The metric geometry nature of Wasserstein spaces of probability measures: On the point of view of submetry projections

Mehmet Yenisey

Submitted to the graduate degree program in the Department of Mathematics and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Jin Feng, Chairperson

David Nualart

Committee members

Yunfeng Jiang

Zhipeng Liu

Huixuan Wu

Date defended: June 6, 2022

The Dissertation Committee for Mehmet Yenisey certifies
that this is the approved version of the following dissertation :

The metric geometry nature of
Wasserstein spaces of probability measures:
On the point of view of submetry projections

Jin Feng, Chairperson

Date approved: June 6, 2022

Abstract

The Wasserstein spaces of probability measures are central objects in optimal transport theory and in probability theory. We formulate a theorem regarding the metric geometry origins of the Wasserstein spaces:

Motivated by certain hidden symmetry considerations arising in the study of Hamilton-Jacobi equations in spaces of probability measures that are associated to a mathematical reinterpretation of certain physical models, for $p \in [1,\infty)$ and a separable Hilbert space E, we set forth to rigorously represent the Wasserstein space $(\mathscr{P}_p(E), W_p)$ as a so-called metric quotient space induced by metric foliation.

Namely, for the following sequence of Lebesgue spaces,

$$L_p^{(n)}(E) := L_p([0,1]^n, \mathscr{B}([0,1]^n), \otimes_{i=1}^n m;\, E), \qquad n \in \mathbb{N},$$

where m denotes Lebesgue measure on $([0,1], \mathscr{B}([0,1]))$, we show that the Wasserstein space $(\mathscr{P}_p(E), W_p)$ is isometrically isomorphic to a metric quotient space induced by a metric foliation of the direct limit metric space, $\varinjlim L_p^{(n)}(E) =: \mathscr{L}_p(E)$, which is defined as a direct limit in a category of metric spaces.

Thereby, we show that the following explicitly described well-defined mapping is a so-called submetry projection from the latter metric space onto the Wasserstein space of probability measures:

$$\Psi : \mathscr{L}_p(E) \to \mathscr{P}_p(E),$$
$$[X] \mapsto \mu_X,$$

where $[X]$ is an equivalence class of random variables in the disjoint union set $\bigsqcup_{n\in\mathbb{N}} L_p^{(n)}(E)$, and μ_X denotes the distribution of any representative X.

The proof uses standard techniques from metric geometry and probability theory. In particular, we rely on certain representations of probability measures as the distribution of a random variable and a lemma about couplings.

As the submetry property, which defines the notion of submetry, ties together the metric geometry structures of the spaces in the domain and the range of a submetry, the mapping Ψ may be used to translate known results set in its domain to prove analogues in the range, and vice versa, to lift a problem formulation set in the Wasserstein space $(\mathscr{P}_p(E), W_p)$, e.g. to study it using Hilbert space techniques in the case $p = 2$.

Namely, the space $\mathscr{L}_2(E)$ that we construct is pre-Hilbert and satisfies certain synthetic lower curvature bound conditions that define the notion of Alexandrov non-negatively curved metric space. We show that so-called Kirszbraun synthetic lower curvature bound conditions are stable under submetries, and altogether are able to exhibit the Wasserstein space $(\mathscr{P}_2(E), W_2)$ as an Alexandrov non-negatively curved metric space explicitly as a metric quotient space of a pre-Hilbert space.

It has already been proved that synthetic lower curvature bound conditions that define Alexandrov curvature bounded below spaces are stable under submetries in Alexander et al. (2022), and that $(\mathscr{P}_2(E), W_2)$ is an Alexandrov non-negatively curved metric space by direct means in Ambrosio et al. (2008). We connect these two results.

We also formulate an analogue of our theorem for the space of all probability measures equipped with the Prokhorov metric by using a relation between Ky Fan and the Prokhorov metrics, which are presented in an appendix.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my Ph.D. thesis advisor Professor Jin Feng for the continued great support, guidance, teachings, inspiration and the immense help that made this work possible. I have been greatly supported by Professor Feng in every stage of my studies and life in Lawrence. It is a great honor for me to be his Ph.D. student.

It is also a great honor for me that Professors David Nualart, Yunfeng Jiang, Zhipeng Liu and Huixuan Wu have been members in my exam committees. I would like to thank them for their support.

I would like to thank the faculty, staff and friends at the University of Kansas for their great support.

I would like to thank my friends and teachers in Türkiye, especially my master's thesis advisor Professor Atilla Yılmaz, who greatly supported me.

I would like to thank longtime family friends for greatly supporting me in Germany and in the United States.

Finally, I would like to thank my whole family for the unconditional great support: I would like to additionally thank my brother Ahmet for his guidance, my aunt Melek for her great support in my days at Boğaziçi Üniversitesi, my nephew Artun for composing nice tunes that I could listen to while finalizing this thesis, and to my wife Şefika for her kindness and the additional great support as a colleague.

Lawrence, June 2022

To my parents Feridun and Dilek, my grandparents, and my family...

Contents

List of Figures

Notational conventions

Unless specified otherwise within the main part of the text, we shall use the following notational conventions.

We set $\mathbb{R}_{>0} := \{x \in \mathbb{R},\ x > 0\}$ and $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R},\ x \geq 0\}$.

For $\kappa \in \mathbb{R}$ we denote by $(\mathbb{M}_\kappa, d_\kappa)$ the so-called model spaces of Alexandrov geometry.

For a topological space (E,τ), we denote by $\mathscr{B}(E)$ the Borel σ-algebra generated by the open sets in E.

We use m to denote Lebesgue measure on $([0,1],\mathscr{B}([0,1]))$ throughout the work.

When working with metric spaces, we shall reserve the symbol d to denote the metrics. E.g. we may write d_X for the metric in a metric space $X = (X, d_X)$, or, write d, d', etc. to distinguish between different metrics.

For a measurable mapping $\Phi : S \to T$ between two measurable spaces $(S,\mathscr{G})$, $(T,\mathscr{H})$, and a measure μ on $(S,\mathscr{G})$, we write $\Phi_\# \mu$ to denote the push-forward of μ under Φ, which is a measure on $(T,\mathscr{H})$.

When (E,d) is a complete separable metric space, $\mathscr{P}(E)$ denotes the set of all probability measures on $(E,\mathscr{B}(E))$, and for $p \in [1,\infty)$, the corresponding Wasserstein space of probability measures on $(E,\mathscr{B}(E))$ shall be denoted as $(\mathscr{P}_p(E), W_p)$.

For $n \in \mathbb{N}$, $p \in [1,\infty)$, and a separable Hilbert space E, the notation $L_p^{(n)}(E)$ shall stand for the Lebesgue space

$$L_p([0,1]^n, \mathscr{B}([0,1]^n), \otimes_{i=1}^n m;\ E)$$

constructed over the probability space $([0,1]^n, \mathscr{B}([0,1]^n), \otimes_{i=1}^n m)$ with state space $(E,\mathscr{B}(E))$.

The distribution of a random variable X may be denoted by μ_X. We also write $X \overset{\mathscr{D}}{=} \mu$ to indicate that the distribution of X should be μ. If X and Y are two random variables, $X \overset{\mathscr{D}}{=} Y$ shall mean that X and Y have the same distribution.

Abbreviations

LDP: Large deviations principle

CBB: Curvature bounded from below

Chapter 1

Introduction

1.1 Wasserstein spaces and quotients of metric spaces

Let $p \in [1,\infty)$ and (E,d) be a complete separable metric space. Consider the following set of probability measures on $(E,\mathscr{B}(E))$,

$$\mathscr{P}_p(E) := \{\mu \in \mathscr{P}(E),\ \exists\, x_0 \in E : \int_E d^p(x_0,y)\mu(dy) < \infty\}.$$

The so-called *Wasserstein distance* in $\mathscr{P}_p(E)$ is defined, for $\mu,\nu \in \mathscr{P}_p(E)$, by

$$W_p(\mu,\nu) := \inf\{\left(\int_{E\times E} d^p(x,y)\pi(dx,dy)\right)^{\frac{1}{p}};\ \pi \in \mathscr{P}(E\times E),\ p_{1\#}\pi = \mu,\ p_{2\#}\pi = \nu\}, \quad (1.1)$$

where $p_i : E\times E \to E$, $i = 1,2$, denote the canonical coordinate projections, respectively. Then, $(\mathscr{P}_p(E), W_p)$ is a complete separable metric space called *p-Wasserstein space*.

Wasserstein spaces are central objects in optimal transport theory and in probability theory. They have been studied from many perspectives.

Indeed, firstly, the definition of the Wasserstein metric itself is immediately tied to Kantorovich's weak formulation of Monge's so-called optimal transport problem. Namely, the weak version of the optimal transport problem is about clarifying the conditions for the infimum in (1.1) to be attained, and then characterizing the minimizing couplings $\pi \in \mathscr{P}(E\times E)$ there, known as *optimal couplings* or *optimal transport plans*.

For instance, if E is locally compact, for each $\mu,\nu \in \mathscr{P}_p(E)$, there exists an optimal coupling

$\pi_0 \in \mathscr{P}(E \times E)$, $p_{1\#}\pi_0 = \mu$, $p_{2\#}\pi_0 = \nu$ with

$$W_p(\mu,\nu) = \left(\int_{E\times E} d^p(x,y)\pi_0(dx,dy) \right)^{\frac{1}{p}}.$$

There is already an extensive amount of research done regarding the geometry of the Wasserstein spaces in optimal transport theory literature. In particular, a calculus has been developed over these spaces. (We refer to Ambrosio et al. (2008), Villani (2009), and Rachev & Rüschendorf (1998), and to the references therein.)

Secondly, the Wasserstein metrics, and hence the Wasserstein spaces, are also central objects in probability theory, since Wasserstein metrics are often used when quantifying the rate of convergence in central limit theorems, as for any sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ in $\mathscr{P}_p(E)$, the metric space convergence $W_p(\mu_n,\mu) \to 0$ implies that the sequence $(\mu_n)_{n\in\mathbb{N}}$ converges weakly to μ in the sense of probability theory.

Let E be a separable Hilbert space. Probability measures on Polish spaces can be represented as the distribution of a random variable defined on a Borel space equipped with a non-atomic probability measure. Applying this result in the Polish space $(E\times E, \mathscr{B}(E\times E))$, the following representation can be proved for the Wasserstein distance:

$$W_p(\mu,\nu) = \inf\{||X-Y||_p;\ X,Y \in L_p([0,1],\mathscr{B}([0,1]),m;\ E),\ X \overset{\mathscr{D}}{=} \mu,\ Y \overset{\mathscr{D}}{=} \nu\}, \tag{1.2}$$

for all $\mu,\nu \in \mathscr{P}_p(E)$, where m denotes Lebesgue measure on $([0,1],\mathscr{B}([0,1]))$ and

$$L_p([0,1],\mathscr{B}([0,1]),m;\ E)$$

denotes a Lebesgue space, and $X \overset{\mathscr{D}}{=} \mu$ means that X, which is actually an equivalence class of a.s.-equal random variables, should have distribution μ.

We compare the representation in display (1.2) with the following general situation: Let (X,d) be a metric space and $\mathscr{F} := \{F_i\}_{i\in I}$ be a partition of the set X into closed subsets $F_i \subset X$, $i \in I$. We

define

$$d(F_i,F_j) := \inf\{d(x,y);\ x \in F_i,\ y \in F_j\}, \qquad F_i,F_j \in \mathscr{F}.$$

If $(\mathscr{F},d)$ is a metric space on its own, it is a "*quotient space*" of (X,d) in a naive sense. So, $(\mathscr{P}_p(E),W_p)$ is isometrically isomorphic to a "quotient space" of $L_p([0,1],\mathscr{B},m;\ E)$ in the above naive sense. Indeed,

$$F_\mu := \{X \in L_p([0,1],\mathscr{B}([0,1]),m;\ E),\ X \overset{\mathscr{D}}{=} \mu\}, \qquad \mu \in \mathscr{P}_p(E),$$

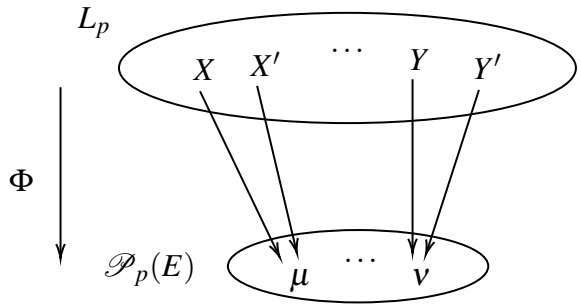
defines a partition, $\mathscr{F}$, of $L_p([0,1],\mathscr{B}([0,1]),m;\ E)$ into closed subsets F_μ, $\mu \in \mathscr{P}_p(E)$, and by (1.2),

$$D(F_\mu,F_\nu) = W_p(\mu,\nu),$$

where we denote by D the notion of distance between subsets in $L_p([0,1],\mathscr{B}([0,1]),m;\ E)$. In the same situation, we can consider a natural projection mapping

$$\Phi:\ L_p([0,1],\mathscr{B}([0,1]),m;\ E) \to (\mathscr{P}_p(E),W_p),\ X \mapsto \mu_X,$$

where μ_X denotes the distribution of the random variable X. We note that $F_\mu = \Phi^{-1}(\mu)$ for all $\mu \in \mathscr{P}_p(E)$, so that the above mentioned partition $\mathscr{F}$ is the collection of all fibers under Φ.



However, the naive sense of quotient space is not immediately useful as it stands, because the notion does not relate well enough the structure of the quotient to the structure of the "covering space".

So-called *metric quotient space induced by metric foliation* is a stronger notion of quotient space, in which the structure of the quotient is tied more strongly to the structure of the covering space through the so-called *equidistance property*. Then, the quotient inherits some key properties from that covering space and problem formulations set in the quotient space may be lifted to be studied in the covering space.

In this thesis, for $p \in [1,\infty)$ and a separable Hilbert space E, we set forth to rigorously represent the Wasserstein space $(\mathscr{P}_p(E), W_p)$ as a metric quotient space induced by a metric foliation. The metric foliation, however, takes place in a so-called direct limit metric space, namely in the direct limit $\mathscr{L}_p(E) := \varinjlim L_p^{(n)}(E)$ of the following sequence of Lebesgue spaces,

$$L_p^{(n)}(E) := L_p([0,1]^n, \mathscr{B}([0,1]^n), \otimes_{i=1}^n m;\ E), \qquad n \in \mathbb{N},$$

where m denotes Lebesgue measure on $([0,1], \mathscr{B}([0,1]))$.

The question whether the Wasserstein space $(\mathscr{P}_p(E), W_p)$ is already a metric quotient space induced by a metric foliation of $L_p([0,1], \mathscr{B}([0,1]), m;\ E)$ remains to be sorted out in its full generality.

Nevertheless, the direct limit metric space approach of our result can be useful to explore relations between Wasserstein spaces and the corresponding Lebesgue spaces. In fact, from a technical view point, the direct limit metric space version of the metric foliation may be more convenient to work with in certain situations.

As an example of its usefulness, if $p = 2$ and E is a separable Hilbert space, we are able to conclude from our result that the Wasserstein space $(\mathscr{P}_2(E), W_2)$ is a so-called Alexandrov non-negatively curved metric space, by showing that the property is inherited from the corresponding Lebesgue spaces that form the direct limit. Indeed, in this case the Lebesgue spaces that we work with are Hilbert, and so in particular non-negatively curved as flat spaces. We point out here that the direct limit metric space $\mathscr{L}_2(E)$ is endowed with a pre-Hilbert structure.

Our interest in relating the structure of the Wasserstein spaces to those of Lebesgue spaces is

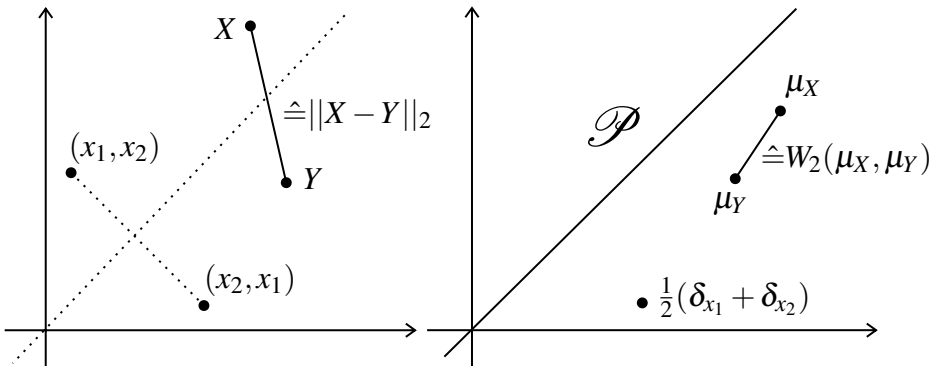


Figure 1.1: An illustration of the space of probability measures on $\mathbb{R}$ of the form $\frac{1}{2}(\delta_{x_1}+\delta_{x_2})$ as a folded Euclidean space

actually motivated by a study of Hamilton-Jacobi equations in spaces of probability measures.

1.2 Some motivational examples

Although it is an interesting task on its own to rigorously represent the Wasserstein spaces as isometrically isomorphic to metric quotient spaces, there is a call for such rigorous results motivated by an abstract approach to the study of Hamilton-Jacobi equations set in spaces of probability measures. Such equations arise in important physical applications.

In this section, we heuristically explain the motivation for our work by firstly mentioning two examples. We explain their relevance to our result, afterwards.

Example 1. **Large deviations for weakly interacting particles**. Let $d \in \mathbb{N}$. For fixed $n \in \mathbb{N}$, consider $\mathbb{R}^d$-valued stochastic processes $(X_1(t))_{t\geq 0}, (X_2(t))_{t\geq 0}, \ldots, (X_n(t))_{t\geq 0}$ satisfying a McKean-Vlasov finite system of SDE:

$$dX_i(t) = -\nabla\Psi(X_i(t))dt - \frac{1}{n}\sum_{j=1}^{n}\nabla\Phi(X_i(t)-X_j(t))dt + dW_i(t), \quad i = 1,...,n$$

Consider the so-called *empirical process* $(\rho_n(t))_{t\geq 0}$, defined via

$$\rho_n(t) := \frac{1}{n}\sum_{k=1}^{n}\delta_{X_{n,k}(t)} \in \mathscr{P}_2(\mathbb{R}^d), \quad t \in [0,\infty),$$

which is an $\mathscr{P}_2(\mathbb{R}^d)$-valued Markov process. As $n \in \mathbb{N}$ varies, a *large deviations principle* (LDP) holds for the sequence of empirical processes, $(\rho_n(t))_{t\geq 0}$, $n \in \mathbb{N}$, on its path space $(\mathscr{P}_2(\mathbb{R}^d))^\infty$. (See, Dawson & Gärtner (1987).)

A general approach for verifying an explicit LDP for sequences of Markov processes at the path space level was devised in Feng & Kurtz (2006), which is applicable in a variety of examples including the above one. In this approach, a non-linear semigroup H_n is associated to the Markov process $(\rho_n(t))_{t\geq 0}$. Then, LDP is proved by identifying a limit, $H := \text{Lim } H_n$. The key step, however, is establishing uniqueness (*comparison principle*) for weak solutions (*viscosity solutions*) of an abstract Hamilton-Jacobi equation set in $\mathscr{P}_2(\mathbb{R}^d)$,

$$f - \alpha H f = h,$$

for a dense class of $h \in C_b(\mathscr{P}_2(\mathbb{R}^d))$ for all $\alpha \in \mathbb{R}_{>0}$. E.g. when the coefficients $\Phi, \Psi \equiv 0$, so that $(X_i(t))_{t\geq 0}$, $i \in \mathbb{N}$, are independent Brownian motions, the Hamiltonian takes the form

$$H(\rho, p) = \langle \rho, \frac{1}{2}\Delta p\rangle + \frac{1}{2}\int |\nabla p|^2 d\rho, \quad p \in C_c^2(\mathbb{R}^d),\ \rho \in \mathscr{P}_2(\mathbb{R}^d),$$

and

$$Hf(\rho) := H(\rho, \frac{\delta f}{\delta \rho}) = \langle -\text{grad}_\rho S, \text{grad}_\rho f\rangle_\rho + \frac{1}{2}||\text{grad}_\rho f||_\rho^2,$$

where $S := \frac{1}{2}\int \rho \log \rho$.

Example 2. **A Hamilton-Jacobi PDE associated to the compressible Euler equations**. (See, Ambrosio & Feng (2014).) Defining

$$U = U_f(t, \rho_0) := \inf_{(\rho, v)} \{\int_0^t \int_{\mathbb{R}^d} \frac{1}{2}|v(s,x)|^2 dx ds + f(\rho(t));\ \rho(0) = \rho_0,\ \partial_t \rho + \text{div}(\rho v) = 0\}, \quad (1.3)$$

the minimizers (ρ, v) of (1.3) solve the *compressible Euler equations*:

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho v) = 0, \\ \partial_t (\rho v) + \operatorname{div}(\rho v \otimes v) = 0. \end{cases}$$

And, U itself solves the following Hamilton-Jacobi PDE in $\mathscr{P}_2(\mathbb{R}^d)$:

$$\begin{cases} \partial_t U(t,\rho) = HU(t,\cdot)(\rho), \\ U(0,\rho) = f(\rho), \end{cases}$$

where

$$HU(\rho) := \frac{1}{2}\int_{\mathbb{R}^d} |\nabla_x \frac{\delta U}{\delta \rho}|^2 \rho(dx) = \frac{1}{2}||\operatorname{grad}_\rho U||_\rho^2.$$

Hamilton-Jacobi equations in Alexandrov spaces. The Hamilton-Jacobi equations presented above can be cast into an abstract framework as Hamilton-Jacobi equations in Alexandrov non-negatively curved metric spaces. Namely, $\mathscr{P}_2(\mathbb{R}^d)$ is an Alexandrov non-negatively curved metric space, and on such space an abstract calculus has been developed, whose notions are mostly in accord with the notions of calculus that have been used in the optimal transport literature for $\mathscr{P}_2(\mathbb{R}^d)$. (See, Alexander et al. (2022) and Ambrosio et al. (2008).)

Moreover, due to hidden symmetries in the problem formulations, models often already are naturally set in a quotient space. Therefore, Hamilton-Jacobi equations of interest can be expected to be possible to be viewed as abstract Hamilton-Jacobi equations set in a metric quotient space of an Alexandrov non-negatively curved metric space. Metric quotient spaces of Alexandrov non-negatively curved metric spaces are Alexandrov non-negatively curved, so quotients are themselves equipped with the abstract Alexandrov space calculus.

In the framework of stochastic differential games, the following approach was envisioned by P.-L. Lions. (See, Cardaliaguet (2010).) If a suitable choice of a covering space can be found in which the metric geometry structure of the metric quotient and the covering space can be related, equations set in the quotient space can be lifted to the covering space, enabling the equations to be

studied in that covering space.

Our main result falls into the point of finding covering spaces for the Wasserstein spaces, that shall in particular enable such a study of abstract Hamilton-Jacobi equations set in $(\mathscr{P}_2(E), W_2)$.

1.3 The outline of the presentation

We conclude this introduction by outlining the structure of the presentation:

In Chapter 2, we have collected preliminaries from probability theory. In particular, we mention here results about representing probability measures as the distribution of a random variable. The definitions of the Lebesgue spaces that we work with and the Wasserstein spaces are to be found in this chapter.

Chapter 3 is devoted to preliminaries from metric space geometry. This is where we provide a definition of the notion of metric quotient space induced by metric foliation, and the definition of the notion of submetry projection.

An attempt has been made to provide the preliminaries in a style ready to be applied in the proofs of the main results. We provide some of the proofs of the preliminary results for sake of completion.

In Chapter 4, we recall the definition of curvature bounded below metric space in the sense of Alexandrov, highlighting the Kirszbraun synthetic lower curvature bound conditions. We prove here that Kirzsbraun lower curvature bound conditions are stable under submetries, which in particular implies that a metric quotient of a non-negatively curved metric space is non-negatively curved.

The main result of this thesis, Theorem in Section 5.3.2, is stated and proved in Chapter 5.

Chapter 2

Preliminaries from probability theory

2.1 Measurable spaces and their morphisms

In this section, we collect some definitions and results concerning Borel isomorphisms, which is a notion of morphism between measurable spaces that is fundamental for a logical construction of some of the rest of the preliminaries that we make use of in this work. We refer to Kallenberg (1997), Dudley (1989), Parthasarathy (1967), or Srivastava (2008) for detailed discussions on Borel isomorphisms and for the proofs of the results that we present here.

Definition 1. (Borel isomorphism). Let $(S,\mathscr{G})$ and $(T,\mathscr{H})$ be two measurable spaces. A mapping $\Phi : S \to T$ is called a *Borel isomorphism*, if and only if,

(i) Φ is measurable,

(ii) Φ is bijective,

(iii) Φ^{-1} is measurable.

$(S,\mathscr{G})$ and $(T,\mathscr{H})$ are called *Borel isomorphic*, if and only if, there exists a Borel isomorphism $\Phi : S \to T$.

Definition 2. (Borel space). A measurable space is called a *Borel space*, if and only if, it is Borel isomorphic to $(A,\mathscr{B}(A))$ for a Borel subset A of $[0,1]$, where $\mathscr{B}(A)$ denotes the Borel σ-algebra induced by the Euclidean distance in A.

Proposition 1. (See e.g. Dudley (1989), Lemma 13.1.3.) Any Polish space endowed with its Borel σ-algebra generated by the open sets, is a Borel space.

Proposition 2. (Borel isomorphism theorem. See, e.g. Parthasarathy (1967), Theorem 2.12.) Let E_1, E_2 be two complete separable metric spaces, and $B_1 \in \mathscr{B}(E_1)$, $B_2 \in \mathscr{B}(E_2)$ be two Borel subsets, respectively. Then, B_1 and B_2 are Borel isomorphic as measurable spaces equipped with induced Borel σ-algebras, if and only if, the sets B_1 and B_2 have the same cardinality.

Corollary 1. Any Polish space with the cardinality $2^{\aleph_0}$, i.e. the cardinality that of the continuum, is Borel isomorphic to $([0,1], \mathscr{B}([0,1]))$.

Proof. The claim immediately follows from Borel isomorphism theorem. We note that the underlying set of any separable metric space has cardinality of at most $2^{\aleph_0}$.

□

2.2 On existence of random variables with given distributions

Let m denote Lebesgue measure on $([0,1], \mathscr{B}([0,1]))$. In this section, we state a representation result due Blackwell and Dubins that extends Skorokhod's representation theorem that was proved in Skorokhod (1957). See, Halmos (1950) and Theorem 3.4.23 in Srivastava (2008) for *the isomorphism theorem* about measure spaces, which enable us to replace the space $([0,1], \mathscr{B}([0,1]), m)$ with a Borel space equipped with a non-atomic probability measure in certain situations.

Definition and Proposition 1. (Fan (1943)). Let (E,d) be a complete separable metric space and $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. Let $R'(\Omega, \mathscr{F}, \mathbb{P};\ E)$ denote the set of all random variables $(\Omega, \mathscr{F}, \mathbb{P}) \to (E, \mathscr{B}(E))$. The *Ky Fan distance* d_F between two random variables $X, Y \in R'(\Omega, \mathscr{F}, \mathbb{P};\ E)$ is defined as

$$d_F(X,Y) := \inf\{\varepsilon \in \mathbb{R}_{>0},\ \mathbb{P}(d(X,Y) \geq \varepsilon) < \varepsilon\},$$

and the following statements hold true:

(1) A sequence $(X_n)_{n\in\mathbb{N}}$ of random variables in $R'(\Omega, \mathscr{F}, \mathbb{P};\ E)$ converges in probability to a random variable $X \in R'(\Omega, \mathscr{F}, \mathbb{P};\ E)$, if and only if, $\lim_{n\to\infty} d_F(X_n, X) = 0$.

(2) The set $R'(\Omega, \mathscr{F}, \mathbb{P};\ E)$ together with the Ky Fan distance α form a pseudo-metric space.

(3) The $\mathbb{P}$-a.s. equivalence of random variables defines an equivalence relation on $R'(\Omega, \mathscr{F}, \mathbb{P};\ E)$. Let $R(\Omega, \mathscr{F}, \mathbb{P};\ E) := \{[X],\ X \in R'(\Omega, \mathscr{F}, \mathbb{P};\ E)\}$ denote the set of all equivalence classes under that equivalence relation, where $[X]$ denotes the equivalence class of $X \in R'(\Omega, \mathscr{F}, \mathbb{P};\ E)$. Then, the set $R(\Omega, \mathscr{F}, \mathbb{P};\ E)$ together with the well-defined distance, denoted again by d_F, defined by

$$d_F([X],[Y]) := d_F(X,Y), \qquad [X],[Y] \in R(\Omega, \mathscr{F}, \mathbb{P};\ E),$$

form a complete metric space $(R(\Omega, \mathscr{F}, \mathbb{P};\ E),\ d_F)$, which we shall refer to as the *Ky Fan space* defined over $(\Omega, \mathscr{F}, \mathbb{P})$.

Proof. See, Fan (1943). □

Definition and Proposition 2. (Prokhorov distance). Let (E,d) be a complete separable metric space and let $\mathscr{P}(E)$ denote the set of all probability measures on $(E, \mathscr{B}(E))$. Then:

(1) The set $\mathscr{P}(E)$, together with the *Prokhorov distance* D_P defined by

$$D_P(\mu,\nu) := \inf\{\varepsilon \in \mathbb{R}_{>0} \mid \forall B \in \mathscr{B}(E):\ \mu(B) \leq \nu(B^\varepsilon) + \varepsilon\}, \qquad \mu,\nu \in \mathscr{P}(E),$$

where $B^\varepsilon := \{x \in E,\ d(x,B) < \varepsilon\}$, form a complete separable metric space $(\mathscr{P}(E), D_P)$.

(2) Moreover, a sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ in $\mathscr{P}(E)$ converges weakly to a probability measure $\mu \in \mathscr{P}(E)$, if and only if, $(\mu_n)_{n\in\mathbb{N}}$ converges to μ in $(\mathscr{P}(E), D_P)$.

Proof. See e.g. Billingsley (2013), or Hofinger (2006). □

Proposition 3. (Blackwell & Dubins (1983)). Let (E,d) be a complete separable metric space. Consider the space of all probability measures on $(E, \mathscr{B}(E))$, denoted by $\mathscr{P}(E)$, as a complete separable metric space endowed with the Prokhorov distance. Then, there exists a Borel measurable mapping

$$F : \mathscr{P}(E) \times [0,1] \to E,$$

such that:

(i) For each fixed $\mu \in \mathscr{P}(E)$, the random variable $F(\mu,\cdot)$ that maps $x \in [0,1] \mapsto F(\mu,x)$, has distribution μ.

(ii) For each fixed $\mu_0 \in \mathscr{P}(E)$, the random mapping $F(\cdot,x)$ that maps $\mu \in \mathscr{P}(E) \mapsto F(\mu,x)$, is m-a.s. continuous at $\mu = \mu_0$.

Remark 1. In particular, we point out that if (E,d) is a complete separable metric space and μ is a probability measure on $(E,\mathscr{B}(E))$, there exists a random variable $X : ([0,1],\mathscr{B}([0,1]),m) \to (E,\mathscr{B}(E))$ with the distribution μ. The same conclusion holds, if we replace the probability space $([0,1],\mathscr{B}([0,1]),m)$ in its role there with any Borel space equipped with a *non-atomic* probability measure, where a probability measure $\mathbb{P}$ on a measurable space is called non-atomic, if and only if, for every measurable set A with $\mathbb{P}(A) > 0$, there exists a measurable proper subset $B \subset A$ with $0 < \mathbb{P}(B) < \mathbb{P}(A)$. (See, Theorem 3.4.23 in Srivastava (2008).)

2.3 Probability kernels and randomization

Definition 3. (Probability kernel). Let $(S,\mathscr{G})$ and $(T,\mathscr{H})$ be measurable spaces. A mapping

$$\mu : S \times \mathscr{H} \to [0,1]$$

is called a *probability kernel* from $(S,\mathscr{G})$ to $(T,\mathscr{H})$, if and only if:

(i) For each fixed $B \in \mathscr{H}$, the mapping $\mu(\cdot,B)$ that maps $s \in S \mapsto \mu(s,B)$, is measurable.

(ii) For each fixed $s \in S$, the mapping $\mu(s,\cdot)$ that maps $B \in \mathscr{H} \mapsto \mu(s,B)$, is a probability measure.

Proposition 4. (Randomization of a probability kernel. Kallenberg (1997), Lemma 2.22.) Let $(S,\mathscr{G})$ be a measurable space, $(T,\mathscr{H})$ a Borel space and $\mu : S \times \mathscr{H} \to [0,1]$ a probability kernel. Let $\Theta : ([0,1],\mathscr{B}([0,1]),m) \to ([0,1],\mathscr{B}([0,1]))$ be a uniformly distributed random variable.

Then, there exists a measurable mapping $f : S \times [0,1] \to T$, such that for each $s \in S$ fixed, $f(s,\Theta)$ has distribution $\mu(s,\cdot)$.

2.4 Regular conditional distributions

In this section, we state another preliminary result taken from Kallenberg (1997) concerning existence of so-called regular conditional distributions.

Proposition 5. (Existence of regular conditional distribution of random variables, and a disintegration. Kallenberg (1997), Theorems 5.3 and 5.4.)

Let $(S,\mathscr{G})$ be a measurable space, $(T,\mathscr{H})$ a Borel space, set $(\Omega,\mathscr{F},\mathbb{P}) := ([0,1],\mathscr{B}([0,1]),m)$ and consider two random variables $X : (\Omega,\mathscr{F},\mathbb{P}) \to (S,\mathscr{G})$ and $Y : (\Omega,\mathscr{F},\mathbb{P}) \to (T,\mathscr{H})$ defined on that underlying probability space .

Then:

(1) There exists a probability kernel $\mu : S \times \mathscr{H} \to [0,1]$, that shall be referred to as a *regular conditional probability of Y given X*, for which the following statement holds true:

$$\forall B \in \mathscr{H} : \ \mu(X(\cdot),B) = \mathbb{P}(Y \in B \mid X) \text{ a.s.}$$

(2) If $g : S \times T \to [0,\infty)$ is a bounded measurable function, with the kernel μ above, we have

$$\mathbb{E}[g(X,Y) \mid X] = \int_T g(X,y)\mu(X,dy) \text{ a.s.},$$

so that

$$\mathbb{E}[g(X,Y)] = \mathbb{E}[\int_T g(X,y)\mu(X,dy)].$$

2.5 Abstract Lebesgue spaces

In this section, we summarize some information regarding Lebesgue spaces. For a detailed construction of these spaces, see Amann & Escher (2001), upon which the following summary is based.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and $(E, |\cdot|)$ a separable Banach space as a vector space over $\mathbb{R}$.

For $p \in [1, \infty)$, we define

$$L'_p(\Omega, \mathscr{F}, \mathbb{P};\, E) := \{X : (\Omega, \mathscr{F}) \to (E, \mathscr{B}(E)) \text{ measurable},\ \mathbb{E}[|X|^p] < \infty\},$$

and

$$||X||_p := (\mathbb{E}[|X|^p])^{\frac{1}{p}}, \quad X \in L'_p(\Omega, \mathscr{F}, \mathbb{P};\, E).$$

Then, $L'_p(\Omega, \mathscr{F}, \mathbb{P};\, E)$ is a vector space with induced pointwise operations and $||\cdot||_p$ is a seminorm on it.

Furthermore, considering the vector subspace

$$\mathscr{N} := \{X \in L'_p(\Omega, \mathscr{F}, \mathbb{P};\, E),\ X = 0 \text{ a.s.}\},$$

we define,

$$L_p(\Omega, \mathscr{F}, \mathbb{P};\, E) := {L'_p(\Omega, \mathscr{F}, \mathbb{P};\, E)}\big/_{\mathscr{N}}$$

as a quotient vector space. Then, the seminorm $||\cdot||_p$ induces a norm on $L_p(\Omega, \mathscr{F}, \mathbb{P};\, E)$, also denoted as $||\cdot||_p$.

Proposition 6. (See, Theorem 4.10 in Amann & Escher (2001).) Let $p \in [1, \infty)$. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and $(E, |\cdot|)$ a separable Banach space as a vector space over $\mathbb{R}$. Then:

(1) $L_p(\Omega, \mathscr{F}, \mathbb{P};\, E)$ equipped with the norm $||\cdot||_p$ is a Banach space.

(2) If E is Hilbert equipped with the scalar product $\langle \cdot, \cdot \rangle_E$, then $L_2(\Omega, \mathscr{F}, \mathbb{P};\, E)$ is Hilbert

equipped with the scalar product

$$\langle \cdot, \cdot \rangle : \ L_2(\Omega, \mathscr{F}, \mathbb{P};\ E) \times L_2(\Omega, \mathscr{F}, \mathbb{P};\ E) \to \mathbb{R},\ \langle X, Y \rangle := \mathbb{E}[\langle X, Y \rangle_E].$$

2.6 Wasserstein spaces

If (E, τ) is a topological space and μ, ν are probability measures on $(E, \mathscr{B}(E))$, then a probability measure π on the product space $(E \times E, \mathscr{B}(E) \otimes \mathscr{B}(E))$ is called a *coupling* of μ, ν, if and only if, $p_{1\#}\pi = \mu$ and $p_{2\#}\pi = \nu$, where $p_i : E \times E \to E$, $i = 1, 2$, denote the canonical coordinate projections respectively. That is, π is a coupling of μ, ν in that order, if and only if,

$$\forall B \in \mathscr{B}(E) : \ \pi(B \times E) = \mu(B),\ \pi(E \times B) = \nu(B).$$

Definition and Proposition 3. (Wasserstein space of probability measures). Let $p \in [1, \infty)$. Let (E, d) be a complete separable metric space, $\mathscr{P}(E)$ the set of all probability measures on $(E, \mathscr{B}(E))$, and let $p_i : E \times E \to E$, $i = 1, 2$, denote the canonical coordinate projections, respectively. Then:

(1) The set

$$\mathscr{P}_p(E) := \{\mu \in \mathscr{P}(E),\ \exists\, x \in E : \int_E d^p(x, y)\mu(dy) < \infty\},$$

together with the so-called *p-Wasserstein distance* W_p, defined for $\mu, \nu \in \mathscr{P}_p(E)$ by

$$W_p(\mu, \nu) := \inf\{\left(\int_{E \times E} d^p(x, y)\pi(dx, dy)\right)^{\frac{1}{p}};\ \pi \in \mathscr{P}(E \times E),\ p_{1\#}\pi = \mu,\ p_{2\#}\pi = \nu\},$$

form a complete separable metric space $(\mathscr{P}_p(E), W_p)$, called *p-Wasserstein space*.

(2) In particular, if E is locally compact, then for each $\mu, \nu \in \mathscr{P}_p(E)$, there exists an *optimal*

coupling $\pi_0 \in \mathscr{P}(E \times E)$ with $p_{1\#}\pi_0 = \mu$, $p_{2\#}\pi_0 = \nu$, in the sense that

$$W_p(\mu,\nu) = \left(\int_{E\times E} d^p(x,y)\pi_0(dx,dy) \right)^{\frac{1}{p}}.$$

The collection of all optimal couplings for $\mu,\nu \in \mathscr{P}_p(E)$ is denoted by $\Gamma_{\text{opt}}(\mu,\nu)$, whereas the collection of all couplings for $\mu,\nu \in \mathscr{P}_p(E)$ is denoted as

$$\Gamma(\mu,\nu) := \{\pi \in \mathscr{P}(E\times E);\ p_{1\#}\pi = \mu,\ p_{2\#}\pi = \nu\}.$$

Proof. See e.g. Ambrosio et al. (2008), Chapters 5 and 7. We also would like to note here that our main result, Theorem 5.3.2, a posteriori implies some of the claims made in this definition and proposition. See, Remark 8.

□

The following Proposition is mentioned in Rachev & Rüschendorf (1998) and Villani (2009). We prove the claim here, as our main result depends on it.

Proposition 7. Let $p \in [1,\infty)$, and E be a complete separable Hilbert space. Then, we have for all $\mu,\nu \in \mathscr{P}_p(E)$,

$$W_p(\mu,\nu) = \inf\{||X-Y||_p;\ X,Y \in L_p([0,1],\mathscr{B}([0,1]),m;\ E),\ X \overset{\mathscr{D}}{=} \mu,\ Y \overset{\mathscr{D}}{=} \nu\}.$$

Proof. Let $\mu,\nu \in \mathscr{P}_p(E)$ be arbitrary.

1. Let $\pi \in \Gamma(\mu,\nu)$ be arbitrary. Since $(E\times E, \mathscr{B}(E\times E))$ is a Polish space, by Remark 1, there exists a random variable

$$Z : ([0,1],\mathscr{B}([0,1]),m) \to E\times E$$

with distribution π. Using the canonical coordinate projections $p_i : E\times E \to E$, $i = 1,2$, consider the random variables $X := p_1 \circ Z$ and $Y := p_2 \circ Z$, also as elements in $L_p([0,1],\mathscr{B}([0,1]),m;\ E)$. Then, $X \overset{\mathscr{D}}{=} \mu$, $Y \overset{\mathscr{D}}{=} \nu$, and $(X,Y) \overset{\mathscr{D}}{=} \pi$, so that

$$\left(\int_{E\times E} d^p(x,y)\pi(dx,dy)\right)^{\frac{1}{p}} = ||X-Y||_p.$$

Hence,

$$\inf\{||X-Y||_p;\ X,Y\in L_p([0,1],\mathscr{B}([0,1]),m;\ E),\ X\overset{\mathscr{D}}{=}\mu,\ Y\overset{\mathscr{D}}{=}\nu\} \leq \left(\int_{E\times E} d^p(x,y)\pi(dx,dy)\right)^{\frac{1}{p}},$$

and so,

$$\inf\{||X-Y||_p;\ X,Y\in L_p([0,1],\mathscr{B}([0,1]),m;\ E),\ X\overset{\mathscr{D}}{=}\mu,\ Y\overset{\mathscr{D}}{=}\nu\} \leq W_p(\mu,\nu).$$

2. Let $X,Y\in L_p([0,1],\mathscr{B}([0,1]),m;\ E)$ be arbitrary with $X\overset{\mathscr{D}}{=}\mu$ and $Y\overset{\mathscr{D}}{=}\nu$. Letting π denote the distribution of (X,Y), we have

$$\left(\int_{E\times E} d^p(x,y)\pi(dx,dy)\right)^{\frac{1}{p}} = ||X-Y||_p,$$

and so,

$$\inf\{\left(\int_{E\times E} d^p(x,y)\pi(dx,dy)\right)^{\frac{1}{p}};\ \pi\in\mathscr{P}(E\times E),\ p_{1\#}\pi=\mu,\ p_{2\#}\pi=\nu\} \leq ||X-Y||_p,$$

i.e. $W_p(\mu,\nu)\leq ||X-Y||_p$. Hence,

$$W_p(\mu,\nu) \leq \inf\{||X-Y||_p;\ X,Y\in L_p([0,1],\mathscr{B}([0,1]),m;\ E),\ X\overset{\mathscr{D}}{=}\mu,\ Y\overset{\mathscr{D}}{=}\nu\}.$$

The claim follows. □

Remark 2. The statement of Proposition 7 and its proof remain valid, if we replace the probability space $([0,1],\mathscr{B}([0,1]),m)$ in its role there with any Borel space equipped with a non-atomic probability measure. (See, Theorem 3.4.23 in Srivastava (2008).)

Chapter 3

Preliminaries from metric geometry

In this chapter, we provide some preliminaries from metric geometry that can e.g. be found in Burago et al. (2001), Alexander et al. (2022) and Bridson & Haefliger (2013).

To set some notations we start with basic notions. Afterwards, for explaining the notion of metric quotient space induced by metric foliation, we define metric foliations and submetries. Then, we present the correspondence of metric foliations and submetries, which was explored in Sosa (2017) and Galaz-García et al. (2018).

3.1 Metric spaces

Definition 4. (Metric space). Let X be a set, and $d : X \times X \mapsto [0,\infty)$ a mapping. Then (X,d) is called a *metric space*, if and only if:

(i) $\forall x,y \in X : d(x,y) = 0 \iff x = y$,

(ii) $\forall x,y \in X : d(x,y) = d(y,x)$,

(iii) $\forall x,y,z \in X : d(x,y) \leq d(x,z) + d(z,y)$.

Definition 5. (Isometric embedding). Let (X,d) and (Y,d') be metric spaces. A mapping $\Phi : X \to Y$ is called an *isometric embedding*, if and only if,

$$\forall\, x_1,x_2 \in X :\ d'(\Phi(x_1),\Phi(x_2)) = d(x_1,x_2).$$

Note that, an isometric embedding as defined above is necessarily injective. A bijective isometric

embedding is called an *isometry*. Two metric spaces are said to be *isometrically isomorphic*, if and only if, there exists an isometry from one of them to the other.

We follow with two variants of the notion of metric space that we shall employ. The systematic naming of these variants is taken from a preliminary version of Alexander et al. (2022).

Definition 6. (Pseudo-metric space, or $[0,\infty)$-metric space). Let X be a set, and $d : X \times X \to [0,\infty)$ be a mapping. (X,d) is called a *pseudo-metric space*, if and only if:

(i) $\forall x \in X :\ d(x,x) = 0,$

(ii) $\forall x,y \in X :\ d(x,y) = d(y,x),$

(iii) $\forall x,y,z \in X :\ d(x,y) \leq d(x,z) + d(z,y).$

The mapping d itself is then called a *pseudo-metric*.

Definition 7. ($(0,\infty]$-Metric space). Let X be a set, and $d : X \times X \to [0,\infty]$ be a mapping. (X,d) is called a $(0,\infty]$-*metric space*, if and only if:

(i) $\forall x,y \in X :\ d(x,y) = 0 \iff x = y,$

(ii) $\forall x,y \in X :\ d(x,y) = d(y,x),$

(iii) $\forall x,y,z \in X :\ d(x,y) \leq d(x,z) + d(z,y).$

The mapping d itself is then called a $(0,\infty]$-*metric*.

Definition 8. (The canonical metric space corresponding to a pseudo-metric space). Let (X,d) be a pseudo-metric space. Consider the equivalence relation on X defined by

$$x \sim y : \iff d(x,y) = 0, \quad x,y \in X,$$

and let $\tilde{X}$ denote the set of all equivalence classes under that equivalence relation. Then, $\tilde{X}$ together with the well-defined distance on it defined by

$$d([x],[y]) := d(x,y), \quad [x],[y] \in \tilde{X},$$

form a metric space $(\tilde{X}, d)$, called the *canonical metric space corresponding to* the pseudo-metric space (X, d).

Definition 9. (Hausdorff convergence of closed subsets in a metric space). Let (X, d) be a metric space, $(F_n)_{n\in\mathbb{N}}$ a sequence of closed subsets and $F \subset X$ a closed subset. The sequence $(F_n)_{n\in\mathbb{N}}$ is said to *converge in the sense of Hausdorff* to the set F, if and only if,

$$\forall x \in X : \; d(x, F) = \lim_{n\to\infty} d(x, F_n).$$

3.2 Length spaces and geodesic spaces

See, e.g. Burago et al. (2001), Lang (2013), Schneider (2001) and Alexander et al. (2022).

Definition 10. (Parameter transformations). Let $\mathbb{I}, \mathbb{I}' \subset \mathbb{R}$ be two intervals in $\mathbb{R}$. A mapping $\tau : \mathbb{I} \to \mathbb{I}'$ is called a *continuous parameter transformation*, if and only if, it is strictly monotone, bijective and continuous.

Definition 11. (Parametrized curve). Let $\mathbb{I} \subset \mathbb{R}$ be an interval and (X, d) be a metric space. A mapping $\gamma : \mathbb{I} \to X$ is called a *parametrized curve* in X, if and only if, it is continuous.

Definition 12. (Curve). Let (X, d) be a metric space. Two parametrized curves $\gamma : \mathbb{I} \to X$ and $\gamma' : \mathbb{I}' \to X$ are said to be *equivalent parametrized curves*, if and only if, there exists a strictly increasing parameter transformation $\tau : \mathbb{I}' \to \mathbb{I}$ such that $\gamma = \gamma' \circ \tau$. Equivalence of parametrized curves defines an equivalence relation on the set of all parametrized curves. A *curve* in X is an equivalence class under this equivalence relation, and any representative of a curve is called a *parametrization* of that curve. We shall not always distinguish notations between curves and their parametrizations.

Definition 13. (Trace of a curve). Let (X, d) be a metric space. The *trace of a curve* in X, is defined as the image of any parametrization of it. This definition is independent of the choice of parametrization.

Definition 14. (Length of a curve). Let (X,d) be a metric space, $[a,b] \subset \mathbb{R}$ a closed interval and $\gamma : [a,b] \to X$ a parametrized curve in X. Then, we define the length of γ as,

$$l(\gamma) := \sup\{\sum_{i=1}^{k} d(\gamma(t_i), \gamma(t_{i-1})) \mid k \in \mathbb{N};\ t_0, \ldots, t_k \in \mathbb{R},\ a = t_0 < t_1 < \cdots < t_k = b\} \in [0,\infty].$$

The *length of a curve* in X is defined as the length of any parametrization of it. This definition is independent of the choice of parametrization.

Definition 15. (The induced length metric). Let (X,d) be a metric space. We define a $(0,\infty]$-metric on X, denoted by d_l and called the *induced length metric* from d, as follows.

$$d_l(x,y) := \inf\{l(\gamma);\ [\gamma] \text{ curve in } X,\ \gamma : [a,b] \to X,\ \gamma(a) = x,\ \gamma(b) = y\} \in [0,\infty], \quad x,y \in X.$$

(X,d_l) is a $(0,\infty]$-metric space, called the *induced length space* or the *induced intrinsic space* from (X,d).

Definition 16. (Length space). A metric space (X,d) is called a *length space*, if and only if, its induced length space is equal to itself, i.e. if:

$$\forall x,y \in X :\ d(x,y) = d_l(x,y).$$

Lemma 1. (Approximate midpoints lemma). A complete metric space (X,d) is a length space, if and only if, for all $x,y \in X$ and all $\varepsilon \in \mathbb{R}_{>0}$, there exists an ε-*midpoint* $z \in X$ in the sense that:

$$d(x,z) \leq \frac{1}{2} d(x,y) + \varepsilon \quad \text{and} \quad d(z,y) \leq \frac{1}{2} d(x,y) + \varepsilon.$$

Proof. See, Lang (2013). □

Definition 17. (Constant speed parametrized geodesic). Let (X,d) be a metric space, $\lambda \in \mathbb{R}$ a real number, $\mathbb{I} \subset \mathbb{R}$ an interval and $\gamma : \mathbb{I} \to X$ a parametrized curve in X. γ is called a *constant speed*

parametrized geodesic with speed λ, if and only if,

$$\forall s,t,\in \mathbb{I}:\ d(\gamma(s),\gamma(t)) = \lambda|s-t|.$$

A speed $\lambda = 1$ parametrized geodesic is called a *unit speed parametrized geodesic*.

Definition 18. (Geodesic space). A metric space (X,d) is called a *geodesic space*, if and only if, for any two points $x,y \in X$, there exists a unit speed parametrized geodesic $\gamma : [a,b] \to X$ such that $\gamma(a) = x$ and $\gamma(b) = y$.

Lemma 2. (Midpoints lemma). A complete metric space (X,d) is a geodesic space, if and only if, for all $x,y \in X$ there exists a *midpoint* $z \in X$ in the sense that:

$$d(x,z) = d(z,y) = \frac{1}{2}d(x,y).$$

Proof. See, Lang (2013). □

Corollary 2. Any complete geodesic space is a length space.

Proof. Indeed, a midpoint is already an ε-midpoint, so that the claim follows from Lemmas 1 and 2 under the completeness assumption. □

3.3 Submetries and metric quotient spaces induced by metric foliation

See, Sosa (2017), Galaz-García et al. (2018) and Alexander et al. (2022).

Definition and Proposition 4. (Metric foliations). Let (X,d) be a metric space. A collection $\mathscr{F} = \{F_i\}_{i\in I}$ of non-empty subsets $F_i \subset X$, $i \in I$, is called a *metric foliation* of the space (X,d), if and only if:

(i) $\forall i \in I:\ F_i \subset X$ is closed,

(ii) $\forall i,j \in I:\ i \neq j \implies F_i \cap F_j = \emptyset$,

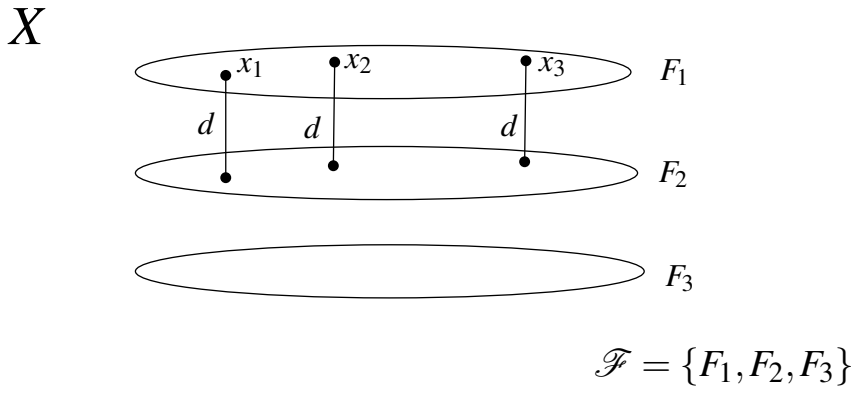


Figure 3.1: Metric foliation illustration

(iii) $X = \bigcup_{i \in I} F_i$,

(iv) *The equidistance property* holds, i.e.,

$$\forall\, i, j \in I\, \forall\, x \in F_i :\ d(F_i, F_j) = d(x, F_j),$$

where $d(F_i, F_j) := \inf\{d(x,y);\ x \in F_i, y \in F_j\}$ and $d(x, F_j) := \inf\{d(x,y),\ y \in F_j\}$.

The set $\mathscr{F}$ together with the notion of distance d between subsets of X as defined above, form a metric space $(\mathscr{F}, d)$ called *the metric quotient space induced by the metric foliation* $\mathscr{F}$.

Definition 19. (Open submetry). Let (X,d), (Y,d') be two metric spaces, and $\Phi : X \to Y$ a mapping. Φ is called a *submetry* (or *submetry projection*), if and only if, the *submetry property* holds, i.e.

$$\forall p \in X\ \forall R \in \mathbb{R}_{\geq 0} : \Phi(B(p,R)_X) = B(\Phi(p), R)_Y,$$

where

$$B(p,R)_X := \{x \in X,\ d(p,x) < R\},$$
$$B(\Phi(p),R)_Y := \{y \in Y,\ d'(\Phi(p), y) < R\},$$

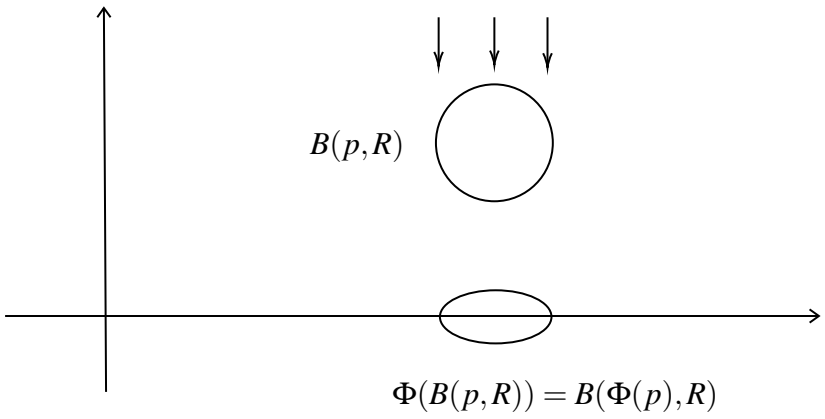


Figure 3.2: Submetry illustration

are defined as open balls in X and Y respectively.

Remark 3. Any submetry $\Phi : X \to Y$ is 1-Lipschitz, and 1-co-Lipschitz: Indeed, the inclusions $\Phi(B(p,R)_X) \subset B(\Phi(p),R)_Y$, $p \in X$, $R \in \mathbb{R}_{\geq 0}$ correspond to the 1-Lipschitzness, and the reversed inclusions $\Phi(B(p,R)_X) \supset B(\Phi(p),R)_Y$ define the notion of 1-co-Lipschitzness for the mapping Φ.

Proposition 8. Let $\Phi : X \to Y$ be a submetry between two metric spaces. Then:

(1) Φ is surjective.

(2) Φ is an open mapping, i.e., image of open sets under Φ are open sets.

(3) Φ is continuous.

Proof. 1. Let $y \in Y$ be arbitrary. Consider an element $x \in X \neq \emptyset$, and pick $R \in \mathbb{R}_{>0}$ with $R > d_Y(\Phi(x), y)$. Then,

$$\Phi(B(x,R)_X) = B(\Phi(x),R)_Y$$

implies that there exists $x' \in X$ with $\Phi(x') = y$. So, Φ is surjective.

2. Each open subset $G \subset X$ can be represented as a union of open balls. Then, by the submetry property, the image of G under Φ is also a union of open balls, hence open in Y.

3. The continuity of Φ follows immediately from the submetry property, as preimages of open balls in Y are open balls in X. □

Proposition 9. Let $\Phi : L \to M$ be a submetry between two metric spaces (L, d_L) and (M, d_M). Then:

(1) L is length space $\implies$ M is length space.

(2) L is complete $\implies$ M is complete.

Proof. 1. Assume that L is length space. Let $x, y \in M$, and $\varepsilon \in \mathbb{R}_{>0}$ be arbitrary. By the submetry property, we have preimages $\hat{x}, \hat{y} \in L$ under Φ of x, y respectively, with

$$d_L(\hat{x}, \hat{y}) < d_M(x, y) + \varepsilon,$$

and, by the assumption that L is length space, there exists a curve γ in L joining $\hat{x}, \hat{y}$ such that its length $l(\gamma)$ approximates $d_L(\hat{x}, \hat{y})$ with

$$d_L(\hat{x}, \hat{y}) \leq l(\gamma) < d_M(x, y) + \varepsilon.$$

Then, $\Phi \circ \gamma$ is a curve in M joining x, y, and

$$d_M(x, y) \leq l(\Phi \circ \gamma) \leq l(\gamma) < d_M(x, y) + \varepsilon.$$

Hence, the distance $d_M(x, y)$ can be arbitrarily well approximated by the length of some curve in M joining $x, y \in M$. So, M is a length space and (1) is established.

2. Assume L is complete and let $(y_n)_{n \in \mathbb{N}}$ be Cauchy in M. We can extract a subsequence $(y'_k)_{k \in \mathbb{N}}$ such that

$$\forall k \in \mathbb{N} : d_M(y'_k, y'_{k+1}) < \frac{1}{2^k}.$$

Then, the submetry property inductively gives rise to a sequence of corresponding preimages $(\hat{y}_k)_{k \in \mathbb{N}}$ in L, $\Phi(\hat{y}_k) = y'_k$ for $k \in \mathbb{N}$, such that

$$\forall k \in \mathbb{N} : d_L(\hat{y}_k, \hat{y}_{k+1}) < \frac{1}{2^{k-1}}.$$

The subsequence $(\hat{y}_k)_{k\in\mathbb{N}}$ thus, is Cauchy in L and therefore converges to, say $\hat{y}\in L$. By the Lipschitzness of Φ, we have $y'_k \to \Phi(\hat{y})$, and so the initial Cauchy sequence $(y_n)_{n\in\mathbb{N}}$ must converge to $\Phi(\hat{y})$. We conclude that M is complete.

□

Proposition 10. (Correspondence of metric foliations and open submetries. See, Sosa (2017), and Lemma 8.4 in Galaz-García et al. (2018).)

(1) If $\mathscr{F}$ is a metric foliation of (X,d), then the *natural projection mapping* $\Phi : X \to \mathscr{F}$, which maps $x\in X$ to the unique subset containing x, denoted here as $F_x \in \mathscr{F}$, is a submetry.

(2) If $\Phi : X \to Y$ is a submetry from a metric space (X,d) to a metric space (Y,d'), then the collection of all fibers over Φ, i.e. the collection $\{\Phi^{-1}(y)\}_{y\in Y}$, is a metric foliation of X.

Proof. 1. Assume that $\mathscr{F}$ is a metric foliation of (X,d) and consider the natural projection mapping $\Phi : X \to \mathscr{F}$, which maps $x\in X$ to the unique subset $F_x \in \mathscr{F}$ containing x. Let $R\in\mathbb{R}_{\geq 0}$ and $x\in X$ be arbitrary.

a) Firstly, assume that $z\in B(x,R)_X$. Then,

$$d'(\Phi(x),\Phi(z)) \leq d(x,z) < R,$$

so that $F_z \in B(F_x,R)_{\mathscr{F}}$, and so $\Phi(B(x,R)_X) \subset B(\Phi(x),R)_{\mathscr{F}}$.

b) Secondly, let $F\in B(\Phi(x),R)_{\mathscr{F}}$, i.e. that $d(F_x,F)<R$. By the equidistance property, we arrive at $d(x,F)=d(F_x,F)<R$. Then, by the definition of infimum in the definition of $d(x,F)$, there exists $z\in F$ such that $d(x,z)<R$. Note that $\Phi(z)=F$, so the previous step shows $F\in \Phi(B(x,R)_X)$, and so $B(\Phi(x),R)_{\mathscr{F}} \subset \Phi(B(x,R)_X)$.

By a) and b), we have $B(\Phi(x),R)_{\mathscr{F}} = \Phi(B(x,R)_X)$ and so, Φ is a submetry.

2. Assume that $\Phi : X \to Y$ is a submetry from a metric space (X,d) to a metric space (Y,d'). Let $\{\Phi^{-1}(y)\}_{y\in Y}$ denote the collection of all fibers over Φ.

i) Since Φ, as a submetry, is a continuous mapping, the preimages of singleton sets are closed as preimages of closed subsets. Hence, the fibers are closed.

ii) The fibers are non-empty since Φ is surjective, and distinct fibers are disjoint subsets in X since Φ is a function.

iii) It remains to verify the equidistance property. Let $y_1, y_2 \in Y$ be arbitrary and consider the fibers $F_1 := \Phi^{-1}(y_1)$, $F_2 := \Phi^{-1}(y_2)$. We have $F_1 \neq \emptyset$, let $x_1 \in F_1$ be an arbitrary member. We set $R := d(F_1, F_2) \in \mathbb{R}_{\geq 0}$. Let $\varepsilon \in \mathbb{R}_{>0}$ be arbitrary. By the submetry property of Φ, we have

$$\Phi(B(x_1, R+\varepsilon)_X) = B(y_1, R+\varepsilon)_Y.$$

Hence, noting that $y_2 \in B(y_1, R+\varepsilon)_Y$, there exists a preimage $x_2 \in X$ such that $\Phi(x_2) = y_2$ and $x_2 \in B(x_1, R+\varepsilon)_X$. Therefore,

$$d(x_1, F_2) \leq d(x_1, x_2) < R+\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$d(x_1, F_2) \leq R = d(F_1, F_2),$$

and so $d(x_1, F_2) = d(F_1, F_2)$. The equidistance property follows.

By (i), (ii) and (iii), the collection $\{\Phi^{-1}(y)\}_{y \in Y}$ is a metric foliation of the space (X, d). □

We conclude this section by mentioning a particular setting, in which a metric foliation naturally arises.

Definition 20. (Group action on a metric space). Let (X, d) be a metric space, and (G, e) a group. G is said to *act on* X, if and only if, the pair is equipped with a mapping

$$G \times X \to X,\ (g, x) \mapsto gx,$$

with the following properties:

(i) $\forall x \in X :\ ex = x$

(ii) $\forall g,h \in G\ \forall x \in X:\ (gh)x = g(hx)$

G is said to *act on X by isometries* if in addition we have:

$$\forall g \in G\ \forall x,y \in X:\ d(gx,gy) = d(x,y)$$

Remark 4. If a group G acts on a metric space (X,d) by isometries, and if the orbits

$$[x] := \{gx,\ g \in G\}, \qquad x \in X,$$

are closed subsets in X, then the equivalence relation on X defined by

$$x \sim y : \Longleftrightarrow\ \exists g \in G : y = gx, \qquad x,y \in X,$$

gives rise to a metric foliation of X, and so induces a corresponding metric quotient space denoted as $(X/G,d)$. Indeed, the group structure of G implies the equidistance property: For $[x],[y] \in X/G$,

$$\begin{aligned} d([x],[y]) &:= \inf\{d(x',y');\ x' \in [x], y' \in [y]\} = \inf\{d(gx,hy);\ g,h \in G\} = \\ &= \inf\{d(x,g^{-1}hy),\ h \in G\} = \inf\{d(x,h'y),\ h' \in G\} =: d(x,[y]). \end{aligned}$$

Chapter 4

Alexandrov curvature bounded below (CBB) metric spaces

After the notion of metric space as a set theoretic way of defining a geometric space was introduced, there had been interest around the beginning of 20th century in recovering common geometric notions within a purely metric space theoretic framework. A. D. Alexandrov defined metric spaces of bounded curvature as metric space generalizations of the model spaces of Riemannian geometry. See, Alexandrov (1957).

Namely, Alexandrov curvature bounded below spaces are metric spaces, which satisfy so-called synthetic lower curvature bound conditions, defined solely in terms of the metric without assuming further analytical structure of the space. Many equivalent versions of these conditions exist in contemporary literature for defining the notion of lower curvature bounds. See, for instance, Alexander et al. (2022). We start this chapter by quickly introducing these spaces, favoring a more recently discovered equivalent lower curvature bound conditions, named Kirszbraun lower curvature bound conditions. (See Alexander et al. (2011). See also, Lang & Schroeder (1997).)

Then, we prove that synthetic lower curvature bounds are stable under submetries by showing that if a Kirszbraun lower curvature bound condition holds in the domain of a submetry, then the same lower curvature bound condition must also hold in its target space. Another proof of the same result that is based on another one of the equivalent lower curvature bound conditions appears already in Alexander et al. (2022).

4.1 The definition of CBB spaces

Definition 21. (Model spaces of Riemannian geometry). For $\kappa \in \mathbb{R}$, by $\mathbb{M}_\kappa$ we shall denote a complete simply connected 2-dimensional Riemannian manifold of constant curvature κ, and by d_κ its induced metric.

Remark 5.
1. For $\kappa > 0$, $(\mathbb{M}_\kappa, d_\kappa)$ is isometrically isomorphic to the intrinsic metric space structure of the surface of a sphere with radius $\frac{1}{\sqrt{\kappa}}$ in the Euclidean space $\mathbb{R}^3$.
2. For $\kappa = 0$, $(\mathbb{M}_\kappa, d_\kappa)$ is isometrically isomorphic to the metric space structure of the Euclidean space $\mathbb{R}^2$.
3. For $\kappa < 0$, $(\mathbb{M}_\kappa, d_\kappa)$ is isometrically isomorphic to the intrinsic metric space structure of the so-called Lobachevsky surface of constant curvature κ in the Euclidean space $\mathbb{R}^3$.

See, e.g. Bär (2010).

Definition 22 (Alexandrov model triangles for a triplet of points in a metric space)**.** Let $\kappa \in \mathbb{R}$, (X,d) be a metric space, and $(p,x,y) \in X^3$ a triplet of points. We set:

$$a := d(x,y),\ b := d(y,p), c := d(p,x).$$

A geodesic triangle in $\mathbb{M}_\kappa$ with side lengths a,b,c, if it exists, shall be called a *model triangle* in $\mathbb{M}_\kappa$ corresponding to the triplet $(p,x,y) \in X^3$ and denoted by $\tilde{\Delta}^\kappa(p,x,y)$, referring back to $p,x,y \in X$. If such triangles exist, then they are all congruent. In case $\kappa \leq 0$ model triangles always exist, whereas in the case $\kappa > 0$ model triangles exist, if and only if,

$$a+b+c < 2 \cdot \operatorname{diam} \mathbb{M}_\kappa = 2 \cdot \frac{\pi}{\sqrt{k}},$$

where $\operatorname{diam} \mathbb{M}_\kappa := \sup\{d_\kappa(\hat{x},\hat{y});\ \hat{x},\hat{y} \in \mathbb{M}_\kappa\}$.

Definition 23 (Alexandrov model angle at a point, for $(1+2)$ points in a metric space)**.** Let $\kappa \in \mathbb{R}$, (X,d) be a metric space, and $(p,x,y) \in X^3$. Assume that a model triangle $\tilde{\Delta}^\kappa(p,x,y)$ corresponding

to the triplet $(p,x,y) \in X^3$ exists, and let $\hat{p},\hat{x},\hat{y} \in \mathbb{M}_\kappa$ be points that realize one such triangle in $\mathbb{M}_\kappa$ as being the vertices, with

$$d_\kappa(\hat{x},\hat{y}) = d(x,y),\ d_\kappa(\hat{y},\hat{p}) = d(y,p),\ d_\kappa(\hat{p},\hat{x}) = d(p,x).$$

Then, the angle $\angle(\hat{p};\hat{x},\hat{y})$, defined in the sense of angle between geodesics in $\mathbb{M}_\kappa$, is called the *model angle* in $\mathbb{M}_\kappa$ for the triplet $(p,x,y) \in X^3$ at p. It is defined independent of the choice of the model triangle, and is denoted by $\tilde{\angle}^\kappa(p;x,y)$. In the case that no such model triangle exists, which is possible as stated above when $\kappa > 0$, we say that the model angle $\tilde{\angle}^\kappa(p;x,y)$ does not exist.

Definition 24 (The $(1+3)$ points angle κ-comparison condition)**.** Let (X,d) be a metric space. A quadruple of points $(p,x_1,x_2,x_3) \in X^4$ is said to satisfy the *(1+3) points angle κ-comparison condition*, if and only if, one of the statements below is true.

(1) All of the model angles $\tilde{\angle}^\kappa(p;x_1,x_2)$, $\tilde{\angle}^\kappa(p;x_2,x_3)$, $\tilde{\angle}^\kappa(p;x_3,x_1)$ are defined, and

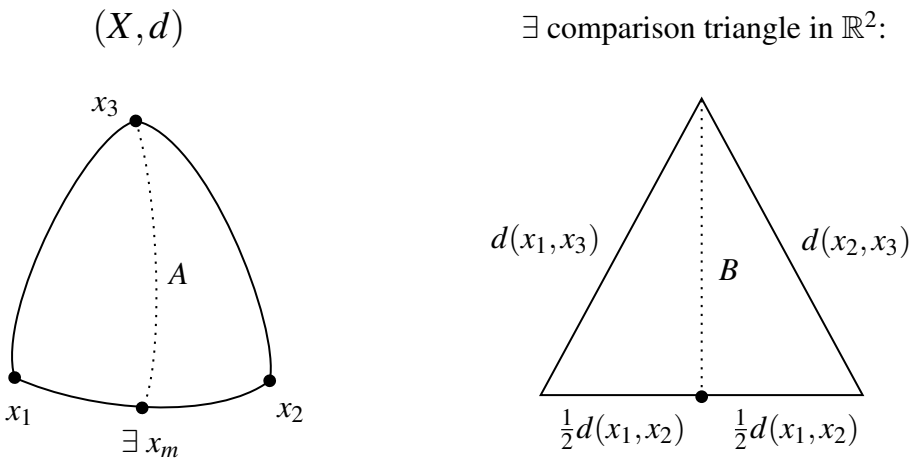
$$\tilde{\angle}^\kappa(p;x_1,x_2) + \tilde{\angle}^\kappa(p;x_2,x_3) + \tilde{\angle}^\kappa(p;x_3,x_1) \leq 2\pi.$$

(2) At least one of the model angles $\tilde{\angle}^\kappa(p;x_1,x_2)$, $\tilde{\angle}^\kappa(p;x_2,x_3)$, $\tilde{\angle}^\kappa(p;x_3,x_1)$ does not exist.

The condition (2) declares the condition in (1) to be regarded trivially verified if there is nothing to check there.

We are now ready to recall the definition of an Alexandrov metric space with curvature bounded below by $\kappa \in \mathbb{R}$. We work under a general complete length space assumption on the space.

Definition 25 (Alexandrov space with curvature bounded below by $\kappa \in \mathbb{R}$)**.** Let $\kappa \in \mathbb{R}$. A complete length space (L,d) is called *Alexandrov, curvature bounded below by* κ, if and only if, all quadruples $(p,x_1,x_2,x_3) \in L^4$ satisfy the $(1+3)$ points angle κ-comparison condition by its definition. Any Alexandrov curvature bounded below by κ metric space will briefly be called a CBB-κ space. A metric space which is CBB-0 is also called *non-negatively curved*.



A geodesic space X is called non-negatively curved, if and only if, $B \leq A$ for all choices of $x_1, x_2, x_3 \in X$.

Figure 4.1: An illustration for a definition of non-negatively curved geodesic spaces

Definition and Proposition 5. (An equivalent definition of CBB-κ spaces.) Let $\kappa \in \mathbb{R}$.

(1) A metric space (X,d) is said to satisfy the *Kirszbraun lower curvature bound κ condition*, if and only if, every 1-Lipschitz mapping $f : S_3 \to \mathbb{M}_\kappa$ which is defined on a subset $S_3 \subset X$ consisting of 3 points, can be extended to any set $S_4 \subset X$, $(S_3 \subset S_4 \subset X)$, that consists of 4 points, as a 1-Lipschitz mapping $F : S_4 \to \mathbb{M}_\kappa$, $F_{|S_3} = f$.

(2) A complete length space, which satisfies the Kirszbraun lower curvature bound κ condition, is Alexandrov curvature $\geq \kappa$.

Proof. See Theorem 3.1 in Alexander et al. (2011). □

Proposition 11. (Kirszbraun's theorem for Hilbert spaces. See, Theorem 1G in Fremlin (2018).) Let H_1 and H_2 be Hilbert spaces. Every Lipschitz mapping $f : S \to H_2$, $S \subset H_1$, possesses a Lipschitz extension $F : H_1 \to H_2$ that preserves the initial Lipschitz constant.

The above theorem was initially proved by Kirszbraun in a Euclidean space framework, where in the places of the Hilbert spaces H_1 and H_2 above, Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ for any $n, m \in \mathbb{N}$ were considered. See Kirszbraun (1934).

Corollary 3. Any Hilbert space satisfies the Kirszbraun lower curvature bound $\kappa = 0$ condition.

Hence, altogether, any pre-Hilbert space satisfies the $(1+3)$ points angle $\kappa = 0$ comparison condition.

4.2 Kirszbraun lower curvature bounds under submetries

The following results were proved in Alexander et al. (2022), we give an alternative proof by favoring the Kirszbraun lower curvature bound conditions.

Proposition 12. Let (X, d_X) and (Y, d_Y) be metric spaces, $\Phi : X \to Y$ a submetry and $\kappa \in \mathbb{R}$.

(1) If X satisfies the Kirszbraun lower curvature bound κ condition, then so does Y.

(2) If X is CBB-κ, then Y is CBB-κ as well.

Proof. Assume that X satisfies the Kirszbraun lower curvature bound κ condition. Let $Y_3 := \{y_1, y_2, y_3\} \subset Y$ be a subset containing exactly three elements, $y_4 \notin Y_3$ be a further element in Y, and $f : \{y_1, y_2, y_3\} \to \mathbb{M}_\kappa$ be 1-Lipschitz.

1. Φ is in particular surjective. We start by picking a preimage $x_4 \in X$ under Φ for the element $y_4 \in Y$, so that $\Phi(x_4) = y_4$.

2. Now, let $\varepsilon \in \mathbb{R}_{>0}$ be arbitrary. By the submetry property, we have

$$\Phi(B(x_4, R_i + \varepsilon)_X) = B(y_4, R_i + \varepsilon)_Y, \quad \text{for } i = 1, 2, 3;$$

where $R_i := d_Y(y_i, y_4)$ for $i = 1, 2, 3$. Hence, there exist preimages, $x_1^\varepsilon, x_2^\varepsilon, x_3^\varepsilon$, under Φ of the values y_1, y_2, y_3 respectively, such that

$$\begin{aligned} d_Y(y_1, y_4) \leq d_X(x_1^\varepsilon, x_4) < d_Y(y_1, y_4) + \varepsilon, \\ d_Y(y_2, y_4) \leq d_X(x_2^\varepsilon, x_4) < d_Y(y_2, y_4) + \varepsilon, \\ d_Y(y_3, y_4) \leq d_X(x_3^\varepsilon, x_4) < d_Y(y_3, y_4) + \varepsilon. \end{aligned}$$

The elements $x_1^\varepsilon, x_2^\varepsilon, x_3^\varepsilon \in X$ and the element $x_4 \in X$ are all distinct, as they reside in different fibers of Φ.

3. The mapping $f \circ \Phi : \{x_1^\varepsilon, x_2^\varepsilon, x_3^\varepsilon\} \to \mathbb{M}_\kappa$ is 1-Lipschitz as the composition of 1-Lipschitz mappings f and Φ. Indeed, for $i, j \in \{1,2,3\}$,

$$d_\kappa((f\circ\Phi)(x_i^\varepsilon),(f\circ\Phi)(x_j^\varepsilon)) = d_\kappa(f(y_i),f(y_j)) \le d_Y(y_i,y_j) = d_Y(\Phi(x_i^\varepsilon),\Phi(x_j^\varepsilon)) \le d_X(x_i^\varepsilon,x_j^\varepsilon),$$

where the metric in $\mathbb{M}_\kappa$ is denoted as d_κ.

4. Now, using the Kirszbraun lower curvature bound κ condition in X, we can extend the mapping $f \circ \Phi$ to

$$\{x_1^\varepsilon, x_2^\varepsilon, x_3^\varepsilon, x_4\}$$

as a 1-Lipschitz mapping, i.e., there exists an $\alpha^\varepsilon \in \mathbb{M}_\kappa$ such that the extension of $f \circ \Phi$ that maps $x_4 \mapsto \alpha^\varepsilon$, is 1-Lipschitz. Thence, for $i = 1,2,3$,

$$d_\kappa(f(y_i), \alpha^\varepsilon) \le d_X(x_i^\varepsilon, x_4) < d_Y(y_i, y_4) + \varepsilon.$$

5. For $n \in \mathbb{N}$, set $\varepsilon_n := \frac{1}{n}$ and $\alpha_n := \alpha_{\varepsilon_n}$. Observe that $(\alpha_n)_{n\in\mathbb{N}}$ is a sequence in the closed ball $\bar{B}(f(y_1),1)$ in $\mathbb{M}_\kappa$. The closed ball $\bar{B}(f(y_1),1)$ as embedded into the Euclidean space $\mathbb{R}^3$ is still closed and bounded, therefore by the Bolzano-Weierstrass theorem in $\mathbb{R}^3$, the sequence $(\alpha_n)_{n\in\mathbb{N}}$ viewed as a sequence in $\mathbb{R}^3$ has a subsequence $(\alpha_{k_n})_{n\in\mathbb{N}}$ that converges in $\mathbb{R}^3$ to a point α on the same constant curvature κ surface. This implies that the same convergence holds in $\mathbb{M}_\kappa$ as well.

6. We define $f(y_4) := \alpha$, which extends f to $\{y_1, y_2, y_3, y_4\}$ as a 1-Lipschitz mapping. Indeed, for $i = 1,2,3$,

$$d_\kappa(f(y_i),\alpha) = \lim_{n\to\infty} d_\kappa(f(y_i),\alpha_{k_n}) \le \lim_{n\to\infty}(d_Y(y_i,y_4) + \frac{1}{k_n}) = d_Y(y_i,y_4).$$

7. Hence, Y satisfies the Kirszbraun lower curvature bound κ condition, which establishes (1). Then, (2) follows from (1) and Proposition 9.

□

By the correspondence of submetries and metric foliations, the following is now immediately true.

Corollary 4. Let $\kappa \in \mathbb{R}$. For a metric foliation of a CBB-κ space, the corresponding induced metric quotient space is CBB-κ as well.

Proof. Indeed, as described in Proposition 10, the natural projection between the covering space and a metric quotient space of it, is a submetry. □

A particular case of the previous result is the next one.

Corollary 5. Let $\kappa \in \mathbb{R}$ and (X,d) be a CBB-κ space, G a group that acts on X by isometries with closed orbits, and $(X/G,d)$ the corresponding metric quotient space. Then, $(X/G,d)$ is CBB-κ as well.

Chapter 5

Wasserstein spaces as metric quotient spaces

In this chapter, for $p \in [1,\infty)$ and a separable Hilbert space E, we exhibit the Wasserstein space $(\mathscr{P}_p(E), W_p)$ as isometrically isomorphic to a metric quotient space induced by a metric foliation.

A glance at the definition of Wasserstein metric suggests that $(\mathscr{P}_p(E), W_p)$ may be interpreted as a metric quotient space of the Lebesgue space $L_p([0,1], \mathscr{B}([0,1]), m;\ E)$ via the projection mapping which maps a random variable to its distribution. However, a rigorous proof of the equidistance property in such context is not to be found in the literature. Indeed, such a result would relate to a corresponding lemma about couplings, limitations of which should be sorted out in its full generality. (Cf. optimal transport theory. See e.g. Villani (2003) or Rachev & Rüschendorf (1998).)

Here, for sake of producing a rigorous metric foliation result that holds under general assumptions, we consider an extension of Lebesgue spaces. Namely, we form a direct limit metric space that extends the Lebesgue spaces. Then, we show that the Wasserstein space is isometrically isomorphic to a metric quotient space induced by a metric foliation of the direct limit that we define, by applying a weaker lemma about couplings present in Kallenberg (1997). See Perrone (2021) for more abstract results that are similar to ours in spirit.

For $p = 2$ and a separable Hilbert space E, the Wasserstein space $(\mathscr{P}_2(E), W_2)$ is a non-negatively curved metric space, i.e. Alexandrov curvature $\geq \kappa = 0$. A direct proof of this fact is presented in Ambrosio et al. (2008), using one of the equivalent conditions that define the lower curvature bound by $\kappa = 0$. However, as an application of our metric foliation result, we are able to exhibit the same fact here as a special occurrence of a more general phenomenon.

Namely, in the case $p = 2$, the direct limit metric space that we define, naturally inherits the

Kirszbraun lower curvature boundedness by $\kappa = 0$ from the Lebesgue spaces that form our direct limit space. Hence, the Wasserstein space $(\mathscr{P}_2(E), W_2)$ also inherits the same property as it is isometrically isomorphic to a metric quotient space of the former space.

5.1 Non-negatively curved spaces

As already mentioned in the previous chapter, a metric space is called non-negatively curved, if it is Alexandrov curvature $\geq \kappa = 0$, i.e. CBB-0.

The following proposition gives another characterization of these spaces under a geodesic space assumption on the space.

Proposition 13. (Cf. Ambrosio et al. (2008), Definition 12.3.1.) Let (X,d) be a complete geodesic space. Then, (X,d) is CBB-0, if and only if, for any three points $x_1, x_2, x_3 \in X$ and any unit speed parametrized geodesic $\gamma : [0,1] \to X$ with $\gamma(0) = x_1$, $\gamma(1) = x_2$, the following statement holds true:

$$\forall\, t \in [0,1] :\ d^2(\gamma(t), x_3) \geq (1-t)d^2(x_1,x_3) + td^2(x_2,x_3) - t(1-t)d^2(x_1,x_2)$$

Proof. See Theorem 8.23 in Alexander et al. (2022), and Sturm (1999). □

Using this characterization along with some knowledge about couplings, for separable Hilbert space E, the Wasserstein space $(\mathscr{P}_2(E), W_2)$ has already been shown to be a non-negatively curved metric space.

Proposition 14. (See, Ambrosio et al. (2008), Theorem 7.3.2.) Let E be a separable Hilbert space. Then, the Wasserstein space $(\mathscr{P}_2(E), W_2)$ is non-negatively curved.

5.2 A lemma on couplings

For proving our metric foliation result we shall make use of the following lemma about couplings. See, Kallenberg (1997). In fact, the formulation of our metric foliation result is tailored upon this lemma.

Here is an unabridged reformulation of the lemma presented in Theorem 5.10 of Kallenberg (1997).

Lemma 3. (On existence of a random variable, which coupled with a given random variable jointly have a prescribed joint distribution. Kallenberg (1997), Theorem 5.10.)

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, $(S, \mathscr{G})$ and $(T, \mathscr{H})$ Borel spaces, X and X' two random variables $\Omega \to S$ having the same distribution, and $Y : \Omega \to T$ be a given random variable. Let $\tilde{X}' : \Omega \times [0,1] \to S$ be defined by

$$\forall\ (\omega, x) \in \Omega \times [0,1] :\ \tilde{X}'(\omega, x) := X'(\omega),$$

and consider $([0,1], \mathscr{B}([0,1]))$ equipped with the Lebesgue measure m.

Then, there exists a random variable $\tilde{Y} : \Omega \times [0,1] \to T$ having the same distribution as Y such that the joint random variables $(\tilde{X}', \tilde{Y})$ and (X, Y), which are defined on different probability spaces, have the same distribution on the product space $(S \times T, \mathscr{G} \otimes \mathscr{H})$.

We include the proof of this lemma here for sake of completeness.

Proof. 1. By Proposition 5, there exists a regular conditional probability of Y given X, i.e. a probability kernel $\mu : S \times \mathscr{H} \to [0,1]$ with

$$\forall B \in \mathscr{H} :\ \mathbb{P}(Y \in B \mid X) = \mu(X(\cdot), B) \text{ a.s.}$$

2. Let Θ be a uniformly distributed random variable $([0,1], \mathscr{B}([0,1]), m) \to [0,1]$ defined on an auxiliary copy of $([0,1], \mathscr{B}([0,1]), m)$, so that it is defined independently of X'.

By Proposition 4, there exists a measurable function $f : S \times [0,1] \to T$, such that $f(s, \Theta)$ has distribution $\mu(s, \cdot)$ for every $s \in S$ fixed. Define, $\tilde{Y} := f(X', \Theta)$.

3. Let $g : S \times [0,1] \to [0, \infty)$ be an arbitrary bounded measurable function. Let the notations $\mathbb{E}$ and $\tilde{\mathbb{E}}$ distinguish between expectations on Ω and $\Omega \times [0,1]$, in that order. Then, by conditioning,

Fubini's theorem, and using that $\tilde{X}' \stackrel{\mathscr{D}}{=} X'$ and $X' \stackrel{\mathscr{D}}{=} X$, we have

$$\tilde{\mathbb{E}}[g(\tilde{X}',\tilde{Y})] = \tilde{\mathbb{E}}[g(\tilde{X}',f(X',\Theta))] = \tilde{\mathbb{E}}[\,\mathbb{E}[g(\tilde{X}',f(X',\Theta))] \mid \Theta] =$$
$$= \int_{[0,1]} \mathbb{E}[g(X',f(X',u))]m(du) = \int_{[0,1]} \mathbb{E}[g(X,f(X,u))]m(du) = \mathbb{E}[\int_{[0,1]} g(X,f(X,u))m(du)],$$

where the independent nature of Θ is used in the third equality.

4. On the other hand, by disintegration

$$\mathbb{E}[g(X,Y)] = \mathbb{E}[\int_T g(X(\cdot),y)\mu(X(\cdot),dy)],$$

and using the fact that $f(s,\Theta) \stackrel{\mathscr{D}}{=} \mu(s,\cdot)$ for all $s \in S$ fixed, we have

$$\forall s \in S: \int_T g(s,y)\mu(s,dy) = \int_{[0,1]} g(s,f(s,u))m(du),$$

and so,

$$\forall \omega \in \Omega: \int_T g(X(\omega),y)\mu(X(\omega),dy) = \int_{[0,1]} g(X(\omega),f(X(\omega),u))m(du),$$

so that

$$\mathbb{E}[\int_T g(X(\cdot),y)\mu(X(\cdot),dy)] = \mathbb{E}[\int_{[0,1]} g(X(\cdot),f(X(\cdot),u))m(du)].$$

5. Altogether, by steps 3 and 4,

$$\tilde{\mathbb{E}}[g(\tilde{X}',\tilde{Y})] = \mathbb{E}[g(X,Y)],$$

and since g was arbitrary, $(\tilde{X}',\tilde{Y}) \stackrel{\mathscr{D}}{=} (X,Y)$.

□

5.3 Wasserstein spaces as metric quotient spaces induced by metric foliation

5.3.1 Direct limit of metric spaces

We shall define a notion of direct limit for certain sequences of metric spaces, as a metric space, by applying the notion of direct limit (a.k.a. inductive limit, or co-limit) from category theory in a category of metric spaces. (See e.g. Jost (2019).)

Informally, we recall that a category consists of a collection of so-called "*objects*", and a collection of so-called "*morphisms*" (or "*arrows*") between the objects equipped with a binary relation called "*composition*", denoted by "$\circ$", such that

(i) If $f : A \to B$ and $g : B \to C$ are morphisms, then there exists a unique morphism $g \circ f$, which is a morphism $A \to C$.

(ii) For each object A, there exists a morphism denoted as $\mathrm{id}_A : A \to A$ such that for any morphism of the forms $g : A \to B$ or $h : C \to A$, we have $g \circ \mathrm{id}_A = g$ and $\mathrm{id}_A \circ h = h$, respectively.

(iii) Composition satisfies associativity.

We can consider the category in which the "objects" are metric spaces and isometric embeddings between metric spaces are identified as the "morphisms". By applying the notions of *direct system* and *direct limit* in this particular example of a category, we arrive at the following definition.

Definition and Proposition 6. (Direct limit of a direct system of metric spaces). Let (M_n, d_n), $n \in \mathbb{N}$ be a sequence of metric spaces equipped with a collection of isometric embeddings,

$$\mathrm{S} := \{\Phi_{n,k} \mid n,k \in \mathbb{N},\ k \geq n\},$$

with the following specification and properties:

(i) $\forall n,k \in \mathbb{N} :\ n < k \implies \Phi_{n,k} :\ M_n \hookrightarrow M_k$

(ii) $\forall n \in \mathbb{N} :\ \Phi_{n,n} = \mathrm{id}$

(iii) $\forall n,l,k \in \mathbb{N}:\ n<l<k \implies \Phi_{l,k}\circ\Phi_{n,l}=\Phi_{n,k}$

Then, (M_n,d_n), $n\in\mathbb{N}$ equipped with the collection S is called a *direct system* of metric spaces. Consider the formal disjoint union of the underlying sets of the metric spaces, i.e. the set

$$\bigsqcup_{n\in\mathbb{N}} M_n,$$

and the equivalence relation on it, denoted $\sim$, which identifies two elements $x\in M_n$, $y\in M_k$, if and only if, there exists $l\in\mathbb{N}$ ($l>n$, $l>k$) such that $\Phi_{n,l}(x)=\Phi_{k,l}(y)$. We equip the quotient set

$$\mathscr{M} := (\bigsqcup_{n\in\mathbb{N}} M_n)\big/\!\sim$$

with the well-defined distance defined by

$$D([x],[y]) := \lim_{l\to\infty} d_l(\Phi_{n,l}(x),\Phi_{k,l}(y)), \quad x\in M_n,\ y\in M_k,$$

where $[z]$ denotes the equivalence class of an element $z\in\bigsqcup_{n\in\mathbb{N}} M_n$, noting that if $k\geq n$, we already have

$$D([x],[y]) = d_k(\Phi_{n,k}(x),y).$$

The pair $(\mathscr{M},D)$ form a metric space called the *direct limit metric space* of the direct system of metric spaces (M_n,d_n), $n\in\mathbb{N}$, under S.

The direct limit metric space shall be briefly denoted also as $\mathscr{M}=\varinjlim M_n$.

Proof. Indeed, as it suffices to work with finitely many elements at a time when verifying the properties of the metric for D, we can choose representatives for those classes and use the corresponding properties for the representatives in a sufficiently large indexed metric space, say (M_N,d_N), that contains all of those representatives as embedded into it. □

5.3.2 Metric foliations of direct limits that induce the Wasserstein spaces

In this subsection, we apply the notion of direct limit metric space presented above in a concrete setting to construct the extension of Lebesgue spaces, that will serve as the covering space in which our metric foliation result takes place.

Let $p \in [1,\infty)$ and E be a separable Hilbert space. We consider the Lebesgue spaces

$$L_p^{(n)}(E) := L_p([0,1]^n, \mathscr{B}([0,1]^n), \otimes_{i=1}^n m;\ E), \qquad n \in \mathbb{N},$$

and view them as metric spaces. For $k,n \in \mathbb{N}$, $k > n$, we have canonical isometric embeddings,

$$\begin{aligned} \Phi_{n,k} :\ L_p^{(n)}(E) &\hookrightarrow L_p^{(k)}(E), \\ X &\mapsto \tilde{X}, \end{aligned}$$

where for $\omega = (\omega_i)_{i=1,\dots,k} \in [0,1]^k$,

$$\tilde{X}(\omega_1,\omega_2,\omega_3,\dots,\omega_n,\omega_{n+1},\dots,\omega_k) := X(\omega_1,\omega_2,\omega_3,\dots,\omega_n),$$

and the identity mappings $\Phi_{n,n} : L_p^{(n)}(E) \to L_p^{(n)}(E)$, $n \in \mathbb{N}$. Then, the sequence $L_p^{(n)}(E)$, $n \in \mathbb{N}$ equipped with the collection $\{\Phi_{n,k} \mid n,k \in \mathbb{N},\ k \geq n\}$ is a direct system of metric spaces, so that we can define the direct limit metric space $\mathscr{L}_p(E) := \varinjlim L_p^{(n)}(E)$, as a metric space $(\mathscr{L}_p(E), D)$, which we briefly denoted as $\mathscr{L}_p(E)$. That is, as a set

$$\mathscr{L}_p(E) := \Big(\bigsqcup_{n\in\mathbb{N}} L_p^{(n)}(E)\Big)\Big/\!\sim,$$

i.e., the disjoint union of the sets $L_p^{(n)}(E)$, $n \in \mathbb{N}$ modulo an equivalence relation, denoted by $\sim$, which identifies two elements $X \in L_p^{(n)}(E)$, $Y \in L_p^{(k)}(E)$, if and only if, there exists $l \in \mathbb{N}$ ($l > n$,

$l > k$) such that $\Phi_{n,l}(X) = \Phi_{k,l}(Y)$. The metric D on $\mathscr{L}_p(E)$ is defined by setting

$$D([X],[Y]) := \lim_{l\to\infty} ||\Phi_{n,l}(X) - \Phi_{k,l}(Y)||, \quad X \in L_p^{(n)}(E),\ Y \in L_p^{(k)}(E),$$

where $[Z]$ denotes the equivalence class of an element $Z \in \bigsqcup_{n\in\mathbb{N}} L_p^{(n)}(E)$, the representation is independent of the choice of representative for the class and

$$D([X],[Y]) = ||\Phi_{n,k}(X) - Y||, \quad \text{if } k > n.$$

$(\mathscr{L}_p(E), D)$ is an extension of each of the spaces $L_p^{(n)}(E)$, $n \in \mathbb{N}$, as a metric space via the isometric embeddings

$$L_p^{(n)}(E) \hookrightarrow \mathscr{L}_p(E),\ X \mapsto [X], \quad n \in \mathbb{N}.$$

Remark 6. In the above situation, the sets $L_p^{(n)}(E)$, $n \in \mathbb{N}$, embedded as closed subsets into the metric space $(\mathscr{L}^p(E), D)$, converge in the sense of Hausdorff to the set $\mathscr{L}^p(E)$. (See, Definition 9.)

Remark 7. In the case $p = 2$, if E is a separable Hilbert space, the direct limit metric space $\varinjlim L_p^{(n)}(E)$ naturally inherits a pre-Hilbert space structure from the Lebesgue spaces that form the direct limit. This amounts to considering the direct limit $\varinjlim L_p^{(n)}(E)$ in a category of pre-Hilbert spaces as a pre-Hilbert space. Considering the same direct limit in a category of Hilbert spaces may require a completion step.

Theorem. *Let $p \in [1,\infty)$ and E be a separable Hilbert space. Then:*

(1) The Wasserstein space $(\mathscr{P}_p(E), W_p)$ is isometrically isomorphic to a metric quotient space induced by a metric foliation of $(\mathscr{L}_p(E), D)$, where the latter metric space is defined as the direct limit metric space

$$\mathscr{L}_p(E) := \varinjlim L_p^{(n)}(E),$$

of the direct system of metric spaces

$$L_p^{(1)}(E) \hookrightarrow L_p^{(2)}(E) \hookrightarrow L_p^{(3)}(E) \hookrightarrow \ldots \hookrightarrow \mathscr{L}_p(E),$$

where for $n \in \mathbb{N}$,

$$L_p^{(n)}(E) := L_p([0,1]^n, \mathscr{B}([0,1]^n), \otimes_{i=1}^n m;\ E)$$

is defined as a Lebesgue space.

(2) The mapping

$$\Psi : \mathscr{L}_p(E) \to \mathscr{P}_p(E),$$
$$[X] \mapsto \mu_X,$$

where μ_X denotes the distribution of any representative of the class $[X]$, is well-defined and a submetry.

Proof. For $\rho \in \mathscr{P}_p(E)$, we define

$$F_\rho := \{[X] \in \mathscr{L}_p(E),\ X \overset{\mathscr{D}}{=} \rho\},$$

i.e. the collection of all equivalence classes $[X] \in \mathscr{L}_p(E)$ that possess a representative X in $\bigsqcup_{n\in\mathbb{N}} L_p^{(n)}(E)$ that has distribution ρ. The definition of F_ρ is consistent in respect to the choice of representatives due to the canonical nature of our constructions.

Also, it is worth noting here that within any of the fixed Lebesgue spaces $L_p^{(n)}(E)$ no issues arise about the definition of F_ρ, because almost surely equal random variables have the same distribution. (Recall that elements in Lebesgue spaces are defined under a.s. equivalence.)

For each ρ, we have $F_\rho \neq \emptyset$, because, since $(E, \mathscr{B}(E))$ is a Polish space there exists a random variable $([0,1], \mathscr{B}([0,1]), m) \to (E, \mathscr{B}(E))$ with the distribution ρ. Define $\mathscr{F} := \{F_\rho,\ \rho \in \mathscr{P}_p(E)\}$.

i) Claim: For each $\rho \in \mathscr{P}_p(E)$, F_ρ is closed in $\mathscr{L}_p(E)$. We postpone the proof of this claim until the end; we shall prove it using the positive definiteness of the Wasserstein metric.

ii) For all $\rho, \gamma \in \mathscr{P}_p(E)$, we have $F_\rho \cap F_\gamma = \emptyset$, because the distribution of a random variable is a uniquely defined object as a set function.

iii) We have

$$\mathscr{L}_p(E) \ = \ \bigcup_{\rho \in \mathscr{P}_p(E)} F_\rho .$$

Indeed, for any element $[X] \in \mathscr{L}_p(E)$ the distribution μ_X is necessarily an element of $\mathscr{P}_p(E)$, so that $[X] \in F_{\mu_X}$. Conversely, for each $\rho \in \mathscr{P}_p(E)$, there exists $Y \in L_p^{(1)}(E)$ with the distribution ρ, so that $[Y] \in F_\rho$.

iv) a) We observe that, using Remark 2, we have for $\rho, \gamma \in \mathscr{P}_p(E)$,

$$\begin{aligned}
&D(F_\rho, F_\gamma) := \inf\{D([X],[Y]);\ X \overset{\mathscr{D}}{=} \rho,\ Y \overset{\mathscr{D}}{=} \gamma\} = \\
&= \inf_{n\in\mathbb{N}} \inf\{||X-Y||_p;\ X,Y \in L_p^{(n)}(E),\ X \overset{\mathscr{D}}{=} \rho,\ Y \overset{\mathscr{D}}{=} \gamma\} = \\
&= \inf_{n\in\mathbb{N}} \inf\{(\mathbb{E}^{(n)}[|X-Y|^p])^{\frac{1}{p}};\ X,Y \in L_p^{(n)}(E),\ X \overset{\mathscr{D}}{=} \rho,\ Y \overset{\mathscr{D}}{=} \gamma\} = \\
&= \inf_{n\in\mathbb{N}} \inf\{\left(\int_{E\times E} |x-y|^p \pi(dx,dy)\right)^{\frac{1}{p}};\ \pi \in \mathscr{P}(E\times E),\ p_{1\#}\pi = \rho,\ p_{2\#}\pi = \gamma\} = \\
&= \inf\{\left(\int_{E\times E} |x-y|^p \pi(dx,dy)\right)^{\frac{1}{p}};\ \pi \in \mathscr{P}(E\times E),\ p_{1\#}\pi = \rho,\ p_{2\#}\pi = \gamma\} = W_p(\rho,\gamma).
\end{aligned}$$

iv) b) Now, for verifying the equidistance property for $(\mathscr{F}, D)$, let $F_\rho, F_\gamma \in \mathscr{F}$ be arbitrary, and $[X] \in F_\rho$ with $X \in L_p^{(n)}(E)$ for some $n \in \mathbb{N}$.

Let $\varepsilon \in \mathbb{R}_{>0}$ be arbitrary, and $\pi_0 \in \Gamma(\rho,\gamma)$ be a coupling with

$$-\varepsilon + \left(\int_{E\times E} |x-y|^p \pi_0(dx,dy)\right)^{\frac{1}{p}} \leq W_p(\rho,\gamma) = D(F_\rho, F_\gamma).$$

There exist random variables $X_0, Y_0 \in L_p^{(1)}(E)$ with $X_0 \overset{\mathscr{D}}{=} \rho$, $Y_0 \overset{\mathscr{D}}{=} \gamma$ and jointly $(X_0,Y_0) \overset{\mathscr{D}}{=} \pi_0$. (Indeed, since $(E\times E, \mathscr{B}(E\times E))$ is a Polish space, there exists a random variable $[0,1] \to E\times E$ that has the distribution π_0. Composing this random variable with the projections to the components of $E\times E$, we obtain the two random variables X_0, Y_0.)

iv) c) We embed the random variables X_0 and Y_0 into $L_p^{(n)}(E)$ using the canonical embedding $\Phi_{1,n}$ from the definition of the direct limit $\mathscr{L}_p(E)$. Using Lemma 3, via randomization there exists a random variable $Y \in L_p^{(n+1)}(E)$ such that $(\Phi_{1,n}(X_0), \Phi_{1,n}(Y_0))$ and $(\Phi_{n,n+1}(X), Y)$ have the same distribution.

iv) d) Finally, defining

$$\begin{aligned}
\tilde{X_0} &:= \Phi_{1,n+1}(X_0) \in L_p^{(n+1)}(E),\\
\tilde{Y_0} &:= \Phi_{1,n+1}(X_0) \in L_p^{(n+1)}(E),\\
\tilde{X} &:= \Phi_{n,n+1}(X) \in L_p^{(n+1)}(E),
\end{aligned}$$

we have $(\tilde{X}, Y) \overset{\mathscr{D}}{=} (\tilde{X_0}, \tilde{Y_0}) \overset{\mathscr{D}}{=} \pi_0$. Altogether,

$$\begin{aligned}
D(F_\rho, F_\gamma) = W_p(\rho,\gamma) \geq -\varepsilon + \left(\int_{E\times E} |x-y|^p \pi_0(dx,dy)\right)^{\frac{1}{p}} =\\
= \left(\mathbb{E}^{(n+1)}[|\tilde{X_0} - \tilde{Y_0}|^p]\right)^{\frac{1}{p}} = \left(\mathbb{E}^{(n+1)}[|\tilde{X} - Y|^p]\right)^{\frac{1}{p}} = D([\tilde{X}],[Y]) = D([X],[Y]),
\end{aligned}$$

so that $D([X], F_\gamma) \leq D(F_\rho, F_\gamma) + \varepsilon$.

Since $\varepsilon \in \mathbb{R}_{>0}$ is arbitrary, we can conclude that $D([X], F_\gamma) \leq D(F_\rho, F_\gamma)$, which immediately implies $D([X], F_\gamma) = D(F_\rho, F_\gamma)$ by definitions. The equidistance property follows for $\mathscr{F}$, since $\rho, \gamma \in \mathscr{P}_p(E)$ and $[X] \in F_\rho$ were chosen arbitrarily.

Proof of the claim in i): Let $([X_n])_{n\in N}$ be an arbitrary sequence in F_ρ that convergences in $\mathscr{L}_p(E)$ to some element $[X]\in\mathscr{L}_p(E)$. Let μ denote the distribution of X, so that $[X]\in F_\mu$. Thence,

$$W_p(\mu,\rho)=D(F_\mu,F_\rho)\leq D([X_n],[X])\to 0,$$

implies that $W_p(\mu,\rho)=0$, so that $\mu=\rho$ using the positive definiteness of the Wasserstein metric. Therefore, $X\in F_\rho$, and F_ρ is closed.

We conclude the following:

(1) By the argumentation presented in (i)-(iv) above, $(\mathscr{F},D)$ is a metric foliation of $\mathscr{L}_p(E)$. By (iv)(a), the mapping $\zeta:\mathscr{P}_p(E)\to\mathscr{F}$, $\rho\mapsto F_\rho$ is an isometric embedding, which is bijective, since the partition $\mathscr{F}$ is indexed by $\mathscr{P}_p(E)$, and so, an isometry.

(2) Since $(\mathscr{F},D)$ is a metric foliation of $\mathscr{L}_p(E)$, the natural projection mapping

$$\begin{aligned}\Phi:\mathscr{L}_p(E)&\to\mathscr{F},\\ [X]&\mapsto F_{\mu_X},\end{aligned}$$

is a submetry. Thence, the mapping

$$\Psi:\mathscr{L}_p(E)\to\mathscr{P}_p(E),\quad \Psi([X]):=\mu_X=\zeta^{-1}(\Phi([X])),$$

is also a submetry as a composition of a submetry and an isometry.

□

Remark 8. In the proof of Theorem 5.3.2, apart from the positive definiteness, we do not make use of the rest of the defining properties of the metric of $(\mathscr{P}_p(E),W_p)$. Hence, when positive definiteness is shown separately, then our results also imply that $(\mathscr{P}_p(E),W_p)$ is a metric space in a logically consistent order, since any metric quotient space induced by a foliation of a metric space is a metric space.

Corollary 6. For a separable Hilbert space E, the Wasserstein space $(\mathscr{P}_2(E), W_2)$ is non-negatively curved.

Proof. For a separable Hilbert space E the Lebesgue spaces $L_2^{(n)}(E)$, $n \in \mathbb{N}$ satisfy the Kirszbraun lower curvature bound $\kappa = 0$ condition as Hilbert spaces, and the direct limit metric space $\mathscr{L}_2(E)$ inherits the Kirszbraun lower curvature bound $\kappa = 0$ condition from the spaces $L_2^{(n)}(E)$, $n \in \mathbb{N}$ that form the direct limit. Indeed, the condition is formulated just using 4 arbitrary points in the space at a time, which for a particular choice of 4 points in $\varinjlim L_2^{(n)}(E)$ can all already be embedded as elements in a common Lebesgue space $L_2^{(N)}(E)$ for a fixed $N \in \mathbb{N}$, in which the condition will readily be satisfied.

Then, using Proposition 12, Proposition 10 and the Theorem, we follow that $(\mathscr{P}_2(E), W_2)$ satisfies the Kirszbraun lower curvature bound $\kappa = 0$ condition as well.

Similarly, $(\mathscr{P}_2(E), W_2)$ is a geodesic space, as it inherits the property from $\mathscr{L}_p(E)$, which inherits that property from the Lebesgue spaces $L_2^{(n)}(E)$, $n \in \mathbb{N}$. Hence, $(\mathscr{P}_2(E), W_2)$ is CBB-0, as it can be verified separately that it is a complete metric space.

□

Remark 9. An analogue of our Theorem may be formulated for certain generalized Wasserstein spaces, where the distance induced by the norm that appears in the integrand in the definition of the Wasserstein metric is replaced with a more general *cost function*, in which case the Lebesgue spaces of our result are to be replaced e.g. with corresponding Orlicz spaces (cf., Sturm (2011)), or depending on the level of generality, with metric space analogues of Lebesgue spaces which are not necessarily vector spaces.

Conclusion

In this thesis, for $p \in [1,\infty)$ and a separable Hilbert space E, the Wasserstein space $(\mathscr{P}_p(E), W_p)$ has been shown to be isometrically isomorphic to a metric quotient space induced by a metric foliation, which is a way of quotient construction that ties together the metric geometry structure of the quotient to that of the covering space.

The space $\mathscr{L}_p(E)$ that we have constructed for its purpose, may be useful when handling certain problems about Wasserstein spaces, as it was introduced as a direct limit of a canonical sequence of Lebesgue spaces and its metric geometry structure induces that of $(\mathscr{P}_p(E), W_p)$ through the submetry projection $\Psi : \mathscr{L}_p(E) \to \mathscr{P}_p(E)$.

As a first example supporting the usefulness of our results, for a separable Hilbert space E, we have applied it to demonstrate that the Wasserstein space $(\mathscr{P}_2(E), W_2)$ is non-negatively curved as a metric quotient space of a non-negatively curved metric space.

References

Alexander, S., Kapovitch, V., & Petrunin, A. (2011). Alexandrov meets Kirszbraun. *Proceedings of the Gökova Geometry-Topology Conference, 88–109*.

Alexander, S., Kapovitch, V., & Petrunin, A. (2022). Alexandrov geometry: Foundations. *Preprint arXiv:1903.08539v2*.

Alexandrov, A. D. (1957). Über eine Verallgemeinerung der Riemannschen Geometrie. *Schriftenreihe des Forschungsinstituts für Mathematik bei der Deutschen Akademie der Wissenschaften zu Berlin 1: 33-84*.

Amann, H. & Escher, J. (2001). *Analysis III*. Birkhäuser.

Ambrosio, L. & Feng, J. (2014). On a class of first order Hamilton-Jacobi equations in metric spaces. *Journal of Differential Equations 256: 2194-2245*.

Ambrosio, L., Gigli, N., & Savaré, G. (2008). *Gradient flows: in metric spaces and in the space of probability measures*. Birkhäuser.

Bär, C. (2010). *Elementary differential geometry*. Cambridge University Press.

Billingsley, P. (2013). *Convergence of probability measures*. John Wiley and Sons.

Blackwell, D. & Dubins, L. E. (1983). An extension of Skorohod's almost sure representation theorem. *Proceedings of the American Mathematical Society 89: 691-692*.

Bridson, M. R. & Haefliger, A. (2013). *Metric spaces of non-positive curvature*. Springer.

Burago, D., Burago, Y., & Ivanov, S. (2001). *A course in metric geometry*. American Mathematical Society.

Cardaliaguet, P. (2010). *Notes on mean field games*. Technical report, Université de Paris - Dauphine.

Dawson, D. A. & Gärtner, J. (1987). Large deviations from the Mckean-Vlasov limit for weakly interacting diffusions. *Stochastics: An International Journal of Probability and Stochastic Processes 20: 247-308*.

Dudley, R. M. (1989). *Real analysis and probability*. Wadsworth and Brooks/Cole.

Fan, K. (1943). Entfernung zweier zufälligen Grössen und die Konvergenz nach Wahrscheinlichkeit. *Mathematische Zeitschrift 49: 681-683*.

Feng, J. & Kurtz, T. G. (2006). *Large deviations for stochastic processes*. American Mathematical Society.

Fremlin, D. H. (2018). Kirszbraun's theorem. *Preprint*.

Galaz-García, F., Kell, M., Mondino, A., & Sosa, G. (2018). On quotients of spaces with Ricci curvature bounded below. *Journal of Functional Analysis 275: 1368-1446*.

Halmos, P. (1950). *Measure theory*. Van Nostrand.

Hofinger, A. (2006). *Assessing uncertainty in linear inverse problems with the metrics of Ky Fan and Prokhorov*. Technical report, Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences.

Jost, J. (2019). *Kategorientheorie: Eine kurze Einführung*. Springer.

Kallenberg, O. (1997). *Foundations of modern probability*. Springer.

Kirszbraun, M. D. (1934). Über die zusammenziehende und Lipschitzsche Transformationen. *Fundamenta Mathematicae 22: 77–108*.

Lang, U. (2013). Length spaces. *Lecture notes, ETH Zürich*.

Lang, U. & Schroeder, V. (1997). Kirszbraun's theorem and metric spaces of bounded curvature. *Geometric and Functional Analysis 7: 535-560*.

Parthasarathy, K. R. (1967). *Probability measures on metric spaces*. Academic Press.

Perrone, P. (2021). Lifting couplings in Wasserstein spaces. *Preprint arXiv:2110.06591v1*.

Rachev, S. T. & Rüschendorf, L. (1998). *Mass transportation problems*. Springer.

Schneider, R. (2001). Analysis. *Lecture notes, Universität Freiburg*.

Skorokhod, A. V. (1957). Limit theorems for stochastic processes with independent increments. *Theory of Probability and Its Applications 2: 138-171*.

Sosa, G. (2017). On symmetric transformations in metric measured geometry. *Ph.D. Thesis, Universität Leipzig*.

Srivastava, S. M. (2008). *A course on Borel sets*. Springer.

Strassen, V. (1965). The existence of probability measures with given marginals. *The Annals of Mathematical Statistics 36: 423-439*.

Sturm, K.-T. (1999). Metric spaces of lower bounded curvature. *Expositiones Mathematicae 17: 35-48*.

Sturm, K.-T. (2011). Generalized Orlicz spaces and Wasserstein distances for convex-concave scale functions. *Bulletin des Sciences Mathématiques 135: 795-802*.

Villani, C. (2003). *Topics in optimal transportation*. American Mathematical Society.

Villani, C. (2009). *Optimal transport: Old and new*. Springer.

Appendix A

Strassen's theorem and the relation between Ky Fan and the Prokhorov metrics

In this appendix we present a relationship between the Ky Fan and the Prokhorov metrics, which we make use of in formulating a metric foliation result in Proposition 17. The following discussion is based on the exposition presented in Hofinger (2006) with some modification.

Let (E,d) be a complete separable metric space. Recall that the *Ky Fan distance* d_F between two random variables $X,Y : (\Omega,\mathscr{F},\mathbb{P}) \to (E,\mathscr{B}(E))$ is defined as

$$d_F(X,Y) := \inf\{\varepsilon \in \mathbb{R}_{>0},\ \mathbb{P}(d(X,Y) \geq \varepsilon) < \varepsilon\},$$

and that the *Prokhorov distance* D_P between two probability measures μ,ν on $(E,\mathscr{B}(E))$ is defined by

$$D_P(\mu,\nu) := \inf\{\varepsilon \in \mathbb{R}_{>0} \mid \forall B \in \mathscr{B}(E) :\ \mu(B) \leq \nu(B^\varepsilon) + \varepsilon\},$$

where $B^\varepsilon := \{x \in E,\ d(x,B) < \varepsilon\}$. Furthermore, let $\overline{B^\varepsilon} := \{x \in E,\ d(x,B) \leq \varepsilon\}$ for $\varepsilon \in \mathbb{R}_{\geq 0}$.

Proposition 15. (Strassen's theorem.) Let (E,d) be a complete separable metric space, μ and ν probability measures on $(E,\mathscr{B}(E))$ and $\alpha,\beta \in \mathbb{R}_{\geq 0}$. Then, the following two statements are equivalent:

(1) $\forall B \in \mathscr{B}(E) :\ \mu(B) \leq \nu(\overline{B^\alpha}) + \beta$

(2) There exists a coupling of μ and ν, i.e. a probability measure π on $(E \times E, \mathscr{B}(E) \otimes \mathscr{B}(E))$

with $p_{1\#}\pi = \mu$ and $p_{2\#}\pi = \nu$, that satisfies

$$\pi(\{(x,y) \in E \times E,\ d(x,y) > \alpha\}) \leq \beta.$$

See, Strassen (1965), Dudley (1989).

Proposition 16. Let (E,d) be a complete separable metric space, $(R([0,1],\mathscr{B}([0,1]),m;\ E),\ d_F)$ denote the Ky Fan space defined over $([0,1],\mathscr{B}([0,1]),m)$, see Definition and Proposition 1, and let D_P denote the Prokhorov metric on $\mathscr{P}(E)$. Then, for any $\mu,\nu \in \mathscr{P}(E)$, we have

$$D_P(\mu,\nu) = \inf\{d_F(X,Y) \mid X,Y \in R([0,1],\mathscr{B}([0,1]),m;\ E),\ X \overset{\mathscr{D}}{=} \mu,\ Y \overset{\mathscr{D}}{=} \nu\}.$$

Proof. 1. Letting $\varepsilon_0 := D_P(\mu,\nu)$, by the definition of Prokhorov metric the following statement holds true.

$$\forall B \in \mathscr{B}(E):\ \mu(B) \leq \nu(B) + \varepsilon_0 \leq \nu(\overline{B^{\varepsilon_0}}) + \varepsilon_0.$$

Then, Strassen's theorem implies that there exists a coupling π on $(E \times E, \mathscr{B}(E) \otimes \mathscr{B}(E))$ with $p_{1\#}\pi = \mu$ and $p_{2\#}\pi = \nu$, that satisfies

$$\pi(\{(x,y) \in E \times E,\ d(x,y) > \varepsilon_0\}) \leq \varepsilon_0. \tag{A.1}$$

2. Furthermore, since $(E \times E, \mathscr{B}(E) \otimes \mathscr{B}(E))$ is Polish, there exists a random variable $\phi : ([0,1],\mathscr{B}([0,1]),m) \to E \times E$ with the distribution π, where m denotes Lebesgue measure on $([0,1],\mathscr{B}([0,1]))$. Then, $X := p_1 \circ \phi$ and $Y := p_2 \circ \phi$ are random variables $([0,1],\mathscr{B}([0,1]),m) \to E$ with distributions μ and ν respectively, and the joint random variable (X,Y) has distribution π. Thence, Inequality A.1 reads

$$m(d(X,Y) > \varepsilon_0) \leq \varepsilon_0 = D_P(\mu,\nu),$$

which implies

$$d_F(X,Y) \leq D_P(\mu,\nu)$$

by the definition of the Ky Fan metric. Thus,

$$\inf\{d_F(X,Y) \mid X,Y \in R([0,1],\mathscr{B}([0,1]),m;\, E),\, X \overset{\mathscr{D}}{=} \mu,\, Y \overset{\mathscr{D}}{=} \nu\} \leq D_P(\mu,\nu).$$

3. For the reversed inequality, let X,Y be random variables $([0,1],\mathscr{B}([0,1]),m) \to E$ with distributions μ and ν respectively, and set $\varepsilon_0 := d_F(X,Y)$. Let π denote the distribution of the joint random variable $(X,Y) : ([0,1],\mathscr{B}([0,1]),m) \to E \times E$. Then, for arbitrary $\delta > 0$, by the definition of Ky Fan metric, we have

$$\pi(\{(x,y) \in E \times E,\, d(x,y) \geq \varepsilon_0 + \delta\}) < \varepsilon_0 + \delta.$$

Strassen's theorem now implies,

$$\forall B \in \mathscr{B}(E) :\ \mu(B) \leq \nu(\overline{B^{\varepsilon_0+\delta}}) + \varepsilon_0 + \delta,$$

so that

$$\forall B \in \mathscr{B}(E) :\ \mu(B) \leq \nu(B^{\varepsilon_0+2\delta}) + \varepsilon_0 + 2\delta.$$

Finally by the definition of Prokhorov metric, we obtain

$$D_P(\mu,\nu) \leq \varepsilon_0 + 2\delta = d_F(X,Y) + 2\delta,$$

which implies

$$D_P(\mu,\nu) \leq \inf\{d_F(X,Y) \mid X,Y \in R([0,1],\mathscr{B}([0,1]),m;\, E),\, X \overset{\mathscr{D}}{=} \mu,\, Y \overset{\mathscr{D}}{=} \nu\},$$

by taking infimum over X,Y and $\delta > 0$. □

Remark 10. The statement of Proposition 16 remains true, if we replace the probability space $([0,1],\mathscr{B}([0,1]),m)$ in its role there with any Borel space equipped with a non-atomic probability measure. (See, Theorem 3.4.23 in Srivastava (2008).)

We formulate an analogue of our main Theorem. Its proof is completely analogous by applying Proposition 16.

Proposition 17. (Metric foliation of Ky Fan spaces induce the Prokhorov space.) Let (E,d) be a complete separable metric space, $\mathscr{P}(E)$ denote the set of all probability measures on $(E,\mathscr{B}(E))$, and D_P denote the Prokhorov metric on $\mathscr{P}(E)$. Then:

(1) The metric space $(\mathscr{P}(E),D_P)$ is isometrically isomorphic to a metric quotient space induced by a metric foliation of $(\mathscr{R}(E),D)$, where the latter metric space is defined as the direct limit metric space

$$\mathscr{R}(E) := \varinjlim R^{(n)}(E),$$

of the direct system of metric spaces

$$R^{(1)}(E) \hookrightarrow R^{(2)}(E) \hookrightarrow R^{(3)}(E) \hookrightarrow \ldots \hookrightarrow \mathscr{R}(E),$$

where for $n \in \mathbb{N}$,

$$R^{(n)}(E) := R([0,1]^n,\mathscr{B}([0,1]^n),\otimes_{i=1}^n m;\ E)$$

is defined as a Ky Fan space over $([0,1]^n,\mathscr{B}([0,1]^n),\otimes_{i=1}^n m)$.

(2) The mapping

$$\mathscr{R}(E) \to \mathscr{P}(E),$$
$$[X] \mapsto \mu_X,$$

where μ_X denotes the distribution of any representative of the class $[X]$, is well-defined and a submetry.