}^n$ and a μ-locally integrable m-vector field ξ such that

$$T(\omega) = \int_{\mathbb{R}^n} \big\langle \omega,\, \xi\big\rangle\, d\mu \qquad \forall \omega \in \mathscr{D}^m.$$

If $m \geqslant 1$, the **boundary** of an m-dimensional current T is defined as

$$\partial T \in \mathscr{D}_{m-1}, \qquad \big(\partial T\big)(\omega) = T(d\omega) \text{ whenever } \omega \in \mathscr{D}^{m-1}.$$

If either $m = 0$, or both T and ∂T are representable by integration, then we will call T **locally normal**. If T is locally normal and compactly supported, then T is called **normal**. The $\mathbf{F}$-closure, in $\mathscr{D}_m$, of the normal currents is the space of **locally flat chains**. Its subspace of currents with finite mass is the $\mathbf{M}$-closure, in $\mathscr{D}_m$, of the normal currents.

To each $\mathscr{L}^n$-measurable m-vector field ξ such that $|\xi|$ is locally integrable there corresponds the current $\mathscr{L}^n \wedge \xi \in \mathscr{D}_m(\mathbb{R}^n)$. If ξ is of class C^1, then this current is locally normal and the divergence of ξ is related to the boundary of the corresponding current by

$$-\partial\big(\mathscr{L}^n \wedge \xi\big) = \mathscr{L}^n \wedge (\operatorname{div}\xi)_{\text{a.c.}},$$

Moreover, if Ω is an open set with C^1 boundary, $\hat{n}$ is its outer unit normal and $d\hat{n}$ the dual of $\hat{n}$, then

$$\partial\big(\mathscr{L}^n \wedge (\chi_\Omega \xi)\big) = -(\mathscr{L}^n \llcorner \Omega) \wedge (\operatorname{div} \xi)_{\text{a.c.}} + (\mathscr{H}^{n-1} \llcorner \partial\Omega) \wedge \big\langle d\hat{n}, \xi \big\rangle. \quad \text{(C.1)}$$

In Chapter 6 we found the analogue of the Green-Gauss Formula (C.1) for the k-dimensional current associated to k-faces, restricted to $\mathcal{D}$-cylinders. In order to do this, we re-defined the function $(\operatorname{div} \xi)_{\text{a.c.}}$ for a less regular k-vector field and this definition is an extension of the above one.

Notations

$\mathbb{N}, \mathbb{N}_0, \mathbb{Q}, \mathbb{R}$	natural numbers, natural numbers with 0, rational numbers, real numbers
$\mathcal{B}$ or $\mathcal{B}(X)$	Borel σ-algebra of the topological space $(X, \mathcal{T})$
$\mathcal{M}(X)$ or $\mathcal{M}(X, \Omega)$	signed measures on a measurable space (X, Ω)
$\mathcal{M}^+(X)$ or $\mathcal{M}^+(X, \Omega)$	positive measures on a measurable space (X, Ω)
$\mathcal{P}(X)$ or $\mathcal{P}(X, \Omega)$	probability measures on a measurable space (X, Ω)
$L(\mu; J)$	μ-measurable maps from the measure space (X, Ω, μ) to $J \subset \mathbb{R} \cup \{\pm\infty\}$
$\mathscr{L}^d$	d-dimensional Lebesgue measure
$\mathscr{H}^d$	d-dimensional Hausdorff outer measure
$L^1_{(loc)}(\mu)$	(locally) integrable functions (w.r.t. μ)
$L^\infty_{(loc)}$	(locally) essentially bounded functions
$C^k_{(c)}$	k-times continuously differentiable functions (with compact support)
$\Pi(\mu_1, \ldots, \mu_I)$	$\pi \in \mathcal{P}(\Pi_{i=1}^I X_i, \otimes_{i=1}^I \Sigma_i)$ with marginals $(P_i)_\sharp \pi = \mu_i \in \mathcal{P}(X_i)$
$\Pi^{\leqslant}(\mu_1, \ldots, \mu_I)$	$\pi \in \mathcal{M}(\Pi_{i=1}^I X_i, \otimes_{i=1}^I \Sigma_i)$, $\pi \geqslant 0$, with $(P_i)_\sharp \pi \leqslant \mu_i \in \mathcal{P}(X_i)$
$\Pi^f(\mu, \nu)$	$\pi \in \Pi(\mu, \nu)$ for which $\mathcal{I}(\pi) \in \mathbb{R}$
$\Pi^{opt}(\mu, \nu)$	$\pi \in \Pi(\mu, \nu)$ for which $\mathcal{I}(\pi)$ is minimal
$P_{i_1 \ldots i_I}$	projection of $x \in \Pi_{k=1,\ldots,K} X_k$ into its $(i_1, \ldots, i_I)$ coordinates, keeping order
$d\mu_2 / d\mu_1$	Radon-Nikodym derivative of (the absolutely continuous part of) μ_2 w.r.t. μ_1
$\mu = \int \mu_\alpha \, d\nu$	disintegration of μ, see Definition 2.6
$\mu \ll \nu$	$\mu(A) = 0$ whenever $\nu(A) = 0$ (absolute continuity of a measure μ w.r.t. ν)
$\|\nu\|$, $(\mu)^+$	the nonnegative measures variation and positive part of $\mu \in \mathcal{M}(X)$
$\mu \wedge \nu$, $\mu \vee \nu$	the measures minimum and maximum of $\mu, \nu \in \mathcal{M}(X)$
π^*	outer measure (B.1)
$\tau_\sharp$	The push forward with a measurable map τ, see Chapter 2
$\mathbf{P}(X)$	power set of X
$c(x, y)$, $\mathcal{I}(\pi)$	cost function and cost functional (7.2); in Chapter 13 $c(x, y) = \tilde{\|} y - x \tilde{\|}$
χ_A	characteristic function of A, $\chi_A = \chi_A : x \mapsto \delta_x(A)$
$\mathbf{I}_A$	indicator function of A, $\mathbf{I}_A(x) = \frac{1 - \chi_A(x)}{\chi_A(x)} \in \{0, +\infty\}$
$A \vartriangle B$	symmetric difference between two sets A, B
$\mathrm{dist}(A, B)$	distance defined in B.2
$\mathrm{graph}(f)$	graph of the function $f : X \to Y$, $\mathrm{graph}(f) = \{(x, y), y = f(x)\} \subset X \times Y$
$\mathrm{epi}(f)$	epigraph of function f, $\mathrm{epi}(f) = \{(x, y), y \geqslant f(x)\} \subset X \times \mathbb{R}$
$\mathbb{I}$, $\mathbb{I}_d$	identity operator on a set and on the space $\mathbb{R}^d$
Λ	measures $\lambda \in \mathcal{M}([0, 1]^d)$ with 0 marginals, see (B.4)
Γ	c-cyclically monotone σ-compact subset of $[0, 1]^2$
$\Gamma(A)$, $\Gamma^{-1}(B)$	the sets $\Gamma(A) = P_2(\Gamma \cap P_1^{-1}(A))$, $\Gamma(B) = P_1(\Gamma \cap P_2^{-1}(B))$
C_n	configuration set of n-cycles (B.4)
D_n	phase set of n-cycles (B.4)
q	projection operator (B.4)
$\tilde{D}_n$	reduced phase set of n-cycles (B.4)
T	cyclical permutation of coordinates, defined in Point (3) at Page 176
A_x, A^x	the sections $\{y : (x, y) \in A\}$, $\{y : (y, x) \in A\}$ for $A \subset X \times Y$

Θ_π	π-completion of the Borel σ-algebra
Θ	$\Pi(\mu,\nu)$-universal σ-algebra (7.1)
xRy, R	a binary relation R over X
$\mathrm{graph}(R)$	graph of the binary relation R, $\mathrm{graph}(R)=\{(x,y):xRy\}\subset X^2$
$x\sim y$ or xEy, E	an equivalence relation over X with graph E
$x^\bullet$	equivalence class of x, $x^\bullet=E_x$
$A^\bullet$	saturated set for an equivalence relation, $A^\bullet=\cup_{x\in A}x^\bullet$
$X^\bullet$, $X/\sim$	quotient space of an equivalence relation
Σ^1_1, $\Sigma^1_1(X)$	the pointclass of analytic subsets of Polish space X, i.e. projection of Borel sets
Π^1_1	the pointclass of coanalytic sets, i.e. complementary of Σ^1_1
Σ^1_n, Π^1_n	the pointclass of projections of Π^1_{n-1}-sets, its complementary
Δ^1_n	the ambiguous class $\Sigma^1_n\cap\Pi^1_n$
$\mathcal{A}$	σ-algebra generated by Σ^1_1
$\mathcal{A}$-function	$f:X\to\mathbb{R}$ such that $f^{-1}((t,+\infty])$ belongs to $\mathcal{A}$
F_h	the set where the function h is finite (10.11)
$\Sigma^1_n(\Pi^1_n,\Delta^1_n)$-function	$f:X\to\mathbb{R}$ such that $f^{-1}((t,+\infty])\in\Sigma^1_n\ (\Pi^1_n,\ \Delta^1_n)$
equivalent	μ is eqivalent to ν if $\mu\ll\nu$ and $\nu\ll\mu$
separated	two sets A and B sets are separated if each is disjoint from the other's closure
perpendicular	A set A is perpendicular to an affine plane H of $\mathbb{R}^d$ if $\exists w\in H$ s.t. $\pi_H(A)=w$
$v\cdot w$	Euclidean scalar product in $\mathbb{R}^n$
$\lvert\cdot\rvert$	Euclidean norm in $\mathbb{R}^n$
$\tilde{\lVert}\cdot\tilde{\rVert}$	A possibly asymmetric norm on $\mathbb{R}^n$ whose unit ball is strictly convex
$(\!\lvert a,b\rvert\!)$	segment in $\mathbb{R}^n$ from a to b, without the endpoints
$[\![a,b]\!]$	segment in $\mathbb{R}^n$ from a to b, including the endpoints
D^*	unit ball $\{x\in\mathbb{R}^n:\tilde{\lVert}x\tilde{\rVert}\leqslant 1\}$
D	dual convex set of D^*: $D=\Big\{\ell:\ \ell\cdot d\leqslant 1\ \forall d\in D^*\Big\}$
∂D	boundary of D
δD	support cone of D at $\ell\in\partial D$: $\delta D=\Big\{d\in\partial D^*:\ d\cdot\ell=1=\sup_{\hat\ell\in\partial D}d\cdot\hat\ell\Big\}$
ϕ	See Definition 4.1 and (13.1)
$\partial_c\phi$	c-subdifferential of ϕ, $\quad\partial_c\phi=\Big\{(x,y):\ \phi(x)-\phi(y)=c(x,y)\Big\}$
$\mathbf{S}^{n-1},\mathbf{B}^n$	Respectively unit sphere and unit ball of $\mathbb{R}^n$
$\mathbf{G}(k,n)$	Grassmaniann of k-dimensional vector spaces in $\mathbb{R}^n$
π_L	orthogonal projection from $\mathbb{R}^d$ to the affine plane $L\subset\mathbb{R}^d$
$\langle\cdot,\cdot\rangle$	pairing, see Appendix C. We denote $\langle\mu,\varphi\rangle=\int\varphi\,d\mu$
$\langle v_1,\dots,v_k\rangle$	linear span of vectors $\{v_1,\dots,v_k\}$ in $\mathbb{R}^n$
$\mathrm{aff}(A)$	affine hull of A, the smallest affine plane containing A
$\mathrm{conv}(A)$	convex envelope of A, the smallest convex set containing A
$\dim(A)$	linear dimension of $\mathrm{aff}(A)$
$\mathrm{ri}(C)$	relative interior of C, the interior of C w.r.t. the topology of $\mathrm{aff}(C)$
$\mathrm{rb}(C)$	relative boundary of C, the boundary of C w.r.t. the topology of $\mathrm{aff}(C)$
R-face	see Definition 5.16
$\mathrm{ext}(C)$	extreme points of a convex set C, i.e. zero-dimensional R-faces of C
$\mathrm{dom}\,g$	the domain of a function g
$\mathrm{graph}\,g$	$\{(x,g(x)):\ x\in\mathrm{dom}\,g\}$ (graph)
$\mathrm{epi}\,g$	$\{(x,t):\ x\in\mathrm{dom}\,g,\ t\geqslant g(x)\}$ (epigraph)
∇g	gradient of g
$\partial^- g$	subdifferential of g, see Page 49
$g\lvert_a$	evaluation of g at the point a

$g\vert_b^a$	the difference $g(b) - g(a)$
$g_{\vert A}$	the restriction of g to a subset A of $\mathrm{dom}\, g$
$f\llcorner A$	The restriction of the function f to a set S
f	a fixed convex function $\mathbb{R}^n \to \mathbb{R}$
$\mathrm{dom}\,\nabla f$	a fixed σ-compact set where f is differentiable, see Subsection 5.1
$\mathrm{Im}\,\nabla f$	$\{\nabla f(x) : x \in \mathrm{dom}\,\nabla f\}$, see Subsection 5.1
face of f	intersection of graph $f_{\vert_{\mathrm{dom}\,\nabla f}}$ with a tangent hyperplane
k-face of f	k-dimensional face of f
F_y	$\nabla f^{-1}(y) = \{x \in \mathrm{dom}\,\nabla f : \nabla f(x) = y\}$
F_y^k	F_y when $\dim(F_y) = k$, $\quad k = 0, \dots, n$
E_y, E_y^k	the sets, respectively, $\mathrm{ri}(F_y)$ and $\mathrm{ri}(F_y^k)$
E^k, F^k	the sets, respectively, $\bigcup_y E_y^k$ and $\bigcup_y F_y^k$
$\bar{\mathcal{P}}$, $\mathcal{P}$	outgoing rays, see Definition 4.4 and Formula (5.16)
$\bar{\mathcal{R}}$, $\mathcal{R}$	rays, Definition 4.4 and $\mathcal{R}(x) = F_{\nabla f(x)}$, for every $x \in \mathrm{dom}\,\nabla f$
$\bar{\mathcal{T}}$, $\bar{\mathcal{T}}_e$, $\mathcal{T}$	the transport sets, Definition 4.2 and $\mathcal{T} = \{x \in \mathrm{dom}\,\nabla f : \mathcal{R}(x) \neq \{x\}\}$
$\mathcal{D}$	multivalued map of unit faces directions, see Formula (5.17)
$\bar{\mathcal{Z}}$, $\mathscr{Z}^k$	sheaf set, see Definitions 4.10, 5.8
Z, Z^k	section (or basis) of a sheaf set, see Definitions 4.10, 5.8
$[v, w]$	segment that connects v to w, i.e. $\{(1-\lambda)v + \lambda w : \lambda \in [0,1]\}$
$\Pi_{i=1}^k [v_i, w_i]$	k-dimensional rectangle in $\mathbb{R}^n$ with sides parallel to $\{[v_i, w_i]\}_{i=1}^k$, equal to the convex envelope of $\{v_i, w_i\}_{i=1}^k$
$\bar{\mathcal{K}}$	d-cylinder, see Definition 4.12
$\mathcal{C}^k(\mathscr{Z}^k, C^k)$	k-dimensional $\mathcal{D}$-cylinder $\mathcal{C}^k$, see Definition 5.10
$\partial\mathcal{C}^k$, $\hat{n}\llcorner_{\partial\mathcal{C}^k}$	border of $\mathcal{C}^k$ transversal to $\mathcal{D}$ and outer unit normal, see Formula (6.8)
$\sigma^{\cdot}(\cdot)$	See Definition 4.12, and also Page 41 for $\sigma^{\cdot}_{d_I}(\cdot)$
σ^{w+te}	parameterization of a $\mathcal{D}$-cylinder $\mathcal{C}^k(\mathscr{Z}^k, C^k)$, see Formula (5.34)
σ^{te}	$\sigma^{te} = \sigma^{0+te}$, where $e \in \mathbf{S}^{n-1}$, $t \in \mathbb{R}$
σ^t	if we write $t = te$ with e a unit direction, then $\sigma^t = \sigma^{0+te}$
α, $\alpha(t,s,x)$	see respectively Lemma 4.22 and Formula (5.73)
$\tilde{\alpha}$	see Corollary 4.24, Lemmata 4.22, 4.25
$c(t,z)$	see Theorem 4.26, Lemma 6.3, (6.3)
$\mathcal{S}$	see Theorems 4.26, 13.2
$\mathrm{div}\, v$	the distributional divergence of $v \in L^1_{\mathrm{loc}}(\mathbb{R}^n; \mathbb{R}^n)$
$(\mathrm{div}\, v)_{\mathrm{a.c.}}$	see Lemmata 6.3-6.7 and Notation 6.2.1, Formula (6.12)
v_i	see Definition 6.8
$(\mathrm{div}\, v_i)_{\mathrm{a.c.}}$	see Formula (6.9)

List of Figures

[AA] G. Alberti and L. Ambrosio. A geometrical approach to monotone functions in $\mathbf{R}^n$. *Math. Z.*, 230(2):259–316, 1999. (Cited on pages 50 and 67.)

[AAC] G. Alberti, L. Ambrosio, and P. Cannarsa. On the singularities of convex functions. *Manuscripta Math.*, 76(3-4):421–435, 1992. (Cited on page 50.)

[AFP] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000. (Cited on pages 2, 25, and 35.)

[AKP1] G. Alberti, B. Kirchheim, and D. Preiss. Personal communication in [AKP2]. (Cited on page 5.)

[AKP2] L. Ambrosio, B. Kirchheim, and A. Pratelli. Existence of optimal transport maps for crystalline norms. *Duke Math. J.*, 125(2):207–241, 2004. (Cited on pages 7, 16, 31, 34, 51, and 194.)

[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. (Cited on pages 7, 14, 15, 16, 105, 139, 156, 160, 162, 164, and 173.)

[BC1] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. Preprint SISSA 46/2009/M. (Cited on pages iii and 2.)

[BC2] S. Bianchini and F. Cavalleti. Monge problem in metric spaces. Preprint SISSA. (Cited on pages 16, 140, and 141.)

[BG] S. Bianchini and M. Gloyer. On the Euler Lagrange equation for a variational problem: the general case II. To appear on Math. Z. (Cited on pages 5, 9, 52, 62, 70, 86, and 160.)

[BGMS] M. Beiglböck, M. Goldstern, G. Maresch, and W. Schachermayer. Optimal and better transport plans. *J. Funct. Anal.*, 256(6):1907–1927, 2009. (Cited on pages 13, 14, 105, 121, 123, 173, and 174.)

[Bia] S. Bianchini. On the Euler-Lagrange equation for a variational problem. *Discrete Contin. Dyn. Syst.*, 17(3):449–480, 2007. (Cited on pages 5, 7, and 9.)

[BS] M. Beiglböck and W. Schachermayer. Duality for borel measurable cost function. Preprint. (Cited on page 122.)

[Car1] L. Caravenna. Construction of a Borel c-envelope relation. Unpublished. (Cited on page 174.)

[Car2] L. Caravenna. A proof of Sudakov theorem with strictly convex norms. Preprint SISSA 64/2008/M, 2008. (Cited on pages iii, 2, 52, 58, 60, 62, 64, 70, 73, and 74.)

[CD] L. Caravenna and S. Daneri. A disintegration of the Lebesgue measure on the faces of a convex body. Preprint SISSA 24/2009/M. (Cited on pages iii and 2.)

[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. (Cited on page 16.)

[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. (Cited on page 16.)

[CP2] T. Champion and L. De Pascale. The Monge problem in $\mathbb{R}^d$. Preprint, 2009. (Cited on page 16.)

[Dou] R. G. Douglas. On extremal measures and subspace density. *Michigan Math. J.*, 11:243–246, 1964. (Cited on pages 12 and 111.)

[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. (Cited on page 16.)

[ELR] G. Ewald, D. G. Larman, and C. A. Rogers. The directions of the line segments and of the r-dimensional balls on the boundary of a convex body in Euclidean space. *Mathematika*, 17:1–20, 1970. (Cited on page 7.)

[Fed] H. Federer. *Geometric measure theory*. Berlin Springer-Verlag, 1969. (Cited on pages 18 and 187.)

[FM1] M. Feldman and R. McCann. Monge's transport problem on a Riemannian manifold. *Trans. Amer. Math. Soc.*, (354):1667–1697, 2002. (Cited on page 7.)

[FM2] M. Feldman and R. J. McCann. Uniqueness and transport density in Monge's mass transportation problem. *Calc. Var. Partial Differential Equations*, 15(1):81–113, 2002. (Cited on page 160.)

[Fre1] D. H. Fremlin. *Measure theory*, volume 3. Torres Fremlin, Colchester, 2004. Measure algebras. Corrected second printing of the 2002 original. (Cited on pages 2, 21, and 27.)

[Fre2] D. H. Fremlin. *Measure theory*, volume 4. Torres Fremlin, Colchester, 2006. Topological measure spaces. Part I, II. Corrected second printing of the 2003 original. (Cited on pages 2, 21, 23, 24, and 107.)

[HJ] J. Hoffmann-Jørgensen. Existence of Conditional Probabilities. *Math. Scand.*, 28:257–264, 1971. (Cited on pages 2, 25, 103, 107, and 171.)

[HW] K. Hestir and S. C. Williams. Supports of doubly stochastic measures. *Bernoulli*, 1(3):217–243, 1995. (Cited on pages 11, 104, 111, 112, 113, 114, 115, 117, and 173.)

[Kan1] L. V. Kantorovich. On the transfer of masses. *Docl. Akad. Nauk. SSSR*, 37:227–229, 1942. (Cited on page 15.)

[Kan2] L. V. Kantorovich. On a problem of Monge. *Uskpekhi Mat. Nauk.*, 3:225–226, 1948. (Cited on page 15.)

[Kel] H. G. Kellerer. Duality theorems for marginals problems. *Z. Wahrcsh. Verw. Gebiete*, 67(4):399–432, 1984. (Cited on pages 17, 121, 173, 174, and 175.)

[KM] V. Klee and M. Martin. Semicontinuity of the face-function of a convex set. *Comment. Math. Helv.*, 46:1–12, 1971. (Cited on page 7.)

[Lar1] 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. (Cited on pages 5, 7, 32, and 34.)

[Lar2] D. G. Larman. On a conjecture of Klee and Martin for convex bodies. *Proc. London Math. Soc. (3)*, 23:668–682, 1971. (Cited on pages 7 and 76.)

[Lin] J. Lindenstrauss. A remark on doubly stochastic measures. *Amer. Math. Monthly*, 72:379–382, 1965. (Cited on pages 12 and 111.)

[LR] D. G. Larman and C. A. Rogers. Increasing paths on the one-skeleton of a convex body and the directions of line segments on the boundary of a convex body. *Proc. London Math. Soc. (3)*, 23:683–698, 1971. (Cited on page 7.)

[Mon] G. Monge. Mémoire sur la Theorie des Déblais et des Remblais. *Histoire de l'Acad. de Sciences de Paris*, pages 666–704, 1781. (Cited on page 15.)

[Mor] F. Morgan. *Geometric Measure Theory: A Beginner's Guide*. Academic Press Inc.,U.S., 2000. (Cited on pages 18 and 187.)

[Mos] Y. N. Moschovakis. *Descriptive set theory*, volume 100 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1980. (Cited on page 173.)

[MS] D. A. Martin and J. R. Steel. A proof of projective determinacy. *J. Amer. Math. Soc.*, 2(1):71–125, 1989. (Cited on page 173.)

[Pra] A. Pratelli. On the sufficiency of c-cyclically monotonicity for optimality of transport plans. *Math. Z.*, 2007. (Cited on pages 13, 105, and 121.)

[PZ] D. Pavlica and L. Zajíček. On the directions of segments and r-dimensional balls on a convex surface. *J. Convex Anal.*, 14(1):149–167, 2007. (Cited on page 7.)

[Roc] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970. (Cited on page 63.)

[RR] S. T. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. II*. Probability and its Applications (New York). Springer-Verlag, New York, 1998. Applications. (Cited on pages 14 and 105.)

[Sri] S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998. (Cited on pages 2, 26, 124, 140, 150, 152, and 173.)

[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). (Cited on pages 16, 32, and 155.)

[TW] N. S. Trudinger and X. J. Wang. On the Monge mass transfer problem. *Calc. Var. PDE*, (13):19–31, 2001. (Cited on pages 7 and 16.)

[Vil] C. Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009. (Cited on page 16.)

[Zaj] L. Zajíček. On the points of multiplicity of monotone operators. *Comment. Math. Univ. Carolinae*, 19(1):179–189, 1978. (Cited on page 67.)