



HAL
open science

Approximation by finitely supported measures

Benoit Kloeckner

► **To cite this version:**

| Benoit Kloeckner. Approximation by finitely supported measures. 2010. hal-00461329v1

HAL Id: hal-00461329

<https://hal.science/hal-00461329v1>

Preprint submitted on 4 Mar 2010 (v1), last revised 2 Dec 2010 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

APPROXIMATION BY FINITELY SUPPORTED MEASURES

by

Benoît Kloeckner

1. Introduction

In this paper we are interested in the following question: given a finite measure μ , at what speed can it be approximated by finitely supported measures? To give a sense to the question, one needs a distance on the space of measures; we shall use the Wasserstein distances W_p , with arbitrary exponent $p \in [1, +\infty)$ (definitions are recalled in Section 2).

The case of exponent $p = 1$ has been previously studied under the name of *location problem*, and the case $p = 2$ is linked with *optimal centroidal Voronoi tessellations*; previous related works are discussed below.

This problem could be of interest for a numerical study of transportation problems, where measures can be represented by discrete ones. One would need to know the number of points needed to achieve some precision in the approximation.

We shall restrict our attention to compactly supported measures on \mathbb{R}^d ; moreover we only consider Borelian measures.

1.1. Statement of the results. — First we show that the order of convergence is determined by the dimension of the measure (see definitions in Section 2); Δ_N denotes the set of all measures supported in at most N points.

Theorem 1.1. — *If μ is compactly supported and Ahlfors regular of dimension $s > 0$, then*

$$W_p(\mu, \Delta_N) \approx \frac{1}{N^{1/s}}.$$

Here we write \approx to say that one quantity is bounded above and below by positive multiples of the other. Examples of Ahlfors regular measures are given by the volume measures on submanifolds, and self-similar measures (see for example [7]). Theorem 1.1, to be proved in a slightly more general and precise form in section 4, is simple and unsurprising; it reminds of ball packing and covering, and indeed relies on a standard covering argument.

In the particular case of absolutely continuous measures, one can give much more finer estimates. First, it is easily seen that if \square^d denotes the uniform measure on a unit cube of dimension d , then there is a constant $\theta(d, p)$ such that

$$W_p(\square^d, \Delta_N) \sim \frac{\theta(d, p)}{N^{1/d}}$$

(Proposition 5.1). Note that determining the precise value of $\theta(d, p)$ seems difficult; known cases are discussed in Section 1.2.

The main result of this paper is the following, where we denote by λ the Lebesgue measure on \mathbb{R}^d .

Theorem 1.2. — *If $\mu = \rho\lambda$ where ρ is a compactly supported function on \mathbb{R}^d , then for all $1 \leq p < \infty$ we have*

$$W_p(\mu, \Delta_N) \sim \frac{\theta(d, p) |\rho|_{\frac{d}{d+p}}^{1/p}}{N^{1/d}}$$

where $|\rho|_\beta = (\int \rho^\beta)^{1/\beta}$ is the L^β “norm”, here with $\beta < 1$ though.

Moreover, if (μ_N) is a sequence of finitely supported measures such that $\mu_N \in \Delta_N$ minimizes $W_p(\mu, \mu_N)$, then the sequence of probability measures $(\bar{\mu}_N)$ that are uniform on the support of μ_N converges weakly to the multiple of $\rho^{\frac{d}{p+d}}\lambda$ that has mass 1.

I discovered during the last part of the redaction of this paper that this result is already known, at least partially (see Section 1.2).

Theorem 1.2 is proved in Section 5. Note that the hypothesis that μ has compact support is obviously needed: otherwise, $|\rho|_{d/(d+p)}$ could be infinite. Even when μ is in $L^{d/(d+p)}$, there is the case where it is supported on a sequence of small balls going to infinity. Then the positions of the

balls is important in the quality of approximation and not only the profile of the density function.

Theorem 1.2 has no real analog for measure of fractional dimension.

Theorem 1.3. — *There is a s -dimensional Ahlfors regular measure κ on \mathbb{R} (namely, κ is the Cantor dyadic measure) such that $W_p(\kappa, \Delta_N)N^{1/s}$ has no limit.*

Section 6 is devoted to this example.

Part of the interest of Theorem 1.2 comes from the following observation, to be discussed in Section 7: when $p = 2$, the support of a distance minimizing $\mu_N \in \Delta_N$ generates a centroidal Voronoi tessellation, that is, each point is the center of mass (with respect to μ) of its Voronoi cell. We thus get the asymptotic repartition of an important family of centroidal Voronoi tessellations, which enables us to prove some sort of principle of energy equidistribution.

1.2. Discussion of previously known results. — There are several previous works closely related to the content of this paper.

The case $p = 1$ is usually called the *location problem*. For this exponent, Theorem 1.2 has been proved by Bouchitté, Jimenez and Rajesh [1] under the additional assumption that ρ is lower semi-continuous.

Mosconi and Tilli in [14] have studied (for any exponent p) the *irrigation problem*, where the approximating measures are supported on connected sets of length $< \ell$ (the length being the 1-dimensional Hausdorff measure) instead of being supported on N points; the order of approximation is then $\ell^{1/(d-1)}$. These authors note that their methods yield the result of [1] without the assumption on ρ . Our approach here gives at least another method to prove Theorem 1.2, and since there is up to our knowledge no completely written down proof we shall give one.

Brancolini, Buttazzo, Santambrogio and Stepanov compare in [2] the location problem with its “short-term planning” version, where the support of $\text{supp } \mu_N$ is constructed by adding one point to that of μ_{N-1} , minimizing the cost only locally in N .

Some values of $\theta(d, p)$ have been determined. First, it is easy to compute them in dimension 1:

$$\theta(1, p) = \frac{(p+1)^{-1/p}}{2}.$$

The case $p = d = 2$ has been solved by Fejes Tóth [8] (and also by Newmann [15] later on) and the case $p = 1, d = 2$ by Morgan and Bolton [13]:

$$\theta(2, 2) = 5\sqrt{3}/54 \quad \theta(2, 1) = 2^{-2/3}3^{-7/4}(4 + \ln 27).$$

Both are obtained by the regular hexagon (that is, the given θ is the distance between the uniform measure on a regular hexagon and a Dirac mass at its center). All other cases are open to our knowledge. For numerical evidence in the case $p = 2, d = 3$ see Du and Wang [6]. Note that in the limit case $p = \infty$, determining θ amounts to determining the minimal density of a ball covering of \mathbb{R}^d , which is arguably as challenging as determining the maximal density of a ball packing, a well-known open problem if $d > 3$.

Concerning the order of convergence, it is worth comparing with the problem of estimating the distance from a measure μ to empirical measures $\bar{\mu}_N = N^{-1} \sum_k \delta_{X_k}$ where X_1, \dots, X_N are independent random variables of law μ . It seems that $\bar{\mu}_N$ is almost optimal in the sense that $W_2(\mu, \bar{\mu}_N) \sim C N^{-1/d}$ almost surely (under moment conditions, but here we take μ compactly supported so this is not an issue); Horowitz and Karandikar have shown in [10] that $W_2(\mu, \bar{\mu}_N)$ has the order at most $N^{-1/(d+4)}$ and the better exponent above is suggested in the Mathematical Review of that paper. A somewhat surprising result is that if $\bar{\mu}'_N$ is another empirical measure of μ , then $W_2(\bar{\mu}_N, \bar{\mu}'_N)$ has the order $N^{-(1-1/d)}$ if $d \geq 3$ (see for example Dobrić and Yukich [4]), so that in high dimension two empirical measures are far closer one to each other than to the original measure.

In the case $p = 2$, the problem is linked to (optimal) *centroidal Voronoi Tessellation*, see Section 7 and [5]. In that paper (Section 6.4.1), the principle of energy equidistribution is given in the 1-dimensional case for smooth density ρ . Our corollary 7.1 in the last section generalize this to non regular densities, every exponent, and all dimension; it is however quite a direct consequence of Theorem 1.2.

It seems that the location problem and the study of optimal CVTs, although both specific case of the problem stated in this paper (for $p = 1$ and $p = 2$ respectively), have been only studied independently. Our hope is that noticing this proximity will encourage progress on each question to be translated in the other.

Last, note that Lipman studies in [12] the problem of approximating measures by finite-supported ones in general metric spaces, with explicit error bounds.

1.3. Related open questions. — The number N of points of the support may be the first measure of complexity of a finitely supported measure that one comes up with, but it is not necessarily the most relevant. Concerning the problem of numerical analysis of transportation problems, numbers are usually encoded in a computer by floating numbers. One could therefore define the complexity of a measure supported on points of decimal coordinates, with decimal quantity of mass at each point as the memory size needed to describe it, and search to minimize the distance to a given μ among measures of given complexity.

Another possible notion of complexity is entropy : one defines

$$h\left(\sum_i m_i \delta_{x_i}\right) = -\sum_i m_i \ln(m_i).$$

A natural question is to search a μ_h that minimizes the distance to μ among the finitely supported measures of entropy at most h , and to study the behavior of μ_h when we let $h \rightarrow \infty$.

Acknowledgements. — I am grateful to Romain Joly, Vincent Munier, Hervé Pajot, Rémy Peyre and Cédric Villani for interesting discussions or comments.

2. Recalls and definitions

2.1. Notations. — Given two sequences $(u_n), (v_n)$ of non-negative real numbers, we shall write:

- $u_n \lesssim v_n$ to mean that there exists a positive real a and an integer N_0 such that $u_n \leq av_n$ for all $n \geq N_0$,
- $u_n \approx v_n$ to mean $u_n \lesssim v_n$ and $u_n \gtrsim v_n$.

2.2. Ahlfors regularity and a covering result. — We denote by $B(x, r)$ the closed ball of radius r and center x ; sometimes, when $B = B(x, r)$ and $k \in \mathbb{R}$, we denote by kB the ball $B(x, kr)$.

Let μ be a finite, compactly supported measure on \mathbb{R}^d and $0 < s < +\infty$. One says that μ is *Ahlfors regular* of dimension s if there is a

constant C such that for all $x \in \text{supp } \mu$ and for all $r \leq \text{diam}(\text{supp } \mu)$, one has

$$C^{-1}r^s \leq \mu(B(x, r)) \leq Cr^s.$$

This is a strong condition, but is satisfied for example by auto-similar measures, see [11, 7] for definitions and Section 6 for the most famous example of the Cantor measure.

Note that if μ is Ahlfors regular of dimension s , then s is the Hausdorff dimension of $\text{supp } \mu$ (and therefore $s \leq d$), see [9, Sec. 8.7].

We shall need the following classical covering result.

Proposition 2.1 (5 δ covering). — *If X is a closed set and \mathcal{F} is a family of balls of uniformly bounded diameter such that $X \subset \bigcup_{\mathcal{F}} B$, then there is a subfamily \mathcal{G} of \mathcal{F} such that:*

- $X \subset \bigcup_{\mathcal{G}} 5B$,
- $B \cap B' = \emptyset$ whenever $B \neq B' \in \mathcal{G}$.

2.3. Wasserstein distances. — Here we recall some basic facts on optimal transportation and Wasserstein distances. For more information, the reader is suggested to look for example at Villani's book [16] which provides a very good introduction to this topic.

First consider the case $p < \infty$, which shall attract most of our attention. A finite measure μ on \mathbb{R}^d is said to have *finite p -th moment* if for some (hence all) $x_0 \in \mathbb{R}^d$ the following holds:

$$\int_{\mathbb{R}^d} d(x_0, x)^p \mu(dx) < +\infty.$$

In particular, any compactly supported finite measure has finite p -th moment for all p .

Let μ_0, μ_1 be two finite measures having finite p -th moment and the same mass. A *transport plan* between μ_0 and μ_1 is a measure Π on $\mathbb{R}^d \times \mathbb{R}^d$ that has μ_0 and μ_1 as marginals, that is : $\Pi(A \times \mathbb{R}^d) = \mu_0(A)$ and $\Pi(\mathbb{R}^d \times A) = \mu_1(A)$ for all Borelian set A . One shall think of Π as a assignement of mass: $\Pi(A \times B)$ represents the mass sent from A to B .

The L^p cost of a transport plan is defined as

$$c_p(\Pi) = \int_{\mathbb{R}^d \times \mathbb{R}^d} d(x, y)^p \Pi(dxdy).$$

One defines the L^p Wasserstein distance by

$$W_p(\mu_0, \mu_1) = \inf_{\Pi} c_p(\Pi)^{1/p}$$

where the infimum is on all transport plan between μ_0 and μ_1 . One can show that there is always a transport plan that achieves this infimum, and that W_p defines a distance on the set of measures with finite p -th moment and given mass.

Moreover, the restriction of W_p to the probability measures *supported on some fixed compact* subset of \mathbb{R}^d metrizes the weak topology. Without this restriction, it defines a finer topology.

Most of the time, one restricts itself to probability measures. Here, we shall use extensively mass transportation between submeasures of the main measures to study, so that we need to consider measures of arbitrary mass. Given positive measures μ and ν , we write that $\mu \leq \nu$ if $\mu(A) \leq \nu(A)$ for all borelian set A , which means that $\nu - \mu$ is also a positive measure.

It is important to notice that $c_p(\Pi)$ is homogeneous of degree 1 in the total mass and of degree p on distances, so that if φ is a similitude of ratio r , we have $W_p(m \varphi_{\#} \mu_0, m \varphi_{\#} \mu_1) = m r W_p(\mu_0, \mu_1)$.

The case $p = \infty$ is obtained as a limit of the finite case, see [3]. Let μ_0 and μ_1 be compactly supported measures of the same mass and let Π be a transport plan between μ_0 and μ_1 . The L^∞ length of Π is defined as

$$\ell_\infty(\Pi) = \sup\{d(x, y) \mid x, y \in \text{supp } \Pi\}$$

that is, the maximal distance moved by some infinitesimal amount of mass when applying Π . The L^∞ distance between μ_0 and μ_1 then is

$$W_\infty(\mu_0, \mu_1) = \inf_{\Pi} \ell_\infty(\Pi)$$

where the infimum is on all transport plan from μ_0 to μ_1 . In a sense, the L^∞ distance is a generalisation to measures of the Hausdorff metric on compact sets. We shall use ℓ_∞ , but not d_∞ . The problem of minimizing $W_\infty(\mu, \Delta_N)$ is a matter of covering $\text{supp } \mu$ (independently of μ itself), a problem with quite a different taste than our.

3. Preparatory results

The following lemmas are usefull tools we shall need; the two first at least cannot pretend to any kind of originality by themselves.

Lemma 3.1 (monotony). — *Let μ and ν be finite measures of equal mass and $\tilde{\mu} \leq \mu$. Then there is a measure $\tilde{\nu} \leq \nu$ (in particular, $\text{supp } \tilde{\nu} \subset$*

supp ν) such that

$$W_p(\tilde{\mu}, \tilde{\nu}) \leq W_p(\mu, \nu).$$

Proof. — Let Π be an optimal transportation plan from μ to ν . We construct a low-cost transportation plan from $\tilde{\mu}$ to $\tilde{\nu}$ by disintegrating Π .

There is family of finite measures $(\eta_x)_{x \in \mathbb{R}^d}$ such that $\Pi = \int \eta_x \mu(dx)$, that is

$$\Pi(A \times B) = \int_A \eta_x(B) \mu(dx)$$

for all Borelian A and B . Define

$$\tilde{\Pi}(A \times B) = \int_A \eta_x(B) \tilde{\mu}(dx)$$

and let $\tilde{\nu}$ be the second factor projection of $\tilde{\Pi}$. Since $\tilde{\Pi} \leq \Pi$, we have $\tilde{\nu} \leq \nu$ and $c_p(\tilde{\Pi}) \leq c_p(\Pi)$; moreover $\tilde{\Pi}$ is a transport plan from $\tilde{\mu}$ to $\tilde{\nu}$ by definition of $\tilde{\nu}$. \square

Lemma 3.2 (summing). — *Let (μ, ν) and $(\tilde{\mu}, \tilde{\nu})$ be finite measures with pairwise equal masses. Then*

$$W_p^p(\mu + \tilde{\mu}, \nu + \tilde{\nu}) \leq W_p^p(\mu, \nu) + W_p^p(\tilde{\mu}, \tilde{\nu}).$$

Proof. — Let Π and $\tilde{\Pi}$ be optimal transport plans between respectively μ and ν , $\tilde{\mu}$ and $\tilde{\nu}$. Then $\Pi + \tilde{\Pi}$ is a transport plan between $\mu + \tilde{\mu}$ and $\nu + \tilde{\nu}$ whose cost is $c_p(\Pi + \tilde{\Pi}) = c_p(\Pi) + c_p(\tilde{\Pi})$. \square

This very simple results have a particularly important consequence concerning our question.

Lemma 3.3 (L^1 stability). — *Let μ and $\tilde{\mu}$ be finite compactly supported measures supported, $\varepsilon \in (0, 1)$ and (μ_N) be any sequence of N -supported measures.*

There is a sequence of N -supported measures $\tilde{\mu}_N$ such that there are at most εN points in $\text{supp } \tilde{\mu}_N \setminus \text{supp } \mu_N$ and

$$W_p^p(\tilde{\mu}, \tilde{\mu}_N) \leq W_p^p(\mu, \mu_{N_1}) + O\left(\frac{|\mu - \tilde{\mu}|_{TV}}{(\varepsilon N)^{p/d}}\right)$$

where N_1 is equivalent to (and at least) $(1 - \varepsilon)N$, $|\cdot|_{TV}$ is the total variation norm and the constant in the O depends only on d and the diameter of a cube supporting both μ and $\tilde{\mu}$.

In particular we get

$$W_p^p(\mu, \Delta_N) \leq W_p^p(\tilde{\mu}, \Delta_{N_1}) + O\left(\frac{|\mu - \tilde{\mu}|_{TV}}{(\varepsilon N)^{p/d}}\right).$$

The name of this result has been chosen to emphasize that the total variation distance between two absolutely continuous measures is the L^1 distance between their densities.

Proof. — We can write $\tilde{\mu} = \mu' + \nu$ where $\mu' \leq \mu$ and ν is a positive measure of total mass at most $|\mu - \tilde{\mu}|_{TV}$. Let C be a cube supporting both μ and $\tilde{\mu}$, and denote its side length by L .

Let $M = \lfloor (\varepsilon N)^{1/d} \rfloor$ and consider the centers x_1, \dots, x_{M^d} of the regular subdivision of C by cubes. By sending each point of C to a closest x_i , one constructs a transport plan between ν and a M^d -supported measure ν_N whose cost is at most $|\mu - \tilde{\mu}|_{TV} (L\sqrt{d}/2M)^p$.

Let $N_1 = N - M^d$. The monotony lemma gives a measure $\mu'_N \leq \mu_{N_1}$ (in particular, μ'_N is N_1 -supported) such that

$$W_p(\mu', \mu'_N) \leq W_p(\mu, \mu_{N_1}).$$

The summing lemma now shows that

$$W_p^p(\tilde{\mu}, \mu'_N + \nu_N) \leq W_p^p(\mu, \mu_{N_1}) + \frac{|\mu - \tilde{\mu}|_{TV} (L\sqrt{d}/2)^p}{M^p}.$$

□

Note that the presence of the $|\mu - \tilde{\mu}|_{TV}$ factor will be crucial in the sequel, but would not be present in the limit case $p = \infty$, which is therefore very different.

Let us end with a result showing that no mass is moved very far away by an optimal transport plan to a N -supported measure if N is large enough.

Lemma 3.4 (localization). — *Let μ be a compactly supported finite measure. If μ_N is a closest N -supported measure to μ in L^p -distance and Π_N is a L^p optimal transport plan between μ and μ_N , then when N goes to ∞ ,*

$$\ell_\infty(\Pi_N) \rightarrow 0.$$

Proof. — Assume on the contrary that there are sequences $N_k \rightarrow \infty$, $x_k \in \text{supp } \mu$ and a number $\varepsilon > 0$ such that Π_{N_k} moves x_k by a distance at least ε . There is a covering of $\text{supp } \mu$ by a finite number of balls of radius $\varepsilon/3$. Up to extracting a subsequence, we can assume that all

x_k lie in one of this balls, denoted by B . Since B is a neighborhood of x_k and $x_k \in \text{supp } \mu$, we have $\mu(B) > 0$. Since Π_{N_k} is optimal, it moves x_k to a closest point in $\text{supp } \mu_{N_k}$, which must be at distance at least ε from x_k . Therefore, every point in B is at distance at least $\varepsilon/3$ from $\text{supp } \mu_{N_k}$, so that $c_p(\Pi_{N_k}) \geq \mu(B)(\varepsilon/3)^p > 0$, in contradiction with $W_p(\mu, \Delta_N) \rightarrow 0$. \square

4. Approximation rate and dimension

Theorem 1.1 is the union of the two following propositions. Note that the estimates given do not depend much on p , so that in fact Theorem 1.1 stays true when $p = \infty$.

Proposition 4.1. — *If μ is a compactly supported probability measure on \mathbb{R}^d and if for some $C > 0$ and for all $r \leq \text{diam}(\text{supp } \mu)$, one has*

$$C^{-1}r^s \leq \mu(B(x, r))$$

then for all N

$$W_p(\mu, \Delta_N) \leq \frac{5C^{1/s}}{N^{1/s}}.$$

Proof. — The 5δ covering proposition above implies that given any $\delta > 0$, there is a subset \mathcal{G} of $\text{supp } \mu$ such that

- $\text{supp } \mu \subset \bigcup_{x \in \mathcal{G}} B(x, 5\delta)$,
- $B(x, \delta) \cap B(x', \delta) = \emptyset$ whenever $x \neq x' \in \mathcal{G}$.

In particular, as soon as $\delta < \text{diam}(\text{supp } \mu)$ one has

$$1 \geq \sum_{x \in \mathcal{G}} \mu(B(x, \delta)) \geq |\mathcal{G}| C^{-1} \delta^s$$

so that \mathcal{G} is finite, with $|\mathcal{G}| \leq C\delta^{-s}$.

Let $\tilde{\mu}$ be a measure supported on \mathcal{G} , that minimizes the L^p distance to μ among those. A way to construct $\tilde{\mu}$ is to assign to a point $x \in \mathcal{G}$ a mass equal to the μ -measure of its Voronoi cell, that is of the set of points nearest to x than to any other points in \mathcal{G} . The mass at a point at equal distance from several elements of \mathcal{G} can be split indifferently between those. The previous discussion also gives a transport plan from μ to $\tilde{\mu}$, where each bit of mass moves a distance at most 5δ , so that $W_p(\mu, \tilde{\mu}) \leq 5\delta$ (whatever p).

Now, let N be a positive integer and choose $\delta = (C/N)^{1/s}$. The family \mathcal{G} obtained from that δ has less than N elements, so that $W_p(\mu, \Delta_N) \leq 5(C/N)^{1/s}$. \square

Proposition 4.2. — *If μ is a probability measure on \mathbb{R}^d and if for some $C > 0$ and for all r , one has*

$$\mu(B(x, r)) \leq Cr^s$$

then for all N ,

$$\left(\frac{s}{s+p}\right)^{1/p} C^{-1/s} \frac{1}{N^{1/s}} \leq W_p(\mu, \Delta_N).$$

Proof. — Consider a measure $\mu_N \in \Delta_N$ that minimizes the distance to μ . For all $\delta > 0$, the union of the balls centered at $\text{supp } \mu_N$ and of radius δ has μ -measure at most $NC\delta^s$. In any transport plan from μ to μ_N , a quantity of mass at least $1 - NC\delta^s$ travels a distance at least δ , so that in the best case the quantity of mass traveling a distance between $\delta < (NC)^{-1/s}$ and $\delta + d\delta$ is $NCs\delta^{s-1}d\delta$. It follows that

$$W_p(\mu, \mu_N)^p \geq \int_0^{(NC)^{-1/s}} sNC\delta^{s-1}\delta^p d\delta$$

so that $W_p(\mu, \mu_N) \geq (s/(s+p))^{1/p}(NC)^{-1/s}$. \square

In fact, Theorem 1.1 applies to more general measures, for example combination of Ahlfors regular ones, thanks to the following.

Lemma 4.3. — *If $\mu = a_1\mu_1 + a_2\mu_2$ where $a_i > 0$ and μ_i are probability measures such that $W_p(\mu_2, \Delta_N) \lesssim W_p(\mu_1, \Delta_N)$ and $W_p(\mu_1, \Delta_N) \lesssim W_p(\mu_1, \Delta_{2N})$ then $W_p(\mu, \Delta_N) \approx W_p(\mu_1, \Delta_N)$.*

Proof. — By the monotony lemma, $W_p(a_1\mu_1, \Delta_N) \leq W_p(\mu, \Delta_N)$ so that $W_p(\mu_1, \Delta_N) \leq a_1^{-1/p}W_p(\mu, \Delta_N)$.

The summing lemma gives

$$W_p(\mu, \Delta_{2N}) \leq (W_p(a_1\mu_1, \Delta_N)^p + W_p(a_2\mu_2, \Delta_N)^p)^{1/p}$$

so that

$$W_p(\mu, \Delta_{2N}) \lesssim W_p(a_1\mu_1, \Delta_N) \lesssim W_p(\mu_1, \Delta_{2N}).$$

Since $W_p(\mu, \Delta_{2N+1}) \leq W_p(\mu, \Delta_{2N})$ we also get

$$W_p(\mu, \Delta_{2N+1}) \lesssim W_p(\mu_1, \Delta_{2N}) \lesssim W_p(\mu_1, \Delta_{4N}) \leq W_p(\mu_1, \Delta_{2N+1})$$

\square

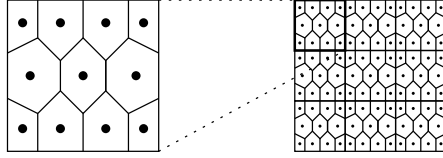


FIGURE 1. An optimal N -supported measure can be used to construct an good $k^d N$ -supported measure for all k .

The following is now an easy consequence of this lemma.

Corollary 4.4. — Assume that $\mu = \sum_{i=1}^k a_i \mu_i$ where $a_i > 0$ and μ_i are probability measures that compactly supported and Ahlfors regular of dimension $s_i > 0$. Let $s = \max_i(s_i)$. Then

$$W_p(\mu, \Delta_N) \approx \frac{1}{N^{1/s}}.$$

5. Absolutely continuous measures

In this section we prove Theorem 1.2. We shall approximate (in the L^1 sense) a measure with density by a combination of uniform measures in squares. Let us start with the uniform measure \square^d on the unit cube of \mathbb{R}^d .

Proposition 5.1. — There is a number $\theta(d, p) > 0$ such that

$$W_p(\square^d, \Delta_N) \sim \frac{\theta(d, p)}{N^{1/d}}.$$

The proof is obviously not new, since it is the same argument that shows that an optimal packing (or covering) of the Euclidean space must have a well-defined density (its upper and lower densities are equal).

Proof. — Let $c(N) = W_p(\square^d, \Delta_N)^p$. We already know that $c(N) \approx N^{-p/d}$, so let $A = \liminf N^{p/d} c(N)$ and consider any $\varepsilon > 0$.

Let N_1 be an integer so that $c(N_1) \leq (A + \varepsilon) N_1^{-p/d}$ and let $\mu_1 \in \Delta_{N_1}$ be nearest to μ . For any integer ℓ , we can write $\ell = k^d + q$ where $k = \lfloor \ell^{1/d} \rfloor$ and q is an integer; then $q = O(\ell^{1-1/d}) = o(\ell)$ where the o depends only on d .

Divide the cube into k^d cubes of side length $1/k$, and consider the element μ_k of $\Delta_{k^d N_1}$ obtained by duplicating μ_1 in each of the cubes, with scaling factor k^{-1} and mass factor k^{-d} (see figure 1). The obvious

transport plan obtained in the same way from the optimal one between \square^d and μ_1 has total cost $k^{-p}c(N_1)$, so that

$$c(\ell N_1) \leq k^{-p} \frac{A + \varepsilon}{N_1^{p/d}} = \left(\frac{\ell}{k^d} \right)^{p/d} \frac{A + \varepsilon}{(\ell N_1)^{p/d}}.$$

But since $k^d \sim \ell$, for ℓ large enough we get

$$c(\ell N_1) \leq \frac{A + 2\varepsilon}{(\ell N_1)^{p/d}}.$$

Now $N \sim \lfloor N/N_1 \rfloor N_1$ so that for N large enough $c(N) \leq (A + 3\varepsilon)N^{-p/d}$. This proves that $\limsup N^{p/d}c(N) \leq A + 3\varepsilon$ for all $\varepsilon > 0$. \square

Note that we used the self-similarity of the cube at many different scales; the result does not hold with more general self-similar (fractal) measures, see Section 6.

Denote by $C \subset \mathbb{R}^d$ any one cube of side length $\delta > 0$ and by \square_δ the restriction of λ to C .

Lemma 5.2. — *Let T_i ($1 \leq i \leq I$) be translations such that the interiors of the cubes $T_i(C)$ do not overlap, and assume that μ is a finite sum*

$$\mu = \sum_{i=1}^I \rho_i (T_i)_\# \square_\delta$$

where ρ_i are positive reals; ρ shall denote the vector (ρ_i) as well as the function that takes the value ρ_i on the interior of $T_i(C)$ and 0 at other points.

Let $\mu_N \in \Delta_N$ be a W_p minimizing sequence and define N_i (with implicit dependency on N) as the number of points of $\text{supp } \mu_N$ that lie on $T_i(C)$, the points lying on a common boundary of two or more cubes being attributed arbitrarily to one of them.

If the vector $(\delta^{-d}N_i/N)_i$ (which gives the density of $\text{supp } \mu_N$ in $T_i(C)$, normalized by N) has a limit $x = (x_i)$ when $N \rightarrow \infty$, then

$$W_p(\mu, \mu_N) \sim \frac{\theta(d, p) F(\rho; x)^{1/p}}{N^{1/d}}$$

where

$$F(\rho; x) = \sum_{i=1}^I \rho_i x_i^{-p/d} \delta^d.$$

Proof. — For simplicity we denote $c_p(N) = W_p(\mu, \Delta_N)^p$. Since $c_p(N) \rightarrow 0$, none of the x_i vanishes: the cost of a transport plan between μ and a measure supported outside $T_i(C)$ is bounded from below by $\rho_i \delta^d (\delta/3)^{p+d}$.

Let ε be any positive, small enough number. By considering optimal transport plans between $(T_i)_\# \square_\delta^d$ and optimal N_i -supported measures, we get that

$$\begin{aligned} c_p(N) &\leq \sum_i \rho_i W_p^p(\square_\delta, \Delta_{N_i}) \\ &\leq \sum_i \rho_i \delta^{d+p} W_p^p(\square_1, \Delta_{N_i}) \\ &\leq \sum_i \rho_i \delta^{d+p} \frac{(\theta(d, p) + \varepsilon)^p}{N_i^{p/d}} \end{aligned}$$

when all N_i are large enough, which happens if N itself is large enough given that $x_i \neq 0$.

For N large enough, the localization lemma ensures that no mass is moved more than $\delta\varepsilon$ by an optimal transport plan between μ and μ_N . This implies that the cost $c_p(N)$ is bounded below by the cost of transporting the mass of the cubes of side length $(1 - \varepsilon)\delta$, same center and same orientation of the $T_i(C)$ to the part of $\text{supp } \mu_N$ contained inside $T_i(C)$. We therefore get

$$c_p(N) \geq \sum_i \rho_i ((1 - \varepsilon)\delta)^{d+p} \frac{\theta(d, p)^p}{N_i^{p/d}}$$

for N large enough.

The two inequalities above give us

$$c_p(N) \sim \frac{\theta(d, p)^p}{N^{p/d}} \sum_i \rho_i x_i^{-p/d} \delta^d$$

and we are done. \square

The study of the functional F is straightforward.

Lemma 5.3. — *Fix a positive vector $\rho = (\rho_1, \rho_2, \dots, \rho_I)$ and consider the simplex $X_\delta = \{(x_i)_i \mid \sum_i x_i = \delta^{-d}, x_i > 0\}$. The function $F(\rho; \cdot)$ has a unique minimizer $x^0 = (x_i^0)$ in X_δ which is proportionnal to $(\rho_i^{\frac{d}{d+p}})_i$,*

with

$$F(\rho; x^0) = \left(\sum_i \rho_i^{\frac{d}{d+p}} \delta^d \right)^{\frac{p+d}{d}} = |\rho|_{\frac{d}{d+p}}.$$

Proof. — First $F(\rho; \cdot)$ goes to ∞ on the boundary of X_δ , so that it must have a minimizer. Any minimizer must be a critical point and therefore satisfy

$$\sum_i \rho_i x_i^{-p/d-1} \eta_i \delta^d = 0$$

for all vector (η_i) such that $\sum_i \eta_i = 0$. This holds only when $\rho_i x_i^{-p/d-1}$ is a constant and we get the uniqueness of x^0 and its expression:

$$x_i^0 = \frac{\rho_i^{\frac{d}{d+p}}}{\sum_j \rho_j^{\frac{d}{d+p}} \delta^d}.$$

The value of $F(\rho; x^0)$ follows. □

As a consequence, we see that with the notations of Lemma 5.2 the density vector $(\delta^{-d} N_i / N)$ must converge to x^0 . Indeed, this is the only accumulation point that asymptotically minimizes the W_p distance.

Proof of Theorem 1.2. — For simplicity, we denote $\beta = d/(d+p)$.

Let C be a cube containing the support of μ . Choose some $\varepsilon > 0$. Let $\tilde{\mu} = \tilde{\rho} \lambda$ be a measure such that $\tilde{\rho}$ is constant on each cube of a regular subdivision of C , is zero outside C , satisfies $|\rho - \tilde{\rho}|_1 \leq \varepsilon^{1+p/d}$ and such that $|\rho - \tilde{\rho}|_\beta \leq \varepsilon |\rho|_\beta$.

The stability Lemma shows that

$$W_p^p(\mu, \Delta_N) \leq W_p^p(\tilde{\mu}, \Delta_{(1-\varepsilon)N}) + O\left(\frac{|\rho - \tilde{\rho}|_1}{(\varepsilon N)^{p/d}}\right)$$

so that, using the hypotheses on $\tilde{\rho}$ and the two previous Lemmas,

$$W_p^p(\mu, \Delta_N) \leq \frac{(\theta(d, p) + \varepsilon)^p |\rho|_\beta (1 + \varepsilon)(1 - \varepsilon)^{-p/d} + O(\varepsilon)}{N^{p/d}}$$

for N large enough.

Symmetrically, we get (again for N large enough)

$$\begin{aligned} W_p(\mu, \Delta_N)^p &\geq W_p^p(\tilde{\mu}, \Delta_{N/(1-\varepsilon)}) - O\left(\frac{|\rho - \tilde{\rho}|_1}{(\varepsilon N)^{p/d}}\right) \\ &\geq \frac{(\theta(d, p) - \varepsilon)^p |\rho|_\beta (1 - \varepsilon)^{1+p/d} - O(\varepsilon)}{N^{p/d}}. \end{aligned}$$

Letting $\varepsilon \rightarrow 0$, the claimed equivalent follows.

Let us now use it to give the asymptotic distribution of the support of any distance minimizing μ_N . Let A be any cube in C . Let x be the limit of the proportion of $\text{supp } \mu_N$ that lies inside A (x exists up to extracting a subsequence). Since the cubes generates the Borel σ -algebra, we only have to prove that $x = \int_A \rho^\beta / \int \rho^\beta$.

Fix any $\varepsilon > 0$ and let A' be the cube homothetic of A with same center and ratio $(1 - \varepsilon)$. Let B' be the complement in C of the cube homothetic of A with same center and ratio $(1 + \varepsilon)$. Then we have by the localization lemma

$$W_p^p(\mu, \mu_N) \geq W_p^p(\mu|_{A'}, \Delta_{(1+\varepsilon)xN}) + W_p^p(\mu|_{B'}, \Delta_{(1+\varepsilon)(1-x)N})$$

for N large enough. Using the first part of Theorem 1.2, we get

$$W_p^p(\mu, \mu_N) \geq \frac{\theta(d, p)^p + \varepsilon}{((1 + \varepsilon)N)^{p/d}} \left(\frac{|\rho|_{A'}|_\beta}{x^{p/d}} + \frac{|\rho|_{B'}|_\beta}{(1-x)^{p/d}} \right)$$

for N large enough.

Letting $\varepsilon \rightarrow 0$, we get that

$$\liminf W_p(\mu, \mu_N) N^{-1/d} \geq \theta(d, p) \left(\frac{|\rho|_A|_\beta}{x^{p/d}} + \frac{|\rho|_{C \setminus A}|_\beta}{(1-x)^{p/d}} \right)^{1/p}$$

A simple computation shows that the unique optimal value of x in the above is given by the expected $|\rho|_A|_\beta / |\rho|_\beta$, and gives the above constant $|\rho|_\beta^{1/p}$ in the order of approximation. If x were different than this value, μ_N could not be optimal. \square

6. The dyadic Cantor measure

In this section we study the approximation problem for the dyadic Cantor measure κ to prove Theorem 1.3.

Let S^0, S^1 be the dilations of ratio $1/3$ and fixed point $0, 1$. The map $\mathcal{S} : \mu \mapsto 1/2 S^0_\# \mu + 1/2 S^1_\# \mu$ is $1/3$ -Lipschitz on the complete metric space of probability measures having finite p -th moment endowed with

the L^p Wasserstein metric. It has therefore a unique fixed point, called the dyadic Cantor measure and denoted by κ . It can be considered as the “uniform” measure on the usual Cantor set.

By convexity of the cost function and symmetry, $c_1 := W_p(\kappa, \Delta_1)$ is realized by the Dirac measure at $1/2$. Using the contractivity of \mathcal{S} , we see at once that $W_p(\kappa, \Delta_{2^k}) \leq 3^{-k} c_1$. Denote by $s = \log 2 / \log 3$ the dimension of κ . We have $W_p(\kappa, \Delta_{2^k}) (2^k)^{1/s} \leq c_1$ for all integer k .

To study the case when the number of points is not a power of 2, and to get lower bounds in all cases, we introduce a notation to code the regions of $\text{supp } \kappa$. Let $I^0 = [0, 1]$ and given a word $w = \epsilon_n \dots \epsilon_1$ where $\epsilon_i \in \{0, 1\}$, define $I_w^n = S_{\epsilon_n} S_{\epsilon_{n-1}} \dots S_{\epsilon_1} [0, 1]$. The *soul* of such an interval is the open interval of one-third length with the same center. The *sons* of I_w^n are the two intervals $I_{\epsilon w}^{n+1}$ where $\epsilon \in \{0, 1\}$, and an interval is the *father* of its sons. The two sons of an interval are *brothers*. Finally, we say that n is the *generation* of the interval I_w^n .

Let N be an integer, and $\mu_N \in \Delta_N$ be a measure closest to κ , whose support is denoted by $\{x_1, \dots, x_N\}$. An interval I_w^n is said to be *terminal* if there is an x_i in its soul. A point in I_w^n is always closer to the center of I_w^n than to the center of its father. This and the optimality of μ_N implies that a terminal interval contains only one x_i , at its center.

Since the restriction of κ to I_w^n is a copy of κ with mass 2^{-n} and size 3^{-n} , it follows that

$$W_p(\kappa, \mu_N)^p = W_p^p \sum_{I_w^n} 2^{-n} 3^{-np}$$

where the sum is on terminal intervals. A simple convexity arguments shows that the terminal intervals are of at most two (successive) generations.

Consider the numbers $N_k = 3 \cdot 2^k$. The terminal intervals of an optimal μ_{N_k} must be of generation $k+1$ (for 2^k of them) and $k+2$ (for 2^{k+1} of them). Therefore

$$W_p(\kappa, \mu_{N_k})^p = c_1^p (3^{-(k+1)p} + 3^{-(k+2)p}) / 2$$

and finally

$$W_p(\kappa, \Delta_{N_k}) N_k^{1/s} = c_1 \left(\frac{1 + 3^{-p}}{2} \right)^{1/p} 3^{\frac{\log 3}{\log 2} - 1}.$$

Note that the precise repartition of the support does not have any importance (see figure 2).

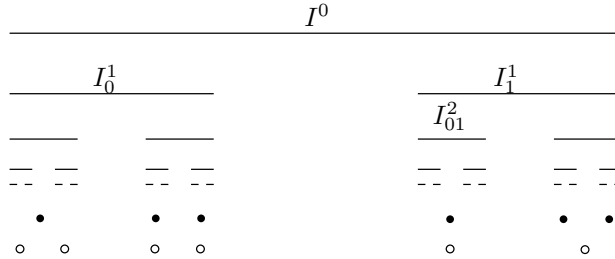


FIGURE 2. The four first steps of the construction of the Cantor set; the Cantor measure is equally divided between the intervals of a given step. The bullets show the supports of two optimal approximation of κ by 6-supported measures. We see that there is no need for the support to be equally distributed between the intervals of the first generation.

To see that $W_p(\kappa, \Delta_N)N^{1/s}$ has no limit, it is now sufficient to estimate the factor of c_1 in the right-hand side of the above formula. First we remark that $\left(\frac{1+3^{-p}}{2}\right)^{1/p}$ is greater than $1 - (1 - 3^{-p})/(2p)$ which is increasing in p and takes for $p = 1$ the value $2/3$. Finally, we compute $2/3 \cdot 3^{\frac{\log 3}{\log 2} - 1} \simeq 1.27 > 1$.

Note that the fundamental property of κ we used is that the points in a given I_w^n are closest to its center than to that of its father. The same method can therefore be used to study the approximation of sparser Cantor measure, or to some higher-dimensionnal analogue like the one generated by four contractions of ratio $1/4$ on the plane, centered at the four vertices of a square.

Moreover, one could study into more detail the variations in the approximations $W_p(\kappa, \Delta_N)$. As said before, here our point was only to show the limitation to Theorem 1.2.

7. Link with Centroidal Voronoi Tessellations

Here we explain the link between our optimization problem and the centroidal Voronoi tessellations (CVTs in short). For a complete account on CVTs, the reader can consult [5] from where all definitions below are taken.

7.1. A quick presentation. — Consider a compact convex domain Ω in \mathbb{R}^d and a density (positive, L^1) function ρ on Ω .

Given a N -tuple $X = (x_1, \dots, x_N)$ of so-called *generating points*, one defines the associated *Voronoi Tessellation* as the collection of convex sets

$$V_i = \{x \in \Omega \mid |x - x_i| \leq |x - x_j| \text{ for all } j \in \llbracket 1, N \rrbracket\}$$

and we denote it by $V(X)$. One says that V_i is the *Voronoi cell* of x_i . It is a tiling of Ω , in particular the cells cover Ω and have disjoint interiors.

Each V_i has a center of mass, equivalently defined as

$$g_i = \frac{\int_{V_i} x \rho(x) dx}{\int_{V_i} \rho(x) dx}$$

or as the minimizer of the energy functional

$$\mathcal{E}_{V_i}(g) = \int_{V_i} |x - g|^2 \rho(x) dx.$$

One says that $(V_i)_i$ is a *centroidal Voronoi tessellation* or *CVT*, if for all i , $g_i = x_i$. The existence of CVTs comes easily by considering the following optimization problem: search for a N -tuple of points $X = (x_1, \dots, x_N)$ and a tiling V of Ω by N sets V_1, \dots, V_N which together minimize

$$\mathcal{E}_V(X) = \sum_{i=1}^N \mathcal{E}_{V_i}(x_i)$$

A compactness argument shows that such a minimizer exists, so let us explain why a minimizer must be a CVT together with its generating set. First, each x_i must be the center of mass of V_i , otherwise one could reduce the total energy by moving x_i to g_i and changing nothing else. But also, V_i should be the Voronoi cell of x_i , otherwise there is a $j \neq i$ and a set of positive measure in V_i whose points are closest to x_j than to x_i . Transferring this set from V_i to V_j would reduce the total cost.

We observe that this optimization problem is exactly that of approximating the measure $\rho\lambda$ in L^2 Wasserstein distance; more precisely, finding the N -tuple x that minimizes $\inf_V \mathcal{E}_V(X)$ is equivalent to finding the support of an optimal $\mu_N \in \Delta_N$ closest to $\rho\lambda$, and then the Voronoi tessellation generated by X gives the mass of μ_N at each x_i and the optimal transport from $\rho\lambda$ to μ_N .

One says that a CVT is *optimal* when its generating set is a global minimizer of the energy functional

$$\mathcal{E}(X) = \mathcal{E}_{V(X)}(X).$$

Optimal CVTs are most important in applications, which include for example mesh generation and image analysis (see [5]).

7.2. Equidistribution of Energy. — The *principle of energy equidistribution* says that if X generates an optimal CVT, the energies $\mathcal{E}_{V_i}^p(x_i)$ of the generating points should be asymptotically independent of i when N goes to ∞ .

Our goal here is to deduce a mesoscopic version of this principle from Theorem 1.2. A similar result holds for any exponent, so that we introduce the L^p energy functionals $\mathcal{E}_{V_i}^p(x_i) = \int_{V_i} |x - x_i|^p \rho(x) dx$, $\mathcal{E}_V^p(X) = \sum_i \mathcal{E}_{V_i}^p(x_i)$ and $\mathcal{E}^p(X) = \inf_V \mathcal{E}_V^p(x) = \mathcal{E}_{V(X)}^p(X)$. In particular, an optimal X for this last functional is the support of an element of Δ_N minimizing the L^p Wasserstein distance to $\rho\lambda$.

Note that for $p \neq 2$ an x minimizing $\mathcal{E}^p(x)$ need not generate a CVT, since the minimizer of $\mathcal{E}_{V_i}^p$ is not always the center of mass of V_i (but it is unique as soon as $p > 1$).

Corollary 7.1. — *Let A be a cube of Ω . Let $X^N = \{x_1^N, \dots, x_N^N\}$ be a sequence of N -sets minimizing \mathcal{E}^p for the density ρ , and denote by $\bar{\mathcal{E}}_A^p(N)$ the average energy of the points of X^N that lie in A . Then $\bar{\mathcal{E}}_A^p(N)N^{(d+p)/d}$ has a limit when $N \rightarrow \infty$, and this limit does not depend on A .*

The cube A could be replaced by any set that is sufficiently regular, but not by any open set. Since the union of the X_N is countable, there are indeed open sets of arbitrarily small measure containing all the points $(x_i^N)_{N,i}$.

Proof. — Fix some $\varepsilon > 0$ and let $A' \subset A$ be the set of points that are at distance at least ε from $\Omega \setminus A$ and by $A'' \supset A$ the set of points at distance at most ε from A .

First, the numbers N', N'' of points of X^N in A', A'' satisfy

$$N' \sim N \frac{\int_{A'} \rho^{d/(d+p)}}{\int_{\Omega} \rho^{d/(d+p)}} \quad N'' \sim N \frac{\int_{A''} \rho^{d/(d+p)}}{\int_{\Omega} \rho^{d/(d+p)}}.$$

The localization lemma implies that the maximal distance by which mass is moved by the optimal transport between $\rho\lambda$ and the optimal X^N -supported measure tends to 0, so that for N large enough the energy of all points in A is at least the minimal cost between $\rho|_{A'}\lambda$ and $\Delta_{N'}$ and at most the minimal cost between $\rho|_{A''}\lambda$ and $\Delta_{N''}$.

Letting $\varepsilon \rightarrow 0$ we thus get that the total energy of all points of X^N lying in A is equivalent to

$$\theta(d, p) \frac{\left(\int_A \rho^{d/(d+p)}\right)^{(d+p)/d}}{\left(N \int_A \rho^{d/(d+p)}\right)^{p/d}} = \theta(d, p) N^{-p/d} \int_A \rho^{d/(d+p)}$$

As a consequence we have $\bar{\mathcal{E}}_A(N) \sim \left(\theta(d, p) \int_\Omega \rho^{d/(d+p)}\right) N^{-(d+p)/d}$. \square

References

- [1] G. BOUCHITTÉ, C. JIMENEZ & M. RAJESH – “Asymptotique d’un problème de positionnement optimal”, *C. R. Math. Acad. Sci. Paris* **335** (2002), no. 10, p. 853–858.
- [2] A. BRANCOLINI, G. BUTTAZZO, F. SANTAMBROGIO & E. STEPANOV – “Long-term planning versus short-term planning in the asymptotical location problem”, *ESAIM Control Optim. Calc. Var.* **15** (2009), no. 3, p. 509–524.
- [3] T. CHAMPION, L. DE PASCALE & P. JUUTINEN – “The ∞ -Wasserstein distance: local solutions and existence of optimal transport maps”, *SIAM J. Math. Anal.* **40** (2008), no. 1, p. 1–20.
- [4] V. DOBRIĆ & J. E. YUKICH – “Asymptotics for transportation cost in high dimensions”, *J. Theoret. Probab.* **8** (1995), no. 1, p. 97–118.
- [5] Q. DU, V. FABER & M. GUNZBURGER – “Centroidal Voronoi tessellations: applications and algorithms”, *SIAM Rev.* **41** (1999), no. 4, p. 637–676 (electronic).
- [6] Q. DU & D. WANG – “The optimal centroidal Voronoi tessellations and the Gershgorin’s conjecture in the three-dimensional space”, *Comput. Math. Appl.* **49** (2005), no. 9-10, p. 1355–1373.
- [7] K. J. FALCONER – *The geometry of fractal sets*, Cambridge Tracts in Mathematics, vol. 85, Cambridge University Press, Cambridge, 1986.
- [8] L. FEJES TÓTH – “Sur la représentation d’une population infinie par un nombre fini d’éléments”, *Acta. Math. Acad. Sci. Hungar* **10** (1959), p. 299–304 (unbound insert).
- [9] J. HEINONEN – *Lectures on analysis on metric spaces*, Universitext, Springer-Verlag, New York, 2001.
- [10] J. HOROWITZ & R. L. KARANDIKAR – “Mean rates of convergence of empirical measures in the Wasserstein metric”, *J. Comput. Appl. Math.* **55** (1994), no. 3, p. 261–273.
- [11] J. E. HUTCHINSON – “Fractals and self-similarity”, *Indiana Univ. Math. J.* **30** (1981), no. 5, p. 713–747.

- [12] Y. LIPMAN – “Approximation of optimal transportation cost”, preprint.
- [13] F. MORGAN & R. BOLTON – “Hexagonal economic regions solve the location problem”, *Amer. Math. Monthly* **109** (2002), no. 2, p. 165–172.
- [14] S. J. N. MOSCONI & P. TILLI – “ Γ -convergence for the irrigation problem”, *J. Convex Anal.* **12** (2005), no. 1, p. 145–158.
- [15] D. J. NEWMAN – “The hexagon theorem”, *IEEE Trans. Inform. Theory* **28** (1982), no. 2, p. 137–139.
- [16] C. VILLANI – *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.