Wissenschaftliche Hausarbeit
zur Ersten Staatsprüfung für das Lehramt an Gymnasien

im Fach: Mathematik

Thema:

Ricci curvature on graphs

vorgelegt von: Ricardo Kalke

geb. am 05.05.1988 in Rodewisch

Erfurt, den 10.08.2012

Abstract. We study Ricci curvature on graphs. This concept was first introduced on metric measure spaces by Ollivier and later applied to graphs by Lin\Yau and Jost\Liu. This curvature is defined by means of optimal transport. To obtain lower bounds it suffices to present a transference plan. To get lower bounds we present the dual formulation of the transportation distance using Lipschitz-continuous functions due to Kantorovich. These two bounds allow us to compute the Ricci curvature of the Platonic solids explicitly.

Contents

1 Introduction

Ricci curvature, which is named after the Italian mathematician Gregorio Ricci-Curbastro (1853-1925), is a fundamental concept in Riemannian geometry. It controls how fast geodesics starting at the same point diverge on average. Ricci curvature has its application in differential geometry, string theory or in general relativity where it occurs in the Einstein field equations. In this work we apply the Ricci curvature on finite and infinite graphs. We utilize a definition of Ricci curvature developed by Ollivier [15] to compute the curvature of Platonic solids. There are important works of Y. Lin and S.-T. Yau [13, 12, 4] concerning Ricci curvature on graphs. They established a Harnack inequality for finite connected graphs with non-negative Ricci curvature and gave a generalization of lower Ricci curvature bound in the framework of graphs. Yau was awarded the Fields medal in 1982 for his work in differential geometry and partial differential equations.

This work starts with an introduction to measure and integration theory. After we introduced the general concepts we apply them on at most countable and finite sets, respectively. We will also give an introduction to metrics and graphs where we restrict to locally finite graphs.

An important role in the definition of Ricci curvature is the notion of optimal transportation which is elaborate in the third section. Mathematicians deal with optimal transportation since over 200 years. It has become an important field in probability theory, economics and optimization. In particular, in the last 20 years there were published a multitude of papers that applied optimal transportation problems in partial differential equations, fluid mechanics, geometry and functional analysis. In this context it is worth mentioning the works of Brenier [3], whose paper paved the way towards an sufficient interplay between the mentioned topics, and Villani [17, 18]. Villani was awarded the Fields medal in 2010 for his work in optimal transport and the kinetic theory of the Boltzmann equation.

In the fourth section we refer closely to Villani [17] and introduce the notion of transportation distance between two probability measures and apply this to locally finite graphs following J. Jost and S. Liu, [8]. In particular, their paper served as an inspiration for this work. Furthermore, we derive a duality result for the transportation distance for graph case which is known as Kantorovich-Rubenstein theorem. We prove the theorem with methods of convex analysis. We use the Legendre-Fenchel transform and get a minimax theorem which is known as Fenchel-Rockafellar duality. Finally, we get an upper and lower bound for the transportation distance. This allows us to compute the exact value of the transportation distance for the Platonic solids and to compute the Ricci curvature. There are many references concerning this topic in various contexts. Especially Kantorovich worked on optimal transportation in economics [9, 10]. He was awarded the Nobel prize together with Tjalling Koopmans for their theory of optimum allocation of resources in 1976.

In the final section we calculate the exact values of the Ricci curvature of the Platonic solids. It turns out that only three of them have positive curvature which comes out as a surprise. This once more shows that the discretizations of continuum concepts is a non-trivial matter. This poses a major challenge for future research.

At this point I want to thank my advisor Dr. Matthias Keller. I am very thankful for his caring support, advices and patience in the last four months.I thank Prof. Dr. Daniel Lenz for giving me the opportunity to participate in a summer school in State College, Pensylvania, together with Marcel Schmidt and Felix Pogorzelski. My special gratitude goes to Sebastian Haeseler who always supported me in the last years and answered an uncountable amount of mathematical questions. Finally, I want to thank my friend Markus Böhm. Studying would not have been half as much fun without him.

2 Mathematical basics

In this first section we introduce the theory of measures and integration which will be fundamental for the following chapters. We first discuss the general concepts and turn to the discrete setting afterwards.

In Subsection 1.1 we start with the axiomatic introduction of a σ-algebra. After we discussed some simple examples, we define the notation of a measure. As a final result of this subsection we establish that for an at most countable set measures are given by non-negative functions.

In the first part of Subsection 1.2 we introduce integrals over general measures spaces. We apply the theory of integration to an at most countable measure space and notice that integrals simplify to a sum.

In Subsection 1.3 we endow an arbitrary set with a structure that permits to determine a distance between two elements of the set. This structure is given by a mapping which is called metric.

The last subsection is devoted to graphs. We define them on at most countable sets which is sufficient for further applications. Following Subsection 1.1 and 1.3 we define a metric and a measure on graphs.

2.1 σ-algebras and measures

In this subsection we explain how to assign a reasonable content to a subset of a given set X. Let 2^X be the power set, i.e., $2^X = \{Y \mid Y \subset X\}$. The notion stems from the fact that in the case where X has n elements the set 2^X has 2^n elements. The map that assign suitable elements $A \in 2^X$ a non-negative number is called measure. This number, the content, can be thought as the volume, weight or energy.

We define $A^c = X \backslash A$. A subset $\mathcal{A}$ of the power set 2^X is called a **σ-algebra** if it satisfies the following properties:

(i) $X \in \mathcal{A}$;

(ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$;

(iii) $A_j \in \mathcal{A}, j \in \mathbb{N} \Rightarrow \bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}$.

A set $A \subseteq \mathcal{A}$ is called **measurable** if $A \in \mathcal{A}$.

Examples 2.1. The following subsets of 2^X are σ-algebras:

1.) $\mathcal{A} = \{\emptyset, X\}$,

2.) $\mathcal{A} = 2^X$,

3.) $\mathcal{A}_0 = \{A \subseteq X | A \text{ or } A^c \text{ is countable}\}$.

Clearly, Examples 1.1.1) and 1.1.2) are σ-algebras. We prove that Example 1.1.3) is a σ-algebra.

Proof. (i) $X \in \mathcal{A}_0$:
If $A = X$, then $A^c = \emptyset$ is countable. So, $X \in \mathcal{A}_0$.

(ii) $A \in \mathcal{A}_0 \Rightarrow A^c \in \mathcal{A}_0$:
If A is not countable, then A^c is countable. So, $A^c \in \mathcal{A}_0$.
If A is countable, so $A = (A^c)^c$ is countable as well. So, $A^c \in \mathcal{A}_0$.

(iii) $A_j \in \mathcal{A}, j \in \mathbb{N} \Rightarrow \bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}$:
The first case is if $\bigcup_{j\in\mathbb{N}} A_j$ is countable, then $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}_0$.
If $\bigcup_{j\in\mathbb{N}} A_j$ is not countable, then there must exist at least one A_j in the union which is not countable, but its complement is countable. It follows that

$$\left(\bigcup_{j\in\mathbb{N}} A_j\right)^c = \bigcap_{j\in\mathbb{N}} A_j^c \qquad \text{is countable.}$$

So, $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}$.

□

In the following, let X be a non-empty set and define $[0, \infty] = [0, \infty) \cup \{\infty\}$.

Let $\mathcal{A}$ be a σ-algebra over X. A map $\mu : \mathcal{A} \to [0, \infty]$ is called a **measure** on $(X, \mathcal{A})$ if $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_{j\in\mathbb{N}} A_j\right) = \sum_{j=1}^{n} \mu(A_j)$$

if $A_j \in \mathcal{A}, j \in \mathbb{N}$, are mutually disjoint.
With the second property the map $\mu : \mathcal{A} \to [0, \infty]$ is also called σ-additive.

The triple $(X, \mathcal{A}, \mu)$ is called a **measure space**. If $\mu(X) = 1$, then we call μ a **probability measure** and $(X, \mathcal{A}, \mu)$ a **probability space**.
A measure is called **finite** if $\mu(x) < \infty$.

The most important rules for working with measures are gathered in the following proposition.

Proposition 2.1. Let $(X, \mathcal{A}, \mu)$ be a measure space. For $A, B \in \mathcal{A}$ we have:

(i) $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.

(ii) $\mu(B \backslash A) = \mu(B) - \mu(A)$ if $A \subseteq B$ and $\mu(A) < \infty$.

(iii) $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$, that is, μ is increasing.

Proof. (i): The set $A \cup B$ can be written as $A \cup B = A \cup (B \backslash A)$ with the property that $A \cap (B \backslash A) = \emptyset$. It follows with the $\sigma-$additivity of a measure that

$$\mu(A \cup B) = \mu(A \cup (B \backslash A)) = \mu(A) + \mu(B \backslash A). \tag{2.1}$$

Analogously, we get from $B = (A \cap B) \cup (B \backslash A)$ and $(A \cap B) \cap (B \backslash A) = \emptyset$ that

$$\mu(A \cap B) + \mu(B \backslash A) = \mu(B). \tag{2.2}$$

By adding Equation 2.1 and Equation 2.2 we have

$$\mu(A \cup B) + \mu(A \cap B) + \mu(B \backslash A) = \mu(A) + \mu(B) + \mu(B \backslash A).$$

If $\mu(B \backslash A)$ is finite, the claim follows. In case of $\mu(B \backslash A)$ is infinite, $\mu(A \cup B) = \infty$ because of Equation 2.1 and $\mu(B) = \infty$ because of Equation 2.2 and the claim is again verified.

(ii): Since $A \subseteq B$ we can write B as $B = A \cup (B \backslash A)$ with $A \cap (B \backslash A) = \emptyset$. Because of (i) we can write

$$\mu(B) = \mu(A \cup (B \backslash A)) = \mu(A) + \mu(B \backslash A).$$

By assumption that $\mu(A) < \infty$ the claim follows.

(iii): Since $A \subseteq B$ we have as in (ii) $B = A \cup (B \backslash A)$ and $\mu(B) = \mu(A) + \mu(B \backslash A)$. This implies

$$\mu(A) \leq \mu(B) \text{ if } A \subseteq B.$$

□

Let X be an at most countable set. For a set $A \subseteq X$ and $f : A \to [0, \infty]$ we define

$$\sum_{x \in A} f(x) = \sup \left\{ \sum_{j=1}^{n} f(x_j) \mid \{x_1, ..., x_n\} \subseteq A, n \geq 1 \right\} \in [0, \infty].$$

In particular, we set

$$\sum_{x \in \emptyset} f(x) = 0.$$

Let the σ-algebra $\mathcal{A}$ over X be given by $\mathcal{A} = 2^X$. The set $M(X)$ denotes the set of measures on $(X, 2^X)$. Then, the following theorem shows exactly the measures on $(X, 2^X)$ given by non-negative functions.

Theorem 2.1. Let X be an at most countable set. The mapping

$$\{m : X \to [0, \infty]\} \to M(X), \ m \mapsto \hat{m}$$

with

$$\hat{m}(A) = \sum_{x \in A} m(x), \ A \subseteq X$$

is a bijection.

Proof. Let $m : X \to [0, \infty]$. First, we show that $\hat{m}$ is a measure on $\left(X, 2^X\right)$.
The number $\hat{m}(A)$ is non-negative because m is a non-negative function for all $A \subseteq X$. Further on, $\hat{m}(\emptyset) = \sum_{x \in \emptyset} m(x) = 0$ because of the convention above.
In the following we show the σ-additivity of $\hat{m}$:
Let $X_0 := \bigcup_{i=1}^{\infty} X_i$ and $X_i, i \in \mathbb{N}$, be pairwise disjoint. Let $Y \subseteq X_0$ be a finite subset and set $Y_i = Y \cap X_i$. Obviously, $Y_i, i \in \mathbb{N}$, are finite and pairwise disjoint. For $Y_i \subseteq X_i$ we get by the definition of $\hat{m}$

$$\hat{m}(X_i) = \sum_{y \in X_i} m(y) = \sum_{y \in Y_i} m(y) + \sum_{x \in X_i \setminus Y_i} m(x) = \hat{m}(Y_i) + \sum_{x \in X_i \setminus Y_i} m(x) \geq \hat{m}(Y_i).$$

So, $\hat{m}(Y_i) \leq \hat{m}(X_i)$. Thereby, it follows that $\hat{m}(Y) = \sum_{i=1}^{\infty} \hat{m}(Y_i) \leq \sum_{i=1}^{\infty} \hat{m}(X_i)$ for any finite subset Y. So, we have

$$\hat{m}(X_0) = \sup_{Y \subseteq X_0 \text{ finite}} \hat{m}(Y) \leq \sum_{i=1}^{\infty} \hat{m}(X_i).$$

On the other hand, for all $n \geq 1$ and finite $Y \subseteq X_0$ we have

$$\hat{m}(X_0) \geq \hat{m}(Y) \geq \sum_{i=1}^{n} \hat{m}(Y_i).$$

As $\hat{m}(X_i) = \sup_{Y, Y_i = Y \cap X_i} \hat{m}(Y_i)$ it follows

$$\hat{m}(X_0) \geq \sum_{i=1}^{n} \sup_{Y, Y_i = Y \cap X_i} \hat{m}(Y_i) = \sum_{i=1}^{n} \hat{m}(X_i).$$

Therefore, taking the limit $n \to \infty$ yields

$$\hat{m}(X_0) = \sum_{i=1}^{\infty} \hat{m}(X_i).$$

Thus, we have shown that $\hat{m}$ is indeed a measure.
In the second step, we show that $m \mapsto \hat{m}$ is a bijection, i.e., the map is surjective and injective.
To check surjectivity, we find for every $\mu \in M(X)$ a function $m : X \to [0, \infty]$ such that $\mu = \hat{m}$. For a given $\mu \in M(X)$ we define $m(x) := \mu(\{x\}), x \in X$. Now, for $A \subseteq X$ we have

$$\mu(A) = \sum_{x \in A} \mu(\{x\}) = \sum_{x \in A} m(x) = \hat{m}(A).$$

Thus, $\mu = \hat{m}$.
To check injectivity, let $m, n : X \to [0, \infty]$ be such that $\hat{m} = \hat{n}$. This implies together with the definition of $\hat{m}$ and $\hat{n}$ for all $x \in X$

$$m(x) = \hat{m}(\{x\}) = \hat{n}(\{x\}) = n(x).$$

Therefore, $m = n$. This finishes the proof. □

2.2 Measurable and integrable functions

In this subsection we explain how to define the integral of a function over a measure space.

Let $(X_1, \mathcal{A}_1, \mu_1), (X_2, \mathcal{A}_2, \mu_2)$ be measure spaces. A function $f : X_1 \to X_2$ is called **measurable** if for every measurable set $A \subseteq X_2$ the set $f^{-1}(A) \subseteq X_1$ is measurable.

Remark 2.1. If $\mathcal{A}_1 = 2^X$, then every function is measurable. This follows by the definition of the power set, which is the set of all possible subsets of X.

Let $(X, \mathcal{A}, \mu)$ be a measure space. A function $\varphi : X \to \mathbb{R}$ is called **simple** if there exists non-negative numbers $a_1, ..., a_n$ and measurable sets $A_1, ..., A_n \in \mathcal{A}$ such that φ can be written as

$$\varphi = \sum_{i=1}^{n} a_i 1_{A_i}$$

where 1_A is equal to 1 on A and zero otherwise.
For example if φ is simple and takes the values $a_1, ..., a_n$, then $A_i \in \mathcal{A}, i = 1, ..., n$, can be chosen as $A_i = \{x \in X \mid \varphi(x) = a_i\}$.

We define the integral $\int_X \varphi d\mu \in [0, \infty]$ for a simple function $\varphi : X \to [0, \infty)$ with $\varphi = \sum\limits_{i=1}^{n} a_i 1_{A_i}$ by

$$\int_X \varphi d\mu = \int_X \varphi(x) d\mu(x) = \sum_{i=1}^{n} a_i \mu(A_i).$$

For a measurable function $f : X \to [0, \infty)$ the integral $\int_X f d\mu \in [0, \infty]$ is defined by

$$\int_X f d\mu = \int_X f(x) d\mu(x) = \sup \left\{ \int_X \varphi d\mu \mid 0 \leq \varphi \leq f,\ \varphi \text{ simple} \right\}. \tag{2.3}$$

A measurable function $f : X \to \mathbb{R}$ is said to be integrable if

$$f_+ = \max\{f, 0\} \qquad \text{and} \qquad f_- = \max\{-f, 0\}$$

satisfy $\int_X f_+ d\mu < \infty$ and $\int_X f_- d\mu < \infty$. The integral $\int_X f d\mu \in \mathbb{R}$ of an integrable function f is defined as

$$\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu. \tag{2.4}$$

Let X be an at most countable set and $\phi : X \to \mathbb{R}$ a function which satisfies $\sum\limits_{x \in X} \phi_+(x) < \infty$ and $\sum\limits_{x \in X} \phi_-(x) < \infty$. Then, we write

$$\sum_{x \in X} \phi(x) = \sum_{x \in X} \phi_+(x) - \sum_{x \in X} \phi_-(x). \tag{2.5}$$

The following theorem shows that the integral in (1.4) can be written as a sum.

Theorem 2.2. Let X be an at most countable set and μ a measure on $(X, 2^X)$. Then, there is a function $m : X \to [0, \infty)$ and the integral of an integrable function $f : X \to \mathbb{R}$ is given by

$$\int_X f(x) d\mu(x) = \sum_{x \in X} f(x) m(x). \tag{2.6}$$

Proof. By Theorem 2.1 we know that for every measure μ on X there is $m : X \to [0, \infty]$ such that $\mu = \hat{m}$. Moreover, because of Remark 2.1, every function is measurable. For every finite subset $A \subseteq X$ the function $f_+ 1_A$ is simple and it satisfies $0 \leq f_+ 1_A \leq f_+$. So, we have

$$\int_X f_+ 1_A d\mu = \sum_{x \in A} f_+(x) \mu(\{x\}) = \sum_{x \in A} f_+(x) m(x).$$

Analogously, we have for the simple function $f_- 1_A$ that satisfies $0 \leq f_- 1_A \leq f_-$

$$\int_X f_- 1_A d\mu = \sum_{x \in A} f_-(x) \mu(\{x\}) = \sum_{x \in A} f_-(x) m(x).$$

Moreover, for every simple function $\varphi : X \to [0, \infty), 0 \leq \varphi \leq f_+$ there is $A \subseteq X$ such that $\varphi \leq f_+ 1_A \leq f_+$. An analogous statement holds for f_-. Therefore, as f is assumed to be integrable, we get by Equation 2.3

$$\int_X f_+ d\mu = \sup_{A \subseteq X \text{finite}} \left\{ \sum_{x \in A} f_+(x) m(x) \right\} = \sum_{x \in X} f_+(x) m(x) < \infty$$

and

$$\int_X f_- d\mu = \sup_{A \subseteq X \text{finite}} \left\{ \sum_{x \in A} f_-(x) m(x) \right\} = \sum_{x \in X} f_-(x) m(x) < \infty.$$

Since m is a positive function, we get for $\phi(x) = (fm)(x) = f(x)m(x)$ by Equation 2.4 and Equation 2.5

$$\begin{aligned}
\int_X f(x) d\mu(x) &= \int_X f_+ d\mu - \int_X f_- d\mu \\
&= \sum_{x \in X} f_+(x) m(x) - \sum_{x \in X} f_-(x) m(x) \\
&= \sum_{X \in X} f(x) m(x)
\end{aligned}$$

which completes the claim. □

2.3 Metric

Definition 2.1. Let X be a non-empty set. The map $d : X \times X \to [0, \infty)$ is called **metric** on X if for all $x, y, z \in X$,

(M1) $d(x, y) = 0$ if, and only if $x = y$;

(M2) $d(x, y) = d(y, x)$;

(M3) $d(x, y) \leq d(x, z) + d(z, y)$.

If d is a metric on X we say (X, d) is a metric space and call the number $d(x, y)$ the distance between x and y.

Examples 2.2. For an element x in the n-dimensional Euclidean space $\mathbb{R}^n$ we write $x = (x_1, ..., x_n)^T$.

1.) Let $p \in \mathbb{N}$ be given. The distance between x and y in the n-dimensional Euclidean space is given by

$$d_p(x, y) = \left[\sum_{i=1}^{n} |x_i - y_i|^p\right]^{\frac{1}{p}}.$$

2.) The Tchebychev distance or maximum metric between the elements x and y in the n-dimensional Euclidean space is given by

$$d_\infty(x, y) = \max_i |x_i - y_i|.$$

3.) Let (X, d) a metric space and Y a non-empty set of X. Then the restriction of d to $Y \times Y, d_Y = d|_{Y \times Y}$ is a metric and (Y, d_Y) a metric space.

Clearly, d_Y in Example 1.2.3) is a metric. In Example 1.2.2) d_∞ can be seen to be a metric either by the argument given in Lemma 2.1 below or via Remark 2.2 below. In the following we prove that Example 1.2.1) is a metric.

Proof. (M1): Clearly, if $x = y$, then $d(x, y) = 0$. Otherwise, the metric takes the value 0 if, and only if $|x_i - y_i|^p = 0$ for all $i = 1, ..., n$. So, $x_i = y_i$ for every $i = 1, ..., n$. It follows that $x = y$.
(M2): Since $|x_i - y_i|^p = |-(-x_i + y_i)|^p = |y_i - x_i|^p$ it follows that $d(x, y) = d(y, x)$.
(M3): The Minkowski inequality [7, Section 59.3] provides that

$$\begin{aligned} \left[\sum_{i=1}^{n} |x_i - y_i|^p\right]^{\frac{1}{p}} &= \left[\sum_{i=1}^{n} |x_i - z_i + z_i - y_i|^p\right]^{\frac{1}{p}} \\ &\leq \left[\sum_{i=1}^{n} |x_i - z_i|^p\right]^{\frac{1}{p}} + \left[\sum_{i=1}^{n} |z_i - y_i|^p\right]^{\frac{1}{p}} \end{aligned}$$

for all $z \in X$. It follows that $d(x, y) \leq d(x, z) + d(z, y)$. □

Remark 2.2. The maximum metric emerges form the Euclidean metric for $p \to \infty$. Between d_p and d_∞ exists the following estimate

$$\max_i |x_i - y_i| \leq \left[\sum_{i=1}^{n} (x_i - y_i)^p\right]^{\frac{1}{p}} \leq n^{-\frac{1}{p}} \max_i |x_i - y_i|. \tag{2.7}$$

Because of the sandwich theorem [19, p. 63] and $\lim_{p\to\infty} n^{-\frac{1}{p}} = 1$ we get

$$\lim_{p\to\infty}\left[\sum_{i=1}^{n}(x_i - y_i)^p\right]^{\frac{1}{p}} = \max_i |x_i - y_i| .$$

Definition 2.2. Two metrics d_1 and d_2 on X are equivalent if there exist two constants $c_1, c_2 \in \mathbb{R}$ with $0 < c_1 < c_2 < \infty$ such that

$$c_1 d_1(x, y) \leq d_2(x, y) \leq c_2 d_1(x, y)$$

for all $x, y \in X$.

Example 1.3. The metrics from Example 1.2.1) and Example 1.2.2) are equivalent. By Equation 2.7 of Remark 2.2 we get $c_1 = n^{\frac{1}{p}}$ and $c_2 = 1$.

Let X be an at most countable set and define $\mathbb{R}^{X\times X} = \{\xi \mid \xi : X \times X \to \mathbb{R}\}$.

Lemma 2.1. The map

$$\rho : \mathbb{R}^{X\times X} \times \mathbb{R}^{X\times X} \to [0, \infty), \rho(\xi, \eta) = \sup_{x,y\in X} |\xi(x, y) - \eta(x, y)|$$

is a metric.

Proof. The axioms (M1) and (M2) are checked precisely as in the proof of Example 1.2.1).
Let $\psi \in \mathbb{R}^{X\times X}$. The triangle inequality for $|\cdot|$ and for suprema provides that

$$\begin{aligned}\sup_{x,y\in X} |\xi(x, y) - \eta(x, y)| &= \sup_{x,y\in X} |\xi(x, y) - \psi(x, y) + \psi(x, y)\eta(x, y)| \\ &\leq \sup_{x,y\in X} \{|\xi(x, y) - \psi(x, y)| + |\psi(x, y) - \eta(x, y)|\} \\ &\leq \sup_{x,y\in X} |\xi(x, y) - \psi(x, y)| + \sup_{x,y\in X} |\psi(x, y) - \eta(x, y)|.\end{aligned}$$

It follows that $\rho(\xi, \eta) \leq \rho(\xi, \psi) + \rho(\psi, \eta)$. □

2.4 Graphs

Graphs are mathematical concepts for net-like structures which we can be found in many applied scientific disciplines such as engineering, computer science or management disciplines. We give a short introduction to this topic in the following subsection.

Let X be an at most countable set and the function $g : X \times X \to [0, \infty)$ be such that

(G1) $g(x, x) = 0$ for $x \in X$; (zero diagonal)

(G2) $g(x, y) = g(y, x)$ for $x, y \in X$; (symmetry)

(G3) the set $\{y \in X \mid g(x, y) > 0\}$ is finite for every $x \in X$ (local finiteness).

We can think of g as a **graph** over X in the following way:
Let X be the set of vertices. Two vertices $x, y \in X$ are connected by an edge of weight $g(x, y)$ whenever $g(x, y) > 0$. In this case we can write $x \sim y$ and say x and y are **adjacent** respectively x and y are called neighbors.

A **path** is a sequence of vertices $(x_0, ..., x_n)$ with $x_{i-1} \sim x_i, i = 1, ..., n$. We say n is the length of the path. A graph is called **connected** if any two vertices can be connected by a path.

An illustration of a graph is given by the following picture. The set of vertices is $X = \{1, 2, 3, 4\}$ and

$$g(1, 2) = g(2, 1) = g(1, 3) = g(3, 1) = g(2, 3) = g(3, 2) = 1$$

and $g(i, 4) = g(4, i) = 0$ for $i = 1, 2, 3$.

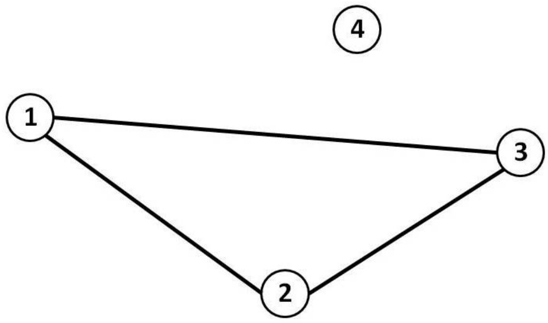

Figure 1: Non-connected graph

The **degree** w_a of the vertex $a \in X$ is defined as

$$w_a = \sum_{y \in X} g(x, a).$$

In the case of $g(x, y) \in \{0, 1\}$ for all $x, y \in X$, the degree w_a is the number of neighbors of a.

Let $d(x, y)$ be the minimal n such that the vertices x and y of a graph g can be connected by a path of length n.

Lemma 2.2. If g is connected, then d is a metric.

Proof. Since g is connected there is a path connecting every two vertices of the graph. Thus $d(x, y) < \infty$ for all $x, y \in X$. On the other hand it is obvious by the definition that d takes only values in $\mathbb{N}_0$. Therefore, $d : X \times X \to \mathbb{N}_0$. Further on, we check d for the properties (M1) to (M3):
Since the path from x to x has the length zero, the distance $d(x, x) = 0$. Clearly, if $x \neq y$, then $d(x, y) > 1$ and thus $d(x, y) \neq 0$.

Let $(x_0, ..., x_n)$ be the shortest path from $x := x_0$ to $y := y_0$. By renumbering $i \mapsto n - i, i = 0, ..., n$, we get the path $(x_n, ..., x_0)$ from y to x. Since the length of the path did not change we have $d(x, y) = d(y, x)$. It is left to show the triangle inequality.
Clearly, if z is not a vertex on a minimal path from x to y, the path becomes longer. It follows that $d(x, y) \leq d(x, z) + d(z, y)$ which finishes the proof. □

To each vertex $a \in X$ we attach the non-negative function

$$m_a : X \to [0, 1],\ x \mapsto \frac{g(x, a)}{w_a}.$$

By Theorem 1.1 the function m_a induces a measure $\hat{m}_a$ via

$$\hat{m}_a(A) = \sum_{x \in A} m_a(x)$$

for every subset $A \subseteq X$ of vertices. For $A = X$ we have

$$\hat{m}_a(X) = \sum_{x \in X} m_a(x) = \sum_{x \in X} \frac{g(x, a)}{w_a} = \frac{1}{w_a} \sum_{x \in X} g(x, a) = \frac{w_a}{w_a} = 1.$$

Thus, $\hat{m}_a$ is a probability measure on $(X, 2^X)$.
If $A \subset X$ is a singleton set, an intuitive illustration of $\hat{m}_a$ is that it models a random walker that sits at a and then chooses amongst the neighbors of a the vertex x with probability $\frac{g(x,y)}{w_a}$.

3 Optimal transportation problem and transportation distance

In this section we introduce the mathematical concept of optimal transportation. It concerns the famous problem to transport a given mass in the best economical way. The notion of optimal mass transportation was developed by Gaspard Monge in 1781. Important contributions to this topic are provided by Kantorovich in the 1940's [9],[10]. In this section we follow closely the work by Villani [17] who worked on the theory of optimal transport and especially its applications. He was awarded the Fields Medal in 2010 "for his proofs of nonlinear Landau damping and convergence to equilibrium for the Boltzmann equation", [21].

3.1 Definition of transportation distance

We can think of the mass as a pile of sand which we want to transport into a hole. Obviously, the pile and the hole must have the same volume. We model the amount of sand and the size of the hole by measures μ and ν, defined on some measure spaces $(X, \mathcal{A}_X)$ and $(Y, \mathcal{A}_Y)$, respectively. Since the real mass of the pile is not important, let us normalize $\mu(X)$ and $\nu(Y)$ to 1. Thus, μ and ν are probability measures. Whenever A and B are measurable subsets of X and Y, respectively, the value $\mu(A)$ gives a measure of how much sand is located inside A and $\nu(B)$ of how much sand can be piled in B.

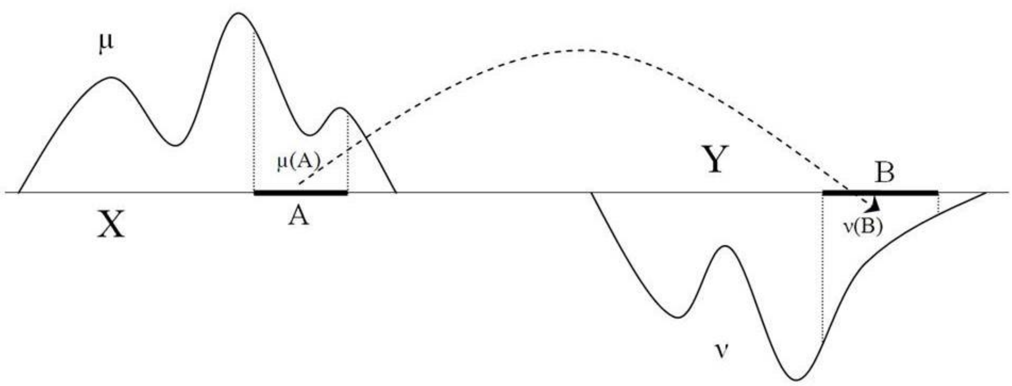

Figure 2: Mass Transportation

The sand moving from a position $A \subseteq X$ to a position $B \subseteq Y$ has to be transported over a distance d. A transference plan indicates how the transport of the mass is managed. We model transference plans by probability measures ζ on the product space $X \times Y$. This means $\zeta(A \times B)$ measures how much sand of position A is filled in at position B of the hole. The space of probability measures is denoted by $P(X \times Y)$.

Definition 3.1. The measure $\zeta \in P(X \times Y)$ is a **transference plan** if it satisfies for $A \subseteq X$

$$\zeta(A \times Y) = \int_{A \times Y} d\zeta(x,y) = \mu(A)$$

and for $B \subseteq Y$

$$\zeta(X \times B) = \int_{X \times B} d\zeta(x,y) = \nu(B).$$

We denote the set of all transference plans as

$$\Pi(\mu,\nu) = \{\zeta \in P(X \times Y) \mid \zeta(A \times Y) = \mu(A),\ \zeta(X \times B) = \nu(B),\ A \in \mathcal{A}_X,\ B \in \mathcal{A}_Y\}.$$

An important special case is $Y = X$. Now, a basic problem is to realize a best transport by the most appropriate transference plan. A measure of how good a transference plan works is the transportation distance which is defined for the general case in the following.

Definition 3.2. Transportation distance (cf. L. V. Kantorovich [9],[10])
For two probability measures μ and ν on a metric space (X,d) the transportation distance between μ and ν is defined as

$$W_1(\mu,\nu) = \inf_{\zeta \in \Pi(\mu,\nu)} \int\limits_{X \times X} d(x,y) d\zeta(x,y).$$

3.2 Transportation distance for graphs

As a next step, we derive the transportation distance for locally finite graphs. The main result of this section is found by J. Jost and S. Liu in [8]. We give a proof with all details.

Let X be an at most countable set and let $g : X \times X \to [0,\infty)$ be a a graph as introduced in Section 2.4. Let $a, b \in X$ be two vertices of g. We recall that in Section 2.4 the function m_a was defined as $m_a(x) = \frac{g(x,a)}{w_a} = \frac{g(a,x)}{w_a}$ and $\hat{m}_a$ are the associated measures. Then $\Pi(\mu,\nu)$ with $\mu = \hat{m}_a$ and $\nu = \hat{m}_b$ reads as

$$\Pi(\hat{m}_a, \hat{m}_b) = \{\zeta \in P(X \times X) \mid \zeta(A \times X) = \hat{m}_a(A),\ \zeta(X \times B) = \hat{m}_b(B),\ A, B \subseteq X\}.$$

We set for $a, b \in X$

$$P_{ab} = \Bigg\{\xi : X \times X \to [0,\infty) \mid \sum_{x \in X} \xi(x,z) = m_b(z), \sum_{y \in X} \xi(z,y) = m_a(z),$$
$$\text{and for all } z \in X \sum_{x,y \in X} \xi(x,y) = 1\Bigg\}.$$

The following theorem give a formula for the transportation distance for graphs.

Theorem 3.1. Let X be an at most countable set and g a graph over X. For all $a, b \in X$ the transportation distance $W_1(\hat{m}_a, \hat{m}_b)$ satisfies

$$W_1(\hat{m}_a, \hat{m}_b) = \inf_{\xi \in P_{ab}} \sum_{x, x \sim a} \sum_{y, y \sim b} d(x,y)\xi(x,y). \tag{3.1}$$

In the case of a finite vertex set X of the graph g, the next theorem shows that the infimum over ξ in the definition of W_1 is attained.

Theorem 3.2. Let X be a finite set and g a graph over X. For all $a, b \in X$ the transportation distance $W_1(\hat{m}_a, \hat{m}_b)$ satisfies

$$W_1(\hat{m}_a, \hat{m}_b) = \min_{\xi \in P_{ab}} \sum_{x, x \sim a} \sum_{y, y \sim b} d(x,y)\xi(x,y). \tag{3.2}$$

3.3 Proof of the infinite graph case

Lemma 3.1. The map

$$P_{ab} \to \Pi(\hat{m}_a, \hat{m}_b), \xi \mapsto \hat{\xi}$$

is a bijection. Moreover, if $\xi \in P_{ab}$, then $\xi(x,y) = 0$ for $x \nsim a$ or $y \nsim b$ and

$$\sum_{x \in X, x \sim a} \xi(x,z) = \frac{g(b,z)}{w_b} \text{ respectivly } \sum_{y \in X, y \sim b} \xi(z,y) = \frac{g(a,z)}{w_a}$$

for all $z \in X$.

Proof. We first show that for $\xi \in P_{ab}$ the measure $\hat{\xi}$ defined by Theorem 2.1 is indeed an element of $\Pi(\hat{m}_a, \hat{m}_b)$.
Let $\xi \in P_{ab}$. Since X is at most countable the set $X \times X$ is at most countable, too. We recall that for a non-negative function ξ on an at most countable set a measure $\hat{\xi}$ is given by Theorem 2.1. Let A and B be subsets of X. Then, we get by definition of P_{ab}

$$\begin{aligned} \hat{\xi}(A \times X) &= \hat{\xi}\left(\{(x,y) \mid x \in A, y \in X\}\right) \\ &= \sum_{(x,y) \in A \times X} \hat{\xi}(\{(x,y)\}) = \sum_{x \in A} \sum_{y \in X} \xi(x,y) = \sum_{x \in A} m_a(x) = \hat{m}_a(A) \end{aligned}$$

and analogously

$$\hat{\xi}(X \times B) = \sum_{(x,y) \in X \times B} \hat{\xi}(\{(x,y)\}) = \sum_{x \in X} \sum_{y \in B} \xi(x,y) = \sum_{y \in B} m_b(y) = \hat{m}_b(B).$$

In particular, $\hat{\xi}(X \times X) = 1$ since $\hat{m}_a$ and $\hat{m}_b$ are probability measures. It follows that $\hat{\xi} \in \Pi(\hat{m}_a, \hat{m}_b)$. The injectivity of the map is a consequence of the injectivity of $\xi \mapsto \hat{\xi}$ by Theorem 2.1.
To check surjectivity, we find by Theorem 2.1 for every measure ζ on $X \times X$ a function $\xi : X \times X \to [0, \infty)$ such that $\zeta = \hat{\xi}$. Let $\zeta \in \Pi(\hat{m}_a, \hat{m}_b)$ and ξ such that

$\hat{\xi} = \zeta$. We check that $\xi \in P_{ab}$. Set $A = \{x\}$ and $B = \{y\}$. Therefore, we get by definition of $\Pi(\hat{m}_a, \hat{m}_b)$

$$m_a(x) = \hat{m}_a(\{x\}) = \zeta(\{x\} \times X) = \sum_{y \in X} \zeta(\{(x,y)\}) = \sum_{y \in X} \hat{\xi}(\{(x,y)\}) = \sum_{y \in X} \xi(x,y)$$

and analogously

$$m_b(y) = \hat{m}_b(\{y\}) = \zeta(X \times \{y\}) = \sum_{x \in X} \zeta(\{(x,y)\}) = \sum_{x \in X} \hat{\xi}(\{(x,y)\}) = \sum_{x \in X} \xi(x,y).$$

Finally, for $A = B = X$ we have by $\hat{\xi} = \zeta$ and by the fact that ζ is a probability measure

$$\sum_{x,y \in X} \xi(x,y) = \sum_{(x,y) \in X \times X} \hat{\xi}(\{(x,y)\} = \zeta(X \times X) = 1.$$

Thus, $\xi \in P_{ab}$.
In particular, if x and a are not neighbors the sum $\sum_{y \in X} \xi(x,y) = m_a(x) = \frac{g(a,x)}{w_a}$ takes the value 0. It follows that if $x \nsim a$, then $\xi(x,y) = 0$ for any $y \in X$. In the same way the sum $\sum_{x \in X} \xi(x,y) = m_b(y) = \frac{g(b,y)}{w_b}$ takes the value 0 if y and b are not neighbors. It follows that if $y \nsim b$, then $\xi(x,z) = 0$ for any $x \in X$. Thus, the sums reduces to a summation over the neighbors of a respectively b. So,

$$\sum_{x, x \sim a} \xi(x,z) = \frac{g(b,z)}{w_b} \text{ and } \sum_{y, y \sim b} \xi(z,y) = \frac{g(a,z)}{w_a} \tag{3.3}$$

for all $z \in X$. This completes the proof. □

Now we turn on to prove Theorem 3.1.

Proof of Theorem 3.1. Let $\zeta \in \Pi(\hat{m}_a, \hat{m}_b)$. By Theorem 2.1 exists a function $\xi : X \times X \to [0, \infty)$ such that $\hat{\xi} = \zeta$. Because of Lemma 3.1 the function ξ is an element of P_{ab}. Further on, by Theorem 2.2 we can write

$$\int_{X \times X} d(x,y)\zeta(x,y) = \sum_{X \times X} d(x,y)\xi(x,y) = \sum_{x \in X} \sum_{y \in X} d(x,y)\xi(x,y).$$

By Lemma 3.1 we just need to sum up over the neighbors of a and b. So,

$$\sum_{x \in X} \sum_{y \in X} d(x,y)\xi(x,y) = \sum_{x, x \sim a} \sum_{y, y \sim b} d(x,y)\xi(x,y).$$

In particular, we get by Lemma 3.1 that

$$W_1(\hat{m}_a, \hat{m}_b) = \inf_{\zeta \in \Pi(\hat{m}_a, \hat{m}_b)} \int_{X \times X} d(x,y)\zeta(x,y) = \inf_{\xi \in P_{ab}} \sum_{x, x \sim a} \sum_{y, y \sim b} d(x,y)\xi(x,y).$$

This completes the proof. □

3.4 Proof of the finite graph case

In Subsection 2.3 we showed by Lemma 2.1 that on $P = \{\xi \mid \xi : X \times X \to [0, \infty)\}$ the map ρ given by

$$\rho(\xi, \eta) = \sup_{x,y \in X} |\xi(x, y) - \eta(x, y)|, \qquad \xi, \eta \in \mathbb{R}^{X \times X}$$

is a metric. Obviously, P_{ab} is a subset of $\mathbb{R}^{X \times X}$ and by Example 1.2.3) ρ a metric on P_{ab} as well. If X is a finite set, then

$$\rho(\xi, \eta) = \max_{x,y \in X} |\xi(x, y) - \eta(x, y)|, \qquad \xi, \eta \in \mathbb{R}^{X \times X}.$$

Lemma 3.2. Let X be a finite set. Then, (P_{ab}, ρ) is a compact metric space.

Proof. The set P_{ab} is a subset of the finite dimensional vector space $\mathbb{R}^{X \times X}$. It is compact under ρ if it is bounded and closed in $\mathbb{R}^{X \times X}$(cf. [5, Theorem 2.3.26]). We begin with boundedness.
Let $\xi, \eta \in P_{ab}$. By definition $\sum_{X \times X} \xi(x, y) = 1$ and $\sum_{X \times X} \eta(x, y) = 1$ and $\xi(x, y), \eta(x, y) \geq 0$. This implies $\xi(x, y), \eta(x, y) \leq 1$ for all $x, y \in X$. Therefore,

$$\rho(\xi, \eta) = \max_{x,y \in X} |\xi(x, y) - \eta(x, y)| \leq 1$$

for arbitrary $\xi, \eta \in P_{ab}$. Thus, P_{ab} is bounded.
It is left to show the closedness of P_{ab} in $\mathbb{R}^{X \times X}$, i.e., we show that the defining properties of elements in P_{ab} are preserved under convergence.
Let ξ_n be an arbitrary sequence in P_{ab} and $\xi \in \mathbb{R}^{X \times X}$ with $\lim_{n \to \infty} \xi_n = \xi$ with respect to ρ. So, for all $\varepsilon > 0$ there is a $N \in \mathbb{N}$ such that $\rho(\xi, \xi_n) < \varepsilon$ for all $n > N$. This implies by $\sum_{x, x \sim a} \xi_n(x, y) = \frac{g(b,y)}{w_b}$ that

$$\begin{aligned} \left| \frac{g(b, y)}{w_b} - \sum_{x, x \sim a} \xi(x, y) \right| &= \left| \sum_{x, x \sim a} \xi_n(x, y) - \sum_{x, x \sim a} \xi(x, y) \right| \\ &\leq \sum_{x, x \sim a} |\xi_n(x, y) - \xi(x, y)| \\ &\leq |X \times X| \rho(\xi_n, \xi) < |X|^2 \varepsilon. \end{aligned}$$

Analogously,

$$\left| \frac{g(a, x)}{w_a} - \sum_{y, y \sim b} \xi(x, y) \right| < |X|^2 \varepsilon$$

and

$$\left| 1 - \sum_{x, y \in X} \xi(x, y) \right| < |X|^2 \varepsilon.$$

Since ε was arbitrary, we have $\xi \in P_{ab}$ which shows the closedness of P_{ab}. $\square$

Lemma 3.3. Let X be finite. The map

$$P_{ab} \to [0, \infty),\ \xi \mapsto \sum_{X \times X} d(x, y)\xi(x, y)$$

is continuous with respect to ρ.

Proof. Let

$$\operatorname{diam}(X) = \max\{d(x, y) \mid x, y \in X, x \sim a, y \sim b\}.$$

Let $\varepsilon > 0$ be arbitrary and choose $\delta = \frac{\varepsilon}{\operatorname{diam}(X) \cdot |X|^2}$. Then, for $\xi, \eta \in P_{ab}$ with $\max_{x,y \in X} |\xi(x, y) - \eta(x, y)| < \delta$ it follows that

$$\begin{aligned} \left| \sum_{x,y \in X} d(x, y)\xi(x, y) - \sum_{x,y \in X} d(x, y)\eta(x, y) \right| &\leq \sum_{(x,y) \in X \times X} d(x, y)|\xi(x, y) - \eta(x, y)| \\ &< \delta \cdot |X|^2 \cdot \operatorname{diam}(X) = \varepsilon. \end{aligned}$$

Thus, the map is continuous. □

Finally, we prove Theorem 3.2.

Proof of Theorem 3.2. We already know by Theorem 3.1 that $W_1(\hat{m}_a, \hat{m}_b)$ can be written as

$$W_1(\hat{m}_a, \hat{m}_b) = \inf_{\xi \in P_{ab}} \sum_{x, x \sim a} \sum_{y, y \sim b} d(x, y)\xi(x, y).$$

By Lemma 3.2 we know that (P_{ab}, ρ) is a compact metric space and by Lemma 3.3 that $\xi \mapsto \sum_{X \times X} d(x, y)\xi(x, y)$ is continuous. Since a continuous function on a compact metric space assumes its minimum (cf. [1, Corollar 3.8]), we conclude

$$W_1(\hat{m}_a, \hat{m}_b) = \min_{\xi \in P_{ab}} \sum_{x, x \sim a} \sum_{y, y \sim b} d(x, y)\xi(x, y)$$

which completes the proof. □

4 Kantorovich duality

In this section we derive a duality result for the transportation distance W_1 which is known as "Kantorovich duality". There is a multitude of references concerning this topic in a general context. Again, we especially mention the works of Villani[17],[18] and Kantorovich [9],[10] concerning this topic. Kantorovich was awarded the Nobel prize in economics together with Tjalling Koopmans "for their contributions to the theory of optimum allocation of resources" in 1976 [20]. This is very much related to what we do in this section.
Jost and Liu mention the Kantorovich duality in [8] for the graph case. We give a proof in details and follow in this section closely [6, Section 1.2] and [17, Section 1.1]. To do so, we use a fundamental duality theorem in convex analysis. As a final result of this work, we have a lower bound for the transportation distance W_1 by the Kantorovich duality and an upper bound by Theorem 2.2.

4.1 The Kantorovich-Rubenstein theorem

Let X be an at most countable set. We define $\mathrm{Lip}_M(X), M \in \mathbb{R}$, as the set of all functions $f : X \to \mathbb{R}$ on a graph g that satisfy $f(x) - f(y) \leq M \cdot d(x, y)$ for all $x, y \in X$. We recall that $\hat{m}_a$ is the associated measure to the function $m_a = \frac{g(x,a)}{w_a}$ defined in Section 2.4.
The Kantorovich dual formulation of the transportation distance W_1 for graph case is given by the following theorem.

Theorem 4.1. (Kantorovich-Rubenstein theorem) Let X be a finite set and $a, b \in X$. Then, the transportation distance W_1 is equal to

$$W_1(\hat{m}_a, \hat{m}_b) = \sup \left\{ \sum_{z \in X, z \sim a} f(z) m_a(z) - \sum_{z \in X, z \sim b} f(z) m_b(z) \mid f \in \mathrm{Lip}_1(X) \right\}. \quad (4.1)$$

In order to prove Theorem 4.1 we employ some notions of convex analysis. A proof can be found in Subsection 4.4. However, the inequality "$\leq$" is rather straight forward. Therefore, we give a separate proof for this inequality where we only assume that X is an at most countable set.

Lemma 4.1. Let X be an at most countable set and $a, b \in X$. Then,

$$W_1(\hat{m}_a, \hat{m}_b) \geq \sup \left\{ \sum_{z \in X, z \sim a} f(z) m_a(z) - \sum_{z \in X, z \sim b} f(z) m_b(z) \mid f \in \mathrm{Lip}_1(X) \right\}.$$

Proof. Let $\xi \in P_{ab}$ and $f \in \mathrm{Lip}_1(X)$. We calculate using the properties of ξ, i.e.,

$\sum_{y\in X}\xi(z,y)=m_b(z)$ and $\sum_{x\in X}\xi(x,z)=m_a(z)$ for all $z\in X$,

$$\begin{aligned}\sum_{x,y\in X}(f(x)-f(y))\xi(x,y) &= \sum_{x,y\in X}f(x)\xi(x,y)-\sum_{x,y\in X}f(y)\xi(x,y)\\ &= \sum_{x\in X}f(x)\sum_{y\in X}\xi(x,y)-\sum_{y\in X}f(y)\sum_{x\in X}\xi(x,y)\\ &= \sum_{x\in X}f(x)m_a(x)-\sum_{y\in X}f(y)m_b(y).\end{aligned}$$

Moreover, since $f\in \mathrm{Lip}_1(X)$ we have $\sum_{x,y\in X}(f(x)-f(y))\xi(x,y)\le\sum_{x,y\in X}d(x,y)\xi(x,y)$. Hence,

$$\sum_{x\in X}f(x)m_a(x)-\sum_{y\in X}f(y)m_a(y)\le\sum_{x,y\in X}d(x,y)\xi(x,y)$$

for all $\xi\in P_{ab}$ and $f\in\mathrm{Lip}_1(X)$. We note the left hand side does not depend on ξ and the right hand sight does not depend on f. Thus, taking the supremum on the left hand side and the infimum on the right hand side yields

$$W_1(\hat{m}_a,\hat{m}_b)\ge\sup\left\{\sum_{z\in X,z\sim a}f(z)m_a(z)-\sum_{z\in X,z\sim b}f(z)m_b(z)\mid f\in\mathrm{Lip}_1(X)\right\}$$

which finishes the proof. □

4.2 Normed vector spaces

Next, we define the notion of a normed vector space and its dual space.

Definition 4.1. Let E be a vector space over $\mathbb{R}$. A function $\|\cdot\|:E\to[0,\infty)$ is called **norm** if for all $x,y\in E$

(N1) $\|x\|=0$ if, and only if $x=0$; (positive definite)

(N2) $\|\lambda x\|=|\lambda|\cdot\|x\|$ for all $\lambda\in\mathbb{R}$; (Positive homogeneity)

(N3) $\|x+y\|\le\|x\|+\|y\|$ (triangle inequality).

We call the pair $(E,\|\cdot\|)$ **normed vector space** if the function $\|\cdot\|:E\to[0,\infty)$ satisfies the properties (N1) - (N3) for all $x,y\in E$. If the norm is clear from the context, we write E instead of $(E,\|\cdot\|)$.
We define the **dual space** E^* of a normed vector space E as the set of all continuous linear mappings $E\to\mathbb{R}$.

Example 4.1. Let X be an at most countable set. The vector space $\mathbb{R}^{X\times X}$ equipped with the norm $\|\cdot\|_\infty$ given by

$$\|\xi\|_\infty=\sup_{x\in X}|\xi(x)|$$

is a normed vector space.

In the case that X is finite, we establish in the following lemma that $\mathbb{R}^{X\times X}$ and its dual space are equal.

Lemma 4.2. For a finite set X and $E = \mathbb{R}^{X\times X}$ equipped with $\|\cdot\|_\infty$ the dual space E^* is given by $\mathbb{R}^{X\times X}$ via

$$\varphi : E \to \mathbb{R}, \varphi(\xi) = \sum_{x,y\in X} \varphi(x,y)\xi(x,y) \tag{4.2}$$

for $\xi \in \mathbb{R}^{X\times X} = E$ and $\varphi \in \mathbb{R}^{X\times X} = E^*$.

Proof. Clearly, every $\varphi \in \mathbb{R}^{X\times X}$ defines a continuous linear mapping $E \to \mathbb{R}$ via Equation 4.2. Continuity is checked along the lines of the proof of Lemma 3.3. On the other hand, let $e_{x,y}, x, y \in X$ be the standard basis of $E = \mathbb{R}^{X\times X}$, i.e.,

$$e_{x,y}(z,w) = \begin{cases} 1 & : \text{for } (x,y) = (z,w), \\ 0 & : \text{otherwise} \end{cases}$$

for all $z, w \in X$. For a given linear and continuous $\varphi : E \to \mathbb{R}$, let $\psi \in \mathbb{R}^{X\times X}$ be such that $\psi(x,y) = \varphi(e_{x,y})$ for $x, y \in X$. Then, for $\xi = \sum\limits_{x,y\in X} \xi(x,y)e_{x,y} \in \mathbb{R}^{X\times X}$ we have by linearity

$$\begin{aligned} \varphi(\xi) &= \varphi\left(\sum_{x,y\in X} \xi(x,y)e_{x,y}\right) = \sum_{x,y} \xi(x,y)\varphi(e_{x,y}) \\ &= \sum_{x,y} \psi(x,y)\xi(x,y). \end{aligned}$$

Thus, φ can be identified with ψ. This finishes the proof. □

Remark 4.1. Since all norms on a finite dimensional vector space are equivalent (cf. [1, Satz 3.12]), we can replace $\|\cdot\|_\infty$ in Lemma 4.2 by any other norm on $\mathbb{R}^{X\times X}$.

4.3 Convex analysis

We go on with a short introduction to convex analysis. We define a convex function and its convex conjugate. Furthermore, we state a duality theorem which is known as "Fenchel-Rockafellar duality".

Definition 4.2. Let E be a normed vector space. A function $f : E \to \mathbb{R} \cup \{\infty\}$ is called **convex** if

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y), \qquad \lambda \in (0,1)$$

for $x, y \in E$.
For a convex function $f : E \to \mathbb{R} \cup \{\infty\}$ we define the convex conjugate by

$$f^* : E^* \to \mathbb{R} \cup \{\infty\}, f^*(\varphi) = \sup_{x\in E} (\varphi(x) - f(x)).$$

Theorem 4.2. (**Fenchel-Rockafellar duality**) Let E be a normed vector space, E^* its dual space and $A, B : E \to \mathbb{R} \cup \{\infty\}$ two convex functions. Assume there exists $u_0 \in E$ such that A is continuous at u_0 and $A(u_0), B(u_0)$ belongs to $\mathbb{R}$. Then,

$$\inf_{u \in E}\{A(u) + B(u)\} = \max_{u^* \in E^*}\{-A^*(-u^*) - B^*(u^*)\},$$

where A^*, B^* denote the convex conjugate of A respectively B.

Proof. For the proof we refer to Theorem 7.15 in [16] or for a short proof using the Hahn-Banach theorem see [17, p.24]. □

The following lemma shows two examples for convex functions which we will use for the proof of Theorem 4.1. For that we define

$$E_0 = \{u \in E \mid \text{there is } f, g : X \to \mathbb{R} \text{ with } u(x, y) = f(x) + g(y), x, y \in X\}.$$

Lemma 4.3. Let X be a finite set and $E = (\mathbb{R}^{X \times X}, \|\cdot\|_\infty)$ be a normed vector space. The functions $A_\varepsilon, B : E \to \mathbb{R} \cup \{\infty\}$ given by

$$A_\varepsilon(u) = \begin{cases} 0 & : u(x, y) \geq -d(x, y) - \varepsilon \text{ for all } x, y \in X \\ \infty & : \text{otherwise} \end{cases}$$

for $\varepsilon > 0$ and

$$B(u) = \begin{cases} \sum\limits_{x \in X} f(x)m_a(x) + \sum\limits_{y \in X} g(y)m_b(y) & : u \in E_0 \\ \infty & : \text{otherwise} \end{cases}$$

are convex.

Proof. We start with the convexity of A_ε. Let $u, v \in E$. We only need to show that $A_\varepsilon(\lambda u + (1 - \lambda)v) = \infty$ and $\lambda A_\varepsilon(u) + (1 - \lambda)A(v) = 0$ is excluded. Assume $u, v \geq -d(x, y) - \varepsilon$. Then, every linear combination of u and v of the form $\lambda u + (1 - \lambda)v, \lambda \in (0, 1)$, satisfies $\lambda u + (1 - \lambda)v \geq -d(x, y) - \varepsilon$. But this is a contradiction to $A_\varepsilon(\lambda u + (1 - \lambda)v) = \infty$. Thus, A_ε is convex.
We continue with the convexity of B. We show again that $B(\lambda u + (1 - \lambda)v) = \infty$ and $\lambda B(u) + (1 - \lambda)B(v) = 0$ is excluded. Assume there are functions $g, h \in E_0$ such that $\lambda u = g(x) + g(y)$ and $(1 - \lambda)v = h(x) + h(y)$. Then, $\lambda u + (1 - \lambda)v = g(x) + g(y) + h(x) + h(y) = g(x) + h(x) + g(y) + h(y)$. But this corresponds to a function $f = g + h$ in E_0 that satisfies $\lambda u + (1 - \lambda)v = f(x) + f(y)$ which is a contradiction to $B(\lambda u + (1 - \lambda)v) = \infty$. Thus, B is convex and the lemma is proven. □

4.4 Proof of Kantorovich-Rubenstein theorem

In this subsection, we let $E = \mathbb{R}^{X\times X}$ be equipped with the norm $\|\cdot\|_\infty$ where the particular choice of the norm is of course not relevant (cf. Remark 4.1). We apply Fenchel's duality to prove Theorem 4.1. We define for $M \in \mathbb{R}$

$$F_M(X) = \{(f,g) \mid f,g : X \to \mathbb{R}, f(x)+g(y) \leq M \cdot d(x,y) \text{ for all } x,y \in X\}.$$

Proposition 4.1. Let X be a finite set and $a, b \in X$. Then,

$$W_1(\hat{m}_a, \hat{m}_b) = \sup\left\{\sum_{z\in X} f(z)m_a(z) + \sum_{z\in X} g(z)m_b(z) \mid (f,g) \in F_1(X)\right\}.$$

Proof. Let $E = (\mathbb{R}^{X\times X}, \|\cdot\|_\infty)$. Since X is assumed to be finite, Lemma 4.2 gives that the dual space E^* is again $\mathbb{R}^{X\times X}$. Let $\varepsilon > 0$ and $A_\varepsilon, B : E \to \mathbb{R}\cup\{\infty\}$ be the convex functions defined in Lemma 4.3. Clearly, $A(0) = 0$ because $0 \geq -d(x,y) - \varepsilon$. For $u = 0$ we find functions $f, g \in E_0$ by $f = g \equiv 0$ such that $0 = f(x) + g(y)$ for all $x, y \in X$. Thus, $B(0) = 0$. Furthermore, A is continuous in 0 because $A_\varepsilon(u) = 0$ for $\|u\|_\infty \leq \varepsilon$. Thus, by Theorem 4.2 we find that

$$\inf_{u\in E}\{A_\varepsilon(u) + B(u)\} = \max_{u^*\in E^*}\{-A_\varepsilon^*(-u^*) - B^*(u^*)\}.$$

To achieve the infimum on the left hand side we can choose u such that there is $(f,g) \in F_{1+\varepsilon}(X)$ with $u(x,y) = f(x) + g(y) \leq d(x,y) + \varepsilon$ for all $x, y \in X$, since otherwise $B(u) = \infty$. Therefore the left hand side is equal to

$$\begin{aligned}
&\inf\{B(u) \mid \exists (f,g) \in F_{1+\varepsilon}(X) \text{ s.t. } u(x,y) = f(x)+g(y)\} \\
&= \inf\left\{\sum_{x\in X} f(x)m_a(x) + \sum_{y\in X} g(y)m_b(y) \mid (f,g) \in F_{1+\varepsilon}(X)\right\} \\
&= -\sup\left\{\sum_{x\in X} (-f(x))m_a(x) + \sum_{y\in X} (-g(y))m_b(y) \mid (f,g) \in F_{1+\varepsilon}(X)\right\} \\
&= -\sup\left\{\sum_{x\in X} f(x)m_a(x) + \sum_{y\in X} g(y)m_b(y) \mid (f,g) \in F_{1+\varepsilon}(X)\right\}.
\end{aligned}$$

Let us turn to the right hand side. We first compute A_ε^* and B^*. Let $\xi \in E^* = \mathbb{R}^{X\times X}$. We restrict ourselves to $u \in E$ that satisfy $A_\varepsilon(u) = 0$ to achieve the supremum. In particular, this implies that $u(x,y) \geq -d(x,y) - \varepsilon$ for all $x, y \in X$. Then, as $\xi(u) = \sum_{x,y\in X} u(x,y)\xi(x,y)$ by Lemma 4.2, we get

$$\begin{aligned}
A_\varepsilon^*(-\xi) &= \sup_{u\in E}\{-\xi(u) - A_\varepsilon(u)\} \\
&= \sup\left\{-\sum_{x,y\in X} u(x,y)\xi(x,y) \mid u \in E \text{ with } u(x,y) \geq -d(x,y) - \varepsilon\right\} \\
&= \sup\left\{\sum_{x,y\in X} u(x,y)\xi(x,y) \mid u \in E \text{ with } u(x,y) \leq d(x,y) + \varepsilon\right\}.
\end{aligned}$$

If there are $x_0, y_0 \in X$ such that $\xi(x_0, y_0) < 0$, then we let $v_r \in E = \mathbb{R}^{X\times X}, r > 0$, be such that $v_r(x_0, y_0) = \frac{r}{\xi(x_0,y_0)}$ and $v(x, y) = 0$ for $(x, y) \neq (x_0, y_0)$. Clearly, $v_r(x_0, y_0) \leq d(x, y) + \varepsilon$. Thus,

$$\sum_{x,y\in X} v_r(x, y)\xi(x, y) = \xi(x_0, y_0)v_r(x_0, y_0) = r > 0.$$

Thus, for ξ that is negative somewhere we infer

$$A^*(-\xi) = \sum_{x,y\in X} v_r(x, y)\xi(x, y) \geq r$$

for all $r > 0$. Letting $r \to \infty$ we find that $A^*(-\xi) = \infty$. If ξ is non negative, then

$$\begin{aligned} A^*(-\xi) &= \sup\left\{ \sum_{x,y\in X} u(x, y)\xi(x, y) \mid u \in E \text{ with } u(x, y) \leq d(x, y) + \varepsilon \right\} \\ &= \sum_{x,y\in X} (d(x, y) + \varepsilon)\xi(x, y). \end{aligned}$$

In summary, we have

$$A_\varepsilon^*(u) = \begin{cases} \sum\limits_{x,y\in X} (d(x, y) + \varepsilon)\xi(x, y) & : \xi \geq 0. \\ \infty & : \text{otherwise.} \end{cases}$$

We turn on to B^*. If $u \in E\backslash E_0$, then $B(u) = \infty$. Otherwise, for $u \in E_0$ we have

$$B(u) = \sum_{x\in X} f(x)m_a(x) + \sum_{y\in X} g(y)m_b(y)$$

for $f, g : X \to \mathbb{R}$ with $f(x) + g(y) = u(x, y)$. Therefore, for any $\xi \in \mathbb{R}^{X\times X}$ we get

$$\begin{aligned} B^*(\xi) &= \sup\{\xi(u) - B(u) \mid u \in E\} \\ &= \sup\left\{ \sum_{x,y\in X} u(x, y)\xi(x, y) - B(u) \mid u \in E_0 \right\} \\ &= \sup\Bigg\{ \sum_{x,y\in X} (f(x) + g(y))\xi(x, y) - \left(\sum_{x\in X} f(x)m_a(x) + \sum_{y\in X} g(y)m_b(y) \right) \\ &\qquad\qquad \mid f, g : X \to \mathbb{R} \Bigg\} \\ &= \sup\Bigg\{ \sum_{x\in X} f(x) \left(\sum_{y\in X} \xi(x, y) - m_a(x) \right) + \sum_{y\in X} g(y) \left(\sum_{x\in X} \xi(x, y) - m_b(y) \right) \\ &\qquad\qquad \mid f, g : X \to \mathbb{R} \Bigg\}. \end{aligned}$$

If $\xi \in P_{ab}$ we find using the properties of ξ that $B^*(\xi) = 0$. Otherwise, there is $x_0 \in X$ such that

$$\sum_{x \in X} \xi(x_0, y) \neq m_a(x_0) \text{ or } \sum_{y \in X} \xi(y, x_0) \neq m_b(x_0).$$

Assume without loss of generality that

$$\sum_{x \in X} (\xi(x_0, y) - m_a(x_0)) = C \neq 0.$$

Let $f_r : X \to \mathbb{R}, r > 0$, be such that $f_r(x_0) = \frac{r}{C}$ and $f(x) = 0$ for $x \neq x_0$ and let $g \equiv 0$. It follows that

$$B^*(\xi) = \sup \left\{ \sum_{x \in X} f(x) \left(\sum_{y \in X} \xi(x, y) - m_a(x) \right) \mid f : X \to \mathbb{R} \right\} \geq r$$

which tends to ∞ as $r \to \infty$. Thus,

$$B^*(\xi) = \begin{cases} 0 & : \text{if } \xi \in P_{ab}. \\ \infty & : \text{otherwise.} \end{cases}$$

Thus the right hand side reads as

$$\begin{aligned} \max\{-A^*_\varepsilon(-\xi) - B^*(\xi) \mid \xi \in E^*\} &= \max\{-A^*_\varepsilon(-\xi) \mid \xi \in P_{ab}\} \\ &= -\min \left\{ \sum_{x,y \in X} (d(x, y) + \varepsilon)\xi(x, y) \mid \xi \in P_{ab} \right\} \\ &= -\min \left\{ \sum_{x,y \in X} d(x, y)\xi(x, y) + \varepsilon \mid \xi \in P_{ab} \right\}. \end{aligned}$$

Summarizing, we obtain

$$\sup \left\{ \sum_{x,y \in X} f(x)m_a(x) + \sum_{x,y \in X} g(y)m_b(y) \mid (f, g) \in F_{1+\varepsilon}(X) \right\} = W_1(\hat{m}_a, \hat{m}_b) + \varepsilon$$

and the result follows by letting $\varepsilon \to 0$. □

Definition 4.3. For $f : X \to \mathbb{R}$ let $f^d, f^{dd} : X \to \mathbb{R} \cup \{\infty\}$ be defined as

$$f^d(y) = \inf_{x \in X} \{d(x, y) - f(x)\}$$

and

$$f^{dd}(x) = \inf_{y \in X} \{d(x, y) - f^d(y)\}.$$

The functions f^{dd} and f^d are said to be d**-conjugate**.

If X is a finite set the infimum can be replaced by a minimum. The next lemma shows that $(f^d, f^{dd}) \in F_1(X)$ if $(f, g) \in F_1(X)$.

Lemma 4.4. Let $(f, g) \in F_1(X)$. Then

$$f^d(x) + f^{dd}(y) \leq d(x, y)$$

and $f^d \geq g$ and $f^{dd} \geq f$.

Proof. Since $(f, g) \in F_1(X)$ we have $g(y) \leq d(x, y) - f(x)$ for all $x, y \in X$. As the left hand side does not depend on y, taking the infimum over x on the right hand side yields

$$g(y) \leq \inf_{x \in X} \{d(x, y) - f(x)\} = f^d(y)$$

for all $y \in X$. Furthermore,

$$\begin{aligned} f(x) + f^d(y) &= f(x) + \inf_{x \in X} \{d(x, y) - f(x)\} \\ &\leq f(x) + d(x, y) - f(x) = d(x, y). \end{aligned}$$

This implies $f(x) \leq d(x, y) - f^d(y)$ and therefore taking the infimum over y on the right hand side yields

$$f(x) \leq f^{dd}(x)$$

for all $x \in X$. Thus, $f^d \geq g$ and $f^{dd} \geq f$. Finally,

$$\begin{aligned} f^{dd}(x) + f^d(y) &= \inf_{y \in X} \{d(x, y) - f^d(y)\} + f^d(y) \\ &\leq d(x, y) - f^d(y) + f^d(y) = d(x, y) \end{aligned}$$

which completes the claim. □

Lemma 4.5. Let $f : X \to \mathbb{R}$. Then, $f^d \in \mathrm{Lip}_1$ and $f^{dd} = -f^d$.

Proof. For a fixed $x_0 \in X$ we define the function $\psi_{x_0} : X \to \mathbb{R}$ by $\psi_{x_0}(x) = d(x_0, x) - f(x_0)$. This function is Lipschitz-continuous on X as by the triangle inequality

$$\psi_{x_0}(x) - \psi_{x_0}(y) = d(x_0, x) - d(x_0, y) \leq d(x, y).$$

For $x, y \in X$ let $(x_n), (y_n)$ be sequences of vertices such that

$$\lim_{n \to \infty} \psi_{x_n}(x) = f^d(x) \text{ and } \lim_{n \to \infty} \psi_{y_n}(y) = f^d(y).$$

Then, picking $\varepsilon > 0$ and $N \in \mathbb{N}$ such that $|\psi_{x_n}(x) - f^d(x)| \leq \frac{\varepsilon}{2}$ and $|\psi_{y_n}(y) - f^d(y)| \leq \frac{\varepsilon}{2}$ for $n \geq N$, we conclude

$$\begin{aligned} |f^d(x) - f^d(y)| &\leq |f^d(x) - \psi_{x_n}(x)| + |\psi_{x_n}(x) - \psi_{y_n}(y)| + |\psi_{y_n}(y) - f^d(y)| \\ &\leq \frac{\varepsilon}{2} + d(x, y) + \frac{\varepsilon}{2} = d(x, y) + \varepsilon. \end{aligned}$$

Since ε was arbitrary, $f^d \in \mathrm{Lip}_1(X)$. Moreover, since $f^d \in \mathrm{Lip}_1(X)$ we have $-f^d(x) \leq d(x, y) - f^d(y)$ and taking the infimum over y on the right hand side provides

$$-f^d(x) \leq \inf_{y \in X} \{d(x, y) - f^d(y)\} = f^{dd}(x) \tag{4.3}$$

for all $x \in X$. Furthermore, $\inf_{y \in X}\{d(x,y) - f^d(y)\} \leq d(x,y) - f^d(y)$ for all $x \in X$ and in particular for $y = x$ we get

$$f^{dd}(x) = \inf_{y \in X}\{d(x,y) - f^d(y)\} \leq -f^d(x). \tag{4.4}$$

Thus, by Equations 4.3 and 4.4 we get $-f^d = f^{dd}$ which completes the claim. □

Theorem 4.1 follows from the next theorem.

Theorem 4.3. Let X be a finite set. Then,

$$\begin{aligned} W_1(\hat{m}_a, \hat{m}_b) &= \sup\left\{ \sum_{x \in X} f(x) m_a(x) + \sum_{y \in Y} g(y) m_b(y) \mid (f,g) \in F_1(X) \right\} \\ &= \sup\left\{ \sum_{x \in X} f^{dd}(x) m_a(x) + \sum_{y \in Y} f^d(y) m_b(y) \mid f : X \to \mathbb{R} \right\} \\ &= \sup\left\{ \sum_{x \in X} f(x) m_a(x) - \sum_{y \in X} f(y) m_b(y) \mid f \in \mathrm{Lip}_1(X) \right\}. \end{aligned}$$

Proof. Let $J(f,g) = \sum_{x \in X} f(x) m_a(x) + \sum_{y \in Y} g(y) m_b(y)$. By Lemma 4.4 and Lemma 4.5 we have

$$J(f,g) \leq J(f^{dd}, f^d) = J(-f^d, f^d)$$

for $(f,g) \in F_1(X)$. By Lemma 4.5 we have $(f^d, f^{dd}) \in F_1$. We infer

$$\begin{aligned} \sup\{J(f,g) \mid (f,g) \in F_1(X)\} &\leq \sup\{J(f^{dd}, f^d) \mid f : X \to \mathbb{R}\} \\ &= \sup\{J(-f^d, f^d) \mid f : X \to \mathbb{R}\} \\ &\leq \sup\{J(-f, f) \mid f \in \mathrm{Lip}_1(X)\} \\ &= \sup\{J(f, -f) \mid f \in \mathrm{Lip}_1(X)\}. \end{aligned}$$

Clearly, $(f, -f) \in F_1$ and thus

$$\sup\{J(f,-f) \mid f \in \mathrm{Lip}_1\} \leq \sup\{J(f,g) \mid (f,g) \in F_1\}.$$

Finally, the statement follows by Proposition 4.1. □

5 Ricci curvature for Platonic solids

In this section we apply the preliminary work to calculate the Ricci curvature for Platonic solids. We restrict to compute the curvature on edges connecting two adjacent vertices. A definition of Ricci curvature is given by Ollivier in [15]. There are a lot of papers that were written in the last years concerning this topic. Especially to mention is Y. Lin and S.-T. Yau who worked on Ricci curvature on graphs [13, 12, 4]. Yau was awarded the Fields medal in 1982 for his for "his contributions to partial differential equations, to the Calabi conjecture in algebraic geometry, to the positive mass conjecture of general relativity theory, and to real and complex Monge-Ampère equations" [22].

Let X be a finite set. As in Subsection 2.4 we let d be the metric on a graph g such that $d(x, y)$ is the minimal n such that the vertices x and y can be connected by a path of length n. We recall that in Subsection 2.4 the function $m_a, a \in X$ was defined as $m_a(x) = \frac{g(x,a)}{w_a}$ and $\hat{m}_a$ the associated measure.

Definition 5.1 (**Ollivier-Ricci curvature**)**.** Let g be a graph over X. For any two distinct points $x, y \in X$, the **Ollivier-Ricci curvature** along the path from x to y is defined as

$$\kappa(x, y) = 1 - \frac{W_1(\hat{m}_x, \hat{m}_y)}{d(x, y)}.$$

In the case of finite X we can calculate the curvature of a graph since we are able to estimate the transportation distance W_1 from above and below. For the upper bound we pick $\xi \in P_{ab}$ and by Theorem 3.2 we get

$$W_1(\hat{m}_a, \hat{m}_b) \leq \sum_{x,x\sim a} \sum_{y,y\sim b} d(x, y)\xi(x, y).$$

For the lower bound we find for $f \in \text{Lip}_1(X)$ by Theorem 4.1

$$W_1(\hat{m}_a, \hat{m}_b) \geq \sum_{z\in X,z\sim a} f(z)m_a(z) - \sum_{z\in X,z\sim b} f(z)m_b(z).$$

This means we aim to give a proper transference plan $\xi \in P_{ab}$ and a proper Lipschitz-continuous function f such that the upper and the lower bound are equal. In particular, the transference plan ξ is a matrix with terms $\xi(x, y)$ representing the mass moving from x to y.

In the following we calculate the Ollivier-Ricci curvature of all Platonic solids along the edge between two adjacent vertices. Let the graph that describes a Platonic solid be such that $g(x, y) \in \{0, 1\}$ for all $x, y \in X$. Therefore, w_a is the number of neighbors of a vertex $a \in X$ and $m_a(x) = \frac{1}{w_a}$ if $x \sim a$ and $m_a(x) = 0$ otherwise.

5.1 Tetrahedron

Let $X = \{1, 2, 3, 4\}$ be the set of vertices which are assigned to the tetrahedron as shown in the following picture, Figure 3.

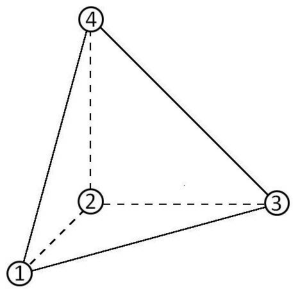

Figure 3: Tetrahedron

We calculate W_1 for the vertices $x = 1$ and $y = 2$ exemplarily. Due to the symmetry of the tetrahedron W_1 is equal for every arbitrary choice of two different vertices. The degree of vertex every is three which means $m_1(x) = \frac{1}{3}$ for all $x \in X \backslash \{1\}$ and 0 for $x = 1$. Analogously, $m_2(x) = \frac{1}{3}$ for all $x \in X \backslash \{2\}$ and 0 for $x = 2$. By Lemma 3.1 we get $\xi(x, 2) = 0$ for all $x \in X$ respectively $\xi(1, y) = 0$ for all $y \in X$ since

$$\sum_{x \in X, x \sim 1} \xi(x, 2) = \frac{g(2,2)}{w_2} = 0 \text{ and } \sum_{y \in X, y \sim 2} \xi(1, y) = \frac{g(1,1)}{w_1} = 0.$$

This means, as ξ is a matrix, the first line and the second column of ξ have only entries of value zero. Let a transference plan ξ be given by

$$\xi = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

An illustration of the moved mass by the transference plan is given in the following picture, Figure 4.

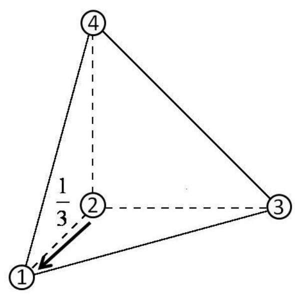

Figure 4: Transference plan for the tetrahedron

Then, by Theorem 3.2

$$\begin{aligned}W_1(\hat{m}_1, \hat{m}_2) &\leq \sum_{x\in X, x\sim 1} \sum_{y \in X, y\sim 2} d(x,y)\xi(x,y) \\ &= d(2,1)\xi(2,1) + d(3,3)\xi(3,3) + d(4,4)\xi(4,4) \\ &= \frac{1}{3}.\end{aligned}$$

Thus, $\frac{1}{3}$ is an upper bound for W_1. Furthermore, for any $f \in \mathrm{Lip}_1(X)$ we have by Theorem 4.1

$$\begin{aligned}W_1(\hat{m}_1, \hat{m}_2) \geq &\sum_{z\in X, z\sim 1} f(z)m_1(z) - \sum_{z\in X, z\sim 2} f(z)m_2(z) \\ =&f(2)m_1(2) + f(3)m_1(3) + f(4)m_1(4) \\ &- f(1)m_2(1) - f(3)m_2(3) - f(4)m_2(4) \\ =&\frac{1}{3}(f(2) - f(1)).\end{aligned}$$

Let $f : X \to \mathbb{R}$ be such that $f(2) = 1$ and $f(x) = 0$ for $x \neq 2$. An illustration of a Lipschitz-continuous function is given in the following picture, Figure 5. The function values of f are associated with the vertices of the tetrahedron.

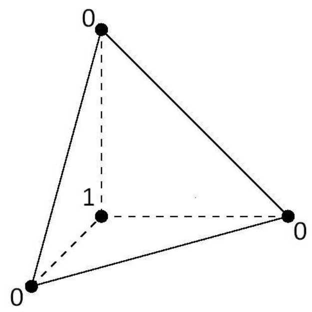

Figure 5: Lipschitz-function for the tetrahedron

With this choice of f we obtain $W_1(\hat{m}_1, \hat{m}_2) \geq \frac{1}{3}$. Thus, we found a proper transference plan ξ and a proper function $f \in \mathrm{Lip}_1(X)$ such that

$$W_1(\hat{m}_1, \hat{m}_2) = \frac{1}{3}.$$

Finally, the Ricci-curvature along the path from $x = 1$ to $y = 2$ is

$$\kappa(1, 2) = \frac{2}{3}.$$

By symmetry we get

$$\kappa(x, y) = \frac{2}{3}$$

for $x \sim y$.

5.2 Hexahedron

Let $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$ be the set of vertices which are assigned to the hexahedron as in the following picture, Figure 6.

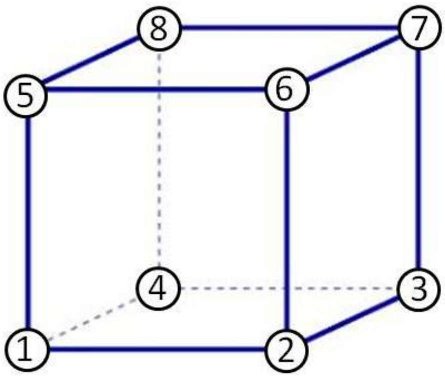

Figure 6: Hexahedron

We calculate W_1 for the vertices $x = 1$ and $y = 2$. Due to symmetry of the hexahedron W_1 is equal for every arbitrary choice of two adjacent vertices. The degree of every vertex is three. Thus, $m_1(x) = \frac{1}{3}$ for $x \in \{2, 4, 5\}$ and 0 otherwise. Analogously, $m_2(x) = \frac{1}{3}$ for $x \in \{1, 3, 6\}$ and 0 otherwise. Let a transference plan be given by

$$\xi(2,1) = \xi(4,3) = \xi(5,6) = \frac{1}{3}$$

and ξ is zero otherwise. An illustration of this transference plan is given in the following picture, Figure 7.

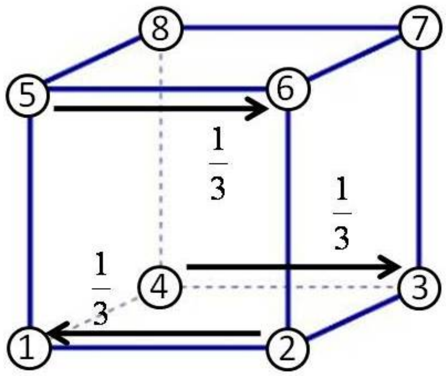

Figure 7: Transference plan for the hexahedron

Therefore,

$$\begin{aligned} W_1(\hat{m}_1, \hat{m}_2) &\leq \sum_{x \in X, x \sim 1} \sum_{y \in X, y \sim 2} d(x,y)\xi(x,y) \\ &= d(2,1)\xi(2,1) + d(4,3)\xi(4,3) + d(5,6)\xi(5,6) \\ &= 1. \end{aligned}$$

For the upper bound let $f : X \to \mathbb{R}$ be such that

$$f(x) = \begin{cases} 1 & : x \in \{2,4,5\}. \\ 0 & : \text{else.} \end{cases}$$

Clearly, $f \in \mathrm{Lip}_1(X)$ and, therefore, we have

$$\begin{aligned} W_1(\hat{m}_1, \hat{m}_2) &\geq \sum_{z \in X, z \sim 1} f(z) m_1(z) - \sum_{z \in X, z \sim 2} f(z) m_2(z) \\ &= \frac{1}{3}(f(2) - f(1) + f(4) - f(3) + f(5) - f(6)) \\ &= 1. \end{aligned}$$

This Lipschitz-continuous function f is illustrated in the following picture, Figure 8.

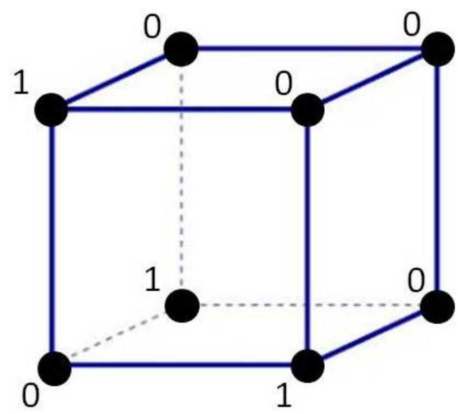

Figure 8: Lipschitz-function for the hexahedron

Therefore, $W_1(\hat{m}_1, \hat{m}_2) \geq 1$ and the transportation distance

$$W_1(\hat{m}_1, \hat{m}_2) = 1.$$

The Ricci-curvature along the edge connecting the vertices 1 and 2 is

$$\kappa(1,2) = 0.$$

By symmetry we get

$$\kappa(x,y) = 0$$

for $x \sim y$.

5.3 Octahedron

Let $X = \{1,2,3,4,5,6,7,8\}$ be the set of vertices which are assigned to the octahedron as in the following picture, Figure 9.

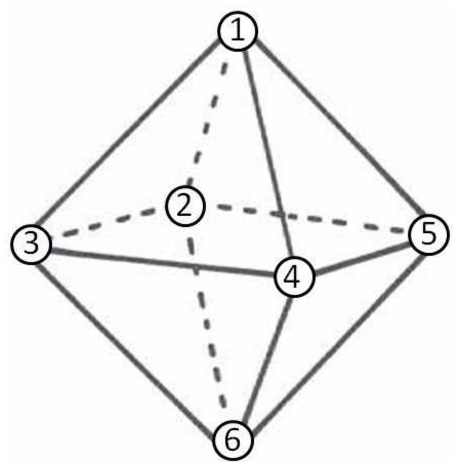

Figure 9: Octahedron

We calculate W_1 for the vertices $x = 1$ and $y = 2$. Due to symmetry of the octahedron W_1 is equal for every arbitrary choice of two adjacent vertices. The degree of every vertex is four. Thus, $m_1(x) = \frac{1}{4}$ for $x \in \{2, 3, 4, 5\}$ and zero otherwise. Analogously, $m_2(x) = \frac{1}{4}$ for $x \in \{1, 3, 5, 6\}$ and zero otherwise. Let a transference plan by given by

$$\xi(2, 1) = \xi(3, 3) = \xi(4, 6) = \xi(5, 5) = \frac{1}{4}$$

and ξ is zero otherwise. An illustration of this transference plan is given in the following picture, Figure 10.

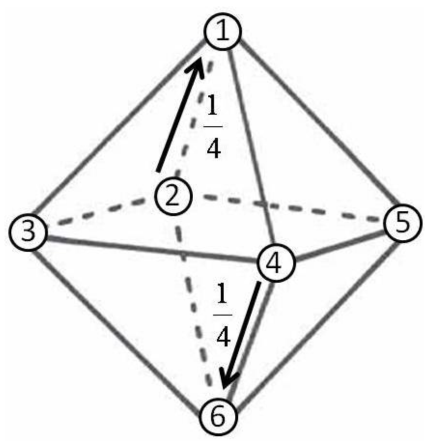

Figure 10: Transference plan for the octahedron

Therefore,

$$\begin{aligned} W_1(\hat{m}_1, \hat{m}_2) &\leq \sum_{x \in X, x \sim 1} \sum_{y \in X, y \sim 2} d(x,y)\xi(x,y) \\ &= d(1,4)\xi(1,4) + d(4,6)\xi(4,6) \\ &= \frac{1}{2}. \end{aligned}$$

Let $f : X \to \mathbb{R}$ be such

$$f(x) = \begin{cases} 1 & : x \in \{2,4\}. \\ 0 & : \text{else}. \end{cases}$$

Obviously, $f \in \mathrm{Lip}_1(X)$ and we obtain

$$\begin{aligned} W_1(\hat{m}_1, \hat{m}_2) &\geq \sum_{z \in X, z \sim 1} f(z)m_1(z) - \sum_{z \in X, z \sim 2} f(z)m_2(z) \\ &= \frac{1}{4}(f(2) - f(1) + f(4) - f(6)) \\ &= \frac{1}{2}. \end{aligned}$$

This Lipschitz-function f is illustrated in the following picture, Figure 11.

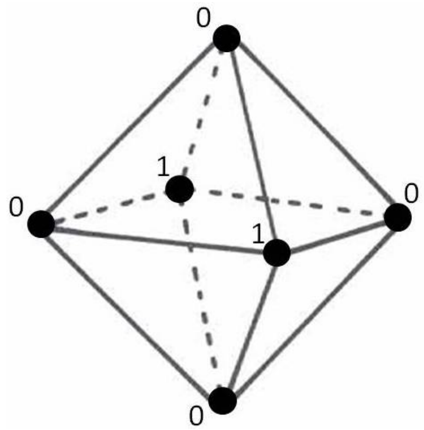

Figure 11: Lipschitz-function for the octahedron

Therefore, $W_1(\hat{m}_1, \hat{m}_2) \geq \frac{1}{2}$ and the transportation distance

$$W_1(\hat{m}_1, \hat{m}_2) = \frac{1}{2}.$$

The Ricci curvature along the edge connecting the vertices 1 and 2 is

$$\kappa(1,2) = \frac{1}{2}.$$

By symmetry we get

$$\kappa(x,y) = \frac{1}{2}$$

for all adjacent vertices $x, y \in X$.

5.4 Dodecahedron

Let $X = \{1, \dots, 20\}$ be a set of vertices of the dodecahedron. In the following picture, Figure 12, we only label the neighbors of the vertices of $x = 1$ and $y = 2$.

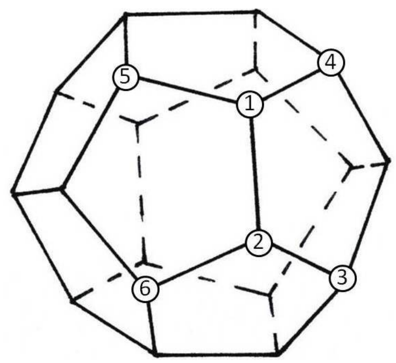

Figure 12: Dodecahedron

We calculate W_1 for the vertices $x = 1$ and $y = 2$. Due to symmetry of the dodecahedron W_1 is equal for every arbitrary choice of two different vertices. The degree of every vertex is three which means $m_1(x) = \frac{1}{3}$ for $x \in \{2, 4, 5\}$ and zero otherwise. Analogously, $m_2(x) = \frac{1}{3}$ for $x \in \{1, 3, 6\}$ and zero otherwise. Let a transference plan be given by

$$\xi(4, 1) = \xi(2, 3) = \xi(5, 6) = \frac{1}{3}$$

and ξ is zero otherwise. An illustration of this transference plan is given in the following picture, Figure 13.

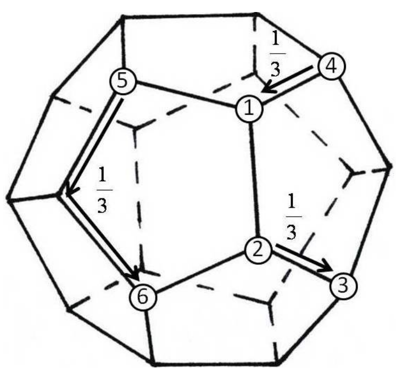

Figure 13: Transference plan for the dodecahedron

Therefore,

$$
\begin{aligned}
W_1(\hat{m}_1, \hat{m}_2) &\leq \sum_{x\in X, x\sim 1} \sum_{y\in X, y\sim 2} d(x,y)\xi(x,y) \\
&= d(4,1)\xi(4,1) + d(3,2)\xi(2,3) + d(5,6)\xi(5,6) \\
&= \frac{4}{3}.
\end{aligned}
$$

Thus, $\frac{4}{3}$ is an upper bound for W_1. Let $f : X \to \mathbb{R}$ be such that

$$
f(x) = \begin{cases} 1 & : x \in \{4,5\}. \\ 0 & : x \in \{1,2\}. \\ -1 & : x \in \{3,6\}. \end{cases}
$$

Clearly, $f \in \mathrm{Lip}_1(X)$ and we have

$$
\begin{aligned}
W_1(\hat{m}_1, \hat{m}_2) &\geq \sum_{z\in X, z\sim 1} f(z)m_1(z) - \sum_{z\in X, z\sim 2} f(z)m_2(z) \\
&= \frac{1}{3}(f(2) - f(1) + f(4) - f(3) + f(5) - f(6)) \\
&= \frac{4}{3}.
\end{aligned}
$$

This Lipschitz-function f is illustrated in the following picture, Figure 14.

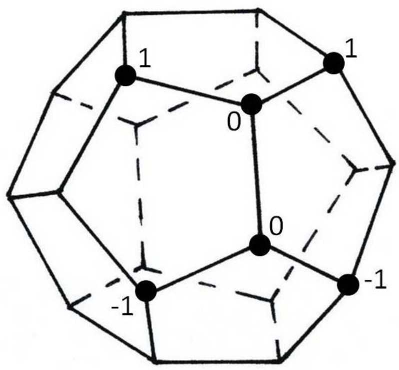

Figure 14: Lipschitz-function for the dodecahedron

Therefore, $W_1(\hat{m}_1, \hat{m}_2) \geq \frac{4}{3}$ and the transportation distance

$$
W_1(\hat{m}_1, \hat{m}_2) = \frac{4}{3}.
$$

The Ricci curvature along the edge connecting the vertices 1 and 2 is

$$
\kappa(1,2) = -\frac{1}{3}
$$

and

$$
\kappa(x,y) = -\frac{1}{3}
$$

for all adjacent vertices $x, y \in X$.

5.5 Icosahedron

Let $X = \{1, \dots, 12\}$ be a set of vertices of the icosahedron. In the following picture, Figure 15, we only label the neighbors of the vertices of $x = 1$ and $y = 2$.

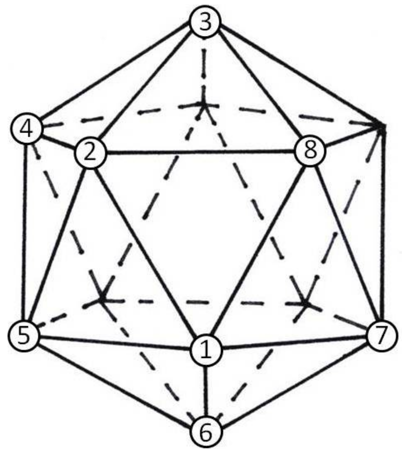

Figure 15: Icosahedron

We calculate W_1 for the vertices $x = 1$ and $y = 2$. Due to symmetry of the icosahedron W_1 is equal to every arbitrary choice of two adjacent vertices. The degree of every vertex is five. Thus, $m_1(x) = \frac{1}{5}$ for $x \in \{2, 5, 6, 7, 8\}$ and zero otherwise. Analogously, $m_2(x) = \frac{1}{5}$ for $x \in \{1, 3, 4, 5, 8\}$ and zero otherwise. Let a transference plan be given by

$$\xi(7,1) = \xi(2,3) = \xi(6,4) = \xi(5,5) = \xi(8,8) = \frac{1}{5}$$

and ξ is zero otherwise. An illustration of this transference plan is given in the following picture, Figure 16.

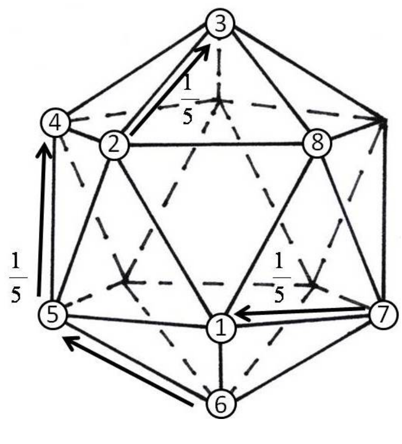

Figure 16: Transference plan for the icosahedron

Therefore,

$$
\begin{aligned}
W_1(\hat{m}_1, \hat{m}_2) \leq & \sum_{x\in X, x\sim 1} \sum_{y \in X, y\sim 2} d(x,y)\xi(x,y) \\
=& d(7,1)\xi(7,1) + d(2,3)\xi(2,3) + d(6,4)\xi(6,4) \\
& + d(5,5)\xi(5,5) + d(8,8)\xi(8,8) \\
=& \frac{4}{5}.
\end{aligned}
$$

Thus, $\frac{4}{5}$ is an upper bound for W_1. Let $f : X \to \mathbb{R}$ be such that

$$
f(x) = \begin{cases} 1 & : x \in \{6,7\}. \\ 0 & : x \in \{1,2,5,8\}. \\ -1 & : x \in \{3,4\}. \end{cases}
$$

Clearly, $f \in \mathrm{Lip}_1(X)$ and we obtain

$$
\begin{aligned}
W_1(\hat{m}_1, \hat{m}_2) \geq & \sum_{z\in X, z\sim 1} f(z)m_1(z) - \sum_{z\in X, z\sim 2} f(z)m_2(z) \\
=& \frac{1}{5}(f(2) - f(1) + f(6) - f(4) + f(7) - f(3)) \\
=& \frac{4}{5}.
\end{aligned}
$$

An illustration of this Lipschitz-continuous function is given in the following picture, Figure 17.

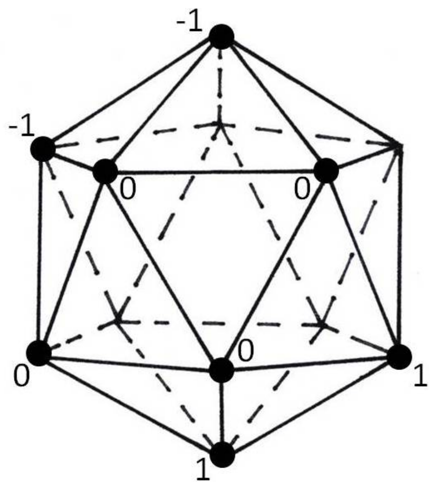

Figure 17: Lipschitz-function for the icosahedron

With this choice of f we obtain $W_1(\hat{m}_1, \hat{m}_2) \geq \frac{4}{5}$ and the transportation distance

$$
W_1(\hat{m}_1, \hat{m}_2) = \frac{4}{5}.
$$

The Ricci curvature along the edge connecting the vertices 1 and 2 is

$$\kappa(1,2) = \frac{1}{5}.$$

By symmetry we get

$$\kappa(x,y) = \frac{1}{5}$$

for $x \sim y$.

5.6 Summary

In the previous subsections we calculated the Ricci curvatures for adjacent vertices in the Platonic solids. Let x and y be two adjacent neighbors in the corresponding Platonic solid. The following table shows the Ricci curvature.

Tetrahedron	$\kappa(x,y) = \frac{2}{3}$
Hexahedron	$\kappa(x,y) = 0$
Octahedron	$\kappa(x,y) = \frac{1}{2}$
Dodecahedron	$\kappa(x,y) = -\frac{1}{3}$
Icosahedron	$\kappa(x,y) = \frac{1}{5}$

Despite of the conceptual strength of this definition of Ricci curvature the above table also displays its major weak point. The Platonic solids can be considered as discretizations of the Euclidean sphere. Therefore, one expects positive curvature intuitively. However, the hexahedron has flat curvature and the dodecahedron even negative curvature. This problem is subject to present research. For instance Lin and Yau propose a modified definition of discrete Ricci curvature in [14]. With this modified definition, the hexahedron enjoys positive curvature as do the tetrahedron, octahedron and icosahedron. Yet, the dodecahedron still has flat curvature. So, there is still room for improvement in future research.

References

[1] H. Amann and J. Escher, *Analysis I*, Birkhäuser Verlag, Basel, Number 2 (2002).

[2] H. Amann and J. Escher, *Analysis III*, Birkhäuser Verlag, Basel, Number 2 (2001).

[3] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris Sér. I Math. 305 (1987), no. 19, 805-808 (French, with English summary).

[4] F. Chung, Y. Lin and S.-T. Yau, *Harnack inequalities for graphs with non-negative Ricci curvature*, arXiv:1207.6612v1 (2012).

[5] K. Fritzsche, *Grundkurs Analysis 1*, Differentiation und Integration in einer Veränderlichen, Spektrum Akademischer Verlag, München (2005).

[6] L. Granieri, *Optimal Transportation and Minimizing Measures*, Monge-Kantorovich, Mather and Geometric Measure Theory, LAP LAMBERT Academic Publishing, Saarbrücken (2010).

[7] H. Heuser, *Lehrbuch der Analysis*, Teil 1, 14. Edition, Teubner Verlag, Stuttgart (2001)

[8] J. Jost and S. Liu, *Ollivier's Ricci curvature, local clustering and curvature dimension inequalities on graphs*, http://arxiv.org/abs/1103.4037 (2011).

[9] L. V. Kantorovich, *On the translocation of masses*, Dokl. Akad. Nauk SSSR, Volume **37**, Numbers 7-8 (1942).

[10] L. V. Kantorovich, *On a problem of Monge*, Uspekhi Mat. Nauk, 3, Number 2 (1948), p. 225 – 226.

[11] M. Keller, *Applications of operator theory: Discrete operators*, http://www.analysis-lenz.uni-jena.de/lsanalysis_multimedia/Matthias/12_DiscreteOperators_keinSkript.pdf, (2.7.2012).

[12] Y. Lin and S.-T. Yau, *Ricci curvature and eigenvalue estimation on Locally Finite Graphs*, Mathematics Research Letters, Vol.17, No.2, (2010), 345-358.

[13] Y. Lin, *Ricci curvature on graphs*, John H. Barrett Memorial Lecture, http://www.math.utk.edu/barrett/2010/talks/YongLinRicci.pdf, (2010).

[14] Y. Lin and S.-T. Yau, *A brief review on geometry and spectrum of graphs*, arXiv:1204.3168, (2012).

[15] Y. Ollivier, *Ricci Curvature of Markov Chains on Metric Spaces*, J. Funct. Anal. 256 (2009), no. 3, 810-864.

[16] J. van Tiel, *Convex Analysis*, An introductory text, John Wiley & Sons Ltd, Chichester (1984).

[17] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, American Mathematical Society Providence, Rhode Island, Volume **58** (2003).

[18] C. Villani, *Optimal Transport*, Old and New, Springer Verlag, Berlin (2009).

[19] W. Walter, *Analysis 1*, 7. Edition, Springer-Verlag, Berlin (2004).

[20] *http://www.nobelprize.org/nobel_prizes/economics/laureates/1975/*, (1.8.2012).

[21] http://www.icm2010.in/prize-winners-2010/fields-medal-cedric-villani, (5.8.2012).

[22] http://en.wikipedia.org/wiki/Shing-Tung_YauPrizes_and_awards, (01.08.2012).

List of Figures

Erklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der gegebenen Literatur und Hilfsmittel verfasst habe.

Sämtliche Stellen, die anderen Werken entnommen sind, wurden unter Angabe der Quellen als Entlehnung kenntlich gemacht.

Jena, den 10.08.2012