Université Paul Sabatier Toulouse 3

Manuscrit

présenté pour l'obtention de

l'Habilitation à Diriger des Recherches

par

Jérôme Bertrand

de l'Institut de Mathématiques de Toulouse

Alexandrov, Kantorovitch et quelques autres.

Exemples d'interactions entre transport optimal et géométrie d'Alexandrov.

Soutenue le 30 novembre 2015,
après avis des rapporteurs

Pr. Robert J. McCann	Université de Toronto
Pr. Takashi Shioya	Université de Tohoku
Pr. Marc Troyanov	E.P.F. de Lausanne

devant le jury formé de

Pr. Franck Barthe	Université de Toulouse
D.R. Gérard Besson	Université Joseph Fourier
Pr. Gilles Carron	Université de Nantes
D.R. Gilles Courtois	Université Pierre et Marie Curie
Pr. Robert J. McCann	Université de Toronto
Pr. Ludovic Rifford	Université de Nice
Pr. Marc Troyanov	E.P.F. de Lausanne

Notre vie est un voyage
Dans l'hiver et dans la Nuit,
Nous cherchons notre passage
Dans le Ciel où rien ne luit.
Chanson des Gardes Suisses
1793

Voyager, c'est bien utile, ça fait travailler l'imagination.
Tout le reste n'est que déceptions et fatigues. Notre voyage à nous est entièrement imaginaire. Voilà sa force.

[...]

Et puis d'abord tout le monde peut en faire autant. Il suffit de fermer les yeux.

C'est de l'autre côté de la vie.[a]

[a]L.-F. Céline, préambule de *Voyage au bout de la nuit.*

Introduction

This manuscript describes my recent research in geometry. Most of my work deals with both optimal mass transport and metric geometry.

In metric geometry, I am primarily interested in Alexandrov spaces with curvature bounded either from below or from above. Alexandrov introduced this concept in the first half of the twentieth century mainly to study the -possibly singular- convex surfaces in three-dimensional Euclidean space. This notion of metric spaces with bounded curvature applies to geodesic spaces and is equivalent to bounded sectional curvature when the space is a smooth Riemannian manifold.

Alexandrov also studied many inverse problems consisting in prescribing the distance (metric with conical singularities), the curvature, or the area measure associated with the underlying convex polyhedron [**Ale05**]. The tremendous amount of work completed by the Russian school of geometry in this field led, among other things, to a *metric* characterization of Riemannian manifolds obtained by Berestovski and Nikolaev [**geo93**]; their result includes an estimate on the regularity of the Riemannian metric which is proved to be locally in $W^{2,p}$ for any $p \geq 1$. To do so, the authors assume the space to be locally compact and to satisfy two-sided curvature bounds, and that, locally, a geodesic can always be extended. More generally, a relevant question is to study the regularity of Alexandrov spaces with curvature bounded below. Indeed, according to Gromov's compactness theorem, the set of Alexandrov spaces with a uniform upper bound on the Hausdorff dimension and on the diameter, whose curvature is uniformly bounded from below, forms a compact set when endowed with the Gromov-Hausdorff distance. However, such an Alexandrov space is not a manifold in general (except in two dimensions); nevertheless, from an analytical point of view, it can be considered as the union of a manifold and a singular subset of codimension at least two. Besides, the distance derives from a Riemannian metric defined almost everywhere whose coefficients read in a map are functions of locally bounded variations -denoted by BV_{loc}; this property follows from Perelman's work [**Per94**] and Otsu and Shioya's paper [**OS94**].

In Chapter 2, we study the regularity of finite dimensional Alexandrov spaces with curvature bounded below. Especially, we show that in two dimensions the metric components are not only BV_{loc} but also Sobolev. In my opinion, this result is a bit surprising since even in the case of convex surfaces in $\mathbb{R}^3$ where the distance and the differential structure are induced by Euclidean space, one gets nothing more than BV_{loc} in general by using the extrinsic structure. In addition to this result, we develop tools in order to improve -if possible- the regularity of the metric in higher dimensions. Our work consists in providing a full second order calculus on finite Alexandrov spaces, building on Perelman's earlier work [**Per94**]. This is joint work with Luigi Ambrosio.

Chapter 3 is devoted to Alexandrov's problem on prescribing the curvature measure of a convex body as well as related topics. Various proofs of Alexandrov's theorem are available: by reduction to the case of polyhedra (Alexandrov's original proof), by reduction to the case of smooth convex bodies (proof by Pogorelov) and the study of a Monge-Ampere type equation. The last equation also appears in optimal mass transport where weak solutions can be obtained by studying the mass transport problem for the quadratic cost and probability measures absolutely continuous with respect to the Lebesgue measure. Indeed, we provide a proof of Alexandrov's curvature prescription problem based only on Kantorovitch's dual problem, a standard tool in optimal mass transport. Using the same approach, we also prove a hyperbolic analogue of this result in Minkowski space. The overall idea is to use classical functions in the theory of convex bodies, namely the support and the radial functions, that can be defined one in terms of the other if and only if the underlying set is convex. Those functions are not very regular in general, say Lipschitz, therefore the soft approach provided by optimal mass transport fits well with Alexandrov's problem. However, both proofs are obtained by studying non-standard forms of Kantorovitch's dual problem. In the Euclidean case, the underlying cost function is not real-valued and the standard theory of optimal mass transport does not apply. On the contrary, in the hyperbolic version, the cost function satisfies standard assumptions in the field but the space on which it is defined is singular in general -precisely, the relevant setting is that of hyperbolic orbifolds. To circumvent the issue due to singularities, we use our generalization to Alexandrov spaces of the classical Brenier-McCann theorem about solutions of the optimal mass transport problem. This generalization applies to hyperbolic orbifolds since they are Alexandrov spaces of curvature at least -1. This generalization is explained in Chapter 2 and is based on the differential structure which covers all but a sufficiently small subset of the Alexandrov space.

Another interesting feature of optimal mass transport is that it can be used to define a distance on the set of probability measures over a given metric space that induces the *-weak convergence of probability measures (plus the convergence of the second order moment if the base space is not compact). This distance is usually called Wasserstein distance or quadratic Wasserstein distance if one wants to emphasize that it corresponds to the cost function $c(x, y) = d^2(x, y)$. Of course, there are other distances on the set of probability measures verifying the above property (like Wasserstein distances relative to $c(x, y) = d^p(x, y)$ for $p \geq 1$ but there are others); however the Wasserstein distance is rather sensitive to the geometry of the base space. A well-known instance of this phenomenon is the fact that the convexity of Boltzman's entropy on the Wasserstein space over a Riemannian manifold is governed by the behaviour of the Ricci curvature of the base space. Roughly, the Hessian of Boltzmann's entropy is bounded from below by k if and only if $Ric \geq k$. This is a very active field of research including work of Cordero-McCann-Schumenckenchleger [**CEMS01**], Lott-Villani [**LV09**], Sturm [**Stu06a, Stu06b, vRS05**], and more recently Ambrosio-Gigli-Savare [**AGS14**]. This list is far from being exhaustive. Under suitable hypotheses, the Wasserstein space itself reveals geometrical properties of the base space. For instance, Sturm [**Stu06a**] proved that the Wasserstein space over a nonnegatively curved Alexandrov space is nonnegatively curved as well. In collaboration with Benoît Kloeckner, we consider Wasserstein space over a *nonpositively* curved base space. More precisely we assume that the base space is $CAT(0)$ which, roughly speaking, is the global version of Alexandrov's definition for nonpositively curved space. Under this assumption on the base space, the Wasserstein space is *not* nonpositively curved simply because a geodesic between two given points is not unique in general. However, we prove that it satisfies properties reminiscent of those available on nonpositively curved spaces, especially we prove the existence of a boundary at infinity. We also study the isometry group of the Wasserstein space over a negatively curved space and we prove that, contrary to the case of Euclidean space, any isometry derives from an isometry of the base space. All these results are discussed in Chapter 4.

The first chapter of this memoir contains no new results. It provides a minimal introduction to the tools, both in optimal mass transport and Alexandrov geometry, that are used constantly in the subsequent chapters.

Contents

Papers described in this manuscript

[1] Jérôme Bertrand and Benoît Kloeckner. A geometric study of Wasserstein spaces: Isometric rigidity. *I.M.R.N.*, 2015. doi: 10.1093/imrn/rnv177.
[2] Jérôme Bertrand. Prescription of gauss curvature using Optimal Mass Transport. 2015. Preprint.
[3] Luigi Ambrosio and Jérôme Bertrand. On the regularity of Alexandrov surfaces with curvature bounded below. *Preprint*, 2015.
[4] Luigi Ambrosio and Jérôme Bertrand. DC Calculus. *Preprint*, 2015.
[5] Jérôme Bertrand. Prescription of Gauss curvature on compact hyperbolic orbifolds. *Discrete Contin. Dyn. Syst.*, 34(4):1269–1284, 2014.
[6] Jérôme Bertrand and Benoît Kloeckner. A geometric study of Wasserstein spaces: Hadamard spaces. *J. Topol. Anal.*, 4(4):515–542, 2012.
[7] Jérôme Bertrand. Existence and uniqueness of optimal maps on Alexandrov spaces. *Adv. Math.*, 219(3):838–851, 2008.

Other papers

[8] Jérôme Bertrand, Aldo Pratelli, and Marjolaine Puel. Kantorovitch potentials and continuity of total cost for relativistic cost functions. 2015. Preprint.
[9] Jérôme Bertrand and Marjolaine Puel. The optimal mass transport problem for relativistic costs. *Calc. Var. Partial Differential Equations*, 46(1-2):353–374, 2013.
[10] Jérôme Bertrand and Benoît R. Kloeckner. A geometric study of Wasserstein spaces: an addendum on the boundary. In *Geometric science of information*, volume 8085 of *Lecture Notes in Comput. Sci.*, pages 405–412. Springer, Heidelberg, 2013.
[11] Erwann Aubry, Jérôme Bertrand, and Bruno Colbois. Eigenvalue pinching on convex domains in space forms. *Trans. Amer. Math. Soc.*, 361(1):1–18, 2009.
[12] Jérôme Bertrand. Pincement spectral en courbure de Ricci positive. *Comment. Math. Helv.*, 82(2):323–352, 2007.
[13] Jérôme Bertrand and Bruno Colbois. Capacité et inégalité de Faber-Krahn dans $\mathbb{R}^n$. *J. Funct. Anal.*, 232(1):1–28, 2006.
[14] Jérôme Bertrand. Stabilité de l'inégalité de Faber-Krahn en courbure de Ricci positive. *Ann. Inst. Fourier (Grenoble)*, 55(2):353–372, 2005.

CHAPTER 1

Background

Convention: in what follows, a geodesic is always assumed to be constant speed and parameterized on $[0,1]$.

1. Optimal Mass Transport

In this part, unless otherwise stated, the space (X,d) is assumed to be a complete separable geodesic space.

1.1. The mass transport problem. Let us start with the concept of optimal transport which consists in studying the *Monge-Kantorovich* problem.

There are the standard data for this problem. We are given a lower semicontinuous and nonnegative function

$$c : X \times X \to \mathbb{R}^+ \cup \{+\infty\}$$

called the *cost function* and two Borel probability measures μ_0, μ_1 defined on X. A *transport plan* (or simply a plan) Π between μ_0 and μ_1 is a probability measure on $X \times X$ whose marginals are μ_0 and μ_1. This means that for any Borel set $A \subset X$

$$\mu_0(A) = \Pi(A \times X) \text{ and } \mu_1(A) = \Pi(X \times A).$$

One should think of a transport plan as a specification of how the mass in X, distributed according to μ_0, is moved so as to be distributed according to μ_1. We denote by $\Gamma(\mu_0, \mu_1)$ the set of transport plans which is never empty (it contains $\mu_0 \otimes \mu_1$) and most of the time not reduced to one element. The Monge-Kantorovich problem is now

$$\inf_{\Pi \in \Gamma(\mu_0,\mu_1)} \int_{X \times X} c(x,y)\, \Pi(dxdy)$$

where the above quantity is assumed to be finite so that the problem makes sense. When it exists, a minimizer is called an *optimal transport plan*. The set of optimal transport plans is denoted by $\Gamma_o(\mu_0, \mu_1)$.

Let us make a few comments on this problem. First, note that under these assumptions, the cost function is measurable (see, for instance, [**Vil03**, p. 26]). Secondly, existence of minimizers follows readily from the lower semicontinuity of the cost function together with the following compactness result.

Theorem 1.1 (Prokhorov's Theorem). *Given a complete separable metric space (X,d), a subset $P \subset \mathcal{P}(X)$ of probability measures on X is totally bounded (i.e. has compact closure) for the weak topology if and only if it is tight, namely for any $\varepsilon > 0$, there exists a compact set K_ε such that $\mu(X \setminus K_\varepsilon) \leq \varepsilon$ for any $\mu \in P$.*

This theorem implies that the set $\Gamma(\mu_0, \mu_1)$ is always compact.

We also mention that, compared to the existence problem, proving the *uniqueness* of minimizers is considerably harder (see [**MRar**]) and requires, in general, additional assumptions. Strongly related to the uniqueness question is this one: how to prove that an optimal transport plan is actually induced by a map? By convention, this is expressed in terms of the initial measure and reads $\Pi = (Id, T)_\sharp \mu_0$, with T being a μ_0-measurable map called optimal transport map and $T_\sharp \mu$ defined, for any Borel set B, by the formula $T_\sharp \mu(B) := \mu(T^{-1}(B))$. The existence of optimal transport map is the original problem proposed by Monge and is much more delicate that the Monge-Kantorovitch problem. Roughly, it means that the optimal way to move mass does not split it: all the mass located at a point x is sent to the same location $T(x)$. In Section 1.3, we recall the Brenier-McCann theorem which gives conditions under which the Monge problem can be solved.

To conclude this part, we state a useful criterion to detect an optimal transport plan among other plans, named *cyclical monotonicity.* A set $\Gamma \subset X \times X$ is said to be *c-cyclically monotone* if for any finite family of pairs $(x_1, y_1), \cdots, (x_m, y_m)$ in Γ, the following inequality holds

$$\sum_{i=1}^{m} c(x_{i+1}, y_i) \geq \sum_{i=1}^{m} c(x_i, y_i) \tag{1}$$

where $x_{m+1} = x_1$.

Note that instead of only considering a shift $(x_i, y_i) \to (x_{i+1}, y_i)$, we could have defined a c-cyclically monotone set by requiring the above inequality to be true for any permutation of $\{1, \cdots, m\}$. The two definitions are equivalent. In particular, when both measures μ_0 and μ_1 are *finitely supported*, a plan is optimal if and only if its support is c-cyclical monotone. Actually, the relevance of this notion goes far beyond this very specific setting. For instance, by approximating c by "nicer" cost functions, it can be proved that an optimal plan Π_0 for c is always concentrated on a c-cyclically monotone set Γ: $\Pi_0(\Gamma) = 1$. More surprisingly, the optimality of a plan can be detected thanks to c-cyclical monotonicity. Precisely, when the cost function $c : X \times X \to \mathbb{R}^+ \cup \{+\infty\}$ is continuous (the set $\mathbb{R}^+ \cup \{+\infty\}$ being endowed with the order topology), or if it is only lower semicontinuous but real-valued, then a transport plan concentrated on a c-cyclically monotone set is optimal. The result for continuous cost is due to Pratelli [**Pra08**] while the other one is due to Schachermayer and Teichmann [**ST09**].

1.2. Wasserstein space. Wasserstein spaces arise in the particular case where $c(x, y) = d^2(x, y)$.

Definition 1.2 (Wasserstein space). *Given a metric space (X, d), its (quadratic) Wasserstein space $\mathscr{W}_2(X)$ is the set of Borel probability measures μ on X with finite second moment:*

$$\int_Y d(x_0, x)^2 \, \mu(dx) < +\infty \qquad \text{for some, hence all } x_0 \in Y.$$

The set $\mathscr{W}_2(X)$ is endowed with the Wasserstein distance defined by

$$\mathrm{W}^2(\mu_0, \mu_1) = \min_{\Pi \in \Gamma(\mu_0, \mu_1)} \int_{X \times X} d^2(x, y) \, \Pi(dx, dy).$$

From now on, the cost c will therefore be $c = d^2$. It is sometimes more convenient to consider $1/2 \, d^2$ instead of d^2; this modification has no impact on the properties described below.

The fact that W is indeed a metric follows from the so-called “gluing lemma” which enables one to prove the triangle inequality, see e.g. [**Vil09**].

The Wasserstein space inherits several nice properties of the base space: first it is complete and separable, it is compact as soon as X is, in which case the Wasserstein metric metrizes the weak topology; but if X is not compact, then $\mathscr{W}_2(X)$ is not even locally compact and the Wasserstein metric induces a topology stronger than the weak one (more precisely, convergence in Wasserstein distance is equivalent to weak convergence plus convergence of the second moment). Another important property is that $\mathscr{W}_2(X)$ is a geodesic space: given any $\mu_0, \mu_1 \in \mathscr{W}_2(X)$, there exists a geodesic curve $(\mu_t)_{t \in [0,1]}$ in $\mathscr{W}_2(X)$ connecting μ_0 to μ_1. Moreover, if $\mathcal{G}(X)$ denotes the set of constant speed geodesics in X parameterized on $[0, 1]$, there exists a measure $\mu \in \mathcal{P}(\mathcal{G}(X))$ called *dynamical transport plan* such that for any $t \in [0, 1]$,

$$\mu_t = e_{t\sharp}\mu$$

with $e_t(\gamma) := \gamma(t)$.

Note that in general μ is not unique. Even in the case where there is a unique optimal plan between μ_0 and μ_1, non uniqueness can occur if (X,d) is branching (consider for instance a crosslike graph, take μ_0 be the empirical measure supported on the two vertices on the left and μ_1 its counterpart supported on the two right-vertices, you get several dynamical plans depending on where you send the mass once it arrives at the center of the cross).

1.3. Kantorovitch's dual problem. The variational problem defined below was introduced by Kantorovitch in order to study the properties of the optimal transport plans. This problem is of primary importance for us in our study of the Gauss curvature prescription problem introduced by Alexandrov. Kantorovitch's dual problem is the problem defined by

$$\mathcal{K} := \sup_{(\phi,\psi)\in\mathcal{A}} \left\{ \int_X \phi(x)d\mu_0(x) + \int_X \psi(y)d\mu_1(y) \right\}. \tag{2}$$

where $\mathcal{A}$ is defined either as

$$\mathcal{A}_{L^1} := \{(\phi,\psi) \in L^1(\mu_0) \times L^1(\mu_1); \forall (x,y) \in N_0 \times X \cup X \times N_1,\ \phi(x) + \psi(y) \leq c(x,y)\}$$

where $\mu_0(N_0) = \mu_1(N_1) = 0$, or

$$\mathcal{A}_C := \{(\phi,\psi) \in C^b(X) \times C^b(X); \forall (x,y) \in X \times X,\ \phi(x) + \psi(y) \leq c(x,y)\}.$$

The objects μ_0, μ_1, and c are defined as in the first paragraph. From now on, the above functional is denoted by $J(\phi,\psi)$. Note that $\mathcal{K}$ a priori depends on the choice of $\mathcal{A}$; precisely, with obvious notation, $\mathcal{K}_C \leq \mathcal{K}_{L^1}$ but it is not clear whether the reverse inequality holds true. As a consequence of the approximation process described below, we get $\mathcal{K}_C = \mathcal{K}_{L^1}$.

Our goal is to relate $\mathcal{K}$ to the Monge-Kantorovitch problem and discuss the existence of maximisers of Kantorovitch's dual problem. Fix $\Pi \in \Gamma(\mu_0,\mu_1)$, and $(\phi,\psi) \in \mathcal{A}_C$ (the argument also works for $\mathcal{A}_{L^1}$). Now, observe that by definition of a transport plan

$$\int_X \phi(x)d\mu_0(x) + \int_X \psi(y)d\mu_1(y) = \int_{X\times X} (\phi(x) + \psi(y))\, \Pi(dxdy)$$

which combined with the definition of $\mathcal{A}_C$ yields

$$\int_X \phi(x)d\mu_0(x) + \int_X \psi(y)d\mu_1(y) \leq \int_{X\times X} c(x,y)\, \Pi(dxdy).$$

Being (ϕ,ψ) and Π arbitrary, we infer

$$\mathcal{K}_C \leq \min_{\Pi\in\Gamma(\mu_0,\mu_1)} \int_{X\times X} c(x,y)\, \Pi(dxdy).$$

With much more work, it can be proved that both quantities actually coincide. Let us outline the argument. The first step consists in proving the equality under the extra assumptions: (X,d) is compact and c is a Lipschitz, bounded, real-valued cost function. In that case, a proof can be obtained by duality in the following sense. First, observe that Riesz's theorem implies that the space of Radon measures (with finite total variation) is the dual space of the space $C(X)$ of continuous functions. Second, note that $J(\phi,\psi)$ is invariant by the transformation $(\phi,\psi) \mapsto (\phi+\lambda, \psi-\lambda)$ with λ being any real number. As a consequence, one can extend the functional J to an upper semicontinous convex functional $\widetilde{J}$ on $C(X \times X)$ defined by $\widetilde{J}(f) = J(\phi,\psi)$ if f can be written $f(x,y) = \phi(x) + \psi(y)$ with $(\phi,\psi) \in \mathcal{A}_C$ and $-\infty$ otherwise. The invariance by translation guarantees that $\widetilde{J}$ is well-defined. Finally, the optimisation problem $\mathcal{K}$ can be rewritten as

$$sup_{f\in C(X\times X)} \left\{ \widetilde{J}(f) - \delta(f; \{f \leq c\}) \right\}$$

where δ stands for the indicatrix function of the set $\{f \leq c\}$, namely

$$\delta(f;\{f \leq c\}) = 0 \text{ if } f \leq c \text{ and } \delta(f;\{f \leq c\}) = +\infty \text{ otherwise.}$$

The result can then be proved as an application of a min-max theorem for the sum of two concave functionals and their Legendre transforms. The fact that we extend J to $C(X \times X)$ yields that the Legendre transform $\mathcal{L}(\widetilde{J})$ of $\widetilde{J}$ is defined on $\mathcal{P}(X \times X)$ as in the Monge-Kantorovitch problem. Moreover, the indicatrix functional implies that $\mathcal{L}(\widetilde{J})$ is finite only if $\Pi \in \Gamma(\mu_0, \mu_1)$. We refer to **[Vil03]** for an exhaustive argument. Note that as a by-product, we get *continuous maximisers* of Kantorovitch's problem. The general case is obtained, first by using a delicate truncation process on the space (X, d) which allows one to remove the compactness assumption on (X, d), and then by approximating the cost function c by a sequence of Lipschitz, bounded, real-valued cost functions c_k. (Note that the lower semicontinuity of c is needed to get the result.) As a by-product, we get that $\mathcal{K}$ remains the same if we replace $\mathcal{A}_C$ by $\mathcal{A}_{L^1}$. We refer to **[Vil03]** for more details.

We emphasize that contrary to the Monge-Kantorovitch problem, the dual problem does not admit maximisers in general. A rather broad setting for which maximisers do exist is described in [**Vil09**, Chapter 5]; note however that non real-valued cost functions do not fit the assumptions. Counter-examples to the existence of maximisers can be found in [**BS11**, Section 4]. In Chapter 3, Section 2.2, we use solution of a Kantorovitch's dual problem relative to a non real-valued cost function to prove a geometrical result. A significant part of the proof is devoted to the existence of such maximisers.

To conclude this part, we add a useful property on the maximisers of J assuming their existence. In this regard, we introduce the c-transform of a function $\phi : X \to \mathbb{R} \cup \{-\infty\}$ asuming that $\phi \not\equiv -\infty$. The definition is the following:

$$\phi^c(a) = \inf_{b \in X} c(a, b) - \phi(b).$$

In the good cases, ϕ^c is continuous (or even Lipschitz) and real-valued. This is for instance true when (X, d) is compact and c is Lipschitz. Note that in general, it is a difficult problem to prove that ϕ^c is merely measurable. Note also that it is customary in the field to write ϕ^{cc} instead of $(\phi^c)^c$. Now, discarding the measurability/integrability issue concerning the c-transform, observe that by definition of $\mathcal{A}_{L^1}$, we have the following inequalities:

$$J(\phi, \psi) \leq J(\phi, \phi^c) \leq J(\phi^{cc}, \phi^c).$$

In particular, if (ϕ, ψ) is a solution of Kantorovitch's dual problem, this gives $\phi = \phi^{cc}$ μ_0-a.e. and $\psi = \phi^c$ μ_1-a.e. Besides, very little is required on ϕ to prove that $\phi^{ccc} = \phi^c$, at least μ_0-a.e. Thus, at least formally, any solution of Kantorovitch's problem coincides almost everywhere with a pair (φ, φ^c) such that $\varphi = \varphi^{cc}$ (where $\varphi = \phi^{cc}$). Such a function φ is said to be c-conjugate. As explained in subsequent chapters (Sections 1.2 and 2.2), when the c-conjugate functions happen to be Lipschitz, they can be used to prove that the Monge-Kantorovitch problem has a unique solution, furthermore this solution is induced by a map T which can be expressed in terms of $\nabla\varphi$. The classical Brenier-McCann theorem provides a particular instance of this phenomenon [**Bre91, McC01**]. On a Riemannian manifold, the statement is as follows (note that the assumptions on the probability measures are not sharp).

THEOREM 1.3. *Let (M, g) be a complete Riemannian manifold. We set $c(x, y) = 1/2\, d^2(x, y)$ the quadratic cost and μ_0, μ_1 two probability measures on M with compact support. We further assume μ_0 to be absolutely continuous with respect to the Riemannian measure and* $\operatorname{supp} \mu_0$ *is connected. Then, Kantorovitch's dual problem admits a solution (ϕ, ϕ^c) with ϕ a c-conjugate function. As a consequence, the mass transport problem admits a unique solution and this solution is induced by a map F. Furthermore, for μ_0 almost every $x \in X$, the map F satisfies*

$$F(x) = exp(-\nabla\phi(x)),$$

The above result is generalised to the setting of Alexandrov spaces in Section 1.2 and a proof is sketched there.

2. Alexandrov spaces

Classical books for the material treated in this part are [**BBI01, BH99**].

2.1. Definition and properties.

DEFINITION 2.1 (Alexandrov space). *Let* S_k^2 *be the 2-dimensional space form of curvature* k *and* δ_k *be the distance induced by the Riemannian metric. A complete geodesic space* (X,d) *is said to be an* Alexandrov space of curvature bounded below by k *if any point is contained in an open set* U *such that whenever a point* p *and a constant speed geodesic* γ *lie within* U*, the following inequality holds. Let* $\overline{\gamma}$ *be a geodesic in* S_k^2 *of same length as that of* γ*, and such that* $d(p,\gamma(0)) = \delta_k(\overline{p},\overline{\gamma}(0))$ *and* $d(p,\gamma(1)) = \delta_k(\overline{p},\overline{\gamma}(1))$*, then for all* $t \in [0,1]$*:*

$$d(p,\gamma(t)) \geq \delta_k(\overline{p},\overline{\gamma}(t)).$$

When $k > 0$*, we further assume that the perimeter of* $p\gamma(0)\gamma(1)$ *is less than* $2\pi/\sqrt{k}$[1]*. Similarly, a complete geodesic space* (X,d) *is said to be an* Alexandrov space of curvature bounded above by k *if any point admits a neighborhood* U *with the same properties as above except that the conclusion is now*

$$d(p,\gamma(t)) \leq \delta_k(\overline{p},\overline{\gamma}(t)).$$

When (X,d) is isometric to a Riemannian manifold (M,g), the above definitions are equivalent to $K_g \geq k$ and $K_g \leq k$ respectively, with K_g being the sectional curvature of g. This justifies the terminology of curvature bounded above or below; in what follows, we will simply write that (X,d) satisfies $Curv \geq k$ or $Curv \leq k$.

To any triple (x,y,z) in a neighborhood U as above can be associated a unique (up to isometry) triples $(\widetilde{x},\widetilde{y},\widetilde{z}) \in S_k^2$ which forms a triangle whose sidelengths are the same as the ones induced by (x,y,z). We set $\widetilde{\measuredangle} yxz$ the angle at $\widetilde{x}$ of this triangle. Now, given two geodesics γ,σ starting from the same point p and contained in an open set U as in the above definitions of bounded curvature, we obtain that $\widetilde{\measuredangle}\gamma(s)p\sigma(t)$ is a nonincreasing function of s and t whenever the curvature of (X,d) is bounded below by k while it is a nondecreasing function of s and t when the curvature is bounded above from k. As a consequence, in both setting, the angle between σ and γ at p can be defined as

$$\measuredangle(\gamma,\sigma) := \lim_{s,t\to 0} \widetilde{\measuredangle}\gamma(s)p\sigma(t).$$

An important property of a space with curvature bounded below is that it is a *strongly non-branching space*: if two geodesics (with the same speed) coincide on an open subinterval of $[0,1]$, then they actually coincide everywhere; moreover the angle between two disctinct geodesics starting at the same point is positive. Spaces with curvature bounded from above do not share this property, for instance a locally finite metric graph is a space of curvature bounded from above by k for any number k.

Another important distinction between these two classes of spaces is that a space with curvature bounded below satisfies a local-to-global property; namely, in the definition, the restriction to a suitable neighborhood can be dropped. For spaces with curvature bounded above this is not true in general as it can be easily seen from the example of a Euclidean plane with an open disc removed. More precisely, equipped with the distance induced by the scalar product, such a space is $Curv \leq 0$ but does not satisfy the assumption globally (arbitrary complementary halfcircles are geodesics between their common endpoints whereas on globally nonpositively curved space, there is a unique geodesic between two given points; see below). Consequently, in what follows we will say that a space is $CAT(k)$ if it satisfies the $Curv \leq k$ condition globally (we refer to Chapter 4 for a precise definition which includes additional metric properties). When $k \leq 0$, there is a very nice characterization of $CAT(k)$-spaces as the *simply connected* $Curv \leq k$ spaces. Let us add two very useful properties of $CAT(0)$ space. First, given two geodesics σ and γ, the function $t \mapsto d(\gamma(t),\sigma(t))$ is convex. This property entails uniqueness of geodesic beetween two given points. Second, given $\mathcal{C}$ a closed convex subset of a $CAT(0)$ space, then the metric projection

$$\begin{array}{rccl} p_{\mathcal{C}}: & X\setminus\mathcal{C} & \longrightarrow & \mathcal{C} \\ & x & \longmapsto & \operatorname{argmin}_{c\in\mathcal{C}} d(x,c) \end{array}$$

[1]Actually there is no triangle whose perimeter is greater than $2\pi/\sqrt{k}$ in a space of curvature at least k.

is well-defined (i.e. realized by a unique point) and is a 1-Lipschitz function.

A relevant question is whether the metric notions $Curv \geq k$ and $Curv \leq k$ are stable with respect to Gromov-Hausdorff Convergence. This is indeed the case for the $Curv \geq k$ spaces and yields, in particular, that any limit of Riemannian manifolds with sectional curvature uniformly bounded below by k is a $Curv \geq k$ space. On the contrary, $Curv \leq k$ is not stable with respect to Gromov-Hausdorff convergence [2]. However, its global counterpart $CAT(k)$ is indeed stable. Actually, $CAT(k)$ is stable with respect to various limit processes; basically what is needed is the stability of a four point condition. We refer to [**BH99**, Part II.3] for more on this subject.

Let us, to conclude this part, remind the reader of the notion of tangent cone. To do so, we introduce an equivalence relation on the set of geodesics starting from a given point p. Two such geodesics are said to be equivalent if they coincide on a subinterval $[0, \varepsilon)$ of $[0, 1]$ (where $\varepsilon > 0$ depends on the geodesics). We set Σ_p the *set of directions* at p, namely the completion of the set of equivalence classes with respect to the metric angle $\measuredangle$. The *tangent cone* $C_p(X)$ is then defined as the Euclidean cone over $[0, +\infty) \times \Sigma_p$ equipped with the distance

$$\delta((t, \sigma), (s, \gamma))^2 = s^2 + t^2 - 2st \cos \measuredangle_p(\sigma, \gamma). \tag{3}$$

At that stage, we need additional assumptions on the space. If (X, d) is a $Curv \geq k$ space, we further assume its Hausdorff dimension is finite, say N. This forces the space to be locally compact. In the same vein, it can be proved that Σ_p is a compact metric space (for any p). Finally, the local compactness allows one to prove that any sequence $(X, \lambda_n\, d, p)$ converges to $C_p(X)$ with respect to the pointed Gromov-Hausdorff convergence whenever $\lambda_n \to +\infty$. Since the lower bound on the curvature behaves well under dilation of the distance, the tangent cone $C_p(X)$ is a N-dimensional nonnegatively curved space for any $p \in X$.

Similarly, for spaces with an upper curvature bound, extra assumptions are needed to get analogous results on the tangent cone. These assumptions are of two kinds: local compactness and geodesic completeness. We recall that geodesic completeness means that any geodesic is the restriction of a *complete geodesic* -or *geodesic line*-, namely a geodesic defined on $\mathbb{R}$. Under these assumptions, the space of directions Σ_p is compact at any point p and any limit of rescaled spaces as above converges to the tangent cone with respect to the pointed Gromov-Hausdorff convergence. The tangent cone $C_p(X)$ is a nonpositively curved space.

2.2. Examples.

2.2.1. *Spaces with curvature bounded below.* The first source of examples is obtained by considering Gromov-Hausdorff limits of smooth Riemannian manifolds with a uniform lower bound on the sectional curvature. This includes as a very particular case, all convex hypersurfaces in Euclidean space.

The product of two spaces with curvature bounded below by k is a space of curvature bounded below by $\min\{0, k\}$.

One also gets a space with nonnegative curvature by considering the Euclidean cone over a metric space (X, d) with curvature at least one. The metric δ on the cone is given by (3). Cone of curvature at least k can be obtained similarly. It suffices to replace the expression of δ in accordance with the expression of the standard distance on S_2^k in polar coordinates. Note that when $k > 0$, the "cones" are actually compact sets, when $k = 1$ they are also called sin-suspension.

Finally, the quotient of an Alexandrov space (X, d) of curvature at least k by a subgroup of the isometry group of X is an Alexandrov space of curvature at least k provided that the orbits are closed. In particular, one gets singular examples of Alexandrov spaces by considering quotient of spaces forms by appropriate isometry subgroups. See Chapter 3 for an application of this result.

[2]Indeed, one can approximate any compact riemannian manifold by a finite ϵ-net where $\epsilon > 0$ is arbitrarily small. Turn this net into a metric graph by using the geodesics of the manifold. Since this graph has finitely many vertices, it is locally a tree thus it has nonpositive curvature. On the other hand, the sequence of graphs converges to the manifold when ϵ goes to 0

2.2.2. *Spaces with curvature bounded above.* As for spaces with curvature bounded below, any GH-limit of $CAT(k)$-spaces is a $CAT(k)$-space.

Spaces with curvature bounded from above are much more flexible than their counterparts with curvature bounded below. For instance, the space obtained by gluing two spaces of curvature bounded above by k along isometric convex subsets is a space of curvature bounded above by k.

Another important class of examples is that of metric simplicial complexes. It consists in gluing along isometric faces simplices contained in a space form of curvature k. When they are only finitely many models of faces (up to isometry), the distance on each simplex (induced by the standard distance of the space form) extends to a geodesic distance on the resulting space. Such a restriction is somehow necessary to guarantee that the pseudo-metric induced by the metrics on the simplices is positive on pair of distinct points. For instance, a two-points metric graph whoses vertices A, B are joined by infinitely many edges σ_n where the length of σ_n is $1/n$ is not a metric space: $d(A, B) = 0$. A metric simplicial complex is $Curv \leq k$ if and only if the link at any vertex is $CAT(1)$. The link at a vertex x is the metric simplicial complex induced by the collection of spaces of directions relative to a simplex to which x belongs.

When $k \leq 0$, the above examples give rise to $CAT(k)$ spaces provided that they are simply-connected.

Another way to produce new $CAT(0)$-spaces is to consider the warped product of two $CAT(0)$ spaces. The resulting space is a $CAT(0)$-space whenever the warping function is concave.

2.3. Analytical properties of spaces with curvature bounded below. In this part, we present analytical results for N-dimensional Alexandrov spaces (X, d) with curvature bounded below. Let us mention that (X, d) is assumed to be boundaryless (see [**BGP92**, 7.19] for the precise definition). These results consist in estimating the Hausdorff dimension or the Hausdorff measure of some special subsets of X. Some of these results can be found in [**BBI01**], the others are proved either in a paper by Burago, Gromov, and Perelman [**BGP92**] or in a paper by Otsu and Shioya [**OS94**].

Let us start with the notion of δ-regular point. Given $\delta > 0$, a point x is said to be δ-regular if there exists N pairs of points $(p_1, q_1), \cdots, (p_N, q_N)$ such that

$$\begin{cases} \widetilde{\measuredangle} p_i x p_j > \frac{\pi}{2} - \delta & \text{for all } i \neq j, \\ \widetilde{\measuredangle} p_i x q_i > \pi - \delta & \text{for all } i \end{cases}$$

The collection of pairs $(p_1, q_1) \cdots (p_N, q_N)$ is called a δ-strainer (at x). By continuity of the comparison angle with respect to x, the set of δ-regular points is open.

The main interest of this notion is twofold. First, it can be proved that for sufficiently small δ (depending on the dimension), the strainer of a δ-regular point x can be used to build a biLipschitz homeomorphism between a neighborhood of x and an open set of $\mathbb{R}^N$. The second important property is that most of the points of X are δ-regular for a small but fixed δ.

More precisely it is proved in [**BGP92**] that for $0 < \delta \leq \delta_N$, if $(p_1, q_1) \cdots (p_N, q_N)$ is a δ-strainer at x, the map $(d_{p_1}, \cdots, d_{p_N})$ is a biLipschitz homeomorphism when restricted to an appropriate neighborhood of x. Besides, the Lipschitz constants can be chosen arbitrary close to 1 provided that δ and the neighborhood are sufficiently small. Moreover, the set X_δ of δ-regular points is dense in X. More precisely, the Hausdorff codimension of X_δ satisfies

$$\dim_{\mathcal{H}}(X \setminus X_\delta) \leq N - 2. \tag{4}$$

We call X_{δ_N} the set of *quasiregular points* of X. From now on, this set is denoted by X^*. According to what precedes, X^* is a N- dimensional Lipchitz manifold whose complement satisfies (4).

Another important susbset is that of *regular points*. It is denoted by $Reg(X)$ and defined as the set of points whose tangent cone is isometric to N-dimensional Euclidean space. In the same way, it is the set of points which are δ-regular for any $\delta > 0$. Its complement set is denoted by

$Sing(X)$; as a consequence of (4), we get

$$\dim_{\mathcal{H}}(Sing(X)) \leq N-2.$$

In particular, it will be useful to us that $\mathcal{H}^{N-1}(Sing(X)) = 0$.

The set of regular points is a *strongly convex* subset as proved by Petrunin [**Pet98**]. Namely, any geodesic which has a regular point as an endpoint is only made of regular points except maybe the other endpoint.

In relation to the study of distance function d_p, two other subsets are particulary relevant. The first one, denoted by V_p, is the subset of $X \setminus \{p\}$ made of points connected to p by a *unique* geodesic. In [**OS94**], the authors prove [3] that the complement of V_p can be covered by countably many sets with finite $\mathcal{H}^{N-1}$-measure. From now on, such a set will be called a σ-finite set with respect to $\mathcal{H}^{N-1}$. The second one is the Cut locus of d_p. It is denoted by C_p and defined as the complement of the set of points which belong to the interior of a geodesic starting at p. The Cut Locus of any point p satisfies

$$\mathcal{H}^N(C_p) = 0.$$

[3]despite a weaker statement, this is the content of Lemma 2.2. The result follows by combining Lemma 2.2 with Proposition 3.3

CHAPTER 2

Regularity of Alexandrov space with curvature bounded below

In this part, we describe the differential structures available on a N-dimensional Alexandrov space with curvature bounded below. In particular, we explain our improvements regarding the second order differential structure introduced by Perelman [**Per94**]. We also discuss the analogue of the classical Brenier-McCann theorem on Alexandrov space. Our results in this part are extracted from [**7, 4, 3**].

1. First order structure and optimal mass transport

1.1. First order differential structure. In this part, we review the Lipschitz structure available on (almost all of) a finite dimensional Alexandrov space. These results come from papers by Burago, Gromov, Perelman, Petrunin, Otsu & Shioya [**BGP92, OS94, Pet98**]. We then use this structure to generalize the Brenier-McCann theorem to the setting of Alexandrov spaces.

An important result in this field is the characterization of points where the distance function d_p is differentiable in an "intrinsic sense" (i.e. without refering to a chart). This property comes from a strong form of the first variation formula due to Otsu and Shioya [**OS94**] which we now recall. For $x \in V_p \cap \mathrm{Reg}(X)$, the following formula holds

$$d_p(y) = d_p(x) - d(x,y) \cos \min_{\uparrow_x^y} \measuredangle(\uparrow_x^p, \uparrow_x^y) + o(d(x,y)) \tag{5}$$

Note also that both assumptions $x \in V_p \cap \mathrm{Reg}(X)$ are somehow necessary; the assumption $x \in \mathrm{Reg}(X)$ guarantees the existence of a linear structure at x while it is easy to check that d_p cannot be differentiable at a point x related to p by several geodesics (recall that geodesics in Alexandrov space are strongly non-branching, namely the angle at x between two distinct geodesics is positive). The drawback of the above formula is that the set of differentiability points of d_p depends on p. Otsu and Shioya get rid of this by averaging the base point of d_p on a small ball. More precisely, they introduce the function

$$\hat{d}_p(x) := \fint_{B(p,\varepsilon)} d_z(x)\, d\mathcal{H}^N(z) = \frac{1}{\mathcal{H}^N(B(p,\varepsilon))} \int_{B(p,\varepsilon)} d_z(x)\, d\mathcal{H}^N(z);$$

where $\varepsilon > 0$ is a small number depending on p. The point is to observe that $x \in V_z$ if and only if $z \in V_x$. Moreover, as recalled in the previous chapter, they also prove that the complement of V_p is $\mathcal{H}^N$-negligible. Consequently, formula (5) applies with $p = z$ for $\mathcal{H}^N$-a.e $z \in B(p,\varepsilon)$ and $\hat{d}_p$ is differentiable on $Reg(X)$ [1].

Let us now explain how these functions can be used to define a *Lipschitz structure* on X^*. We say that a set S admits a Lipschitz structure if there exists an atlas whose chart domains cover S and whose transition maps are biLipschitz homeomorphisms, in other terms S is a Lipschitz manifold.

By definition, around any fixed point $x \in X^*$, there exists N distance functions $d_{p_1}, \cdots, d_{p_N}$ such that the mapping $(d_{p_1}, \cdots, d_{p_N})$ is a biLipschitz homeomorphism when restricted to a small open neighborhood of x. Otsu and Shioya prove this result remains true if the distance functions are

[1] Note that according to (5), $z \mapsto o(d(x,y))/d(x,y)$ is integrable with respect to $\mathcal{H}^N \llcorner B(p,\varepsilon)$. Therefore $\fint_{B(p,\varepsilon)} o(d(x,y))\, d\mathcal{H}^N(z) = o(d(x,y))$

replaced by their average counterparts, provided that ε is chosen sufficiently small (meaning small compared to the number $\delta = \delta(N)$ in the definition of X^*) and $x \in \mathrm{Reg}(X)$. As a consequence, there exists an open set $\mathcal{M}$ where $X^* \supset \mathcal{M} \supset \mathrm{Reg}(X)$, which admits a Lipschitz structure. To summarize, we have the following result

THEOREM 1.1 (biLipschitz structure). *On a finite dimensional Alexandrov space (X, d), there exists an open set $\mathcal{M}$ such that $X^* \supset \mathcal{M} \supset \mathrm{Reg}(X)$ and $\mathcal{M}$ is a Lipschitz manifold. Moreover, the biLipschitz transition maps are differentiable at each point belonging to the image of* $\mathrm{Reg}(X)$ *and the differential depends continuously of the base point when restricted to the image of* $\mathrm{Reg}(X)$.

Therefore, an Alexandrov space can be seen as the union of a Lipschitz manifold and a "singular" part $X \setminus \mathcal{M}$ whose Hausdorff dimension is small:

$$dim_{\mathcal{H}}(X \setminus \mathcal{M}) \leq dim_{\mathcal{H}}(X \setminus \mathrm{Reg}(X)) \leq N - 2.$$

Moreover, compared to a general Lipschitz manifold, locally Lipschitz functions can be defined in two equivalent ways: either in terms of the distance on X or through the charts using the Euclidean metric. Note also that the property for a function to be differentiable at a *regular* point is intrinsic (i.e. does not depend on the choice of the chart). Last, recall that a finite dimensional Alexandrov space is locally compact. By combining all these properties together, we get

THEOREM 1.2 (Rademacher theorem). *Let $f : \Omega \subset X \to \mathbb{R}$ with Ω an open set and X a finite dimensional Alexandrov space. Assume that f is a Lipschitz function, then f is differentiable $\mathcal{H}^N$-almost everywhere.*

Following the construction of a Riemannian metric on a smooth manifold, one ends up with a Riemannian metric with measurable and locally bounded components when the space is equipped with a Lipschitz structure. For Alexandrov space, better results can be proved. They are also due to Otsu and Shioya [**OS94**]. Stronger results relative to a second order differential structure are discussed in the next section.

THEOREM 1.3 (Riemannian structure on $\mathrm{Reg}(X)$). *A N-dimensional Alexandrov space (X, d) admits a locally bounded Riemannian metric g defined everywhere on* $\mathrm{Reg}(X)$. *Morever, the metric varies continuously with respect to the base point $x \in \mathrm{Reg}(X)$. Finally, the metric g is compatible with the Alexandrov distance:*
i) The tangent cone based at a point $x \in Reg(X)$, endowed with the cone distance induced by the angle, is isometric to (T_xX, g_x).
ii) The distance induced by the Riemannian metric coincides with the original distance on X^2.
iii) The volume form induced by g coincides with the N-dimensional Hausdorff measure.

1.2. The Brenier-McCann theorem on Alexandrov space. Using the Lipschitz structure above, one can generalize the Brenier-McCann theorem to the setting of Alexandrov spaces. The statement reads as follows.

THEOREM 1.4. *Let (X, d) be a N-dimensional Alexandrov space. We set $c(x, y) = \frac{1}{2} d^2(x, y)$ the quadratic cost and μ_0, μ_1 two probability measures on X with compact support. We further assume μ_0 to be absolutely continuous with respect to $\mathcal{H}^N$. Then, Kantorovitch's dual problem admits a solution (ϕ, ϕ^c) with ϕ a c-conjugate function. As a consequence, the mass transport problem admits a unique solution and this solution is induced by a map F. Furthermore, for μ_0 almost every $x \in X$, the map F satisfies*

$$F(x) = exp_x(-\nabla\phi(x)),$$

meaning that $F(x) \in V_x$ and that $-\nabla\phi(x)$ is the direction of the geodesic from x to $F(x)$.

REMARK 1.5. *Observe that in the above theorem, the lower bound on the curvature does not appear explicitly in the statement. Consequently, our result also applies to any compact Riemannian manifold and allows us to give another proof of McCann's theorem* [**McC01**].

By means of the Arzela-Ascoli theorem, the compactness of the supports of μ_0 and μ_1 guarantees the existence of continuous solutions (ϕ, ψ) to Kantorovitch's variational problem. Moreover, the double complexification trick (recalled in Chapter 1) implies that ϕ is a c-conjugate function and ψ coincides μ_1-a.e. with ϕ^c. Besides, the compactness of $\operatorname{supp}\mu_1$ gives us that ϕ is Lipschitz.

The next step is to notice that if ϕ is differentiable at x and y is such that

$$\phi(x) = \frac{1}{2} d^2(x, y) - \phi^c(y)$$

then the first variation formula yields that d_y^2 is differentiable at x. As a by-product this gives $y \in V_x$ and

$$y = exp_x(-\nabla\phi(x)).$$

Now, thanks to the Rademacher theorem on Alexandrov space, the c-subdifferential of ϕ:

$$\partial_c\phi = \{(x, y) \in \operatorname{supp}\mu_0 \times \operatorname{supp}\mu_1; \phi(x) + \phi^c(y) = c(x, y)\}$$

coincides with the graph of F out of a set $\mathcal{N} \times \operatorname{supp}\mu_1$ where $\mathcal{H}^N(\mathcal{N}) = 0$.

To conclude, consider $\Pi_0 \in \Gamma_0(\mu_0, \mu_1)$ an optimal plan. By definition,

$$\int_X \phi \, d\mu_0 + \int_X \phi^c \, d\mu_1 = \int_{X\times X} c \, d\Pi_0$$

which can be rewritten (using that ϕ, ϕ^c are Lipschitz and the definition of a transport plan)

$$\int_{X\times X} (c - \phi - \phi^c) \, d\Pi_0 = 0$$

The fact that the above integrand is always nonnegative yields $\Pi_0(\partial_c\phi) = 1$. Finally, the assumption on μ_0 is used to discard the set $\mathcal{N} \times \operatorname{supp}\mu_1$. Therefore Π_0 is the plan $(Id, F)_\sharp\mu_0$ induced by the map F.

In the case of Riemannian manifolds, the assumption on μ_0 can be weakened. For instance, the Brenier-McCann theorem holds whenever μ_0 does not give mass to $(N-1)$-dimensional subsets [**McC95**]. The sharp statement on μ_0 can be described in terms of graphs of *DC* functions of $N-1$ variables. *DC* functions are discussed in details in the next subsection. In particular, we state the optimal statement for μ_0 there and explain why the same result holds on an Alexandrov space.

2. Second order differential structure: *DC* Calculus

In this part, we describe the second order differential structure available on Alexandrov space. These results have been initiated by Perelman in [**Per94**]. In collaboration with Luigi Ambrosio, we are continuing this study. The results described in this part are contained in [**3, 4**]. We start with a brief review of *BV* functions as a central tool in this section and refer to [**AFP00**] for a more detailed account on the subject.

2.1. Review of *BV* functions. Given $f : \Omega \subset \mathbb{R}^N \to \mathbb{R}$ with Ω an open set and $f \in L^1_{loc}(\Omega)$. the function f is said to have locally bounded variation (in what follows this will be denoted by $f \in BV_{loc}(\Omega)$) if its distributional derivative $Df = (\partial_{x_1} f, \cdots, \partial_{x_N} f)$ is a vector-valued Radon measure[2]. Moreover the total variation measure $|Df|$ of Df is supposed to be locally finite, meaning that for any bounded open subset Ω' of Ω, $|Df|(\Omega') < +\infty$. When $|Df|(\Omega) < +\infty$ then f is said to have finite bounded variation -this will be denoted by $f \in BV(\Omega)$. For a vector-valued function, a similar definition is given by arguing componentwise. Let us recall the definition of the total variation measure $|\mu|$ of a $\mathbb{R}^M$-valued Radon measure μ. Given a Borel set E,

$$|\mu|(E) := \sup_{(A_i);\cup A_i = E} \sum_{i\in I} ||\mu(A_i)||$$

[2]Especially, Df remains unchanged if f is modified on a negligible subset.In short, a *BV* function is defined up to a negligible subset. However, due to our geometric context, we will mainly have to deal with *BV* functions that are defined everywhere.

where (A_i) is a finite or countable partition of E into Borel sets and $||\cdot||$ is the Euclidean norm on $\mathbb{R}^M$. According to the *Polar decomposition*, there exists a unique $\mathbb{S}^{M-1}$-valued function $\rho \in L^1(|\mu|)$ such that

$$\mu = \rho\,|\mu|.$$

Now, let us recall some properties relative to the derivative of a BV_{loc} function f. The distributional derivative Df can be written as

$$Df = D^{ac}f + D^{ju}f + D^{ca}f,$$

where $D^{ac}f$ is the absolutely continuous part w.r.t. Lebesgue measure, $D^{ju}f$ is the jump part, and $D^{ca}f$ is the Cantor part. The jump part of the derivative is concentrated on a set σ-finite w.r.t. $\mathcal{H}^{N-1}$ (i.e. a countable union of sets with finite $\mathcal{H}^{N-1}$-measure) while the Cantor part is concentrated on a $\mathscr{L}^N$-negligible set and vanishes on sets with finite $\mathcal{H}^{N-1}$ measure.

2.2. On the regularity of the Riemannian metric. Let us start with the notion of semiconcave function on Alexandrov space. A locally Lipschitz function f defined on an open set $\Omega \subset X$ is said to be *semiconcave* if for any point x in Ω, there is a neighborhood of x and a number λ such that for any geodesic γ in this neighborhood, $f \circ \gamma$ is λ-concave, namely

$$t \mapsto f(\gamma(t)) - \lambda/2\, t^2$$

is a concave function.

Let us review some examples of semiconcave function. First, on any space form, a simple computation of the Hessian shows that a distance function d_p is semiconcave away from p. More generally, note that the definition of Alexandrov space of curvature at least k can be rephrased in terms of semiconcavity: locally, any distance function d_p^2 read along a geodesic is more concave than it would be if the space were of constant curvature k. By combining the two preceding facts, one gets that d_p^2 is semiconcave which in turn implies the semiconcavity of d_p away from p. This latter property is easily generalized to distance function from a closed set. Average of distance functions is also an example of semiconcave function. This is an important fact since (average) distance functions are used to build the aforementioned Lipschitz structure. Now, mimicking the Euclidean definition, one can introduce the set of DC functions as the set of functions which, locally, can be written as the difference of two semiconcave functions. The definition is extended to mapping by arguing componentwise. The following striking result of Perelman allows one to improve the differential structure. Let Φ be a chart defined as above (whose components are either distance functions or average distance functions), then any function F well-defined on the domain of Φ is DC if and only if $F \circ \Phi^{-1}$ is DC in the Euclidean sense. Furthermore, the set of Euclidean DC mappings is known to be stable with respect to composition. An important consequence of this result is that any transition map relative to any of the two differential structures defined above is a Euclidean DC function. As concave functions from which they are built, Euclidean DC functions admit second order derivatives $\mathscr{L}^N$-almost everywhere. Moreover, their second distributional derivatives form a matrix-valued Radon measure. This result opens the way toward a second order calculus on Alexandrov space. From now on, we always use the charts built from average distance functions. As a consequence of his result on DC functions, Perelman proves that the Riemannian metric components are functions of locally bounded variations. His idea is to expand the equality $g(\nabla d_p, \nabla d_p) = 1$ which holds $\mathcal{H}^N$-a.e. using the coordinates. This gives

$$\sum_{i=1}^{N} g^{ij} \frac{\partial f_p}{\partial x_i} \frac{\partial f_p}{\partial x_j} = 1$$

where g^{ij} are the components of the matrix inverse of (g_{ij}), f_p is the distance function d_p read in the chart -thus a Euclidean DC function. Then, consider the above equality as a linear equation with g^{ij} being the variables. Since $x \in \mathrm{Reg}(X)$, one can find (as in the case of $\mathbb{R}^N$) sufficiently many distance functions d_{p_i} so that the combination of the equalities relative to f_{p_i}, forms a linear system whose matrix is invertible. As a consequence, the g^{ij} can be expressed as rational functions of first derivatives of f_{p_i} thus they are locally BV by Perelman's theorem.

Before we further discuss the features of a space covered by an atlas with DC transition maps, let us first ask the following question: is BV the best regularity of the Riemannian metric one can hope in the setting of Alexandrov spaces?

This is certainly not the case if we further assume the existence of a metric upper bound on the curvature and if we ask the space to be locally geodesically complete (i.e. any point admits an open neighborhood such that any geodesic contained in this neighborhood and defined on, say, $[0, T]$ can be extended to a geodesic defined on $[0, T + \varepsilon]$ where ε depends on the geodesic). Under this set of assumptions (including finite Hausdorff dimension), Berestovski and Nikolaev proved that the metric space is isometric to a manifold endowed with a $C^{2,\alpha}$ atlas (for all $\alpha \in [0, 1)$) and a Riemannian metric whose components are in the Sobolev space $W^{2,p}$ for any $p \geq 1$ [**geo93**]. Let us add that as far as the regularity of the space is concerned, the assumption of local geodesic completeness is more important than the upper bound on the curvature. Indeed, building on Otsu-Shioya results [**OS94**], Berestovski later proved that a locally geodesically complete Alexandrov space is actually a manifold endowed with a 1/2-Hölder Riemannian metric [**Ber94**]; unfortunately this paper is available in Russian only. Note however that local geodesic completeness is a very restrictive condition. For instance, a locally geodesically complete Alexandrov surface has no singular point [3]. Indeed, Petrunin proved [**Pet98**] that the interior of a geodesic containing at least one regular point is actually made of regular points. For general Alexandrov surfaces -for which singular points can form a dense subset, we prove the following regularity result.

THEOREM 2.1. *Let (S, d) be a closed surface with curvature bounded below. Then S is a topological surface and the distance d derives from a Riemannian metric g. Furthermore, for all $p \in [1, 2)$ there exists a discrete set $\mathfrak{S}_p \subset S$ such that the components of g read in a local chart belong to $W^{1,p}_{\text{loc}}(S \setminus \mathfrak{S}_p, \mathcal{H}^2)$. In particular, the metric components are in $W^{1,1}_{\text{loc}}$ around any $x \in Reg(S)$.*

This result applies to convex surfaces in Euclidean space. Note that using the differential structure induced by the ambient space, you can only prove that the metric components are functions of locally bounded variation in general (the charts as well as their inverse functions are convex, their first derivatives are then in BV_{loc}). Therefore, even in this simple case the above result leads to something new.

The proof of it is specific to the case of surfaces. It is based on the fact that the curvature viewed as a Radon measure, is well-defined on a surface with curvature bounded below. This is due to Alexandrov [**Ale06**]. Basically, the point is that we know how to define the curvature measure at a (conical) point, on a geodesic, and for a geodesic triangle thanks to the Gauss-Bonnet formula. Then, through a long and technical process, Alexandrov proves that such a surface can be approximated by "nice" triangulated surfaces on which the curvature measure is well-defined thanks to the previous remark. The existence of curvature measure on the initial space is then proved by passing to the limit in the approximation process. Later, Alexandrov and Zalgaller made this process formal and gave birth to the notion of surfaces with bounded integral curvature (i.e. surfaces on which the curvature measure is a Radon measure that can be approximated nicely by considering triangulated surfaces). Finally, Reshetnyak proved that a surface with bounded integral curvature can also be approximated in a nice way by smooth Riemannian surfaces [**geo93**]. As a consequence of his study, he proved the existence of *isothermal coordinates* on these surfaces. Our result applies to this general setting (and the above theorem is actually a corollary of the result below. It is based on the fact that an Alexandrov surface is a particular instance of surface with bounded integral curvature). It reads

THEOREM 2.2. *Let (S, d) be a closed surface of bounded integral curvature ω and let*

$$\Omega = \left\{ z \in S : \ \omega^+(\{z\}) < 2\pi \right\} \tag{6}$$

[3] Actually, this is true in any dimension as a consequence of the splitting theorem.

where ω^+ is the nonnegative part of ω.

Then, for all $z \in \Omega$ there exist a chart (U, ϕ) with $z \in U$ and a Riemannian metric g defined on $V = \phi(U) \subset \mathbb{R}^2$ by the formula

$$g(x_1, x_2) = \lambda(x_1, x_2)(dx_1^2 + dx_2^2)$$

such that the distance induced by the Riemannian metric coincides with the distance d. Setting for $q \geq 4$

$$\Omega_q = \left\{ z \in S : \ \omega^+(\{z\}) < \frac{2\pi}{q} \right\},$$

if $z \in \Omega_q$ we can choose (U, ϕ) in such a way that:

(a) *the metric components g_{ij} and the volume form $\sqrt{\det(g)}$ belong to $L^q(V, dx_1 dx_2)$;*
(b) *the distributional derivatives $\frac{\partial g_{ij}}{\partial x_k}$ belong to $L^p(V, \sqrt{\det(g)}\, dx_1 dx_2)$ where $p = 2 - 6/(q+2)$.*
(c) *the Christoffel symbols Γ_{ij}^k belong to $L^p(V, \sqrt{\det(g)}\, dx_1 dx_2)$ where $p = 2 - 2/q$.*

We emphasize that the existence of the Riemannian metric g above is due to Reshetnyak. Our contribution consists in studying its regularity. Such a metric g is said to be *subharmonic* in the sense that $\lambda(x_1, x_2) = \exp(-2u_+(x_1, x_2) + 2u_-(x_1, x_2))$ where $u_\pm$ are Euclidean subharmonic functions whose distributional Laplacians coincide with the positive and negative part of the curvature measure, denoted by ω^+ and ω^- respectively. The regularity results about the metric then follows from the regularity of these subharmonic functions. The latter can be achieved by studying the *logarithmic potentials* of ω^+ and ω^- as a consequence of Weyl's lemma. To proceed, we make use of estimates due to Troyanov [**Tro91, Tro**].

2.3. ***DC*** **Calculus and measure-valued tensors on Alexandrov spaces.** Let us now come back to the *DC* structure induced by average distance functions. First, it is important to notice that the transition maps are not only *DC*, the first partial derivatives of their components also satisfy a weak continuity property. More precisely, according to Otsu and Shioya results, the first derivatives of a transition map exist at (the image of) any regular point and depend continuously of the point when it varies in $\mathrm{Reg}(X)$. We use Perelman's notation and call DC_0 map a *DC* map which satisfies this weak continuity property. Before we give further details about this seemingly unimportant point, let us set our strategy out to develop a well-defined tensor calculus on the manifold part of X. Due to the low regularity, we use the old-fashioned approach consisting in defining local tensors in charts and imposing *compatibility conditions* between them. We proceed in the same way to define the covariant derivative of tensor. Checking the compatibility condition when the involved objects are Radon measure requires a careful analysis. Below, without being too technical, we present what we think to be the more delicate points in this process. As a guideline, we consider the problem of defining the *Hessian* of a *DC* function h as a (symmetric) measure-valued tensor. In what follows, $F : \Omega \to \Theta$ stands for a transition map, (y_i) is a coordinate system on Θ, and (x_i) a coordinate system on Ω. For simplicity, we assume that h is defined on Θ, so we want to check that h and $h \circ F$ induce compatible tensors. We consider dh as a one form, namely

$$dh = \sum \frac{\partial h}{\partial y_i} dy_i \tag{7}$$

so that

$$D(dh) = \sum_{i,j} A_{ij} dy_i \otimes dy_j, \tag{8}$$

with

$$A_{ij} = \frac{\partial}{\partial y_i}\left(\frac{\partial h}{\partial y_j}\right) - \sum_m \frac{\partial h}{\partial y_k} \widetilde{\Gamma}_{ij}^k$$

where $\widetilde{\Gamma}_{ij}^{k}$ are the Christoffel symbols. Note that A_{ij} is a linear combination of derivatives of BV function thus, in particular, it gives no mass to $\mathcal{H}^{N-1}$-negligible sets (see the reminder on BV function in Section 2.1); the set of such Radon measures is denoted by $\mathcal{GM}$. Note also that the expression of the Christoffel symbols in terms of derivatives of the metric, and the continuity of the metric components on $\mathrm{Reg}(X)$ implies that the Christoffel symbols do not give mass to sets with *finite* $\mathcal{H}^{N-1}$-measure (extra explanations are given below). This subset of $\mathcal{GM}$ is denoted by $\mathcal{GM}_0$.

In order to relate the above expression to the one of $f \circ F$ in the coordinate system (x_i), one can mimick the proof in the smooth case (where some of the terms are now measures multiplied by densities). To get the compatibility result, one basically needs two properties. The first one is the chain rule formula

$$\frac{\partial}{\partial x_i}(h \circ F) = \sum_{s=1}^{N} \frac{\partial F_s}{\partial x_i} \left(\frac{\partial h}{\partial y_s} \circ F \right) \tag{9}$$

and the second one is the Leibnitz rule.

Let us start with the Leibnitz rule. For general BV functions, the Leibnitz rule does not hold because of the jump part of the derivative. Roughly speaking, the jump part is the set of points where the function is not "continuous" in the sense of Geometric Measure Theory (the correct term being approximately continuous). In particular, if a BV function $f : \Omega \to \mathbb{R}$ is continuous when restricted to $\Omega \setminus \mathfrak{S}$ with $\mathfrak{S}$ a $\mathcal{H}^{N-1}$-negligible set then Df has no jump part. Consequently, any first derivative of a DC_0 function as well as any component of the Riemannian metric have no jump part in their derivatives. Combining this property with the fact that Leibnitz rule holds true provided that *at least* one of the two BV functions has no jump part in its derivative highlights the importance of having DC_0 transition maps instead of mere DC ones.

Now, let us discuss the validity of (9). First notice that despite F being locally biLipschitz, it is not clear that $h \circ F$ is a BV function. Indeed, if we approximate h by a sequence of smooth functions f_n in L^1, one needs to uniformly bound from above the L^1-norm of $|D(f_n \circ F)|$ in order to conclude that $h \circ F$ is BV. To proceed, one makes use of the change of variable formula which involves $|\det dF|$. Consequently, discrepancy can occur when the sign of $\det dF$ is not constant $\mathscr{L}^N$-a.e. Once again, the fact that transition maps are DC_0 allows us to show the required property on the determinant sign. The next step is to give a consistent meaning to

$$\frac{\partial h}{\partial y_s} \circ F \tag{10}$$

when h is a BV function. Inspired by the case where the partial derivative is absolutely continuous with respect to the Lebesgue measure, we define

$$\langle F^*(\mu), \psi \rangle = \int_{\Theta} \psi \circ F^{-1} |\det dF^{-1}| \, \mu(dx),$$

where ψ is a compactly supported continuous function. Then when $\mu = \rho \mathscr{L}^N$, the change of variable formula gives us $F^*(\rho \mathscr{L}^N) = \rho \circ F \mathscr{L}^N$ so we define (10) using F^*. The last important point is that partial derivatives are continuous only in a weak sense, therefore, given ψ a compactly supported Lipschitz function, it is not clear whether the equality below is true

$$\int \psi \frac{\partial F_s}{\partial x_i} \left(\frac{\partial h}{\partial y_s} \right) = - \int \frac{\partial}{\partial y_s} \left(\psi \frac{\partial F_s}{\partial x_i} \right) h, \tag{11}$$

F_s being a component of F. However, the fact that $\frac{\partial F_s}{\partial x_i}$ is continuous out of a $\mathcal{H}^{N-1}$-negligible set while $\frac{\partial h}{\partial y_s}$ does not give mass to such a set, allows us to prove the above formula. We can also prove the same formula when the density (i.e. $\frac{\partial F_s}{\partial x_i}$ in the example above) is continuous out of a σ-finite set with respect to $\mathcal{H}^{N-1}$ provided that $\frac{\partial h}{\partial y_s}$ has no jump part.

Finally, we end up with a well-defined notion of tensors S with $\mathcal{GM}$-Radon measure components. Besides, we can also define the covariant derivative DS of a tensor S provided it has $\mathcal{BV}$

components where $f \in \mathcal{BV}$ roughly means that f is a bounded function of locally finite variation which is continuous out of a σ-finite set with respect to $\mathcal{H}^{N-1}$. The space $\mathcal{BV}$ is larger than the space BV_0 introduced by Perelman [4] (in our paper, we use the notation $\mathcal{BV}_0$ instead). Moreover, up to considering the orientable double cover, it is a standard algebraic matter to extend our result to differential forms with $\mathcal{BV}$ components and recover Perelman's result. Last, a similar approach can be performed for vector fields, it leads to the notion of $\mathcal{BV}$ vector field X as well as its covariant derivative DX, a tensor with $\mathcal{GM}$ components.

As mentioned above, this degree of generality permits us to define the Hessian of a DC function f as the covariant derivative of df. This includes the case of distance function d_p. Note that even on a smooth Riemannian manifold, the distributional Hessian of a distance function $\operatorname{Hess} d_p$ is *not*, in general, absolutely continuous with respect to the volume measure, see for instance [**MMU14**]. The extra term in $\operatorname{Hess} d_p$ is of jump type, namely it is concentrated on the Cut Locus of p which is $(N-1)$-dimensional in general, N being the dimension of the manifold. For example, this phenomenon arises on the real projective space. In our opinion, distance functions are central objects in the theory of Alexandrov spaces, and this justifies the fine analysis needed to get a setting encompassing them, unlike DC_0.

We end this part with some of the properties of the Hessian that can be generalized to Alexandrov spaces. We refer to [**4**] for a more exhaustive picture. To proceed, the crucial technical tool is to prove that a tensor with $\mathcal{GM}$ components can be evaluated by $\mathcal{BV}$ vector fields giving rise to a Radon measure. For instance, $\operatorname{Hess} d_p(X, X)$ makes sense as a Radon measure for any $\mathcal{BV}$ vector field X. Note that there is an apparent discrepancy in the latter definition since, by what precedes, $\operatorname{Hess} d_p$ is a measure-valued tensor which, in general, gives mass to $(N-1)$-dimensional sets while $\mathcal{BV}$ functions are only continuous out of a σ-finite set with respect to $\mathcal{H}^{N-1}$. In such a case, we use the *precise representative* of a BV function which is a specific representative defined everywhere out of a $(N-1)$-negligible set. However, despite $\operatorname{Hess} d_p(X, X)$ being well-defined, we cannot a priori hope for an integration by part formula like (11) at this level of generality. By a fine study of the jump part of the derivatives, we are nonetheless capable of proving the formula

$$\operatorname{Hess} f(X, Y) = D(df(Y))(X) - df(D_X Y)$$

for $f \in DC$, and X, $Y \in \mathcal{BV}$. We also establish

$$D_{\nabla^g \phi} \nabla^g \phi = 1/2\, \nabla^g |\nabla^g \phi|^2$$

for ϕ a DC function.

This extra work leads to interesting consequences however. For instance if $\phi = d_p$ is a distance function, then the modulus of its gradient equals 1 $\mathcal{H}^N$-a.e. Therefore, the above formulas give us

$$D_{\nabla^g d_p} \nabla^g d_p = 0 \text{ and } \operatorname{Hess} f(\nabla^g d_p, \nabla^g d_p) = \nabla^g d_p \cdot (\nabla^g d_p \cdot f).$$

Another corollary is the following integration by part formula for the Hessian, similar to that appearing in the Γ_2 calculus developed by Bakry [**Bak94b**]. For v a DC function, u a DC_0 one, and ψ a compactly supported Lipschitz function, it reads

$$\int \psi \operatorname{Hess} v(\nabla u, \nabla u) = \\ - \int \psi\, g(\nabla v, \nabla u) \Delta^g u - \frac{1}{2} \int \psi\, g(\nabla v, \nabla |\nabla u|_g^2) - \int g(\nabla v, \nabla u) g(\nabla u, \nabla \psi)\, dv_g.$$

To conclude, let us mention a related (and, in my opinion, delicate) open question: does a semiconcave function on Alexandrov space satisfy locally

$$\operatorname{Hess} f \le \lambda \mathcal{H}^N,$$

[4]See in particular paragraph 4.3 on the covariant derivative of a vector field with BV_0 components

λ being related to the constant appearing in the definition of semiconcavity? This question is part of my ongoing research in collaboration with Luigi Ambrosio. Its validity would have significant consequences on the regularity theory of Alexandrov spaces with curvature bounded below.

2.4. Application of *DC* Calculus to Optimal Mass Transport on Alexandrov space. In this part, we present an improved version of the Brenier-McCann theorem on an Alexandrov space. On Euclidean space, the proof is due to Gangbo-McCann [**GM96**] while on Riemannian manifold, it is proved by Gigli in [**Gig11**]. In the same paper, Gigli also proves that the assumption on the initial measure is sharp in the setting of Riemannian manifolds -see below for more details. The proof of the Brenier-McCann theorem sketched above highlights the connection between the set of non-differentiability points of a c-conjugate function and the initial measure: the optimal plan is unique and induced by a map whenever the initial measure does not give mass to the set of non differentiability points of an arbitrary c-conjugate map.

Concerning the regularity of a c-conjugate function ϕ, recall that it satisfies for any $x \in \operatorname{supp} \mu_0$

$$\phi(x) = \min_{y \in \operatorname{supp} \mu_1} 1/2\, d^2(x,y) - \phi^c(y).$$

Using the compactness of $\operatorname{supp} \mu_1$, we infer

$$\phi(x) = 1/2\, d^2(x, y_x) - \phi^c(y_x) \text{ and } \phi(z) \leq 1/2\, d^2(z, y_x) - \phi^c(y_x)$$

for all z. Now, recall that d_y^2 is (locally) semiconcave on all of X. More precisely, on a given bounded neighborhood and for y in a compact set K, the function d_y^2 is λ-concave where λ is uniform with respect to $y \in K$. By definition of semiconcavity, we infer from the above formulas that ϕ is semiconcave around any $x \in \operatorname{supp} \mu_0$.

Thus, in the Euclidean case, the set of non differentiability points of a c-conjugate function is that of a concave function. The latter set is well-understood as recalled below. We first introduce the following definition.

Definition 2.3 (c-c hypersurface in $\mathbb{R}^N$). *A set $\Omega \subset \mathbb{R}^N$ is a $c-c$ hypersurface if, up to a permutation of the indices, there exist two convex functions $f, g : \mathbb{R}^{N-1} \to \mathbb{R}$ such that Ω is the graph of $f-g$, i.e.*

$$\Omega = \{(x,t) \in \mathbb{R}^N; t = f(x) - g(x)\}.$$

In [**Zaj79**], Zajícek proved

Theorem 2.4. *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a concave function. Then the set of points where f is not differentiable is contained in the union of countably many $c-c$ hypersurfaces. Conversely, if a set $\Omega \subset \mathbb{R}^N$ can be covered by countably many $c-c$ hypersurfaces, then there exists a convex function $f : \mathbb{R}^d \to \mathbb{R}$ which is not differentiable at all the points in Ω.*

Therefore, it makes sense to introduce the following analogue of $c-c$ hypersurfaces on X, taking into account that the Hausdorff dimension of the complement of $\operatorname{Reg}(X)$ is at most $N-2$.

Definition 2.5 ($c-c$ hypersurface in X). *A set $\Omega \subset X$ is a countable $c-c$ hypersurface if, up to a set of Haudorff dimension at most $N-2$, it can be covered by chart domains on each of which it can be covered by the image of a countable number of $c-c$ hypersurfaces on $\mathbb{R}^N$ through the inverse of the chart.*

We infer from Perelman's *DC* theorem that the set of non differentiability points of a semiconcave function is a countable $c-c$ hypersurface. Indeed, using a partition of unity made of compactly supported DC_0 functions (see [**4**, Lemma 3.7] for a proof), we can without loss of generality assume that the semiconcave function is defined on the domain of a chart. The image of the semiconcave function through the chart is then a Euclidean *DC* function. We get the result by applying Zajícek's theorem to it. Moreover, out of $Sing(X)$, we know that the property of being differentiable does not depend on the chart. As a consequence, we get the following sharp version of the Brenier-McCann theorem on Alexandrov space.

THEOREM 2.6. *Let (X,d) be a N-dimensional Alexandrov space and $\mathcal{H}^N$ be the corresponding Hausdorff measure. We set $c(x,y)=1/2\,d^2(x,y)$ and μ_0,μ_1 two probability measures on X with compact supports. We further assume that μ_0 does not give mass to countable $c-c$ hypersurfaces. Then, Kantorovitch's dual problem admits a solution (ϕ,ϕ^c) with ϕ a c-conjugate function. As a consequence, the mass transport problem admits a unique solution and this solution is induced by a map F. Further, the map F satisfies for μ_0 almost every $x\in X$,*

$$F(x)=exp_x(-\nabla\phi(x)).$$

To conclude, let us add a few words concerning the sharpness of the assumption on μ_0. First, let us mention that, in the case of Riemannian manifold, Gigli's definition does not include the possibility of discarding a "small" set. The sharpness is then intended as follows. Let μ_0 be compactly supported and such that for any compactly supported μ_1, there is a unique plan between μ_0 and μ_1, optimal relative to c, and, moreover, this plan is supposed to be induced by a map. Gigli proves that this property holds iff μ_0 does not give mass to countable $c-c$ hypersurfaces (actually his statement holds for measures with finite second order moments). His proof is based on the second statement in Zajícek 's Theorem 2.4. In our setting, we cannot adapt the argument out of the manifold part of X. Maybe, this could be done if we had the stronger property $\mathcal{H}^{N-2}(X\setminus \mathrm{Reg}(X))<+\infty$ but this property is unknown.

CHAPTER 3

Prescription of Gauss curvature

In this part, we describe results related to a classical theorem by Alexandrov on the Gauss curvature prescription of Euclidean convex bodies [**Ale42, Ale05**]. These results are contained in the papers [**5, 2**]. Here "Gauss curvature" is intended in a generalized measure-theoretic sense. We thus start with describing the Gauss curvature measure introduced by Alexandrov himself.

Consider a convex body (*i.e.* a closed bounded convex set whose interior is nonempty) $\mathcal{C}$ in $\mathbb{R}^{m+1}$ and assume that the origin of $\mathbb{R}^{m+1}$ is located *within* $\mathcal{C}$. Let us call $\mathcal{G} : \partial\mathcal{C} \rightrightarrows \mathbb{S}^m$ the Gauss multivalued map which maps a point $c \in \partial\mathcal{C}$ onto the set of all outward unit normal vectors at c and consider $\sigma(\mathcal{G}(\cdot))$ which is the *pull-back* of the uniform probability measure through the Gauss map. This object is indeed a Borel measure supported on $\partial\mathcal{C}$ thanks to the following fact

$$\sigma\left(\{n \in \mathbb{S}^m;\, \exists\, c_1 \neq c_2 \in \partial\mathcal{C};\, n \in \mathcal{G}(c_1) \cap \mathcal{G}(c_2)\}\right) = 0. \tag{12}$$

(see [**Bak94a**, Lemma 5.2] for a proof).

Alexandrov's problem consists in prescribing the shape of a convex body knowing its curvature. The above measure being supported on the boundary of $\mathcal{C}$ makes the question easy to solve. The Gauss curvature measure is thus defined as the pull-back of the above measure onto the unit sphere centered at the origin through the following homeomorphism

$$\overrightarrow{\rho} : \begin{array}{ccc} \mathbb{S}^m & \longrightarrow & \partial\mathcal{C} \\ x & \longmapsto & \rho(x)x \end{array} \tag{13}$$

where ρ is the *radial function* defined by $\rho(x) = \sup\{s; sx \in \mathcal{C}\}$). Note that ρ is a Lipschitz function bounded away from 0 thanks to the assumption on the origin location. In the rest of the chapter the Gauss curvature measure is denoted by

$$\mu := \sigma(\mathcal{G} \circ \overrightarrow{\rho}(\cdot)).$$

Note also that the curvature measure depends on the location of the origin within the convex body and is invariant under homotheties about that point. For instance, if the underlying convex body is a polyhedron, then the curvature measure is a finite convex combination of Dirac masses $\sum a_i \delta_{x_i}$ where x_i are the directions determined by the vertices of the polyhedron and a_i the exterior normal angles.

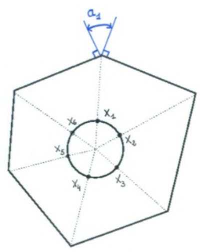

FIGURE 1. Gauss curvature measure of a convex polyhedron.

Let us comment a little bit on the terminology "curvature measure". When the underlying convex body $\mathcal{C}$ is smooth, the Borel measure $\sigma(\mathcal{G}(\cdot))$ defined on $\partial\mathcal{C}$ is absolutely continuous with respect to the surface measure of $\partial\mathcal{C}$ (equivalently, the m-Hausdorff measure restricted to this set) and the density is nothing but the standard Gauss curvature (up to a multiplicative normalization factor). There is another natural generalization of the Gauss curvature to arbitrary convex body. It consists in pushing the surface measure through the generalized Gauss map. Once again, the fact that $\mathcal{G}$ is multivalued in general is no trouble thanks to (12). In order to compare to the curvature measure, note that when $\mathcal{C}$ is a convex polyhedron, the corresponding measure is a finite combination of Dirac masses $\sum a_i\delta_{x_i}$ where the (x_i) are the outward unit directions normal to the polyhedron faces and a_i are the area measures of the corresponding faces. In the literature, this measure is called *area measure*, see for instance Schneider's book on convex bodies [**Sch93**]. The area measure can also be studied by variational methods as in the paper [**Car04**], see also [**McC95**].

Now that the Gauss curvature has been defined, let us come back to Alexandrov's problem. The assumption on the origin location implies the existence of $\varepsilon > 0$ such that the open ball $B(0,\varepsilon) \subset \mathcal{C}$. This fact yields the following bound

$$\measuredangle x, n < \pi/2 - \varepsilon' \tag{14}$$

on the angle between x and any outward normal vector n to a supporting hyperplane at $\vec{\rho}(x)$.

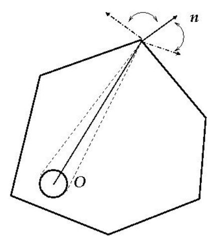

FIGURE 2. Upper bound on the angle.

Recall that a convex subset of $\mathbb{S}^m$ is defined as the intersection of a convex cone in $\mathbb{R}^{m+1}$ with $\mathbb{S}^m$. As a consequence of (14), we get that for all non-empty spherical convex set $\omega \subsetneq \mathbb{S}^m$, the curvature measure satisfies

$$\mu(\omega) < \sigma(\omega_{\pi/2}) \tag{15}$$

where $\omega_{\pi/2} = \{x \in \mathbb{S}^m; d(x,\omega) < \pi/2\}$ and $d(\cdot,\cdot) = \measuredangle \cdot, \cdot$ is the standard distance on $\mathbb{S}^m$.

In particular the above formula for ω a closed hemisphere tells us that μ cannot be supported in ω. Alexandrov's theorem states that (15) is actually a sufficient condition for μ arising from this construction. More precisely

THEOREM 0.1 (Alexandrov). *Let σ be the uniform probability measure on $\mathbb{S}^m$ and μ be a Borel probability measure on $\mathbb{S}^m$ such that for any non-empty convex set $\omega \subsetneq \mathbb{S}^m$,*

$$\mu(\omega) < \sigma(\omega_{\pi/2}).$$

Then, there exists a unique convex body in $\mathbb{R}^{m+1}$ containing 0 in its interior (up to homotheties) whose μ is the Gauss curvature measure.

In a subsequent section, we describe a variational proof of Theorem 0.1 [**2**]. After a detour to the relativistic heat equation, a quite different subject which however can be treated with similar tools, we describe an hyperbolic analogue of Alexandrov's theorem. To proceed, we consider Alexandrov's result as an embedding result for manifolds homeomorphic to the sphere with a given Gauss curvature. Of course, the Gauss formula prevents from embedding a hyperbolic manifold as the boundary of Euclidean convex set, however one can circumvent this problem by making use of the Minkowski spacetime. We postpone the precise statement to the dedicated chapter.

The next part contains an outline of Alexandrov's original arguments.

1. Alexandrov's proof

A common feature in Alexandrov's proofs about convex sets is that most of them are obtained by treating the case of convex polyhedra first and then generalizing the result to all convex sets by an approximation argument. The proof of the result discussed in the preceeding part follows this strategy.

We first outline the proof and then comment and add more details about some specific points. The first step is to approximate the given measure by a sequence $(\mu_k)_k$ of finitely supported measures which satisfy Alexandrov's condition (this is not very difficult, we also use this argument in our paper; a proof of this fact can be found there). Then, taking for granted that the theorem is proved in the case of finitely supported measures, we consider a sequence of convex polyhedra $(P_k)_k$ whose curvature measures are $(\mu_k)_k$. Noticing that the curvature measure is invariant by dilations about the origin, we can further assume that all the P_k are contained in a fixed ball. Consequently, this sequence is a compact set with respect to Hausdorff distance; thus up to extracting a subsequence, we can assume that P_k converges to a convex body. To conclude, Alexandrov shows that the Gauss curvature measure is a continuous mapping with respect to the Hausdorff convergence and the $*$-weak convergence. To summarize, the two main ingredients are: solving the problem when the underlying convex body is a polyhedron and showing that the Gauss curvature measure is continuous.

The solution of the problem in the polyhedron case is a classical result so we only sketch it (details can be found in Alexandrov's book [**Ale05**]). If $\sum_{i=1}^{k} a_i \delta_{x_i}$ is a finite measure satisfying Alexandrov's condition then the vertices of the polyhedron are on the half-lines determined by the x_i. The set $\mathcal{P}_{ol}$ of such convex polyhedra is parameterized by $(d_1, \cdots, d_k)$, being d_i the positive distance between the vertex relative to x_i and the origin. It can be proved that $\mathcal{P}_{ol}$ is a convex cone and, being the curvature measure invariant by dilations, it is natural to restrict our attention to the subset $\mathcal{P}_{ol1} \subset \mathcal{P}_{ol}$ whose elements satisfy $\sum d_i = 1$. Similarly, the set $\mathcal{A}_l$ of a_i for which $\sum_{i=1}^{k} a_i \delta_{x_i}$ satisfies Alexandrov's condition is a convex set (indeed, in the polyhedral case, the condition can be rewritten as finitely many affine inequalities). The point is that both sets $\mathcal{A}_l$ and $\mathcal{P}_{ol1}$ have the same *dimension.* Finally, Alexandrov proves that mapping a convex polyhedron in $\mathcal{P}_{ol1}$ onto its curvature measure is a continuous, injective (and proper) map; in the rest of this part, we shall call this map "Gauss curvature mapping". Essentially, this follows from easy considerations on how the polar cone P° behaves when the initial polyhedron P varies (including $P_1 \subset P_2 \Rightarrow P_2^\circ \subset P_1^\circ$). Then, it suffices to apply this monotonicity result at each vertex of the polyhedron (the cone generated by the spherical image at the vertex v is the polar cone of the inward cone at v determined by the convex body). As a consequence of these properties, the Gauss curvature mapping is a homeomorphism onto its image which is a closed subset of $\mathcal{A}_l$. The fact that both sets are connected and have the same dimension then implies by the *invariance domain theorem* that the mapping is surjective; this completes the proof.

Concerning the continuity of the Gauss curvature measure, Alexandrov's proof is only available in Russian so we provide a proof based on Schneider's book [**Sch93**, Chapter 4]. To this end, we denote by $U_{\mathcal{C}}$ the set of points $n \in \mathbb{S}^m$ such that $n \in \mathcal{G}(x)$ for a *unique* vector $x \in \mathcal{C}$; According to (12), $U_{\mathcal{C}}$ carries all the mass of σ. Also recall that for $n \in \mathbb{S}^m$ and $x \in \partial\mathcal{C}$, $n \in \mathcal{G}(x)$ iff $h_{\mathcal{C}}(n) = \langle n, x\rangle$, being $h_{\mathcal{C}}$ the support function of $\mathcal{C}$ (the definition is recalled in the next section). Besides, given K_j a sequence of convex bodies converging to $\mathcal{C}$, the sequence of corresponding support functions $(h_{K_j})_j$ converges (uniformly on compact sets) to that of $\mathcal{C}$ -the converse property also holds. The proof of the continuity of the Gauss curvature measure is then as follows. Fix $\Theta \subset \mathbb{S}^m$ an open set and $n \in \mathcal{G}(x) \cap U_{\mathcal{C}}$ where $x \in \overrightarrow{\rho}_{\mathcal{C}}(\Theta)$. There exists a sequence $x_j \in \partial K_j$ such that $n \in \mathcal{G}(x_j)$; by compactness, we can further assume that x_j converges to a point z. By what precedes, we get that $h_{\mathcal{C}}(n) = \langle z, n\rangle$; moreover being $n \in U_{\mathcal{C}}$, we infer $z = x$. Therefore, since $\Theta \subset \mathbb{S}^m$ is an open set, $x_j \in \overrightarrow{\rho}_{K_j}(\Theta)$ for large j and we get

$$\mu_{\mathcal{C}}(\Theta) = \sigma(\mathcal{G}_{\mathcal{C}}(\overrightarrow{\rho}_{\mathcal{C}}(\Theta))) = \sigma(\mathcal{G}_{\mathcal{C}}(\overrightarrow{\rho}_{\mathcal{C}}(\Theta)) \cap U_{\mathcal{C}}) \leq \liminf \sigma(\mathcal{G}_{K_j}(\overrightarrow{\rho}_{K_j}(\Theta))).$$

Being Θ arbitrary, we conclude by using that the total mass of curvature measure does not depend on the convex body, therefore it is a standard result in measure theory that μ_{K_j} weakly converges to $\mu_{\mathcal{C}}$.

It remains to prove the uniqueness part of the statement, this cannot be done by the above argument. In what follows, we call *smooth point* a point on the boundary of a convex body which has a tangent hyperplane and denote by $\mathcal{S}_{m,\mathcal{C}}$ the set of smooth points of $\partial\mathcal{C}$. Given two convex bodies K_1 and K_2 having the same curvature measure, Alexandrov proves that if for any $x_1 \in \mathcal{S}_{m,K_1}$, $x_2 \in \mathcal{S}_{m,K_2}$ such that $0, x_1, x_2$ are aligned, the unique supporting hyperplanes at x_1 and x_2 are parallel then K_1 and K_2 are homothetical. This fact does not seem completely obvious to us, so we provide an argument for completeness. Let us first assume that both K_1 and K_2 are polyhedra. Then, take two adjacent m-faces F_1, F_2 of K_1 and the corresponding parallel faces $\widetilde{F}_1, \widetilde{F}_2$ of K_2. If the distances between the two pairs of parallel faces were distinct then, by easy geometric considerations, the traces of $F_1 \cap F_2$ and $\widetilde{F}_1 \cap \widetilde{F}_2$ onto $\mathbb{S}^m$ would be disjoint and this would give us two distinct vertices $v_i \in K_i$, $i \in \{1, 2\}$; in particular the curvature measures of K_1 and K_2 could not be equal. To prove the general case, approximate K_1 and K_2 by convex polyhedra P_1^n and P_2^n respectively, defined as the intersection of finitely many halfspaces supported at aligned smooth points for both K_1 and K_2 (more precisely, construct the polyhedra with the help of a dense subset of such smooth points; the fact that

$$0 \in \overset{\circ}{K}_1 \cap \overset{\circ}{K}_2$$

guarantees that the limit of P_i^n is not larger that K_i for $i \in \{1, 2\}$). By what precedes $P_1^n = \lambda_n P_2^n$, further the coefficients λ_n form a bounded sequence; this gives us the result up to extracting a subsequence.

To conclude, let us explain why the supporting hyperplanes must be parallel at smooth points as above. The proof is by contradiction. Assume this is not true for $x_1 \in \mathcal{S}_{m,K_1}$ and $x_2 \in \mathcal{S}_{m,K_2}$. Up to dilating one of the bodies, we can further assume that $x_1 = x_2$. Next, we partition ∂K_1 as

$$\partial K_1 = \big(\partial K_1 \setminus K_2\big) \sqcup \big(\partial K_1 \cap \overset{\circ}{K}_2\big) \sqcup \big(\partial K_1 \cap \partial K_2\big)$$

and we call F_{11}, F_{12}, F_{13} the elements of the partition (in the same order). We partition similarly ∂K_2 and call F_{21}, F_{22}, F_{23} the corresponding elements, in particular $F_{13} = F_{23}$. We are going to prove that $\sigma(\mathcal{G}(F_{22})) < \sigma(\mathcal{G}(F_{11}))$ contradicting $\mu_{K_1} = \mu_{K_2}$ since the traces of F_{22} and F_{11} on the unit sphere are the same. Notice that being $Conv(\{0\} \cup F_{22}) \subset Conv(\{0\} \cup F_{11})$, the corresponding spherical images satisfy

$$\mathcal{G}(F_{22}) \supset \mathcal{G}(F_{11}).$$

Some extra work is needed to get a strict inequality between the measures of the above sets. First, note that for $i \in \{1, 2\}$, $F_{i2} \cup F_{i3}$ is a closed set thus $\mathcal{G}(F_{i2} \cup F_{i3})$ is closed as well. Consequently, the set

$$\kappa := \big(\mathbb{S}^m \setminus \mathcal{G}(F_{12} \cup F_{13})\big) \cap \big(\mathbb{S}^m \setminus \mathcal{G}(F_{22} \cup F_{23})\big)$$

is open. Now, by definition of a partition,

$$\mathbb{S}^m \setminus \mathcal{G}(F_{12} \cup F_{13}) \subset \mathcal{G}(F_{11})$$

and, up to a σ-negligible set N (12),

$$\big(\mathbb{S}^m \setminus \mathcal{G}(F_{22} \cup F_{23})\big) \cap \big(\mathcal{G}(F_{21}) \setminus N\big) = \emptyset.$$

From these inequalities, it is easy to infer the claim provided the open set κ is nonempty. This last property can be proved as follows. Let us call n_1 and n_2 the two outward unit vectors at the smooth point $x = x_1 = x_2$. Now take the bisector vector n of n_1 and n_2 in $Vect(n_1; n_2)$ and set P its orthogonal hyperplane passing through x. The hyperplane P determines two caps C_1 and C_2

bounded by F_{11} and F_{22} respectively, each of which contains a point y_i such that $n \in \mathcal{G}(y_i)$. By definition, these points belong to $\mathcal{G}(F_{11})$ and $\mathcal{G}(F_{22})$ respectively; a closer look actually gives us $n \in \kappa$.

2. A variational approach

2.1. The variational problem. Besides the radial function

$$\rho(x) = \sup\{s; sx \in \mathcal{C}\},$$

another important function that characterizes a convex body $\mathcal{C}$ is its *support function* whose definition is

$$h(n) = \sup_{x \in \mathbb{S}^m} \{\rho(x)\langle x, n\rangle\} \tag{16}$$

Note that we consider these functions only on $\mathbb{S}^m$ instead of $\mathbb{R}^{m+1}$. Moreover

$$h(n) = \rho(x)\langle x, n\rangle$$

if and only if the hyperplane orthogonal to n through $\overrightarrow{\rho}(x)$ supports the convex $\mathcal{C}$. In other words, this equality amounts to

$$n \in \mathcal{G} \circ \overrightarrow{\rho}(x).$$

More generally, we can use the same formulas to define the radial and support functions of a star-shaped set with respect to 0. For simplicity, we further assume that ρ is continuous and bounded away from 0 and ∞. Let us call F a star-shaped set as above and consider its *polar set*

$$F^\circ = \{n \in \mathbb{R}^{m+1}; \forall x \in F\ \langle x, n\rangle \leq 1\}$$

The set F° is a convex body containing 0 in its interior. Moreover if we apply the polar transform twice, the set $F^{\circ\circ}$ is the *convex hull* of F. Since F is star-shaped, the equality $F = F^{\circ\circ}$ holds if and only if $\partial F = \partial F^{\circ\circ}$. Equivalently, F as above is convex if and only if

$$\rho_{F^{\circ\circ}} = \rho_F \tag{17}$$

while only $\geq$ is true in general. An interesting feature of the polar transform is the following relation

$$\frac{1}{h_F} = \rho_{F^\circ}$$

which follows from the definition. By using it, we can rewrite (17) as

$$\rho_F = \inf_{n \in \mathbb{S}^m; \langle \cdot, n\rangle > 0} \left\{\frac{h(n)}{\langle \cdot, n\rangle}\right\} \tag{18}$$

while only $\leq$ is true in general.

It is then a nice idea due to Oliker **[Oli07]** to introduce the following transformations $\phi = \ln(1/h)$ and $\psi = \ln(\rho)$ which converts (16) and (18) into the more symmetric relations

$$\begin{aligned} \psi(x) &= \min_{n \in \mathbb{S}^m} c(n, x) - \phi(n) \\ \phi(n) &= \min_{x \in \mathbb{S}^m} c(n, x) - \psi(x) \end{aligned}$$

where c is defined by the formula

$$c(n, x) = \begin{cases} -\log\langle n, x\rangle = -\log\cos d(n, x) & \text{if } d(n, x) < \pi/2 \\ +\infty & \text{otherwise} \end{cases}$$

The relations between ψ and ϕ are well-known in optimal mass transport. The functions ϕ and ψ are said to be c-conjugate, one also uses the term c-concave map for ϕ (or ψ). This is due to the fact that the variational problem

$$\sup_{(\phi,\psi) \in \mathcal{A}} \left\{\int_{\mathbb{S}^m} \phi(n) d\sigma(n) + \int_{\mathbb{S}^m} \psi(x) d\mu(x)\right\}, \tag{19}$$

where $\mathcal{A}$ stands for the set of pairs (ϕ, ψ) of Lipschitz functions defined on $\mathbb{S}^m$ that satisfy $\phi(n) + \psi(x) \leq c(n, x)$ for all $x, n \in \mathbb{S}^m$, is very useful to solve the optimal mass transport problem

$$\min_{\Pi \in \Gamma(\sigma,\mu)} \int_{\mathbb{S}^m \times \mathbb{S}^m} c(n, x)\, d\,\Pi(n, x).$$

It is also well-known that when solutions to (19) do exist in L^1, they have to coincide with a pair of c-conjugate functions (up to modifications on σ and μ negligible sets respectively) in order to saturate the constraint. However, while solution does exist when the cost function c is, say, Lipschitz regular and bounded, the fact that our cost function assume infinite values may prevent the variational problem to have solution. We refer to Chapter 1 for a review of optimal mass transport and to [**BS11**, Section 4] for a detailed discussion of the latter point.

Nonetheless, in this particular case, it can be proved that solutions exist and are essentially unique in the sense that if $(\phi, \psi), (\widetilde{\phi}, \widetilde{\psi})$ are solutions then $\widetilde{\phi} - \phi = \psi - \widetilde{\psi}$ is a constant. If we rephrase this uniqueness property in terms of $\rho = e^{\psi}$ and $h = e^{-\phi}$, we get, among other things, that ρ and h satisfy (16) and (18). Therefore, the solutions of Kantorovitch's problem determine convex bodies, besides the convex bodies relative to the solutions are the dilations $\lambda\mathcal{C}$ of a single convex body $\mathcal{C}$.

To conclude, one main property remains to be proved: the fact that the curvature measure of $\mathcal{C}$ is indeed μ. There is a simple way to proceed based on Kantorovitch's duality

$$\max_{(\phi,\psi)\in\mathcal{A}} \left\{ \int_{\mathbb{S}^m} \phi(n) d\sigma(n) + \int_{\mathbb{S}^m} \psi(x) d\mu(x) \right\} = \min_{\Pi \in \Gamma(\sigma,\mu)} \int_{\mathbb{S}^m \times \mathbb{S}^m} c(n, x)\, d\,\Pi(n, x)$$

which holds as a consequence of existence of solutions to (19), see (22) below for the argument. We fix (ϕ_0, ψ_0) a solution of Kantorovitch's variational problem and Π_0 an optimal plan. The above equality reads

$$\int (\phi_0 + \psi_0)\, d\Pi_0 = \int c\, d\Pi_0.$$

As in our proof of the Brenier-McCann theorem 1.4, we infer

$$\Pi_0(\{(n, x) \in (\mathbb{S}^m)^2; \phi_0(n) + \psi_0(x) = c(n, x)\}) = 1.$$

With the latter property at our disposal, we can show that μ is indeed the curvature measure of the convex body determined by Kantorovitch's variational problem. In what follows, U is an arbitrary Borel set of $\mathbb{S}^m$.

$$\begin{aligned}
\mu(U) &= \Pi_0(\mathbb{S}^m \times U) \\
&= \Pi_0(\mathbb{S}^m \times U \cap \{(n, x) \in (\mathbb{S}^m)^2; \phi_0(n) + \psi_0(x) = c(n, x)\}) \\
&= \Pi_0(\mathbb{S}^m \times U \cap \{(n, x) \in (\mathbb{S}^m)^2; n \in \mathcal{G}(\overrightarrow{\rho}(x))\}) \\
&= \Pi_0(\mathcal{G} \circ \overrightarrow{\rho}(U) \times U \cap \{(n, x); n \in \mathcal{G}(\overrightarrow{\rho}(x))\}) \\
&\underset{(12)}{=} \Pi_0(\mathcal{G} \circ \overrightarrow{\rho}(U) \times \mathbb{S}^m \cap \{(n, x); n \in \mathcal{G}(\overrightarrow{\rho}(x))\}) \\
&= \Pi_0(\mathcal{G} \circ \overrightarrow{\rho}(U) \times \mathbb{S}^m) \\
&= \sigma(\mathcal{G} \circ \overrightarrow{\rho}(U)).
\end{aligned}$$

2.2. Solving the variational problem. In this part, we describe the strategy to prove existence of solutions to the variational problem as well as their uniqueness up to additive constant[1]. The uniqueness part is classical provided that *all* the solutions are differentiable σ-almost everywhere. The argument we use is the one that appears in the proof that the mass transport problem admits a unique solution and that this solution is induced by a map, being the map defined almost everywhere in terms of the gradient of ϕ. Uniqueness then follows from the fact: Let $f : M \to \mathbb{R}$

[1]Oliker also solved Kantorovitch's problem for this cost function. His proof is more in the spirit of Alexandrov's one. It consists in solving the problem for finitely supported measures/ convex polyhedra first and then proving the general case by approximation. See [**2**] for more comments.

be a locally Lipschitz function defined on a closed (i.e. connected, compact, without boundary) Riemannian manifold. Suppose $\nabla f = 0$ a.e. then f is constant.

The question of the existence is more delicate to solve. Indeed, being c infinite on the set of points (n, x) at distance at least $\pi/2$, the set of functions $\mathcal{A}$ is not compact. Our approach is based on a generalization of a construction by Rockafellar which is a classical tool in optimal mass transport. Rockafellar's construction consists in building from a cyclically monotone set (i.e. a set Γ such that for any pairs $(x_1, y_1), \cdots, (x_k, y_k) \in \Gamma$, $\sum_{i=1}^{k} \langle x_{i+1} - x_i, y_i \rangle \leq 0$ where $x_{k+1} = x_1$) a lower semi-continuous convex function. The corresponding notion in optimal mass transport is that of c-cyclical monotonicity (see Chapter 1 for a definition). Roughly speaking, when the cost function c is real-valued and continuous, Rockafellar's construction can be generalized and gives a c-concave map from a c-cyclically monotone subset.

This approach is relevant in our problem provided that we take for granted that the mass transport problem is well-posed

$$\min_{\Pi \in \Gamma(\sigma,\mu)} \int_{\mathbb{S}\times\mathbb{S}} c(n, x)\, d\Pi(n, x) < +\infty. \tag{20}$$

Indeed, when the cost function is continuous (and possibly assume infinite value), it is known that the support of an optimal plan Π_0 is a c-cyclically monotone set. Therefore, the set

$$\Gamma := \operatorname{supp} \Pi_0 \cap \{c < +\infty\}$$

is a c-cyclically monotone set (as a subset of $\operatorname{supp} \Pi_0$) of full Π_0-measure (as a consequence of (20)). The generalization of Rockafellar's construction is given by

$$\phi(n) = \inf_{s\in\mathbb{N}} \inf_{(n_i,x_i)_{i=1}^{s} \in \Gamma^s} \Big\{ \sum_{i=0}^{s} c(n_{i+1}, x_i) - \sum_{i=0}^{s} c(n_i, x_i) \Big\} \tag{21}$$

where $n_{s+1} = n$ and $(n_0, x_0) \in \Gamma$ is fixed.

Our choice of Γ guarantees that the term inside the brackets belongs to $\mathbb{R} \cup \{+\infty\}$ thus ϕ is well-defined and belongs to $\overline{\mathbb{R}}$. The main issue is to prove that the image of ϕ is actually in $\mathbb{R} \cup \{-\infty\}$. Indeed, if it is the case then $\phi^c(x) = \inf_{n\in\mathbb{S}^m} c(n, x) - \phi(n)$ is well-defined and one can prove that

$$\Gamma \subset \{(n, x) \in \mathbb{S}^m \times \mathbb{S}^m; \phi(n) + \phi^c(x) = c(n, x)\}.$$

The above property then implies (we discard the question of integrability/regularity of ϕ and ϕ^c which are only of a technical nature)

$$\begin{aligned}
\sup_{(\phi,\psi)\in\mathcal{A}} \int_{\mathbb{S}^m} \phi\, d\sigma + \int_{\mathbb{S}^m} \psi\, d\mu \quad &\leq \quad \min_{\Pi\in\Gamma(\sigma,\mu)} \int_{\mathbb{S}^m\times\mathbb{S}^m} c\, d\Pi \\
&\leq \quad \int_{\Gamma} c\, d\Pi_0 = \int_{\Gamma} \phi + \phi^c\, d\Pi_0 = \int_{\mathbb{S}^m} \phi\, d\sigma + \int_{\mathbb{S}^m} \phi^c\, d\mu
\end{aligned}$$

which in turns gives us that (ϕ, ϕ^c) is a solution of Kantorovitch's variational problem.

To conclude this part, we sketch our argument that gives $\mathrm{Im}(\phi) \subset \mathbb{R} \cup \{-\infty\}$. When the (first) term inside the brackets in (21) is finite, one says that there is a finite chain from n_0 to n of length at most s. We are done if we can prove that there is a finite chain from n_0 to any point n. To do so, we call C_k the set of points that can be linked to n_0 by a chain of length at most k. We also define A_k the set of "last intertwining links" of chains of length at most k. In more mathematical terms, $C_k = (A_k)_{\pi/2}$ by definition of the cost function. The point is to show that as long as $C_k \subsetneq \mathbb{S}^m$, the μ-measure of A_{k+1} keeps growing with a *uniform* lower bound on the growth. Since $\mu(\mathbb{S}^m) < +\infty$, this can only happen finitely many times, i.e. $C_k = \mathbb{S}^m$ for k sufficiently large.

The existence of this uniform lower bound follows from a self-improvement of Alexandrov's assumption (15). Roughly, this improvement is a simple consequence of the compactness of the set of closed convex subsets of $\mathbb{S}^m$ relative to the Haussdorff distance. It reads as follows: there exists $\varepsilon > 0$ such that

$$\mu(\omega) \leq \sigma(\omega_{\pi/2}) - \epsilon \tag{22}$$

for any closed convex subset $\emptyset \neq \omega \subsetneq \mathbb{S}^m$.

The sets A_i are defined by induction by the following formulas

$$\begin{aligned} A_0 &= \{x_0\} \\ A_{i+1} &= p_x(p_n^{-1}((A_i)_{\pi/2}) \cap \Gamma) \end{aligned}$$

where p_n and p_x stand for the projections on the n and x coordinates respectively. To estimate the measures of the A_i, we fix an optimal transport plan Π_0 in $\Gamma_0(\sigma, \mu)$ and write

$$\begin{aligned} \mu(A_{i+1}) \underset{\Pi_0 \text{ is a plan}}{=} \Pi_0(p_x^{-1}(A_{i+1})) &\underset{\text{def. of } A_{i+1}}{\geq} \Pi_0(p_n^{-1}((A_i)_{\pi/2}) \cap \Gamma) \\ \underset{\Pi_0(\Gamma)=1}{=} \Pi_0(p_n^{-1}((A_i)_{\pi/2})) &\underset{\Pi_0 \text{ is a plan}}{=} \sigma((A_i)_{\pi/2}). \end{aligned}$$

Besides, note that $\mathbb{S}^m \setminus (A_i)_{\pi/2} = \{x \in \mathbb{S}^m; \forall z \in A_i,\ \langle x, z\rangle \leq 0\}$ is defined by linear inequalities. Thus

$$\mathbb{S}^m \setminus (A_i)_{\pi/2} = \mathbb{S}^m \setminus (Conv(A_i))_{\pi/2})$$

and, as long as $\mathbb{S}^m \setminus (A_i)_{\pi/2} \neq \emptyset$, the set $Conv(A_i)$ satisfies the hypothesis in (22) which gives

$$\mu(A_{i+1}) \geq \sigma((A_i)_{\pi/2}) = \sigma(Conv(A_i))_{\pi/2}) \geq \mu(Conv(A_i)) + \epsilon \geq \mu(A_i) + \epsilon.$$

Finally, we mention that the techniques described above can also be used to the solve the mass transport problem relative to the cost c. The statement is the following.

THEOREM 2.1. *Let $f\sigma$ and μ two probability measures on $\mathbb{S}^m$ such that there exists $\Pi \in \Gamma(f\sigma, \mu)$ for which $c \in L^\infty(\Pi)$. Then, the mass transport problem*

$$\min_{\Pi \in \Gamma(f\sigma,\mu)} \int_{\mathbb{S}^m \times \mathbb{S}^m} c(n, x)\, d\,\Pi(n, x) < +\infty \tag{23}$$

admits a unique solution Π_0. Moreover, $\Pi_0 = (Id, T)_\sharp f\sigma$ where for σ-a.e. $n \in \mathbb{S}^m$,

$$T(n) = \exp_n\left(\frac{-\arctan|\nabla\phi(n)|}{|\nabla\phi(n)|}\nabla\phi(n)\right)$$

being ϕ a Lipschitz c-concave function.

The assumption we make (i.e. the existence of a plan Π such that $c \in L^\infty(\Pi)$) is stronger than the standard assumption (23). It guarantees that

$$\mu(\omega) < (f\sigma)(\omega_{\pi/2})$$

for all convex sets ω. Indeed, (23) only implies the corresponding non-strict inequality and we don't know whether the result holds under this weaker assumption.

3. A relativistic detour

In the previous section, we construct a Kantorovitch potential by following the generalized Rockafellar construction. The main tool is the existence of finite chain between any pair of points (in the image of Γ). The method seems to strongly depend on the relation between the measures μ and σ. In this part, we describe a much broader setting in which this method can be applied to get the existence of a Kantorovitch potential. In the Monge-Kantorovitch problem, beside the standard assumption concerning absolute continuity, we will additionally require the support of the initial measure to be *connected.* Moreover, we will also assume the cost function to be *relativistic.* This class of cost functions is modelled on the relativistic heat cost defined by Brenier [**Bre03**] in order to study the relativistic heat equation.

The relativistic heat equation is defined by the formula

$$\partial_t \rho = \mathrm{div}\left(\frac{\nabla \rho}{\sqrt{1+\frac{1}{s^2}|\frac{\nabla\rho}{\rho}|^2}}\right) = \mathrm{div}(\rho \nabla h^*(\nabla(\log\rho))) \tag{24}$$

where s bounds the propagation speed. Note that when $s \to +\infty$, we formally recover the standard heat equation.

In the above formula, h^* is the Legendre transform of the function h defined by

$$h(z) = \begin{cases} 1-\sqrt{1-|z|^2} & |z| \le 1 \\ \infty & |z| > 1. \end{cases} \tag{25}$$

Brenier introduced the relativistic heat cost $c(x,y) = h(x-y)$ to obtain the relativistic heat equation as a gradient flow of the Boltzmann entropy via the JKO time discrete scheme [**AGS08**]. Recall that the heat equation can be obtained by computing the time discrete solution at the time step i in the following way

(26)

$$\rho^i = \arg\inf_{\rho\in\mathscr{W}_2(\mathbb{R}^d)} \int \rho\log\rho + \varepsilon\, W^2_{2,\varepsilon}(\rho^{i-1},\rho) \text{ with } W^2_{2,\varepsilon}(\rho^{i-1},\rho) = \min_{\pi\in\Gamma(\rho^{i-1},\rho)} \int \left(\frac{|x-y|}{\varepsilon}\right)^2 d\pi(x,y)$$

and by passing to the limit when the time step ε goes to zero. Analogously, a Cauchy result has been proved by McCann and Puel in [**MP09**] for the relativistic heat equation replacing $W^2_{2,\varepsilon}$ in (26) by

$$W_c = \min_{\pi\in\Gamma(\rho^{i-1},\rho)} \int h\left(\frac{x-y}{\varepsilon}\right) d\pi(x,y) \qquad \text{where } c(x,y) = h(x-y) \text{ is the relativistic heat cost.}$$

More generally, we say that $c(x,y) = h(x-y)$ is a *relativistic cost* if $h : \mathbb{R}^N \to [0,+\infty]$ is a strictly convex function on $h^{-1}([0,+\infty))$, furthermore we assume that $\mathcal{C} := h^{-1}([0,+\infty))$ is the closure of a strictly convex and bounded open set on which h is bounded. Besides, we require that $h(0) = 0$. For simplicity, we also assume that h is differentiable[2] on $h^{-1}([0,+\infty))$ and that $|\nabla h(x)| \to +\infty$ when $x \to \partial\mathcal{C}$.

In order to study the Monge-Kantorovich problem relative to a relativistic cost function c, it is convenient to introduce an additional parameter called the speed (of light) which takes into account the relativistic behaviour of c. We set

$$c_t(x,y) = h\left(\frac{x-y}{t}\right)$$

for any positive number t.

Now, given two probability measures μ_0 and μ_1 with compact support and a relativistic cost function, we can always change the speed of light t so that the minimum in the Monge-Kantorovich problem relative to c_t is finite. The study of the minimum variation in terms of the speed also gives us some useful information. Thus, we define the total cost function

$$C(t) = \min_{\pi\in\Gamma(\mu_0,\mu_1)} \int_{\mathbb{R}^{2d}} c\left(\frac{x-y}{t}\right) d\pi(x,y).$$

The overall idea is that the total cost function is infinite when t is small (at least if $\mu_0 \neq \mu_1$), whereas for large t, the transportation plans are not affected by the discontinuity of the cost function and we get, as in the standard Brenier theorem, existence and uniqueness of the optimal transport map. More precisely, it can be proved that there exists a threshold T from which the total cost is finite:

There exists a speed T such that for any $t \ge T$, $C(t) < +\infty$ whereas $C(t) = +\infty$ otherwise.

We call T the critical speed.

[2]Differentiability is not necessary for our purpose, however it gives a convenient way to express the fact that the slope of h blows up on the boundary of $\mathcal{C}$. The later property is necessary in some of our results.

For $t \geq T$, we prove the uniqueness of the optimal plan moreover this plan is induced by a measurable map T **[9]**. To proceed, we adapt a method due to Champion, De Pascale, and Juutinen [**CDPJ08**] to our setting. However, the study of the optimal map regularity cannot be obtain by this approach. As a first step in this direction we prove, for supercritical speed, the existence of a Kantorovitch potential **[8]**.

THEOREM 3.1. *Let μ_0 and μ_1 be two probability measures with compact support on $\mathbb{R}^N$, c_t be a relativistic cost function, and assume that μ_0 is absolutely continuous with respect to the Lebesgue measure and has a connected support. Then, for any supercritical speed $t > T$ there exists a Kantorovich potential φ_t. Especially, the optimal transport map F_t is defined μ_0-almost everywhere as*

$$F_t(x) = x - t\nabla h^*\big(\widetilde{\nabla}\varphi_t(x)\big)$$

where $\widetilde{\nabla}\varphi_t$ is the approximate gradient of φ_t.

Roughly speaking, the first step of the proof is to show that no point is moved at maximal distance whenever the speed is supercritical.

THEOREM 3.2. *Under the assumptions of Theorem 3.1, let γ be an optimal plan with respect to c_t. Then*

$$\gamma\Big(\Big\{(x,y) \in \mathbb{R}^N \times \mathbb{R}^N : x - y \in t\,\partial\mathcal{C}\Big\}\Big) = 0$$

Combining these results with the lemma below, a Kantorovitch potential can be built following the method sketched in the previous section.

LEMMA 3.3. *Under the assumptions of Theorem 3.1, we set γ_t the optimal transport plan for the cost c_t, $\Gamma \subset \{(x,y) \in \mathbb{R}^{2n}; y - x \in t\overset{\circ}{\mathcal{C}}\}$ a c-cyclically monotone set on which γ_t is concentrated, and let $(x_0, y_0) \in \Gamma$. Then, for every $(x,y) \in \Gamma$ there exists a finite sequence of points $(x_i, y_i) \in \operatorname{supp}\gamma_t$, $1 \leq i \leq k$, such that $(x_k, y_k) = (x,y)$ and for every $0 \leq i < k$ one has $x_{i+1} - y_i \in t\mathcal{C}$.*

Before we give a proof of this lemma, we would like to point out that part of the proof of Theorem 3.2 is also based on an abstract chain lemma for measures. We need to introduce some notation in order to state the lemma. Given $\gamma \in \Pi(\mu,\nu)$, a measure $\alpha = d_\alpha\mu$, and a measure $\beta = d^\beta\nu$, we define the measures

$$\overrightarrow{\Phi}(\alpha) := (p_y)_\sharp(d_\alpha\gamma)$$
$$\overleftarrow{\Phi}(\beta) := (p_x)_\sharp(d^\beta\gamma).$$

These measures describe how a portion of mass of μ (resp. ν) distributed according to α (resp. β) is transported by γ. Given two measures θ and θ', we say that $\theta \leq \theta'$ if $\theta = d_\theta\theta'$ where $0 \leq d_\theta \leq 1$. The statement of the lemma is the following.

LEMMA 3.4 (finite chain lemma). *Given γ, $\gamma' \in \Gamma(\mu,\nu)$, and $\gamma_0 \leq \gamma$ a non-zero measure, we define $\mu_0 = (p_x)_\sharp\gamma_0 \leq \mu$ and $\nu_0 = (p_y)_\sharp\gamma_0 \leq \nu$ the marginals of γ_0. Define recursively the measures*

$$\nu_{i+1} = \overrightarrow{\Phi}'(\mu_i)\,, \qquad\qquad \mu_{i+1} = \overleftarrow{\Phi}(\nu_{i+1})\,,$$

where $\overleftarrow{\Phi}$ and $\overrightarrow{\Phi}'$ are the functions defined above relative to γ and γ' respectively. Then, there exists $j > 0$ such that $\nu_j \wedge \nu_0 > 0$.

We end this part with a proof of Lemma 3.3 which is more geometrical than the one given in **[8]**.

PROOF. As recalled above, there exists a c-cyclically monotone set Γ on which γ_t is concentrated. Moreover, according to Theorem 3.2, we can further assume that $\Gamma \subset \{(x,y) \in \mathbb{R}^{2n}; x - y \in t\overset{\circ}{\mathcal{C}}\}$. Therefore, it suffices to prove the result for $(x,y) \in \Gamma$ and $x - y \in (t-\varepsilon)\mathcal{C}$ for $\varepsilon > 0$ small. Let us introduce some notations. Up to slightly shrink $\varepsilon > 0$, fix $T < t' < t$ such that $\min(t'-T, t-t') > \varepsilon$.

The overall idea is to discretize the target measure and reduce the proof to the case where ν is a finite measure. This is possible thanks to the supercritical regime assumption.

Let $\widetilde{y}_0, \cdots, \widetilde{y}_m \in \operatorname{supp}\nu$ be such that $\cup_{i=0}^m B(\widetilde{y}_i, \widetilde{\varepsilon}) \supset \operatorname{supp}\nu$ with $\widetilde{\varepsilon} = r_m\,\varepsilon/2$, where $r_m > 0$ is such that $B(0, r_m) \subset \mathcal{C}$. We assume this cover to be minimal in the sense that any strict subcollection of $(B(\widetilde{y}_i, \widetilde{\varepsilon}))_{0\le i\le m}$ is not a cover anymore. Now, consider the partition

$$\widetilde{B}(\widetilde{y}_i, \widetilde{\varepsilon}) = B(\widetilde{y}_i, \widetilde{\varepsilon}) \setminus \bigcup_{s=0}^{i-1} B(\widetilde{y}_s, \widetilde{\varepsilon})$$

for $i \in \{0, \cdots, m\}$ and let us assume that $y_0 \in \widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon})$. Define

$$\nu_d = \sum_{i=0}^m \nu(\widetilde{B}(\widetilde{y}_i, \widetilde{\varepsilon}))\delta_{\widetilde{y}_i}.$$

Note that the critical speed T_d for the mass transport problem involving μ, ν_d and the cost c is at most $T + \varepsilon/2$. Indeed, denoting by p the map defined by

$$\begin{array}{rccll} p: & \operatorname{supp}\nu & \longrightarrow & \operatorname{supp}\nu_d & \\ & y & \longmapsto & \widetilde{y}_i & \quad \text{if } y \in \widetilde{B}(\widetilde{y}_i, \widetilde{\varepsilon}), \end{array}$$

it is then easy to check that $\gamma = (Id, p)_\sharp \gamma_T \in \Gamma(\mu, \nu_d)$ and $\int c_{T+\varepsilon/2}\, d\gamma < +\infty$. Using any transport plan with finite c_{T_d}-cost, the property $T_d \le T + \varepsilon/2$ yields for any closed set A,

$$\nu_d(A) \le \mu(A + (T + \varepsilon/2)\mathcal{C}). \tag{27}$$

For simplicity, we set $\widetilde{T} = T + \varepsilon/2$. First, consider the case where

$$(\{\widetilde{y}_0\} + \widetilde{T}\mathcal{C})^c \cap \operatorname{supp}\, \mu = \emptyset \tag{28}$$

which means that for all $x \in \operatorname{supp}\mu$, $x - \widetilde{y}_0 \in \widetilde{T}\mathcal{C}$, thus $x - y_0 \in t\mathcal{C}$, and the proof is complete in this case. For the rest of the proof, we assume

$$(\{\widetilde{y}_0\} + \widetilde{T}\mathcal{C})^c \cap \operatorname{supp}\, \mu \ne \emptyset. \tag{29}$$

We claim there exists y_{i_1} such that

$$\begin{aligned} (x_{i_1}, y_{i_1}) \in \Gamma \text{ with } p(y_{i_1}) = \widetilde{y}_{i_1} \ne \widetilde{y}_0, \\ x_{i_1} - \widetilde{y}_0 \in t'\mathcal{C}. \end{aligned} \tag{30}$$

Again, this yields $x_{i_1} - y_0 \in t\mathcal{C}$. To prove the claim, we first consider the case where

$$\mu((\{\widetilde{y}_0\} + \widetilde{T}\mathcal{C}) \setminus (\Gamma_0^t)^{-1}(\widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon})) > 0$$

where $(\Gamma_0^t)^{-1}(\widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon})) := \{x; \exists\, y \in \widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon}); (x, y) \in \Gamma\}$. Using $\mu(p_x(\Gamma)) = 1$, we get the result in this case. Let us now assume that

$$\mu((\{\widetilde{y}_0\} + \widetilde{T}\mathcal{C}) \setminus (\Gamma_0^t)^{-1}(\widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon})) = 0.$$

As a consequence, we claim that

$$\mu((\Gamma_0^t)^{-1}(\widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon})) \cap (\{\widetilde{y}_0\} - \widetilde{T}\mathcal{C})^c) = 0. \tag{31}$$

Let us prove this. First, we have

$$\mu((\Gamma_0^t)^{-1}(\widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon})) = \nu(\widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon})) = \nu_d(\{\widetilde{y}_0\})$$

since γ_t is induced by a map; besides (27) applied with $A = \{\widetilde{y}_0\}$ gives $\nu_d(\{\widetilde{y}_0\}) \le \mu(\{\widetilde{y}_0\} + \widetilde{T}\mathcal{C})$. Finally, we use the fact that $\mu(A \setminus B) = 0$ and $\mu(B) \le \mu(A)$ implies $\mu(A^c \cap B) = 0$ to get (31).

Therefore, according to (29) and the fact that $(\{\widetilde{y}_0\} + \widetilde{T}\mathcal{C})^c$ is an open set, we can find $\widetilde{T} < t" < t'$ close to $\widetilde{T}$ such that $(\{\widetilde{y}_0\} + t"\mathcal{C})^c \cap \operatorname{supp}\mu \ne \emptyset$. Then, using the connectedness of $\operatorname{supp}\mu$, we obtain

$$((\{\widetilde{y}_0\} + t"\mathcal{C}) \cap \operatorname{supp}\mu) \bigcap \overline{((\{\widetilde{y}_0\} - t"\mathcal{C})^c \cap \operatorname{supp}\mu)} \ne \emptyset$$

where none of the subsets is empty (indeed $(\{\widetilde{y}_0\} + t"\mathcal{C}) \cap \operatorname{supp}\mu = \emptyset$ would imply $\nu(B(\widetilde{y}_0, \widetilde{\varepsilon})) = 0$ by (27), contradicting $\widetilde{y}_0 \in \operatorname{supp}\nu$). As a consequence, we can find z and $\alpha > 0$ small such that

$$B(z,\alpha) \subset (\{\widetilde{y}_0\} + \widetilde{T}\mathcal{C})^c \cap (\{\widetilde{y}_0\} + t'\mathcal{C})$$

and $\mu(B(z,\alpha)) > 0$. Using (31) and $\mu(p_x(\Gamma)) = 1$, we have proved claim (30) in both cases.

Repeating the whole argument with $A = \{\widetilde{y}_0, \widetilde{y}_{i_1}\}$ instead of $\{\widetilde{y}_0\}$ and $\widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon}) \cup \widetilde{B}(\widetilde{y}_{i_1}, \widetilde{\varepsilon})$ instead of $\widetilde{B}(\widetilde{y}_0, \widetilde{\varepsilon})$ and so on, yields after a finite number of steps: 1) the full support of ν_d can be reached by a finite chain or 2) $\operatorname{supp}\mu \subset \{\widetilde{y}_0, \cdots, \widetilde{y}_{i_s}\} + \widetilde{T}\mathcal{C}$ (as in (28) for instance). Note that, in both cases, the graph obtained by adding an edge between y_{i_j} and $y_{i_{j+1}}$ is a tree; use this tree to produce a chain from x_0 to an arbitrary $x \in \operatorname{supp}\mu$. The argument is complete if 2) occurs. To conclude in the case 1), we notice that $p(y) = \widetilde{y}_k$ can be reached by a finite sequence of (x_i, y_i) as above and that the assumption $x - y \in (t-\varepsilon)\mathcal{C}$ yields the missing estimate $x - y_k \in t\mathcal{C}$. □

4. A hyperbolic analogue

4.1. The Minkowski spacetime and discrete subgroups of hyperbolic isometries. In this part, we recall results in Lorentzian geometry in connection with hyperbolic manifolds. We refer to the books **[O'N83, Rat06]** for more on Lorentzian geometry.

The Minkowski spacetime in $m+1$ dimensions, denoted by $\mathbb{R}^m_1$, is $\mathbb{R}^{m+1}$ endowed with the quadratic form

$$q(x) = \sum_{i=1}^{m} x_i^2 - x_{m+1}^2.$$

For simplicity, we use the same notation for the associated bilinear form. This quadratic form is non-degenerate and admits isotropic vectors. The isotropic cone is divided into the future cone and the past cone. We will only use the future cone defined as

$$C_f = \{z \in \mathbb{R}^m_1;\ q(z) < 0 \text{ and } z_{m+1} > 0\}.$$

Throughout this part, we will identify the hyperbolic space $\mathbb{H}^m$ with the following subset of C_f:

$$\{x \in \mathbb{R}^m_1; q(x) = -1 \text{ and } x_{m+1} > 0\},$$

namely the unit sphere of C_f relative to q. More generally, we denote by

$$S(r) = \left\{x \in \mathbb{R}^m_1; q(x) = -r^2 \text{ and } x_{m+1} > 0\right\}$$

the sphere of radius r. Besides, let us also recall that the distance d induced by the Riemannian metric of $\mathbb{H}^m$ is related to q in the following way:

$$\forall x, n \in \mathbb{H}^m \quad q(n,x) = -\cosh(d(n,x)). \tag{32}$$

The above formula emphasizes the relation between the isometry group of the quadratic form q and that of the Riemannian manifold $\mathbb{H}^m$. More precisely, denoting by $Isom(\mathbb{R}^m_1)$ the group of isometries of the Minkowski spacetime, the Riemannian isometry group of $\mathbb{H}^m$ is homeomorphic to $Isom^+(\mathbb{R}^m_1)$, the subset of $Isom(\mathbb{R}^m_1)$ whose elements preserve the future cone C_f. Therefore, being a *compact hyperbolic manifold* covered by the hyperbolic space, it can be identified with $\mathbb{H}^m/\Gamma$ where Γ is a discrete, torsion-free, cocompact subgroup of $Isom^+(\mathbb{R}^m_1)$. Note that this identification depends on a *representation* ρ of $\Pi_1(M)$ into $Isom^+(\mathbb{R}^m_1)$. However, throughout this part, we assume that both the Riemannian manifold (M,g) and ρ are given once for all.

For completeness, let us recall the following definitions.

Definition 4.1 (Discrete, cocompact, and torsion-free subgroups of $Isom(\mathbb{H}^m)$). *Let Γ be a subgroup of $Isom(\mathbb{H}^m)$. Γ is said to be*

- *discrete if the induced topology (by that of $Isom(R^m_1)$) is the discrete topology. In this setting, it can be shown that this is equivalent to require that Γ is discontinuous.*

- *discontinuous if for any compact subset $K \subset \mathbb{H}^m$, the set $K \cap gK$ is nonempty only for finitely many $g \in \Gamma$.*
- *torsion-free if for each $x \in \mathbb{H}^m$, the stabilizer $\{g \in \Gamma; g.\, x = x\} = \{1\}$ is trivial.*
- *cocompact if $\mathbb{H}^m/\Gamma$ endowed with the quotient topology is a compact space.*

Beside compact hyperbolic manifolds, we will also consider more singular objects called *hyperbolic orbifolds*. Such a space is defined as a quotient $\mathbb{H}^m/\Gamma$ where Γ is a discrete, cocompact subgroup of $Isom^+(\mathbb{R}^m_1)$. This is not a smooth manifold in general; however, this is also a commonly studied metric space. The distance on such a space is defined for all $[x], [y] \in \mathbb{H}^m/\Gamma$ by

$$d_\Gamma([x],[y]) := \inf_{\gamma \in \Gamma} d(x, \gamma.\, y)$$

and coincides with the Riemannian distance when $\mathbb{H}^m/\Gamma$ is a smooth manifold.

A hyperbolic orbifold is known to be an Alexandrov space of curvature at least -1, see Chapter 1. As a consequence, we will see that hyperbolic orbifold naturally fits in our setting. In some books, the definition of hyperbolic orbifold differs from ours, this is for instance the case in [**Rat06**]. However, both definitions are actually equivalent as explained in [**Rat06**, Chapter13].

Various statements in the next section involve *equivariant* maps. We recall that, given Γ a subgroup of $Isom^+(\mathbb{R}^m_1)$ and A, B two subsets of the future cone C_f such that $\Gamma.A = A$, a map $f : A \to B$ is said to be Γ-equivariant (or just equivariant if there is no ambiguity) if for all $x \in A$, $\gamma \in \Gamma$, $f(\gamma.x) = \gamma.f(x)$.

The last tool we need is the existence of a proper *fundamental domain* relative to Γ as above. Namely, there exists an open convex set D in $\mathbb{H}^m$ such that

(1) the elements of $\{g.\, D; g \in \Gamma\}$ are pairwise disjoint
(2) $\mathbb{H}^m = \cup_{g\in\Gamma}\, g.\, \overline{D}$.
(3) $\sigma(\cup_{g\in\Gamma}\, g.\partial D) = 0$.

The last property allows us to lift an arbitrary Lipschitz map defined on $\mathbb{H}^m/\Gamma$ to a Γ-equivariant Lipschitz map defined on $\mathbb{H}^m$. The fundamental domain also permits us to build a measure $\sigma_{\mathbb{H}^m/\Gamma}$ on $\mathbb{H}^m/\Gamma$ defined by $\sigma_{\mathbb{H}^m/\Gamma}([U]) := \sigma((\Gamma \cdot U) \cap D)$ for any Borel set $[U] \subset \mathbb{H}^m/\Gamma$. When $\mathbb{H}^m/\Gamma$ is a smooth manifold, $\sigma_{\mathbb{H}^m/\Gamma}$ coincides with the canonical Riemannian measure on $\mathbb{H}^m/\Gamma$.

4.2. Looking for a hyperbolic analogue of Alexandrov's theorem: Fuchsian convex body. Our aim is to prove an analogue of Alexandrov's theorem for hyperbolic orbifolds. Here, we consider Alexandrov's result as an embedding result for manifolds homeomorphic to the sphere as in (13), with a given Gauss curvature measure defined on $\mathbb{S}^m$. For hyperbolic manifolds, the Gauss formula prevents from embedding such a manifold as the boundary of Euclidean convex set. The first step is thus to determine a suitable notion of Gauss curvature measure defined on the hyperbolic manifold (M, g). More generally, our construction applies to any compact hyperbolic orbifold.

Our starting point is a paper by Labourie and Schlenker [**LS00**] in which they prove (the full statement is more general than the one cited below):

Theorem 4.2. *Let (S, g) be a closed Riemannian surface of genus at least 2 where g is a negatively curved Riemannian metric. Then, there exists a unique equivariant isometric embedding Φ of the hyperbolic plane into $\mathbb{R}^3_1$ (modulo an element of $Isom(\mathbb{R}^3_1)$) (and a unique representation of the fundamental group $\pi_1(S)$ into $Isom(\mathbb{R}^3_1)$). Moreover, the image $\Phi(\mathbb{H}^2)$ is contained in C_f.*

In order to relate this result to our own theorem below, let us add some comments on the above statement. First, recall that g is conformal to a (unique) hyperbolic metric g_0. Second, it is true in this setting that the Gauss curvature of $\Phi(\mathbb{H}^2)$ is $-K_g$, being K_g the sectional curvature of g. Especially, this implies that $\Phi(\mathbb{H}^2)$ bounds a Euclidean convex set. Therefore, this result can be seen as an (equivariant) embedding result of (the universal cover of) a compact hyperbolic surface (S, g_0) where the Gauss curvature of the image is prescribed. From this point of view, it is somewhat similar to Alexandrov's theorem.

Let us come back to the question of defining the Gauss curvature measure of an embedded compact hyperbolic orbifold. As recalled above, a compact hyperbolic orbifold can be characterized by the datum of a cocompact discrete subgroup of hyperbolic isometries (recall that the representation of the fundamental group is fixed). We first define a *Fuchsian convex set* as a non-empty (Euclidean) closed convex set which lies on the open future cone C_f and which is invariant under the action of a given cocompact discrete subgroup Γ of $Isom^+(\mathbb{R}^m_1)$, namely

$$\Gamma.\mathcal{C} = \mathcal{C}.$$

It is not easy to figure out what a Fuchsian convex set looks like. A particular case is that of Fuchsian convex polyhedron which was already known before we introduce this general notion of Fuchsian convex set (the paternity of this notion is in a way shared with François Fillastre. I first introduced an equivalent definition which was then simplified by Fillastre in [**Fil13**].).

Given Γ as above, a set F is said to be a Γ-*Fuchsian convex polyhedron* if there exists $x_1, \cdots, x_k \in \mathbb{H}^m$ pairwise non-collinear and $\lambda_1, \cdots, \lambda_k > 0$ such that

$$F = \{z \in C_f; q(z - (1/\lambda_i)\gamma.x_i, \gamma.x_i) \leq 0 \ \ \forall i = 1, \cdots, k, \ \forall \gamma \in \Gamma\}.$$

We concede that even a Fuchsian convex polyhedron is not easy to draw. However, Fuchsian convex sets satisfy properties similar to those satisfied by their Euclidean counterparts. These analogies are our guideline to state a hyperbolic version of Alexandrov's theorem. Here are the main ones, we refer to [**5**] for proofs.

The boundary $\partial\mathcal{C}$ of a Fuchsian convex set $\mathcal{C}$ can be parameterized by the hyperbolic space. The homeomorphism is given by the radial projection onto $\mathbb{H}^m$:

$$\begin{array}{rccc} p: & \partial\mathcal{C} & \longrightarrow & \mathbb{H}^m \\ & x & \longmapsto & \frac{x}{\sqrt{-q(x)}} \end{array}.$$

Besides, given any supporting hyperplane to the Fuchsian convex set (as a Euclidean convex set), its inward unit normal vector (relative to q) belongs to $\mathbb{H}^m$. As a consequence, we get a well-defined multivalued Gauss map $\mathcal{G} : \partial\mathcal{C} \rightrightarrows \mathbb{H}^m$, moreover $\mathcal{G}(\partial\mathcal{C}) = \mathbb{H}^m$.

In order to define the Gauss curvature measure as a mesure on $\mathbb{H}^m/\Gamma$, recall that both p and $\mathcal{G}$ are *equivariant*. Therefore, $\mathcal{G} \circ p^{-1}$ factorizes to a endomorphism $\mathcal{G}_\Gamma \circ p_\Gamma^{-1}$ on $\mathbb{H}^m/\Gamma$ where p_Γ and $\mathcal{G}_\Gamma$ are the maps on the quotient spaces induced by p and $\mathcal{G}$ respectively. As a consequence, we define the *Gauss curvature measure* by the formula

$$\mu(U) = \sigma_{\mathbb{H}^m/\Gamma}(\mathcal{G}_\Gamma \circ p_\Gamma^{-1}(U))$$

where U is an arbitrary Borel subset of $\mathbb{H}^m/\Gamma$.

The same argument as in the Euclidean case guarantees that μ is indeed a measure on $\mathbb{H}^m/\Gamma$. Moreover, being $\mathcal{G}$ a surjective map, the total mass of μ coincides with that of $\sigma_{\mathbb{H}^m/\Gamma}$, namely the volume of M. We can now state the main result of this part.

THEOREM 4.3. *Let $M = \mathbb{H}^m/\Gamma$ be a closed hyperbolic orbifold and μ be a Borel probability measure on M. Then, there exists a unique (up to homotheties) Γ-equivariant homeomorphism $\Phi : \mathbb{H}^m \longrightarrow \mathbb{R}^m_1$ onto its image $\Phi(\mathbb{H}^m)$ which is the boundary of a Γ-Fuchsian convex set and such that $Vol(M)\mu$ is the Gauss curvature measure of M relative to this Fuchsian convex body. Especially, if μ is a finite sum of Dirac masses then the corresponding convex set is an equivariant polyhedron.*

This result generalizes a theorem by Ishkakov [**Isk00**] which proves the result whem $m = 2$ and the underlying convex body is a Fuchsian polyhedron.

The outline of the proof is the same as the one relative to Euclidean convex bodies explained in Section 2. We consider $\theta : \mathbb{H}^m \to (0, +\infty)$ the *radial function* of $\mathcal{C}$ defined by the formula

$$\theta(x) = \sup\{s > 0; sx \in \mathcal{C}^c\}.$$

By analogy, we consider the set of *future starshaped sets* as the closed sets in the future cone whose *complement* is starshaped with respect to 0 and whose radial function (defined as above) is real-valued, continuous and bounded away from 0. Similarly, the *support function* $h : \mathbb{H}^m \to (-\infty, 0)$ of a future starshaped set is defined by

$$h(n) = \sup_{x \in \mathbb{H}^m} \theta(x) q(x, n) \tag{33}$$

(recall that $q(n, x) < 0$ for $n, x \in \mathbb{H}^m$) and

$$h(n) = \theta(x_0) q(x_0, n) \text{ if and only if } n \in \mathcal{G}(\theta(x)x). \tag{34}$$

The hyperbolic analogue of the polar transform is given by

$$S^\circ = \{x \in C_f; \forall n \in S,\, q(x, n) \leq -1\},$$

being S a future starshaped set. Elementary computations guarantee that $S^{\circ\circ} \supset S$ is the convex hull of S.

When S is a Fuchsian convex body, the radial and support functions are Γ-invariant and Lipschitz regular. They induce Lipschitz functions $\theta_{\mathbb{H}^m/\Gamma}$ and $h_{\mathbb{H}^m/\Gamma}$ both defined on $\mathbb{H}^m/\Gamma$. The relations (33) and (34) remain true for $\theta_{\mathbb{H}^m/\Gamma}$ and $h_{\mathbb{H}^m/\Gamma}$ in place of θ and h, provided that $q(n, x)$ is replaced by $-\cosh(d_\Gamma([n], [x])$ (this follows from (32)).

It is then natural to restrict our attention to Γ-equivariant future starshaped sets and the induced quotient maps defined on $\mathbb{H}^m/\Gamma$. Then, the transformations $\phi = \ln(-h_{\mathbb{H}^m/\Gamma})$ and $\psi = -\ln \theta_{\mathbb{H}^m/\Gamma}$ lead us to consider the cost function

$$c([n], [x]) = \ln \circ \cosh(d_\Gamma([x], [n])).$$

As in the Euclidean case, S is a Fuchsian convex body if and only if the underlying maps ϕ and ψ are c-conjugate (and Lipschitz regular). Consequently, to get Theorem 4.3, we are left with proving a strong form of Kantorovitch's duality (i.e. existence and uniqueness of Kantorovitch's variational problem up to an additive constant). This is known to be true when $\mathbb{H}^m/\Gamma$ is a smooth compact manifold as a consequence of McCann's theorem [**McC01**][3]. If not, using that a hyperbolic orbifold is an Alexandrov space of curvature at least -1, we can use our generalization to Alexandrov space of McCann's theorem (see Chapter 2) to conclude.

4.3. Perspective: yet another hyperbolic analogue. Another generalization of Alexandrov's curvature prescription problem would be to extend the result to the convex bodies in the other space forms. In [**Ale05**, p. 395], Alexandrov mentioned the case of convex polyhedra in $\mathbb{H}^m$. As in Euclidean space, the curvature measure is then a finite sum of Dirac masses determined by the vertices and whose weights are given by the exterior normal angles. Some differences worth to be highlighted however. First, the total mass of the curvature measure is not fixed anymore and the curvature measure is no longer invariant by dilations of the convex body. Especially, Alexandrov's proof for general Euclidean convex bodies does not apply in this case, see Section 1. Second, notice that because of the curvature is non-zero, the parallel transport is not trivial anymore; this makes the general definition of curvature measure a bit cumbersome.

In collaboration with Philippe Castillon, we are currently working on this problem. Our approach is based on a non-linear analogue of Kantorovitch's variational problem.

[3]See the beginning of Section 2.2 for more details

CHAPTER 4

Wasserstein space over $CAT(0)$ space

In this chapter, we describe some geometric properties of the Wasserstein space over a $CAT(0)$ space X. A $CAT(0)$ space is a geodesic space which satisfies the Alexandrov nonpositive curvature condition in the *large*, namely for *all* geodesic triangles. We further assume X to be *locally compact* and, in some places, *geodesically complete* which means that any geodesic can be extended to a *geodesic line*, i.e. a geodesic defined on $\mathbb{R}$ (recall that a geodesic in our setting corresponds to a minimizing geodesic when X is a smooth Riemannian manifold). Some results require the space to be negatively curved, we then assume the space to be $CAT(k)$ where $k < 0$. Recall that for $k \leq 0$, a space is $CAT(k)$ if and only if it has curvature bounded from above by k and is simply connected. We refer to Chapter 1 for a review of Alexandrov spaces and optimal mass transport and to the books by Ballmann [**Bal95**] and by Bridson & Haefliger [**BH99**] for all the results on CAT spaces used in this chapter.

The results of this chapter can be summarized as follows. First, we show that, given a $CAT(0)$ space X, its Wasserstein space $\mathscr{W}_2(X)$ admits a *boundary at infinity* which shares various properties with the boundary at infinity of a $CAT(0)$ space (note that up to the case $X = \mathbb{R}$, $\mathscr{W}_2(X)$ is not $CAT(0)$). We then prove that a $CAT(k)$ space ($k < 0$) has rank 1. We finally prove that the isometry group of $\mathscr{W}_2(X)$ coincides with that of X when X is a $CAT(k)$ ($k < 0$) geodesically complete space. We prefer to postpone all the necessary definitions to the core of the chapter since it appears to us that these results are better understood once having in mind the corresponding properties of $\mathscr{W}_2(\mathbb{R}^N)$. These properties are recalled in the first section.

These results come from the papers [**6, 1**] written in collaboration with Benoît Kloeckner.

1. Wasserstein space over Euclidean space

Since the pioneering work of Otto [**Ott01**], it has been known that the Wasserstein space over Euclidean space or, more generally, over a smooth Riemannian manifold, can be considered as a Riemannian manifold of infinite dimension. In this part, we recall some properties of $\mathscr{W}_2(\mathbb{R}^N)$ as a model space for our study of Wasserstein space over $CAT(0)$ space.

The Wasserstein space over $\mathbb{R}^N$ is a geodesic, complete and separable space. From the topological point of view, it is contractible -this remains true for $CAT(0)$ space- and it admits a Euclidean cone structure with respect to any Dirac mass. It is also an infinite dimensional space with nonnegative Alexandrov curvature as a particular instance of a result proved by Sturm [**Stu06a**] asserting that nonnegative curvature is preserved from the base space to its Wasserstein space. In the case of the line, $\mathscr{W}_2(\mathbb{R})$ is even a flat space since it can be isometrically embedded onto a convex subspace of $L^2([0,1])$ using the (right-)inverse of its distribution function. The isometric character of this embedding follows from the classical Hoeffding-Fréchet theorem. In higher dimension, non uniqueness of geodesics between well-chosen measures in $\mathscr{W}_2(\mathbb{R}^N)$ prevents the space to be flat[1]

An interesting feature of the Wasserstein space is that it always contains an isometric copy of the base space (where, to be safe, the base space is assumed to be complete and separable). The

[1]indeed, notice that the comparison triangle determined by a geodesic and the midpoint of another geodesic with the same endpoints is degenerate. If the space were flat, the distance between the two midpoints in $\mathscr{W}_2(\mathbb{R}^N)$ would be 0, a contradiction.

isometric embedding is obtained via the mapping

$$\begin{array}{ccc} X & \longrightarrow & \mathscr{W}_2(X) \\ x & \longmapsto & \delta_x \end{array}$$

Especially, one can isometrically embedd $\mathbb{R}^N$ into $\mathscr{W}_2(\mathbb{R}^N)$. Despite its multiple cone structures, it has been proved by Kloeckner [**Klo10**] that contrary to $\mathbb{R}^N$, $\mathbb{R}^{N+1}$ cannot be isometrically embedded into $\mathscr{W}_2(\mathbb{R}^N)$. Therefore N is the maximal dimension of a Euclidean space that can be isometrically embedded into $\mathscr{W}_2(\mathbb{R}^N)$. The space $\mathscr{W}_2(\mathbb{R}^N)$ is said to have *rank* N. We point out that there are other notions of rank available in the literature, for instance the condition of being isometric can be weakened to biLipschitz. However, most of them are infinite for $\mathscr{W}_2(X)$ whenever X contains a geodesic line. We refer to [**6**, Section 5] for a detailed discussion.

Another interesting result concerns the isometry group of $\mathscr{W}_2(\mathbb{R}^n)$. Once again, the definition of Wasserstein space guarantees that the isometry group of X can be embedded into that of $\mathscr{W}_2(X)$ through the following map (here, X is a merely complete and separable metric space)

$$\begin{array}{ccc} \mathrm{Isom}(X) & \longrightarrow & \mathrm{Isom}(\mathscr{W}_2(X)) \\ \phi & \longmapsto & \phi_\sharp \end{array} \tag{35}$$

and one could expect that both groups coincide. But this is not the case for Euclidean space where there are *non trivial* isometries Φ that fix the set of Dirac masses *pointwise*. We refer the interested reader to [**Klo10**] where the isometry group of $\mathscr{W}_2(\mathbb{R}^n)$ is entirely determined. Let us add the following general fact proved in [**1**]: the set of Dirac masses is *globally* preserved by an isometry (at least if the metric space is locally compact and geodesically complete).

2. Wasserstein space over a $CAT(0)$ space

In what follows, we describe the boundary at infinity of Wasserstein space over a $CAT(0)$ space. Then, we characterize the isometry group of $\mathscr{W}_2(X)$ where X is a $CAT(k)$ space and $k < 0$.

2.1. Boundary at infinity of Wasserstein space.

The *boundary at infinity* of a (non compact) geodesic space X is defined as follows. We consider $\mathcal{R}(X)$ the set of constant speed geodesic rays and $\mathcal{R}_1(X)$ its subset made of unit speed rays. Two geodesic rays $\gamma(t), \sigma(t)$ are said to be *asymptotic* if they are at bounded distance:

$$\gamma \sim \sigma \quad \text{if and only if} \quad \sup_{t\geq 0} d(\gamma(t), \sigma(t)) < +\infty.$$

In particular, note that two asymptotic rays must have the same speed.

The set ∂X of *boundary points* (or *points at infinity*) is the set of equivalence classes of *unit speed* geodesic rays. The equivalence class of a unit speed geodesic ray σ is denoted by σ_∞. The set of equivalence classes $\mathcal{R}(X)/\sim$ of the full set $\mathcal{R}(X)$ can be identified with the cone $Con(\partial X)$ over ∂X, the radial term being given by the speed $s(\gamma)$ of the geodesic ray γ.

When X is a locally compact $CAT(0)$ space, the space $X \cup \partial X$ can be equipped with a *cone topology* for which both ∂X and $X \cup \partial X$ are compact sets, moreover the topology induced on X coincides with the one induced by its metric. This topology is known to be compatible with the topology of uniform convergence on compact sets for continuous curves lying on X; as a consequence, the cone topology actually extends to $Con(\partial X)$.

To perform the construction of the boundary at infinity, the $CAT(0)$ assumption on X is used through the convexity of $t \mapsto d(\gamma(t), \sigma(t))$. The latter property is also useful to introduce a distance on $Con(\partial X)$ defined by

$$d_\infty(\gamma_\infty, \sigma_\infty) := \lim_{t\to+\infty} \frac{d(\gamma(t), \sigma(t))}{t}.$$

The topology induced by d_∞ does not coincide in general with the cone topology and is always finer than the latter (for a $CAT(0)$ space satisfying the *visibility* property, the topology induced by d_∞ is the discrete one; see below for more on the visibility property). An important feature from the point of view of optimal mass transport is that d_∞ is a lower semicontinous mapping with respect to the cone topology.

Given X a locally compact $CAT(0)$ space, our goal is to characterize the boundary at infinity $\partial\mathscr{W}_2(X)$ defined as above in terms of ∂X. As recalled in the relevant chapter, a geodesic (μ_t) in $\mathscr{W}_2(X)$ is induced by a probability measure μ on the set of geodesics in X called a dynamical plan, in the sense that $\mu_t = e_{t\sharp}\mu$ with $e_t(\gamma) = \gamma(t)$. A diagonal argument allows us to extend this property to geodesic ray in $\mathscr{W}_2(X)$ and gives us a dynamical plan μ supported on $\mathcal{R}(X)$ but not on $\mathcal{R}_1(X)$ in general. We set $e_\infty : \mathcal{R}(X) \to Con(\partial X)$ the quotient map, using the identification

$$Con(\partial X) \sim \mathcal{R}(X)/\!\sim . \tag{36}$$

We can then define the *asymptotic measure* of the geodesic ray (μ_t) as

$$\mu_\infty := (e_\infty)_\sharp\mu.$$

Note that when (μ_t) is unit speed, we also have

$$\int s^2(\gamma)\, d\mu_\infty(\gamma) = 1.$$

The set of probability measures on $Con(\partial(X))$ satisfying the above relation is denoted by

$$\mathcal{P}_1(Con(\partial X)).$$

This set reflects the asymptotic behaviour of geodesic rays and is therefore a good candidate to be the boundary of $\mathscr{W}_2(X)$. However, it is important to notice that the asymptotic measure is defined through the dynamical plan μ not the ray itself. This is harmless when the base space X is non branching since, in that case, it can be proved that μ is uniquely defined. But this is no longer true without this assumption. To circumvent this difficulty, a crucial tool is the following formula which, among other things, guarantees that the asymptotic measure depends only on the geodesic ray.

THEOREM 2.1 (asymptotic formula, **[6]**). *Consider two geodesic rays $(\mu_t)_{t\geq 0}$ and $(\sigma_t)_{t\geq 0}$, let μ and σ be any of their dynamical plans and μ_∞, σ_∞ be the corresponding asymptotic measures. Then (μ_t) and (σ_t) are asymptotic if and only if $\mu_\infty = \sigma_\infty$. Moreover we have*

$$\lim_{t\to\infty} \frac{\mathrm{W}(\mu_t,\sigma_t)}{t} = \mathrm{W}_\infty(\mu_\infty,\sigma_\infty)$$

where W_∞ stands for the quadratic Wasserstein space involving the distance d_∞.

The asymptotic formula also allows us to overcome the lack of convexity of $t \mapsto \mathrm{W}(\mu_t,\sigma_t)$ and perform the construction of the cone topology on $\mathscr{W}_2(X) \cup \partial\mathscr{W}_2(X)$. The cone topology retains the properties known in the $CAT(0)$ case except that of compactness. This non-compactness feature is consistent with the non local compactness of $\mathscr{W}_2(X)$ which is known to be true whenever X is non compact. We prove that the spaces below are homeomorphic.

$$\partial\mathscr{W}_2(X) \underset{homeo.}{\sim} \mathcal{P}_1(Con(\partial X))$$

where the latter set is endowed with the topology of weak convergence.

It is known that the cone topology on $Con(\partial X)$ is metrizable. This allows us to rephrase the above result in a more (sym-)metric way using (36):

$$Con(\partial\mathscr{W}_2(X)) \text{ is isometric to } \mathscr{W}_2(Con(\partial X)).$$

In words, the cone over the boundary of Wasserstein space is isometric to the Wasserstein space of the cone over the boundary of the base space. This phenomenon is reminiscent of the one linking geodesic rays and dynamical transport plans.

The boundary $\partial\mathscr{W}_2(X)$ can be used jointly with the asymptotic formula to prove a non embedding result into a $CAT(0)$ space satisfying the *visibility* property. Such a space, called *visibility space*, is a $CAT(0)$ space for which any pair of points in the boundary at infinity can be linked by a geodesic line in X. For example, any $CAT(k)$ space with $k < 0$ satisfies this property. Using the formula expressing d_∞ in terms of the angle metric, it can be proved that, on a visibility space, the range of the metric d_∞ is $\{0, 2\}$.

THEOREM 2.2. *If X is a visibility space, then the Euclidean plane cannot be isometrically embedded into $\mathscr{W}_2(X)$.*

In other terms, the Wasserstein space of a visibility space has rank 1 as the space itself. This result leaves open the question of the rank of the Wasserstein space over a $CAT(0)$ space of rank $k > 1$. This would require a better understanding of $\partial \mathscr{W}_2(X)$.

The scheme of proof is as follows. Given that $d_\infty(\gamma_\infty, \sigma_\infty)$ is 2 as soon as $\gamma_\infty \neq \sigma_\infty$, it is not too difficult to prove that the *restriction* of W_∞ to $\mathcal{P}(\partial X)$ coincides with the *square root* of the total variation norm (the mass stays where it is as much as possible). Especially, the restriction of W_∞ is a "snowflake" metric and cannot contain any rectifiable curve (this follows from the definition of length and the concavity of $x \mapsto x^{1/2}$). On the other hand, if Euclidean plane could be isometrically embedded into $\mathscr{W}_2(X)$ then the asymptotic formula would give us an isometric embedding of the boundary at infinity of $\mathbb{R}^2$ -i.e. the unit circle- into $\partial \mathscr{W}_2(X)$. Finally, some computations only based on the triangle inequality imply that the asymptotic measure of (the restriction to $t \geq 0$ of) a geodesic line μ_t has to be supported in $\partial X = \{1\} \times \partial X \subset Con(\partial X)$. The unit circle being rectifiable, we get a contradiction. Note that the above argument can be used to prove other nonembedding results involving open cones (invariant by $x \mapsto -x$) or plane endowed with an arbitrary norm instead of Euclidean plane.

2.2. Isometry group of $CAT(k)$ space, $k < 0$. The goal of this part is to determine the isometry group of $\mathscr{W}_2(X)$ where X is a locally compact, geodesically complete $CAT(k)$ space. As recalled above, the isometry group $\mathrm{Isom}\,\mathscr{W}_2(X)$ always contains an isometric copy of that of X. We show that for negatively curved spaces, both isometry groups actually coincide. Note that the result of this section is valid under slightly weaker assumption than $CAT(k)$.

In Section 1, it is recalled that under very weak assumptions on X, the set of Dirac masses is always globally preserved by an isometry. Moreover, up to composing this isometry with an isometry coming from the base space (35), we can further assume that the set of Dirac masses is fixed *pointwise*. The proof of the main result then follows from two ingredients. First, we prove that such an isometry fixes *pointwise* any element of $\mathscr{W}_2(X)$ supported on a geodesic line. Second, we introduce the mapping (called "Radon transform" by analogy with the standard Radon transform on Euclidean space)

$$\mathscr{R} : \mu \mapsto ((p_\gamma)_\# \mu)_{\gamma \in \mathcal{R}_\mathbb{R}(X)}$$

which maps $\mu \in \mathscr{W}_2(X)$ onto the collection of all measures $(p_\gamma)_\# \mu$, with p_γ being the projection on the geodesic line γ (γ is a convex set thus the projection onto γ is well-defined since X is $CAT(0)$). We then prove the Radon transform $\mathscr{R}$ is an injective map. As a consequence of these two claims, the isometry we consider has to be the identy map and the proof is complete.

Below, we give elements of the proof of these two facts in order to show how properties of $CAT(k)$ spaces intertwine with optimal mass transport.

We start with the first property. Again, the proof is in two steps. First, we prove that the measures supported on a geodesic line γ are globally preserved by an isometry. Then, using that γ is isometric to $\mathbb{R}$, we combined the description of $\mathrm{Isom}(\mathscr{W}_2(\mathbb{R}))$ obtained by Benoît Kloeckner with the negative curvature assumption to prove that the measures supported on a geodesic line are actually preserved pointwise. The proof of both steps are similar in spirit so we restrict ourself to sketching the first one. The point is to find a way to distinguish between a measure supported on a geodesic line and a measure which is not. The idea is the following: take a measure μ supported on γ and pick up two points $x_0, x_1 \in \gamma$ and set μ_t the geodesic from μ to δ_{x_1} and $\delta_{x(t)}$ the geodesic from δ_{x_0} to δ_{x_1}. Since γ is a convex subset of X isometric to $\mathbb{R}$, the Wasserstein distance between μ_t and $\delta_{x(t)}$ behaves as on $\mathscr{W}_2(\mathbb{R})$: it decreases linearly to 0 when t varies from 0 to 1. On the contrary, if μ is *not* supported on γ, then we can find a set of points z_0 of positive μ-measure such that the points x_0, x_1, z are not aligned; the negative curvature then implies

$$d(z(t), x(t)) < td(z_0, x_0)$$

(where $z(t)$ is the geodesic from z_0 to $z_1 = x_1$). This implies $\mathrm{W}(\mu_t, \delta_{x(t)}) < t\,\mathrm{W}(\mu, \delta_{x_0})$ and the proof is complete.

Now, we sketch the proof that the Radon transform is injective. We first notice that it is sufficient for our purpose to prove a slightly weaker result. Indeed, by continuity, in order to prove that an isometry Φ is the identity map, it suffices to prove $\Phi(\nu) = \nu$ for ν in a dense subset of $\mathscr{W}_2(X)$. Having in mind this remark, it suffices to prove

$$\mathscr{R}\,\mu = \mathscr{R}\,\nu \Rightarrow \mu = \nu$$

for all $\mu \in \mathscr{W}_2(X)$ and all $\nu \in A$, A being a dense subset of $\mathscr{W}_2(X)$.

Let us further assume for a while that X is a smooth Riemannian manifold and define A as the set of atomic measures. Write $\nu = \sum m_i \delta_{x_i}$ where $\sum m_i = 1$. First, we claim that μ must be supported in the (x_i). Let x be another point; consider a geodesic γ such that $\gamma_0 = x$ and its tangent vector $\dot\gamma_0$ is not orthogonal to any of the geodesics (xx_i). Then for all i, $p_\gamma(x_i) \neq x$ and there is an $\varepsilon > 0$ such that the neighborhood of size ε around x in γ does not contain any of these projections. It follows that $\mathscr{R}\,\nu(\gamma)$ is supported outside this neighborhood, and so does $\mathscr{R}\,\mu(\gamma)$. On a $CAT(0)$ space, the projection on γ is 1-Lipschitz, so that μ must be supported outside the ε-ball at x in X. This gives $x \notin \operatorname{supp}\mu$.

Now, if γ is a geodesic containing x_i, then $\mathscr{R}\,\nu(\gamma)$ is finitely supported with a mass at least m_i at x_i. For a generic γ, the mass at x_i is exactly m_i. It follows immediately, since μ is supported on the x_i, that its mass at x_i is m_i.

To deal with the general case, we make use of results by Lytchak and Nagano [**LN**] which describe the local structure of a geodesically complete, locally compact space of curvature bounded from above. We restrict our attention to the spaces with nonpositive curvature. Recall that it is rather easy to glue together two spaces of nonpositive curvature in such a way that the resulting space remains nonpositively curved. Especially, it is possible to glue together (smooth) spaces of various dimensions as an half-line and a Euclidean plane along a point for instance. More generally, one can consider metric polyhedral complexes all of whose faces are (smooth) spaces of constant Hausdorff dimension and $CAT(0)$. A crude description of Lytchak-Nagano results is that *all* locally compact, geodesically complete $CAT(0)$ spaces look like metric polyhedral complexes as above, up to small subsets (where "small" means small relative to the "local" Hausdorff dimension of the space which is proved to exist and to be an integer). More precisely, as a consequence of their results, one has

THEOREM 2.3 (Lytchak-Nagano). *Let U be a relatively compact open subset of X. Then, there exists an integer $n \in \mathbb{N}$ such that*

$$\mathcal{H}^n(U) \in (0, +\infty).$$

Moreover, the Hausdorff dimension of U is related to those of spaces of directions through the formula

$$\dim_{\mathcal{H}}(U) = \sup_{z\in U} \dim_{\mathcal{H}} \Sigma_z + 1.$$

Let $R_n(U) = \{z \in U; \Sigma_z$ is isometric to $\mathbb{S}^{n-1} * Z\}$ where (Z, d) is a metric space and $\mathbb{S}^{n-1} * Z$ is the spherical join of $\mathbb{S}^{n-1}$ and Z. Then, the Hausdorff dimension of $U \setminus R_n(U)$ satisfies

$$\dim_{\mathcal{H}}(U \setminus R_n(U)) \leq n - 1.$$

As a corollary of these results, it is possible to define the *local dimension* of a point x by observing that all balls being relatively compact, the Hausdorff dimension of $B(x, 1/k)$ has to be constant when k is sufficiently large. Let us call $n_x \in \mathbb{N}$ the local dimension of x and assume that $n_x \geq 2$. We can also prove that if $z \in R_{n_x}(B(x, 1/k))$ for large k then Σ_z is isometric to $\mathbb{S}^{n_x-1}$, equivalently the tangent cone at z is isometric to Euclidean space of dimension n_x.

With this result at our disposal and discarding the one dimensional subspace of X which requires more precise informations, one can adapt the argument for smooth $CAT(k)$ spaces to the general case. The suitable dense subset A of $\mathscr{W}_2(X)$ is then made of atomic measures supported on the subset of regular points, namely the points whose tangent cone is isometric to Euclidean space (of any dimension at least 2). As explained above, Theorem 2.3 guarantees that the set of regular points is dense in X which, in turns, gives us the density of A in $\mathscr{W}_2(X)$. The main point of the proof is that angles are well-defined on $CAT(0)$ space and the characterization of $p_\gamma(x)$ in terms of angle retains its validity whenever the argument is applied to a point whose tangent cone is isometric to Euclidean space (of dimension at least 2).

Bibliography

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[AGS08] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[AGS14] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014.

[Ale42] A.D. Aleksandrov. Existence and uniqueness of a convex surface with a given integral curvature. *C. R. (Dokl.) Acad. Sci. URSS, n. Ser.*, 35:131–134, 1942.

[Ale05] A. D. Alexandrov. *Convex polyhedra*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005. Translated from the 1950 Russian edition by N. S. Dairbekov, S. S. Kutateladze and A. B. Sossinsky, With comments and bibliography by V. A. Zalgaller and appendices by L. A. Shor and Yu. A. Volkov.

[Ale06] Alexander D. Alexandrov. *A. D. Alexandrov selected works. Part II*. Chapman & Hall/CRC, Boca Raton, FL, 2006. Intrinsic geometry of convex surfaces, Edited by S. S. Kutateladze, Translated from the Russian by S. Vakhrameyev.

[Bak94a] Ilya J. Bakelman. *Convex analysis and nonlinear geometric elliptic equations*. Springer-Verlag, Berlin, 1994. With an obituary for the author by William Rundell, Edited by Steven D. Taliaferro.

[Bak94b] Dominique Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994.

[Bal95] Werner Ballmann. *Lectures on spaces of nonpositive curvature*, volume 25 of *DMV Seminar*. Birkhäuser Verlag, Basel, 1995. With an appendix by Misha Brin.

[BBI01] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[Ber94] V. N. Berestovskiĭ. Manifolds with an intrinsic metric with one-sided bounded curvature in the sense of A. D. Aleksandrov. *Mat. Fiz. Anal. Geom.*, 1(1):41–59, 1994.

[BGP92] Yuri Burago, Misha Gromov, and Grigori Perel′man. A. D. Aleksandrov spaces with curvatures bounded below. *Uspekhi Mat. Nauk*, 47(2(284)):3–51, 222, 1992.

[BH99] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[Bre91] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[Bre03] Yann Brenier. Extended Monge-Kantorovich theory. In *Optimal transportation and applications (Martina Franca, 2001)*, volume 1813 of *Lecture Notes in Math.*, pages 91–121. Springer, Berlin, 2003.

[BS11] Mathias Beiglböck and Walter Schachermayer. Duality for Borel measurable cost functions. *Trans. Amer. Math. Soc.*, 363(8):4203–4224, 2011.

[Car04] G. Carlier. On a theorem of Alexandrov. *J. Nonlinear Convex Anal.*, 5(1):49–58, 2004.

[CDPJ08] Thierry Champion, Luigi De Pascale, and Petri Juutinen. The ∞-Wasserstein distance: local solutions and existence of optimal transport maps. *SIAM J. Math. Anal.*, 40(1):1–20, 2008.

[CEMS01] Dario Cordero-Erausquin, Robert J. McCann, and Michael Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2):219–257, 2001.

[Fil13] François Fillastre. Fuchsian convex bodies: basics of Brunn-Minkowski theory. *Geom. Funct. Anal.*, 23(1):295–333, 2013.

[geo93] *Geometry. IV*, volume 70 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 1993. Nonregular Riemannian geometry, A translation of ıt Geometry, 4 (Russian), Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1989 [MR1099201 (91k:53003)], Translation by E. Primrose.

[Gig11] Nicola Gigli. On the inverse implication of Brenier-McCann theorems and the structure of $(P_2(M), W_2)$. *Methods Appl. Anal.*, 18(2):127–158, 2011.

[GM96] Wilfrid Gangbo and Robert J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.

[Isk00] I. Iskhakov. On hyperbolic surfaces tesselations and equivariant spacelike convex polyhedra. *PhD thesis, Ohio State University*, 2000.

[Klo10] Benoît Kloeckner. A geometric study of Wasserstein spaces: Euclidean spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 9(2):297–323, 2010.

[LN] A. Lytchak and K. Nagano. Geodesically complete spaces with an upper curvature bound. *Preprint*.

[LS00] François Labourie and Jean-Marc Schlenker. Surfaces convexes fuchsiennes dans les espaces lorentziens à courbure constante. *Math. Ann.*, 316:465–483, 2000.

[LV09] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

[McC95] Robert J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80(2):309–323, 1995.

[McC01] Robert J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[MMU14] Carlo Mantegazza, Giovanni Mascellani, and Gennady Uraltsev. On the distributional Hessian of the distance function. *Pacific J. Math.*, 270(1):151–166, 2014.

[MP09] Robert J. McCann and Marjolaine Puel. Constructing a relativistic heat flow by transport time steps. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(6):2539–2580, 2009.

[MRar] Robert J. McCann and Ludovic Rifford. The intrinsic dynamics of optimal transport. *Journal de l'École Polytechnique - Mathematiques*, To appear.

[Oli07] Vladimir Oliker. Embedding $\mathbf{S}^n$ into $\mathbf{R}^{n+1}$ with given integral Gauss curvature and optimal mass transport on $\mathbf{S}^n$. *Adv. Math.*, 213(2):600–620, 2007.

[O'N83] Barrett O'Neill. *Semi-Riemannian geometry*, volume 103 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1983. With applications to relativity.

[OS94] Yukio Otsu and Takashi Shioya. The Riemannian structure of Alexandrov spaces. *J. Differential Geom.*, 39(3):629–658, 1994.

[Ott01] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[Per94] Grigori Perel′man. DC Strucutre on Alexandrov Space. *Preprint*, 1994.

[Pet98] Anton Petrunin. Parallel transportation for Alexandrov space with curvature bounded below. *Geom. Funct. Anal.*, 8(1):123–148, 1998.

[Pra08] Aldo Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.*, 258(3):677–690, 2008.

[Rat06] John G. Ratcliffe. *Foundations of hyperbolic manifolds*, volume 149 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2006.

[Sch93] Rolf Schneider. *Convex bodies: the Brunn-Minkowski theory*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.

[ST09] Walter Schachermayer and Josef Teichmann. Characterization of optimal transport plans for the Monge-Kantorovich problem. *Proc. Amer. Math. Soc.*, 137(2):519–529, 2009.

[Stu06a] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[Stu06b] Karl-Theodor Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.

[Tro] Marc Troyanov. Les surfaces à courbure intégrale bornée au sens d'Alexandrov. *available on arXiv:0906.3407v1.*

[Tro91] M. Troyanov. Un principe de concentration-compacité pour les suites de surfaces riemanniennes. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 8(5):419–441, 1991.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vil09] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[vRS05] Max-K. von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

[Zaj79] Luděk Zajíček. On the differentiation of convex functions in finite and infinite dimensional spaces. *Czechoslovak Math. J.*, 29(104)(3):340–348, 1979.