Optimal transport between random measures

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch–Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von

Martin Otto Josef Huesmann

aus

Coesfeld

Bonn, Februar 2012

Angefertigt mit der Genehmigung der Mathematisch–Naturwissenschaftlichen–Fakultät der Rheinischen Friedrich–Wilhelms–Universität Bonn

1. Gutachter: Prof. Dr. Karl-Theodor Sturm
2. Gutachter: Prof. Dr. Patrik Ferrari

Tag der Promotion: 11. Juni 2012
Erscheinungsjahr: 2012

Abstract

We study couplings $q^\bullet$ of two equivariant random measures $\lambda^\bullet$ and $\mu^\bullet$ on some Riemannian manifold (M,d,m), i.e. measure valued random variables $\omega \mapsto q^\omega$ such that for almost every ω the measure q^ω on $M \times M$ is a coupling between the measures λ^ω and μ^ω. We assume that M admits a group G of isometries acting properly discontinuously, cocompactly and freely. We ask for a minimizer of the *mean transportation cost* defined by

$$\mathfrak{C}(q^\bullet) \;:=\; \sup_{B \in \mathrm{Adm}(M)} \frac{1}{m(B)} \mathbb{E}\left[\int_{M\times B} \vartheta(d(x,y))\; q^\bullet(dx,dy)\right],$$

for some continuous increasing function $\vartheta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{r\to\infty} \vartheta(r) = \infty$ and a collection of sets $\mathrm{Adm}(M)$ that can be written as finite unions of fundamental regions.

If the mean transportation cost are finite and $\lambda^\omega \ll m$ for a.e. ω, there is a unique equivariant minimizer. This minimizer is called *optimal coupling*. Moreover, it is induced by a transportation map, $q^\bullet = (id, T)_*\lambda^\bullet$. If the group G satisfies some strong form of amenability, we can approximate the optimal coupling by solutions to a transportation problem on bounded regions. In particular, in the case of $M = \mathbb{R}^d$, $\lambda^\bullet \equiv \mathcal{L}$ the Lebesgue measure and $\mu^\bullet$ a simple point process, the optimal coupling induces a fair factor allocation. If we consider the cost function $\vartheta(r) = r^2$, the optimal coupling constitutes a Laguerre tessellation, a random tiling of $\mathbb{R}^d$ by convex polytopes of volume one.

If we transport the Lebesgue measure into a Poisson point process we have rather sharp estimates on the mean transportation cost. Considering cost functions $\vartheta(r) = r^p$ the optimal mean transportation cost are finite iff $p < d/2$ in dimensions $d = 1$ or $d = 2$ and for all p in dimensions $d \geq 3$.

Furthermore, we get similar results in the more general case of *optimal semicouplings*. A semicoupling $q^\bullet$ between two equivariant random measures $\lambda^\bullet$ and $\mu^\bullet$ of intensity one and $\beta \in (0,1]$ is a measure valued random variable $\omega \mapsto q^\omega$ such that q^ω is a coupling between $\rho\lambda^\omega$ and $\mu^\bullet$ for some density $\rho : \Omega \times M \to \mathbb{R}$. Analogously, we can define semicouplings with $\beta \geq 1$ as couplings between $\lambda^\bullet$ and $\rho\mu^\bullet$.

To show how these ideas can be extended to a more general setting we consider the regular k-tree. We show that there is an unique optimal coupling, too. Moreover, we construct an equivariant coupling with finite mean transportation cost.

Finally, we study stability properties of optimal couplings under vague convergence on $M \times M \times \Omega$.

Acknowledgements

First of all, I would like to thank Prof Dr Karl - Theodor Sturm for supervising this thesis. This work and especially myself benefited a lot from numerous discussions and his constant support and encouragement. I have had a wonderful time working with him and his group.

I would like to thank Prof Patrik Ferrari, Prof Wolfgang Lück and Prof Rolf Klein for being part of the commission.

Thanks are also due to my colleagues in the Stochastic Analysis and Probability group for creating such a friendly and open atmosphere. I would like to mention especially my office mates Kazumasa Kuwada and Matthias Erbar for so many short and long discussions and for sharing so much tea. I also thank Matthias for his help with the nice pictures.

I owe a big thanks to my family and friends, especially to my wife Katharina, for many things.

Finally, I would like to thank BIGS and the SFB 611 for its financial and non-financial support.

Contents

Chapter 1

Introduction

Nowadays, allocation and matching problems appear in various fields of everyday life, e.g. assigning students to schools or universities, young doctors to hospitals or resources to companies. We think of allocations and matchings as transportation problems where a distribution of a certain good is transported to a distribution of agents. In this thesis, we study transportation problems between random measures under some minimization condition.

For an illustration of a typical transportation problem consider the following example:
The government has to divide a certain area of land between farmers. Each farmer is to receive a prespecified share of the land under the constraint that the typical squared distance between a farmer's house and a randomly chosen point of his land is minimized. A reason for this constraint might be a minimization of fuel expenses or carbon dioxide emission. Moreover, to be able to work efficiently, the land of each farmer should be connected and its shape should not be too irregular to make the use of their machines easier.

To put this into mathematical terms consider two distributions λ and μ, representing land and farmers respectively. Finding an allocation means, we are looking for a *coupling* between λ and μ. We interpret this coupling as a rule to partition λ and transport the different pieces to their respective targets. Transporting a piece of mass from x to y produce cost of an amount $c(x, y)$. Hence, we search for minimizers of the transportation cost

$$\int c(x,y)\ q(dx,dy)$$

among all couplings q of λ and μ. The case of finite measures λ and μ is extensively covered by the theory of optimal transportation. In this work, we are interested in the case that λ and μ have infinite mass. This will typically result in infinite transportation cost. A more reasonable quantity to consider is therefore the *mean transportation cost,* the transportation cost per unit volume. Moreover, we will work with *random measures* $\lambda^\bullet$ and $\mu^\bullet$ on some Riemannian manifold M that are equivariant, that is they satisfy some invariance properties under a flow of the probability space. In particular, this implies the invariance of their joint distribution under the diagonal action of some group of isometries of M. We will show the following

- If the mean transportation cost is finite, there always is at least one equivariant minimizer of the mean transportation cost.
- If $\lambda^\omega \ll m$, the Riemannian volume measure, for $\mathbb{P}-$a.a. ω there is at most one equivariant minimizer of the mean transportation cost. This minimizer is induced by a transportation map, i.e. $q^\bullet = (id, T^\bullet)_* \lambda^\bullet$.
- The unique equivariant minimizer can be approximated by solutions to the minimization problem between $\lambda^\bullet$ and $\mu^\bullet$ restricted to bounded sets.

The/an equivariant minimizer will be called *optimal coupling* of $\lambda^\bullet$ and $\mu^\bullet$.

Before we describe our results in more detail, we would like to put them into the context of the existing mathematical literature. First, we give a short expository on allocations and matchings and their connection to the theory of point processes. Then, we recall the definition of Voronoi and some related tessellations. Thirdly, we briefly explain the basic ideas of optimal transportation between probability measures. Finally, we come back to the transportation problem between equivariant random measures and explain our results in detail. Lastly, we will give a short overview of the different chapters.

1.1 Survey of relevant mathematical literature

Allocations and matchings

Allocations and matchings are a very wide field of current research. In allocation problems one is asked to divide some continuous quantity between a set of "agents", e.g. land to farmers or resources to companies. In matching problems one has to match discrete sets with discrete sets, e.g. children to schools or kidneys to patients. A very important example of an early work on matchings is the nice article [GS62] by Gale and Shapley on the stability of marriage. They consider two sets of equal size, men and women. Every man ranks the women according to his preferences and every woman ranks the men according to her preferences. Gale and Shapley ask whether it is possible to match these sets in such a way that there are no two pairs, say (Anna and Albert) and (Berta and Balduin), such that Anna prefers Balduin over Albert and Balduin prefers Anna over Berta. The existence of such a pair would result in an unstable marriage. They found two algorithms solving this problem, the first one is the best possible for the men and the second one the best possible for the women.
This triggered an enormous amount of research in different directions. In economics, there is an extensive literature on related algorithms, e.g. for the kidney problem [RSUU05]. In Boston, there is even an algorithm how children are matched with their schools of choice (or not) [APRS05]. For an extensive bibliography we refer to [Rot], a web page maintained by Alvin Roth.

On the other hand Alexander Holroyd and Yuval Peres [HHP06] used a generalization of the stable marriage algorithm to construct an allocation between Lebesgue measure and a Poisson point process $\mu^\bullet$ in which for almost every realization μ^ω of the process every Poisson point gets an equal amount of mass, namely a set of Lebesgue measure one (we will be more precise later). The construction of this allocation was a key step in establishing a very deep link between the theory of point processes on the one side and allocations on the other side. To explain this link, we first need to introduce the concept of Palm measures.

Let $\mu^\bullet$ denote a random measure on $\mathbb{R}^d$ defined on a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ (see section 2.3 for the definition of random measures). Its first moment measure M_1 is defined by $M_1(A) = \mathbb{E}[\mu^\bullet(A)]$ for Borel sets A. A finer picture of the behavior of $\mu^\bullet$ provides the *Campbell measure.* It is a measure on $\mathbb{R}^d \times \Omega$ given by

$$C_\mu^\bullet(A \times B) \;=\; \int_B \mu^\omega(A)\mathbb{P}(d\omega),$$

for measurable sets $A \subset \mathbb{R}^d, B \in \mathfrak{A}$. Clearly, $C_\mu^\bullet \ll M_1$. Therefore, we can conclude by the Radon-Nikodym Theorem the existence of measurable functions $\mathbb{P}_x(B)$ such that

$$C_\mu^\bullet(A \times B) \;=\; \int_A \mathbb{P}_x(B) M_1(dx).$$

By the theory of regular conditional probability measures, we can choose the family $\{\mathbb{P}_x(B)\}$ such that

i) for each fixed $B \in \mathfrak{A}$ $\mathbb{P}_x(B)$ is a measurable function of x which is M_1 integrable on bounded subsets of $\mathbb{R}^d$.

ii) for each fixed $x \in \mathbb{R}^d$ $\mathbb{P}_x(\cdot)$ is a probability measure on $B \in \mathfrak{A}$.

Each such measure is called a *local Palm measure* of $\mu^\bullet$ and the whole family $\{\mathbb{P}_x\}$ is called *Palm kernel* associated with $\mu^\bullet$. If $\mu^\bullet$ happens to be *translation invariant*, i.e. $\mu^\bullet(U) \overset{d}{=} \mu^\bullet(U+y)$ for any $U \in \mathcal{B}(\mathbb{R}^d), y \in \mathbb{R}^d$, the different Palm measures become translated versions of each others

$$\mathbb{P}_x(U+x) = \mathbb{P}_0(U) =: \mathbb{P}'(U) \quad \mathcal{L} \text{ a.s..}$$

In other words, we factor out the translation invariance and look at the random measure from the point x. If $\mu^\bullet$ has unit intensity, $\mathbb{P}'$ can be chosen to be a probability measure on $(\Omega, \mathfrak{A})$. In the case of a simple point process the Palm measure $\mathbb{P}_x$ can be viewed as $\mu^\bullet$ conditioned to have a point at x. Therefore, Palm measures are an extremely useful tool if one is interested in the "typical behavior" of a function of the process. It is used in numerous works in stochastic geometry, e.g. if one studies the behavior of the typical cell of a Voronoi tessellation. For more details and applications as well as references on Palm measures we refer to chapter 13 of [DVJ07].

Given a translation invariant simple point process $\mu^\bullet$ of unit intensity on $\mathbb{R}^d$, that is $\mu^\omega(x) \in \{0,1\}$ and $M_1 = \mathcal{L}$, the Lebesgue measure, put $\Xi(\omega) = \mathrm{supp}(\mu^\omega)$. A *fair allocation* for $\mu^\bullet$ is a measurable map $\Psi_\bullet : \mathbb{R}^d \to \mathbb{R}^d$ such that for $\mathbb{P}$ almost every ω

i) $\mathcal{L}(\mathbb{R}^d \setminus \Psi_\omega^{-1}(\Xi(\omega))) \;=\; 0.$

ii) For all $\xi \in \Xi(\omega)$ we have $\mathcal{L}(\Psi_\omega^{-1}(\xi)) \;=\; 1.$

If the second property does not hold, we simply speak of allocations for $\mu^\bullet$. Each $\xi \in \Xi(\omega)$ is called *center* and $\Psi_\omega^{-1}(\xi)$ is called the *cell* with center ξ. Translations on $\mathbb{R}^d$ induce an action on Ω by translating the support of $\mu^\bullet$. For $z \in \mathbb{R}^d$ set $\mu^{\theta_z\omega}(A) = \mu^\omega(A-z)$. If $\Psi_\bullet$ satisfies

$$\Psi_\omega(x) = y \;\Rightarrow\; \Psi_{\theta_z\omega}(x+z) = y+z \;\forall z \in \mathbb{R}^d$$

we say that $\Psi_\bullet$ is an invariant allocation. An *extra head rule* for $\mu^\bullet$ is a shift coupling of $\mu^\bullet$ and its Palm version. More precisely, it is a $\mathbb{R}^d$ valued random variable X such that $\tau_{-X}\mu^\bullet$ has law $\mathbb{P}'$, where $\tau_z\mu^\omega(A) = \mu^\omega(A+z)$. Holroyd and Peres proved the following remarkable result

Theorem 1.1 (Theorem 13 in [HP05])**.** *Let Ψ be an invariant allocation for $\mu^\bullet$, an ergodic translation invariant simple point process of unit intensity. Then, the random variable $Y = \Psi(0)$ is an extra head scheme if and only if Ψ is fair.*

Moreover, a converse of this statement holds as well. Given an extra head rule one can cook up a fair allocation for $\mu^\bullet$. This is technically more difficult and we refer for the full statement to Theorem 16 of [HP05].

A very natural question is to ask how good can this shift coupling or the (fair) allocation be? What means good? Thinking about the land partition problem of the farmers again, a couple of natural measures of goodness arise:

i) Are the cells connected? What is the tail behavior of $|Y| = |\Psi(0)|$, i.e. how fast does $\mathbb{P}[|Y| \geq R]$ decay?

ii) Is the allocation a factor of $\mu^\bullet$? Is there a rule how to allocate the land to the farmers if we just know the position of their houses?

iii) Does the allocation have the same invariance properties as $\mu^\bullet$?

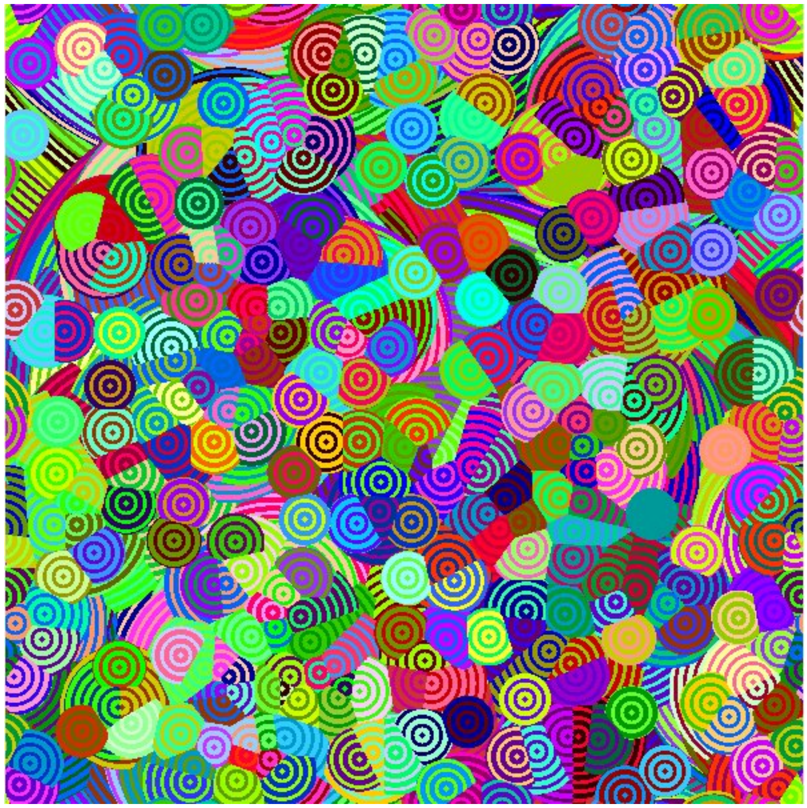

Figure 1.1: A realization of the stable marriage between the Lebesgue measure and a Poisson point process on the torus. Simulation and picture by Alexander Holroyd.

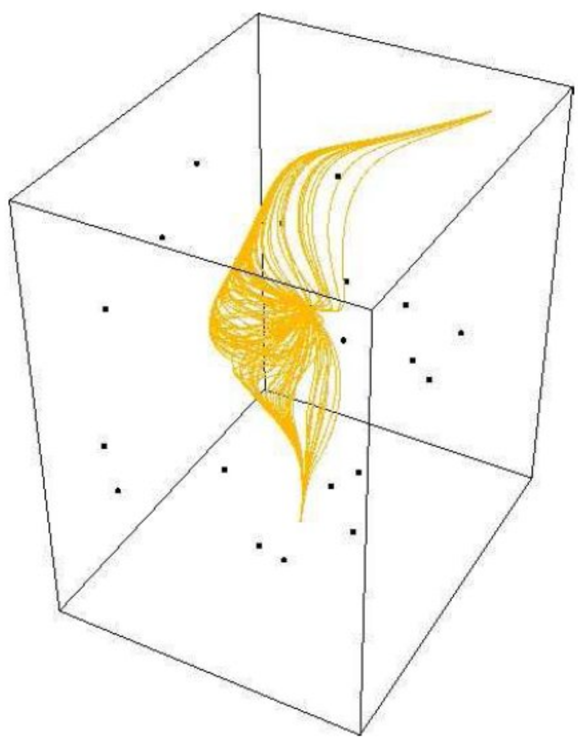

Figure 1.2: A cell of the gravitational allocation. Picture from [CPPRa]

In [HP05, HHP06], Holroyd and Peres respectively Hoffman, Holroyd and Peres constructed a fair allocation of the Lebesgue measure $\mathcal{L}$ to an ergodic simple point process using a variant of the stable marriage algorithm. This algorithm has the nice feature of being "stable"(in the sense of Gale and Shapley) and being a factor of the point process. However, it has the disadvantage of producing disconnected cells. Moreover, in the case of $\mu^\bullet$ being a Poisson point process, Y does not have good integrability properties. Indeed, it holds that $\mathbb{E}[|Y|^d] = \infty$ for dimension $d \geq 3$ and $\mathbb{E}[|Y|^{d/2}] = \infty$ for $d = 1, 2$, see Theorem 7 in [HHP06]. Also in [HP05], Holroyd and Peres showed the existence of a *randomized* allocation of Lebesgue to Poisson in dimension $d \geq 3$ satisfying $\mathbb{E}[\exp(c|Y|^d)] < \infty$ for some positive constant c. The construction is based on an early transportation cost estimate by Talagrand [Tal94]. They naturally asked whether it is possible to also construct factor allocations with this integrability properties and / or connected cells.

In [CPPRa], Chaterjee, Peled, Peres and Romik constructed a wonderful allocation for a Poisson point process in dimensions $d \geq 3$, the gravitational allocation. The idea is the following. Consider the gravitational force field exerted on a point x by the points of the Poisson process

$$F^\omega(x) = \sum_{\xi \in \mathrm{supp}(\mu^\omega), |x-\xi|\uparrow} \frac{\xi - x}{|\xi - x|}.$$

Let Y be the integral curve of F, that is Y solves $\dot{Y}(t) = F(Y(t))$. For every $z \in \mathbb{R}^d$ let Y_z be the integral curve starting at $Y_z(0) = z$. This curve exists up to some maximal time S_z. For $\xi \in \mathrm{supp}(\mu^\omega)$ we define its *basin of attraction* $B(\xi) = \{z \ : \ \lim_{t \to S_z} Y_z(t) = \xi\}$. This allows to define the allocation rule by

$$\Psi_\omega^{Grav}(z) = \begin{cases} \xi & z \in B(\xi) \\ \eth & z \notin \bigcup_{\xi \in \mathrm{supp}(\mu^\omega)} B(\xi) \end{cases}$$

for some cemetery state $\eth$. The astonishing result is, that this is indeed a *fair allocation*. Almost surely every cell has volume one. Moreover, by construction all the cells are connected and contain their centers. In [CPPRb], the same authors show that $\mathbb{P}[|Y_0| \geq R] \leq \exp(-cR^{g_d})$ with $g_3 = 1$ and $g_d = 1 + 1/(d-1)$ for $d \geq 4$. They get similar estimates for the diameter of the cells. However, the tail behavior is still not as good as in the randomized allocation. In [HS10], Huesmann and Sturm managed to close this gap and constructed a non-randomized factor allocation with $\mathbb{P}[|Y| \geq R] \leq \exp(-cR^d)$ in dimensions $d \geq 3$ (see Theorems 1.5 and 1.7). In a very recent work Markó and Timar [MT11] constructed an allocation with optimal tail behavior of the diameter of the cells in dimensions $d \geq 3$. We will explain this construction in section 7.4. Other interesting examples of allocations are [Kri07, NSV07].

A very related concept is the question of matching two invariant simple point processes with unit intensity $\mu_1^\bullet$ and $\mu_2^\bullet$. The goal is to find an invariant bijective map $\Phi_\omega : \mathrm{supp}(\mu_1^\omega) \to \mathrm{supp}(\mu_2^\omega)$. Again, one can ask the question of best possible integrability or geometric properties of a given matching. Most of these questions are answered in [Hol09, HPPS09]. However, there are still a couple of tantalizing open questions (see [Hol09]).

Tessellations

Every allocation for a simple point process induces a partition of the euclidean space. Depending on the allocation the cells that are associated to a given point might be connected, contractible or even very irregular. In stochastic geometry there are many people who study special partitions, called tessellations. A tessellation of $\mathbb{R}^d$ is a partition into convex sets which is generated by a collection of hyperplanes. Each convex set, each cell, is defined by an intersection of halfspaces. Tessellations are used to model for example polycrystalline materials, foams or biological tissues. The probably most famous tessellations is the Voronoi tessellation. Given a discrete set of points

$\mathcal{P} = \{p_i : i \in I\}$ for some at most countable index set I, we associate to every $p_i \in \mathcal{P}$ a cell C_i by

$$C_i = \{x \in \mathbb{R}^d \ : \ |x - p_i| \leq |x - p_j| \ \forall j \in I\}.$$

This defines a partition of $\mathbb{R}^d$ into convex sets, the Voronoi tessellation. There are a couple of straightforward generalizations. Assume that each point is assigned a weight w_i. We can then define the cell associated to p_i by

$$C_i = \{x \in \mathbb{R}^d \ : \ |x - p_i|^q + w_i \leq |x - p_j|^q + w_j \ \forall j \in I\},$$

for any positive q. Two cases are particularly nice. If we take $q = 2$ all the cells will be convex polytopes, in the case $q = 1$ the cells will be starlike wrt to their center. The former partition is called *Laguerre tessellation* the latter *Johnson Mehl diagram.* In fact, we will come back to this in section 4.1. For more details on Voronoi tessellations and their generalizations we refer to [Aur91], for a detailed study and applications of Laguerre tessellations we refer to [LZ08] and references therein.

Optimal transportation

By now we saw different methods of partitioning space and allocating it to targets. However, so far we have not explicitly worked with the minimization constraint. This leads us to the theory of optimal transportation.

As early as in 1781 Gaspard Monge considered in [Mon81] a slightly different transportation problem than transporting children to schools. He wanted to transport a certain amount of rubble to different locations in order to build up a fortification. He was interested in minimizing the total transportation distance. To put it into mathematical terms, given two probability measures λ and μ Monge wanted to minimize

$$\int_{\mathbb{R}^3} |x - T(x)| \ \lambda(dx)$$

over all maps $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that the image measure of λ under the map T equals μ, that is $T_*\lambda = \mu$. This problem turns out to be rather difficult to solve. For example if λ happens to be a single Dirac measure and μ is absolutely continuous there is no map with the desired property. Hence, for a general solution we need to impose additional assumptions on λ and μ.

In [Kan06], Kantorovich studied a relaxation of this problem. He was not looking for minimal maps but for minimal *couplings* of λ and μ, i.e. measures on the product space, which have marginals λ and μ. This has the huge advantage that there is always at least one candidate, the product measure $\lambda \otimes \mu$. If we consider two probability measures λ and μ on a Polish space X together with a lower semicontinuous cost function $c : X \times X \to \mathbb{R}$ the *Monge-Kantorovich problem* is to minimize the transportation cost

$$C(q) \ = \ \int c(x, y) \ q(dx, dy)$$

over all couplings q of λ and μ. As the set of all these couplings $\Pi(\lambda, \mu)$ is compact, there is always a minimizer of this quantity (if the optimal(=infimal) transportation cost is infinite we can take any coupling). Minimizers are called *optimal couplings* and we denote the minimal transportation cost by $C(\lambda, \mu)$.

Kantorovich's result already produced lots of research and there are many applications of his result in mathematics and economics. However, it took another 50 years until the starting question of Monge could be answered. Independently of each other, Brenier [Bre91] and Rüschendorf [RR90] proved the following theorem

Theorem 1.2 ([Bre91]). *Let λ and μ be two compactly supported probability measures on $\mathbb{R}^d$ such that λ is absolutely continuous wrt to the Lebesgue measure, $\lambda \ll \mathcal{L}$. Consider the Monge problem with cost function $c(x, y) = |x-y|^2/2$. Then, there exists a unique optimal transportation map T that solves the Monge-Kantorovich problem. Moreover, $T = \nabla\phi$ for some convex function $\phi : \mathbb{R}^d \to \mathbb{R}$.*

This theorem was later generalized to other cost functions by several authors. McCann managed to prove this result for compact Riemannian manifolds [McC01]. A key step in establishing these results were the dual formulation of the problem. The dual formulation of the Monge-Kantorovich problem is to look for maximizer of

$$\int \phi(x)\ \lambda(dx)\ +\ \int \psi(y)\ \mu(dy)$$

among all pairs of bounded and continuous functions (ϕ, ψ) satisfying $\phi(x) + \psi(y) \leq c(x, y)$ for all $x, y \in X$. Given one such pair, it is always possible to improve it by considering ϕ together with its c-transform

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

Hence, looking for maximizers in the dual problem we can restrict ourselves to pairs of c-concave functions, that is functions satisfying $(\phi^c)^c = \phi$. If q is a minimizer of the primal (original) problem and (ϕ, ϕ^c) is a maximizer of the dual problem, assuming that

$$\int c(x, y)\ q(dx, dy)\ =\ \int \phi(x)\ \lambda(dx) + \int \phi^c(y)\ \mu(dy)\ =\ \int (\phi(x) + \phi^c(y))\ q(dx, dy)$$

we must have

$$\phi(x) + \phi^c(y)\ =\ c(x, y) \quad q \text{ almost everywhere.} \tag{1.1}$$

Therefore, q is concentrated on the $c-$superdifferential of ϕ, $\partial^c\phi(x) = \{y \in X : \phi(x) + \phi^c(y) = c(x, y)\}$.
In the case of $X = \mathbb{R}^d$ and $c(x, y) = -x \cdot y$ the c-transform of a function ϕ becomes the Legendre transform of $(-\phi)$. ϕ is known to be convex iff its double Legendre transform ϕ^{cc} equals ϕ. Thus, considering the cost function $c(x, y) = |x - y|^2/2$ a function ϕ is c-concave iff $\phi(x) - |x|^2/2$ is concave. A concave function is locally Lipschitz and therefore differentiable almost everywhere. In this case the c-superdifferential coincides with the usual gradient, $\partial^c\phi(x) = \nabla\phi(x)$, where $\nabla\phi(x)$ is single valued for all but countably many x. Take $\lambda \ll \mathcal{L}$. The optimal coupling q and the optimal pair (ϕ, ϕ^c) of *Kantorovich potentials* satisfies (1.1). In particular, the set where $\nabla\phi(x)$ is not single valued is a λ null set. Therefore, we must have $q = (id, x + \nabla\phi)_*\lambda$. This shows, that q is given by a transportation map. But then it must also be unique. Indeed, given two optimal couplings q_1 and q_2. Then, $q_3 = \frac{1}{2}(q_1 + q_2)$ is optimal as well. Hence, all of them are concentrated on the graph of a convex function. This is only possible if all of them coincide almost everywhere.

Considering more general cost functions and measures, the dual formulation still gives us an important insight. Any optimal coupling must be concentrated on a cyclically monotone set. A subset $A \subset X \times X$ is called cyclically monotone if for all $N \in \mathbb{N}$ and $(x_1, y_1), \ldots, (x_N, y_N) \in A$ we have

$$\sum_{i=1}^N c(x_i, y_i)\ \leq\ \sum_{i=1}^N c(x_i, y_{i+1}),$$

with $y_{N+1} = y_1$. Moreover, it can be shown that if the cost function is sufficiently nice, e.g. continuous, cyclical monotonicity is already sufficient for optimality. This allows to deduce a stability result for optimal couplings.

Proposition 1.3. *Let $(\lambda_n)_{n\in\mathbb{N}}, (\mu_n)_{n\in\mathbb{N}}$ be two sequences of probability measures on some Polish space X converging to probability measures λ and μ respectively. Let $c(x,y)$ be some continuous cost function. Assume that there is a constant κ such that for all $n\in\mathbb{N}$ the optimal transportation cost satisfies $C(\lambda,\mu)\le\kappa$. Let q_n be an optimal coupling between λ_n and μ_n. Then, there is a subsequence $(q_{n_k})_{k\in\mathbb{N}}$ converging to an optimal coupling q of λ and μ.*

For more details we refer to [RR98, Vil03]. In the last decade the theory of optimal transportation had a major revival. Especially the uniqueness and representation result by Brenier attracted a lot of interest and has produced an enormous amount of deep results, striking applications and stimulating new developments, among others in PDEs (e.g. [Ott01], [AGS08]), evolution semigroups (e.g. [OV00], [ASZ09], [OS09]) and geometry (e.g. [Stu06a, Stu06b], [LV09], [Vil09], [Oht09]).

Several authors studied allocation and matching problems of independently distributed points in the unit cube in terms of transportation cost. The algorithm by Ajtai, Komlós and Tusnády [AKT84] invented for the two dimensional problem was recently used in [MT11] to produce an allocation of optimal tail behavior as remarked earlier. The very remarkable cost estimates by Talagrand [Tal94] in dimensions $d\ge 3$ allowed Holroyd and Peres to produce the randomized allocation for Poisson points mentioned above. Decreusefond studied in [Dec08] the usual Wasserstein distance between point processes. His results mostly apply to the case of finite intensity measure, that is $\mathbb{E}[\mu^\bullet(\mathbb{R}^d)]<\infty$. In the case of point processes with infinite intensity measure he only considers transportation problems between $\mu^\bullet$ and an L^2 perturbed version of $\mu^\bullet$ which has by construction finite transportation distance. Last and Thorisson [LT08] studied invariant transport kernels between random measures on locally compact Abelian groups without additional constraints.

1.2 Main results

In this thesis, we extend the theory of optimal transportation to the case of measures with infinite mass, namely equivariant random measures $(\lambda^\bullet,\mu^\bullet)$. Let (M,d.m) be a connected smooth non-compact Riemannian manifold with Riemannian distance d, and Riemannian volume m. Assume that there is a group G of isometries of M acting properly discontinuous, cocompactly and freely on M. A random measure $\lambda^\bullet$ on M is a measure valued random variable modeled on some probability space $(\Omega,\mathfrak{A},\mathbb{P})$. We assume that the probability space admits a measurable flow $(\theta_g)_{g\in G}$ which we interpret as “shifts” of the support of λ^ω. A random measure $\lambda^\bullet$ is called equivariant if

$$\lambda^{\theta_g\omega}(g\ \cdot)=\lambda^\omega(\cdot)\quad\text{for all }\omega\in\Omega.$$

We will assume that $\mathbb{P}$ is stationary, that is $\mathbb{P}$ is invariant under the flow θ. In particular, this implies that $\lambda^\bullet(B)\overset{d}{=}\lambda^\bullet(gB)$ for any $g\in G$ and Borel set B.

Given two equivariant random measures $(\lambda^\bullet,\mu^\bullet)$ of equal intensity on M, we are interested in couplings $q^\bullet$ of $\lambda^\bullet$ and $\mu^\bullet$, i.e. measure valued random variables $\omega\mapsto q^\omega$ such that for any $\omega\in\Omega$ the measure q^ω on $M\times M$ is a coupling of λ^ω and μ^ω. We look for minimizers of the *mean transportation cost*

$$\mathfrak{C}(q^\bullet)\ :=\ \sup_{B\in\mathrm{Adm}(M)}\frac{1}{m(B)}\mathbb{E}\left[\int_{M\times B}c(x,y)\ q^\bullet(dx,dy)\right],$$

where the supremum is over all bounded Borel sets that can be written as the union of translates of fundamental regions (see section 2.7). For example for $M=\mathbb{R}^d, G=\mathbb{Z}^d$ acting by translation, a typical set would be a cube of integer side length. We always consider cost functions of the form $c(x,y)=\vartheta(d(x,y))$ for some continuous strictly increasing function $\vartheta:\mathbb{R}_+\to\mathbb{R}_+$ with $\vartheta(0)=0$ and $\lim_{r\to\infty}\vartheta(r)=\infty$. Additionally, we assume that the Monge problem between two compactly supported probability measures λ and μ with $\lambda \ll m$ has a unique solution.

A coupling $q^\bullet$ of $\lambda^\bullet$ and $\mu^\bullet$ is called *optimal* if it minimizes the mean transportation cost and if it is equivariant. We will show that there always is at least one optimal coupling as soon as the optimal mean transportation cost is finite. A natural question is in which cases we can say more about the optimal coupling. When is it unique? Is it possible to construct it? Can we say something about its geometry? From the discussion above, there are basically three interesting cases, transporting absolutely continuous measures to absolutely continuous measures, transporting absolutely continuous measures to discrete measures and transporting discrete measures to discrete measures. We will study the first two cases in detail and get positive answers to all of the three questions, uniqueness, construction and the geometry. The first main result states

Theorem 1.4. *Let $(\lambda^\bullet, \mu^\bullet)$ be two equivariant random measures on M. If the optimal mean transportation cost is finite*

$$\mathfrak{c}_\infty \;=\; \inf_{q^\bullet \in \Pi(\lambda^\bullet, \mu^\bullet)} \mathfrak{C}(q^\bullet) \;<\; \infty$$

and λ^ω is absolutely continuous to the volume measure m for almost all ω, then there is a unique optimal coupling between $\lambda^\bullet$ and $\mu^\bullet$.

The unique optimal coupling $q^\bullet$ can be represented as $q^\omega = (id, T^\omega)_*\lambda^\omega$ for some measurable map $T^\omega : \mathrm{supp}(\lambda^\omega) \to \mathrm{supp}(\mu^\omega)$ measurably only dependent on the $\sigma-$algebra generated by $(\lambda^\bullet, \mu^\bullet)$. In particular, considering $M = \mathbb{R}^d$, $\lambda^\bullet = \mathcal{L}$ being the Lebesgue measure and $\mu^\bullet$ a point process on $\mathbb{R}^d$ the optimal transportation map T^ω defines a *fair factor allocation* for $\mu^\bullet$. The inverse map of T^ω assigns to each point ("center") ξ of μ^ω a set ("cell") of volume $\mu^\omega(\xi)$. If the point process is simple, all the cells will have mass one. In the case of quadratic cost $c(x, y) = |x - y|^2$ all cells will be convex polytopes of volume one, they constitute a Laguerre tessellation. In the case of linear cost $c(x, y) = |x - y|$ all cells will be starlike with respect to their center, the allocation becomes a Johnson-Mehl diagram. In the light of these results one might interpret the optimal coupling as a generalized tessellation. If $\mu^\bullet$ is even invariant under the action of $\mathbb{R}^d$ the optimal cost between $\mathcal{L}$ and $\mu^\bullet$ is given by

$$\mathfrak{c}_\infty \;=\; \mathbb{E}[\vartheta(|T(0)|)],$$

recovering the quantity studied by Peres et alii in the context of allocations.

Moreover, we prove that the optimal coupling Q^∞, if it is unique, can be obtained as the limit of optimal couplings of $\lambda^\bullet$ and $\mu^\bullet$ restricted to bounded sets. For the construction we need to impose an additional assumption on the group G. The assumptions on the group action imply that G is finitely generated. Let S be a generating set und consider the Cayley graph of G with respect to S, $\Delta(G, S)$. Let Λ_r denote the closed 2^r neighbourhood of the identity of $\Delta(G, S)$. We will assume that G satisfies some strong kind of amenability, namely

$$\lim_{r\to\infty} \frac{|\Lambda_r \triangle g\Lambda_r|}{|\Lambda_r|} \;=\; 0,$$

for all $g \in G$, where $|\cdot|$ denotes the cardinality and $\triangle$ the symmetric difference. Let B_0 be a fundamental region and $B_r = \Lambda_r B_0$. Let Q_{B_r} be the unique optimal *semicoupling* between $\lambda^\bullet$ and $1_{B_r}\mu^\bullet$, that is the unique optimal coupling between $\rho\cdot\lambda^\bullet$ and $1_{B_r}\mu^\bullet$ for some optimal choice of density ρ. Put

$$\tilde{Q}^r_g \;:=\; \frac{1}{|\Lambda_r|} \sum_{h \in g\Lambda_r} Q_{hB_r}.$$

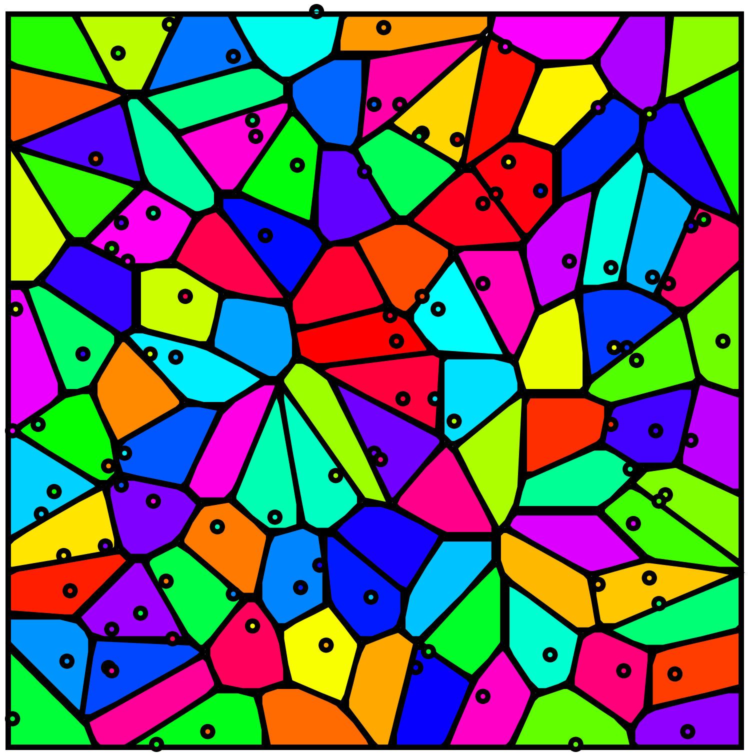

Figure 1.3: Coupling of Lebesgue and 100 points in the cube with $c(x, y) = |x - y|^2$.

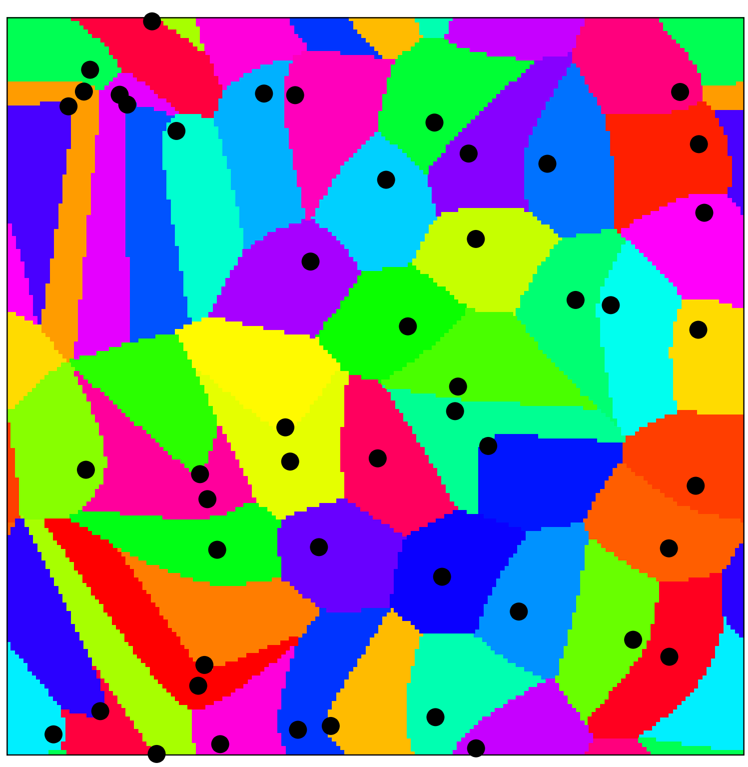

Figure 1.4: Coupling of volume measure and 49 points on a torus with cost function $c(x, y) = d(x, y)$.

Theorem 1.5. *Let $(\lambda^\bullet, \mu^\bullet)$ be two equivariant random measures on M, such that the optimal mean transportation cost are finite, $\mathfrak{c}_\infty < \infty$. Assume, that λ^ω is absolutely continuous to the volume measure m for almost all ω. Then, for every $g \in G$*

$$\tilde{Q}_g^r \to Q^\infty \quad \textit{vaguely}$$

in $\mathcal{M}(M \times M \times \Omega)$.

For the proof of this theorem the assumption of absolute continuity is only needed to ensure uniqueness of Q_{gB_r} and Q^∞. If we do not have absolute continuity but uniqueness of Q_{gB_r} and Q^∞ the same theorem with the same proof holds.

In the case of absolute continuity we can even say a bit more and get rid of the mixing. The unique optimal coupling is given by a map, that is

$$Q^\infty = (id, T)_* \lambda^\bullet.$$

Moreover, the optimal semicoupling Q_{gB_r} is given by

$$Q_{gB_r} = (id, T_{g,r})_* (\rho_{g,r} \lambda^\bullet),$$

for some measurable map $T_{g,r}$ and some density $\rho_{g,r}$. Then, we have

Theorem 1.6. *For every $g \in G$*

$$T_{g,r} \to T \quad \textit{in } \lambda^\bullet \otimes \mathbb{P} \textit{ measure}.$$

Analogous results will be obtained in the more general case of optimal semicouplings between $\lambda^\bullet$ and $\mu^\bullet$ where $\lambda^\bullet$ has intensity one and $\mu^\bullet$ has intensity $\beta \in (0, \infty)$. In the case $\beta \leq 1$, $\lambda^\bullet$ is allowed to not transport all of its mass. There will be some areas from which nothing is transported and the $\mu^\bullet$ mass can choose its favorite $\lambda^\bullet$ mass. In the case $\beta \geq 1$ the situation is the opposite. There is too much $\mu^\bullet$ mass. Hence, $\lambda^\bullet$ can choose its favorite $\mu^\bullet$ mass and some part of the $\mu^\bullet$ mass will not be satisfied, that is they will not get enough or even any of the $\lambda^\bullet$ mass.

If we do not assume this strong form of amenability the methods of our proof break down. In order to show, how non-amenable spaces and spaces beyond the Riemannian setting can be treated we consider the problem on the regular k-tree. There, we show uniqueness and construct an invariant coupling between the "Lebesgue" measure and a Poisson point process with finite mean transportation cost. However, it is not clear if this coupling is optimal or not.

If $M = \mathbb{R}^d, \lambda^\bullet = \mathcal{L}$ the Lebesgue measure and $\mu^\bullet$ is a Poisson point process of intensity β we have rather sharp estimates for the mean transportation cost to be finite.

Theorem 1.7. *(i) Assume $d \geq 3$ (and $\beta \in (0, \infty)$) or $\beta \neq 1$ (and $d \geq 1$). Then there exists a constant $0 < \kappa < \infty$ s.t.*

$$\limsup_{r \to \infty} \frac{\log \vartheta(r)}{r^d} < \kappa \quad \Longrightarrow \quad \mathfrak{c}_\infty < \infty \quad \Longrightarrow \quad \liminf_{r \to \infty} \frac{\log \vartheta(r)}{r^d} \leq \kappa.$$

(ii) Assume $d \leq 2$ and $\beta = 1$. Then for any concave $\hat{\vartheta} : [1, \infty) \to \mathbb{R}$ dominating ϑ

$$\int_1^\infty \frac{\hat{\vartheta}(r)}{r^{1+d/2}} dr < \infty \quad \Longrightarrow \quad \mathfrak{c}_\infty < \infty \quad \Longrightarrow \quad \liminf_{r \to \infty} \frac{\vartheta(r)}{r^{d/2}} = 0.$$

The first implication in assertion (ii) is new. Assertion (i) in the case $\beta = 1$ is due to Holroyd and Peres [HP05], based on a fundamental result of Talagrand [Tal94]. The first implication in assertion (i) in the case $\beta \neq 1$ was proven by Hoffman, Holroyd and Peres [HHP06]. The second implication in assertion (ii) is due to [HL01].
Now let us consider the particular case of L^p transportation cost, i.e. $\vartheta(r) = r^p$.

Corollary 1.8. *(i) For all $d \in \mathbb{N}$, all $\beta \in (0,\infty)$ and $p \in (0,\infty)$ the mean L^p-transportation cost $\mathfrak{c}_\infty$ is finite if and only if*

$$p < \overline{p} := \begin{cases} \infty, & \text{for } d \geq 3 \text{ or } \beta \neq 1; \\ \frac{d}{2}, & \text{for } d \leq 2 \text{ and } \beta = 1. \end{cases}$$

(ii) If $\beta = 1$ then for all $p \in (0,\infty)$ there exist constants $0 < k \leq k' < \infty$ s.t. for all $d > 2(p \wedge 1)$

$$k \cdot d^{p/2} \ \leq\ \mathfrak{c}_\infty \ \leq\ k' \cdot d^{p/2}.$$

The radical shift in the behavior of the mean transportation cost between dimensions two and three might seem strange at first. It is due to the following observation. By the central limit theorem, there are $N^d \pm N^{d/2}$ Poisson points in a cube of side length N. To get finite transportation cost the Lebesgue measure close to the cube needs to compensate these fluctuations. In a one-neighbourhood of the cube there is N^{d-1} Lebesgue mass. Therefore, the "condition" for finite mean transportation cost is

$$N^{d/2} \leq N^{d-1}.$$

In dimension one this condition is violated, dimension two is borderline and for dimension $d \geq 3$ everything is fine.

Stability is not as easy to get as in the classical theory because cyclical monotonicity is not sufficient for uniqueness. For example, the transport $z \mapsto z + 42$ transporting $\sum_{z\in\mathbb{Z}} \delta_z$ to itself is cyclically monotone for quadratic cost but certainly not optimal. However, we get at least close to it.

Proposition 1.9. *Let $(\lambda_n^\bullet)_{n\in\mathbb{N}}$ and $(\mu_n^\bullet)_{n\in\mathbb{N}}$ be two sequences of equivariant random measures. Let $q_n^\bullet$ be the unique optimal coupling between $\lambda_n^\bullet$ and $\mu_n^\bullet$. Assume that $\lambda_n^\bullet\mathbb{P} \to \lambda^\bullet\mathbb{P}$ vaguely, $\mu_n^\bullet\mathbb{P} \to \mu^\bullet\mathbb{P}$ vaguely and $\sup_n \mathfrak{C}(q_n^\bullet) \leq c < \infty$. Then, there is an equivariant coupling $q^\bullet$ of $\lambda^\bullet$ and $\mu^\bullet$ and a subsequence $(q_{n_k}^\bullet)_{k\in\mathbb{N}}$ such that $q_{n_k}^\bullet\mathbb{P} \to q^\bullet\mathbb{P}$ vaguely, the support of $q^\bullet$ is cyclically monotone and*

$$\mathfrak{C}(q^\bullet) \leq \liminf_{n\to\infty} \mathfrak{C}(q_n^\bullet).$$

A consequence of this statement is the continuity of the Wiener mosaic. Start with a Poisson point process of unit intensity in $\mathbb{R}^d$ with $d \geq 3$. Let each of the Poisson points evolve according to independent Brownian motions. For any t let $q_t^\bullet$ be the unique optimal coupling between $\mathcal{L}$ and $\mu_t^\bullet$ wrt to the cost function $c(x,y) = |x-y|^2$. For any t this is a Laguerre tessellation. The stability result tells us that $q_t^\bullet$ is continuous in t. Hence, we get a continuously moving mosaic.

We would very much like to be able to say more about the optimal transport problem between two discrete random measures. It is clear that for example the question of uniqueness is in general false. If you consider the deterministic measures $\lambda = \sum_{z\in\mathbb{Z}} \delta_z$ and $\mu = \sum_{z\in\mathbb{Z}} \delta_{z+1/2}$, there is no uniqueness. This is no surprise and exactly what one would expect from the classical theory. However, if you throw n iid uniformly distributed red points in the unit cube and n iid uniformly distributed blue points in the same cube, almost surely there is a unique optimal coupling which then must be a matching (for our choice of cost function). Therefore, it is reasonable to study optimal couplings of two independent Poisson point processes and ask whether they are unique. If they are unique they must automatically be a matching because matchings are the extreme points in the convex set of couplings of these point processes. Unfortunately, we are not yet able to prove anything in this direction.

In the case of $M = \mathbb{R}^d, \lambda^\bullet = \mathcal{L}$ and $\mu^\bullet$ being a point process of intensity $\beta \leq 1$ Theorems 1.4, 1.5, 1.6 and also Theorem 1.7 and Corollary 1.8 were obtained jointly with Karl-Theodor Sturm in [HS10].

1.3 Overview

In chapter 2 we explain the setting and introduce the objects we will work with. At the end of that chapter we derive the general existence result of optimal semicouplings by a compactness argument. The important existence and uniqueness result for semicouplings on bounded sets will be shown in chapter 3. This enables us to prove Theorem 1.4, the uniqueness statement, in chapter 4. Moreover, the representation of the optimal semicoupling will allow us to draw some conclusions on the geometry of the induced allocations. In chapter 5 we will prove Theorems 1.5 and 1.6. In chapter 4 and 5 we only deal with the case of the second marginal having intensity $\beta \leq 1$. In chapter 6 we show which parts in the arguments have to be changed in order to get the results also for $\beta \geq 1$. The missing parts of Theorem 1.7 will be derived in chapter 7. Moreover, we will show similar estimates for the case of a compound Poisson process. In chapter 8 we show how one could deal with non-amenable spaces by considering the regular k-tree. Finally in chapter 9 we show some metric properties for the case of L^p cost and prove the stability result.

Chapter 2

Set-Up and Basic Concepts

In this chapter we will explain the general set-up, some basic concepts and derive the first result, a general existence result by a compactness argument.

2.1 The setting

From now on we will always assume to work in the following setting. (M, d, m) will denote a complete connected smooth non-compact Riemannian manifold with Riemannian distance d and Riemannian volume measure m. The Borel sets on M will be denoted by $\mathcal{B}(M)$. Given a map S and a measure ρ we denote the push forward of ρ under S by $S_*\rho$, i.e. $S_*\rho(A) = \rho(S^{-1}(A))$ for any Borel set A. Given any product $X = \Pi_{i=1}^n X_i$ of measurable spaces, the projection onto the i–th space will be denoted by π_i. Given a set $A \subset M$ its complement will be denoted by $\complement A$ and the indicator function of A by 1_A.

Moreover, we will assume that there is a group G of isometries acting on M. For a set $A \subset M$ we write $\tau_g A := gA = \{ga : a \in A\}$. For a point $x \in M$ its *orbit* under the group action of G is defined as $Gx = \{gx : g \in G\}$. Its *stabilizer* is defined as $G_x = \{g \in G : gx = x\}$ the elements of G that fix x.

Definition 2.1 (Group action)**.** *Let G act on M. We say that the action is*

- properly discontinuous *if for any $x \in M$ and any compact $K \subset M$ $gx \in K$ for only finitely many $g \in G$.*
- cocompact *if M/G is compact in the quotient topology.*
- free *if $gx = x$ for one $x \in M$ implies $g = id$, that is the stabilizer for every point is trivial.*

We will assume that the group action is properly discontinuous, cocompact and free. By Theorem 3.5 in [Bow06] this already implies that G is finitely generated and therefore countable.

Definition 2.2 (Fundamental region)**.** *A measurable subset $B_0 \subset M$ is defined to be a fundamental region for G if*

i) $\bigcup_{g \in G} gB_0 = M$

ii) $B_0 \cap gB_0 = \emptyset$ *for all* $id \neq g \in G$.

The family $\{gB_0 : g \in G\}$ is called tessellation of M.

There are many different choices of fundamental regions. We will choose a special one, namely a certain subset of the Dirichlet region with respect to some fixed point p. However, each fundamental region has the same volume and therefore defines a tiling of M in pieces of equal volume. Indeed, we have the following Lemma.

Lemma 2.3. *Let F_1 and F_2 be two fundamental regions for G. Assume $m(F_1) < \infty$. Then $m(F_1) = m(F_2)$.*

Proof. As $F_1 \cap gF_2$ and $F_1 \cap hF_2$ are disjoint for $g \neq h$ by the defining property of fundamental regions we have

$$m(F_1) = \sum_{g \in G} m(F_1 \cap gF_2) = \sum_{g \in G} m(g^{-1}F_1 \cap F_2) = m(F_2).$$

□

By scaling of the volume measure m we can assume that $m(B_0) = 1$. This assumption is just made to simplify some notations.
As G is finitely generated, there are finitely many elements $a_1, \dots, a_k \in G$ such that every $g \in G$ can be written as a word in these letters and their inverses. The set $S = \{a_1, \dots, a_k\}$ is called a generating set. The generating set is not unique, e.g. $\mathbb{Z}$ is generated by $\{1\}$ but also by $\{2, 3\}$. We will fix one finite generating set for G. It does not matter which one as the results will be independent from the specific choice.
Given the generating set S. We can construct a graph $\Delta = \Delta(G, S)$ as follows. Put $V(\Delta) = G$ as the vertices. For each $g \in G$ and $a \in S$ we connect g and ag by a directed edge labeled with a. The same edge with opposite orientation is labeled by a^{-1}. This gives a regular graph of degree $2|S|$. We endow Δ with the word metric d_Δ which coincides with the usual graph distance.

Definition 2.4 (Cayley graph)**.** *If S is a generating set of G, then $\Delta(G, S)$ is called Cayley graph of G with respect to S.*

We denote the closed 2^r neighbourhood of the identity element in Δ by Λ_r, that is $\Lambda_r = \{g \in G : d_\Delta(1, g) \leq 2^r\}$. The boundary of Λ_r is defined as $\partial\Lambda_r = \{h \notin \Lambda_r : \exists g \in \Lambda_r \text{ s.t. } d_\Delta(h, g) = 1\}$. By B_r we denote the range of the action of Λ_r on the fundamental domain B_0, that is $B_r = \bigcup_{g \in \Lambda_r} gB_0$.

We will need to control the mass that is close to the boundary of B_r, that is the growth of B_r. To this end we introduce the notion of *amenability*. A linear functional on $L^\infty(G)$ is called a mean if it maps the constant function $\mathbf{1}$ to 1 and nonnegative functions to nonnegative numbers. G acts on $L^\infty(G)$ from the left by $L_g f(h) = f(gh)$ for $f \in L^\infty(G)$ and $g, h \in G$. A mean ρ is called invariant if $\rho(L_g f) = \rho(f)$ for all $f \in L^\infty(G)$ and $g \in G$.

Definition 2.5 (Amenability)**.** *G is called amenable iff it admits an invariant mean.*

A more geometric characterization which we will use is due to Følner. It can be found for example in Theorem 4.13 of [Pat00].

Theorem 2.6. *The following statements are equivalent:*

i) G is amenable.

ii) For any nonempty and compact $C, L \subset G$ and $\epsilon > 0$ we can find a nonempty $K \subset G$ with $L \subset K$ such that

$$h(CK \triangle K)/h(K) < \epsilon,$$

where h denotes the Haar measure on G and $\triangle$ the symmetric difference. The sets K are called Følner sets.

Several times we will use a rather simple but very powerful tool, the *mass transport principle*. It already appeared in the proof of Lemma 2.3. It is a kind of conservation of mass formula for invariant transports.

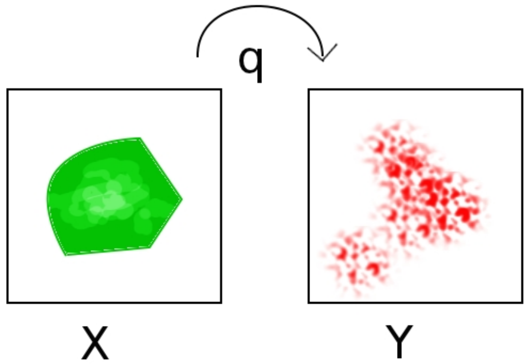

Figure 2.1: Concept of semicoupling with finite mass: Choose a density $f \in [0,1]$ (green) such that $f\lambda$ and μ (red) have the same mass. Then, choose a coupling between them.

Lemma 2.7 (mass transport principle)**.** *Let $f : G \times G \to \mathbb{R}_+$ be a function which is invariant under the diagonal action of G, that is $f(u,v) = f(gu, gv)$ for all $g, u, v \in G$. Then we have*

$$\sum_{v \in G} f(u,v) \;=\; \sum_{v \in G} f(v,u).$$

Proof.

$$\sum_{v \in G} f(u,v) \;=\; \sum_{g \in G} f(u,gu) \;=\; \sum_{g \in G} f(g^{-1}u,u) \;=\; \sum_{v \in G} f(v,u).$$

□

For a more general version we refer to [BLPS99] and [LT08]. The first reference is also a nice example for a probabilistic equivalence to amenability.

2.2 Couplings and Semicouplings

For each Polish space X (i.e. complete separable metric space) the set of Radon measures on X – equipped with its Borel σ-field – will be denoted by $\mathcal{M}(X)$. Given any ordered pair of Polish spaces X, Y and measures $\lambda \in \mathcal{M}(X), \mu \in \mathcal{M}(Y)$ we say that a measure $q \in \mathcal{M}(X \times Y)$ is a *semicoupling* of λ and μ, briefly $q \in \Pi_s(\lambda, \mu)$, iff the (first and second, resp.) marginals satisfy

$$(\pi_1)_* q \leq \lambda, \qquad (\pi_2)_* q = \mu,$$

that is, iff $q(A \times Y) \leq \lambda(A)$ and $q(X \times B) = \mu(B)$ for all Borel sets $A \subset X, B \subset Y$. The semicoupling q is called *coupling*, briefly $q \in \Pi(\lambda, \mu)$, iff in addition

$$(\pi_1)_* q = \lambda.$$

Existence of a coupling requires that the measures λ and μ have the same total mass. If the total masses of λ and μ are finite and equal then the 'renormalized' product measure $q = \frac{1}{\lambda(X)} \lambda \otimes \mu$ is always a coupling of λ and μ.
If λ and μ are Σ-finite, i.e. $\lambda = \sum_{n=1}^{\infty} \lambda_n$, $\mu = \sum_{n=1}^{\infty} \mu_n$ with finite measures $\lambda_n \in \mathcal{M}(X)$ mutually singular, $\mu_n \in \mathcal{M}(Y)$ mutually singular – which is the case for all Radon measures – and if both of them have infinite total mass then there always exists a Σ-finite coupling of them. (Indeed, then the λ_n and μ_n can be chosen to have unit mass and $q = \sum_n (\lambda_n \otimes \mu_n)$ does the job.)
See also [Fig10] for the related concept of *partial coupling*.

2.3 Random measures on M

The set of all Radon measures on M will be denoted by $\mathcal{M}(M)$ which we will endow with the vague topology. The vague topology is defined by duality with continuous functions f with compact support ($f \in C_c(M)$), this means that a sequence of measures $(\lambda_n)_{n\in\mathbb{N}}$ in $\mathcal{M}(M)$ converges to some measure $\lambda \in \mathcal{M}(M)$ iff $\int f d\lambda_n \to \int f d\lambda$ for all $f \in C_c(M)$. The next Lemma summarizes some basic facts about vague topology (e.g. see [Kal97] or [Bau01])

Lemma 2.8 (vague topology)**.** *Let X be a locally compact second countable Haussdorff space. Then,*

i) $\mathcal{M}(X)$ is a Polish space in the vague topology.

ii) $A \subset \mathcal{M}(X)$ is vaguely relatively compact iff $\sup_{\mu\in A} \mu(f) < \infty$ for all $f \in C_c(X)$.

iii) If $\mu_n \xrightarrow{v} \mu$ and $B \subset X$ relatively compact with $\mu(\partial B) = 0$ then $\mu_n(B) \to \mu(B)$.

The action of G on M induces an action of G on $\mathcal{M}(M \times \ldots \times M)$ by push forward with the map τ_g:

$$(\tau_g)_*\lambda(A_1, \ldots, A_k) \;=\; \lambda((g^{-1}(A_1), \ldots, g^{-1}(A_k)) \quad \forall A_1, \ldots A_k \in \mathcal{B}(M), k \in \mathbb{N}.$$

Recall the disintegration theorem for finite measures (e.g. see Theorem 5.1.3 in [AGS08] or III-70 in [DM78]).

Theorem 2.9 (Disintegration of measures)**.** *Let X, Y be Polish spaces, and let γ be a finite Borel measure on $X \times Y$. Denote by μ and ν the marginals of γ on the first and second factor respectively. Then, there exist two measurable families of probability measures $(\gamma_x)_{x\in X}$ and $(\gamma_y)_{y\in Y}$ such that*

$$\gamma(dx, dy) = \gamma_x(dy)\mu(dx) = \gamma_y(dx)\nu(dy).$$

A random measure on M is a random variable $\lambda^\bullet$ (the notation with the "$\bullet$" is intended to make it easier to distinguish random and non-random measures) modeled on some probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ taking values in $\mathcal{M}(M)$. It can also be regarded as a kernel from Ω to M. Therefore, we write either $\lambda^\omega(A)$ or $\lambda(\omega, A)$ depending on which property we want to stress. For convenience, we will assume that Ω is a compact metric space and $\mathfrak{A}$ its completed Borel field. These technical assumptions are only made to simplify the presentation.

A point process is a random measure $\mu^\bullet$ taking values in the (vaguely closed) subset of all locally finite counting measures on M. It is called simple iff $\mu^\omega(\{x\}) \in \{0, 1\}$ for every $x \in M$ and a.e. $\omega \in \Omega$. We call a random measure $\lambda^\bullet$ absolutely continuous iff it is absolutely continuous to the volume measure m on M for a.e. $\omega \in \Omega$. It is called diffusive iff it has no atoms almost surely. The *intensity measure* of a random measure $\lambda^\bullet$ is a measure on M defined by $A \mapsto \mathbb{E}[\lambda^\bullet(A)]$. The class of all relatively compact sets in $\mathcal{B}(M)$ will be denoted by $\hat{\mathcal{B}}$. For a random measure $\lambda^\bullet$ its class of stochastic continuity sets is defined by $\hat{\mathcal{B}}_{\lambda^\bullet} = \{A \in \hat{\mathcal{B}} : \lambda^\bullet(\partial A) = 0 \ a.s.\}$. Convergence in distribution and tightness in $\mathcal{M}(M)$ can be characterized by

Lemma 2.10 (tightness of random measures)**.** *Let $\lambda_1^\bullet, \lambda_2^\bullet, \ldots$ be random measures on M. Then the sequence $(\lambda_n^\bullet)_{n\in\mathbb{N}}$ is relatively compact in distribution iff $(\lambda_n^\bullet(A))_{n\in\mathbb{N}}$ is tight in $\mathbb{R}_+$ for every $A \in \hat{\mathcal{B}}$.*

Theorem 2.11 (convergence of random measures)**.** *Let $\lambda^\bullet, \lambda_1^\bullet, \lambda_2^\bullet, \ldots$ be random measures on M. Then, these conditions are equivalent:*

i) $\lambda_n^\bullet \xrightarrow{d} \lambda^\bullet$

ii) $\lambda_n^\bullet(f) \xrightarrow{d} \lambda^\bullet(f)$ *for all* $f \in C_c(M)$

iii) $(\lambda_n^\bullet(A_1), \ldots, \lambda_n^\bullet(A_k)) \xrightarrow{d} (\lambda^\bullet(A_1), \ldots, \lambda^\bullet(A_k))$ *for all* $A_1, \ldots, A_k \in \hat{\mathcal{B}}_{\lambda^\bullet}, k \in \mathbb{N}$.

If $\lambda^\bullet$ *is a simple point process or a diffusive random measure, it is also equivalent that*

iv) $\lambda_n^\bullet(A) \xrightarrow{d} \lambda^\bullet(A)$ *for all* $A \in \hat{\mathcal{B}}_{\lambda^\bullet}$.

For the proof of these statements we refer to Lemma 14.15 and Theorem 14.16 of [Kal97]. Just as in Lemma 11.1.II of [DVJ07] we can derive the following result on continuity sets of a random measure $\lambda^\bullet$:

Lemma 2.12. *Let* $\lambda^\bullet$ *be a random measure on* M, $A \in \mathcal{B}(M)$ *be bounded and* $(A)_r$ *be the* r*-neighbourhood of* A *in* M. *Then for all but a countable set of* $r \in \mathbb{R}_+$ *we have* $(A)_r \in \hat{\mathcal{B}}_{\lambda^\bullet}$.

A random measure $\lambda^\bullet : \Omega \to \mathcal{M}(M)$ is called G-invariant or just invariant if the distribution of $\lambda^\bullet$ is invariant under the action of G, that is, iff

$$(\tau_g)_*\lambda^\bullet \overset{(d)}{=} \lambda^\bullet$$

for all $g \in G$. A random measure $q^\bullet : \Omega \to \mathcal{M}(M \times M)$ is called invariant if its distribution is invariant under the diagonal action of G.
If $(\Omega, \mathfrak{A})$ admits a measurable flow $\theta_g : \Omega \to \Omega, g \in G$, that is a measurable mapping $(\omega, g) \mapsto \theta_g\omega$ with θ_0 the identity on Ω and

$$\theta_g \circ \theta_h = \theta_{gh}, \quad g, h \in G,$$

then a random measure $\lambda^\bullet : \Omega \to \mathcal{M}(M)$ is called G-equivariant or just equivariant iff

$$\lambda(\theta_g\omega, gA) = \lambda(\omega, A),$$

for all $g \in G, \omega \in \Omega, A \in \mathcal{B}(M)$. We can think of $\lambda(\theta_g\omega, \cdot)$ as $\lambda(\omega, \cdot)$ shifted by g. Indeed, let $\mathfrak{M}$ be the cylindrical $\sigma-$algebra generated by the evaluation functionals $A \mapsto \mu(A), A \in \mathcal{B}(M), \mu \in \mathcal{M}$. As in example 2.1 of [LT08], consider the measurable space $(\mathcal{M}, \mathfrak{M})$ and define for $\mu \in \mathcal{M}, g \in G$ the measure $\theta_g\mu(A) = \mu(g^{-1}A)$. Then, $\{\theta_g, g \in G\}$ is a measurable flow and the identity is an equivariant measure. A random measure $q^\bullet : \Omega \to \mathcal{M}(M \times M)$ is called equivariant iff

$$q^{\theta_g\omega}(gA, gB) = q^\omega(A, B),$$

for all $g \in G, \omega \in \Omega, A, B \in \mathcal{B}(M)$.

Example 2.13. Let $q^\bullet$ be an equivariant random measure on $M \times M$ given by $q^\omega = (id, T^\omega)_*\lambda^\omega$ for some measurable map $T^\bullet$ and some equivariant random measure $\lambda^\bullet$. The equivariance condition

$$\int_A 1_B(y)\delta_{T^{\theta_g\omega}(gx)}(d(gy))\lambda^{\theta_g\omega}(dx) = q^{\theta_g\omega}(gA, gB) = q^\omega(A, B) = \int_A 1_B(y)\delta_{T^\omega(x)}(dy)\lambda^\omega(dx),$$

translates into an equivariance condition for the transport maps:

$$T^{\theta_g\omega}(gx) = gT^\omega(x).$$

A probability measure $\mathbb{P}$ is called *stationary* iff

$$\mathbb{P} \circ \theta_g = \mathbb{P}$$

for all $g \in G$. Given a measure space $(\Omega, \mathfrak{A})$ with a measurable flow $(\theta_g)_{g \in G}$ and a stationary probability measure $\mathbb{P}$ any equivariant measure in automatically invariant. The advantage of this definition is that the sum of equivariant measures is again equivariant, and therefore also invariant. The sum of two invariant random measures does not have to be invariant (see Remark 2.22).

We say that a random measure $\lambda^\bullet$ has subunit intensity iff $\mathbb{E}[\lambda^\bullet(A)] \le m(A)$ for all $A \in \mathcal{B}(M)$. If equality holds in the last statement we say that the random measure has unit intensity. An invariant random measure has subunit (or unit) intensity iff its intensity

$$\beta \;=\; \mathbb{E}[\lambda^\bullet(B_0)]$$

is ≤ 1 (or $= 1$ resp.). Given a random measure, the measure $(\lambda^\bullet\mathbb{P})(dy, d\omega) := \lambda^\omega(dy)\,\mathbb{P}(d\omega)$ on $M \times \Omega$ is called *Campbell measure* of the random measure $\lambda^\bullet$.

Example 2.14. i) The Poisson point process with intensity measure m. It is characterized by

- for each Borel set $A \subset M$ of finite volume the random variable $\omega \mapsto \mu^\omega(A)$ is Poisson distributed with parameter $m(A)$ and
- for disjoint sets $A_1, \dots A_k \subset M$ the random variables $\mu^\omega(A_1), \dots, \mu^\omega(A_k)$ are independent.

It can be written as

$$\mu^\omega = \sum_{\xi \in \Xi(\omega)} \delta_\xi$$

with some countable set $\Xi(\omega) \subset M$ without accumulation points.

ii) The compound Poisson process is a Poisson process with random weights instead of unit weights. It is compounded with another distribution giving the weights of the different atoms. It can be written as

$$\mu^\omega = \sum_{\xi \in \Xi(\omega)} X_\xi \delta_\xi$$

for some iid sequence $(X_\xi)_{\xi \in \Xi(\omega)}$ independent of the Poisson point process. For example one could take X_ξ to be a Poisson random variable or an exponentially distributed random variable. If X_ξ has distribution γ we say $\mu^\bullet$ is a $\gamma-$compound Poisson process.

iii) The volume measure m is a constant random measure which is equivariant.

iv) An example of a non-equivariant random measure is a Poisson point process on $\mathbb{R}^d$ with intensity measure a standard normal distribution.

From now on we will always assume that we are given two equivariant random measures $\lambda^\bullet$ and $\mu^\bullet$ modeled on some probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ admitting a measurable flow $(\theta_g)_{g \in G}$ such that $\mathbb{P}$ is stationary. We will assume that Ω is a compact metric space. Moreover, we will assume that $\lambda^\bullet$ is absolutely continuous and $\lambda^\bullet$ and $\mu^\bullet$ are almost surely not the zero measure. Note that the invariance implies that $\mu^\omega(M) = \lambda^\omega(M) = \infty$ for almost every ω (e.g. see Proposition 12.1.VI in [DVJ07]).

2.4 Semicouplings of $\lambda^\bullet$ and $\mu^\bullet$

A semicoupling of the random measures $\lambda^\bullet$ and $\mu^\bullet$ is a measurable map $q^\bullet : \Omega \to \mathcal{M}(M \times M)$ s.t. for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$q^\omega \;\text{ is a semicoupling of } \lambda^\omega \text{ and } \mu^\omega.$$

Its Campbell measure is given by $Q = q^\bullet \mathbb{P} \in \mathcal{M}(M \times M \times \Omega)$. Q is a semicoupling between the Campbell measures $\lambda^\bullet \mathbb{P}$ and $\mu^\bullet \mathbb{P}$ in the sense that

$$Q(M \times \cdot \times \cdot) \;=\; \mu^\bullet \mathbb{P} \;\text{ and }\; Q(\cdot \times M \times \cdot) \;\leq\; \lambda^\bullet \mathbb{P}.$$

Q could also be regarded as semicoupling between $\lambda^\bullet \mathbb{P}$ and $\mu^\bullet \mathbb{P}$ on $M \times \Omega \times M \times \Omega$ which is concentrated on the diagonal of $\Omega \times \Omega$. It could be interesting to relax this last condition on Q and allow different couplings of the randomness. However, we will not do so and only consider semicouplings of $\lambda^\bullet \mathbb{P}$ and $\mu^\bullet \mathbb{P}$ that are concentrated on the diagonal of $\Omega \times \Omega$. We will always identify these semicouplings with measures on $M \times M \times \Omega$.
Given such a semicoupling $Q \in \mathcal{M}(M \times M \times \Omega)$ we can disintegrate (see Theorem 2.9) Q to get a measurable map $q^\bullet : \; \Omega \to \mathcal{M}(M \times M)$ which is a semicoupling of $\lambda^\bullet$ and $\mu^\bullet$.

According to this one-to-one correspondence between $q^\bullet$ — semicoupling of $\lambda^\bullet$ and $\mu^\bullet$ — and $Q = q^\bullet \mathbb{P}$ — semicoupling of $\lambda^\bullet \mathbb{P}$ and $\mu^\bullet \mathbb{P}$ — we will freely switch between them. And quite often, we will simply speak of *semicouplings of* $\lambda^\bullet$ *and* $\mu^\bullet$.

We denote the set of all semicouplings between $\lambda^\bullet$ and $\mu^\bullet$ by $\Pi_s(\lambda^\bullet, \mu^\bullet)$. The set of all equivariant semicouplings between $\lambda^\bullet$ and $\mu^\bullet$ will be denoted by $\Pi_{is}(\lambda^\bullet, \mu^\bullet)$.

A *factor* of some random variable X is a random variable Y which is measurable with respect to $\sigma(X)$. This is equivalent to the existence of a deterministic function f with $Y = f(X)$. In other words, a factor is a rule such that given X we can construct Y. A *factor semicoupling* is a semicoupling of $\lambda^\bullet$ and $\mu^\bullet$ which is a factor of $\lambda^\bullet$ and $\mu^\bullet$.

2.5 Cost functionals

Throughout this thesis, ϑ will be a strictly increasing, continuous function from $\mathbb{R}_+$ to $\mathbb{R}_+$ with $\vartheta(0) = 0$ and $\lim\limits_{r \to \infty} \vartheta(r) = \infty$. Given a *scale function* ϑ as above we define the *cost function*

$$c(x, y) = \vartheta\left(d(x, y)\right)$$

on $M \times M$, the *cost functional*

$$\mathsf{Cost}(q) = \int_{M \times M} c(x, y)\, q(dx, dy)$$

on $\mathcal{M}(M \times M)$ and the *mean cost functional*

$$\mathfrak{Cost}(Q) = \int_{M \times M \times \Omega} c(x, y)\; Q(dx, dy, d\omega)$$

on $\mathcal{M}(M \times M \times \Omega)$. Moreover, we will always assume that the cost function is such that the optimal transportation problem between $\rho \cdot m$ and μ, with a compactly supported probability density ρ and a compactly supported arbitrary probability measure μ has a unique solution which is induced by a map. To be more precise, the optimal transportation problem between $\rho \cdot m$ and μ for the cost function c is to minimize

$$\int_{M \times M} c(x, y) q(dx, dy)$$

over all couplings q of $\rho \cdot m$ and μ. We will assume, that the cost function is such that the solution of this problem, the unique optimal coupling, is given as $q = (id, T)_*(\rho \cdot m)$ for some measurable map T. For conditions on ϑ such that this assumption is satisfied we refer to chapter 3.

We have the following basic result on existence and uniqueness of optimal semicouplings the proof of which is deferred to chapter 3. The first part of the theorem, the existence and uniqueness of an optimal semicoupling, is very much in the spirit of an analogous result by Figalli [Fig10] on existence and (if enough mass is transported) uniqueness of an optimal partial coupling. However, in our case the second marginal is arbitrary whereas in [Fig10] it is absolutely continuous.

Theorem 2.15. *(i) For each bounded Borel set $A \subset M$ there exists a unique semicoupling Q_A of $\lambda^\bullet \mathbb{P}$ and $(1_A \mu^\bullet)\mathbb{P}$ which minimizes the mean cost functional $\mathfrak{Cost}(.)$.*
(ii) The measure Q_A can be disintegrated as $Q_A(dx, dy, d\omega) := q_A^\omega(dx, dy)\, \mathbb{P}(d\omega)$ where for $\mathbb{P}$-a.e. ω the measure q_A^ω is the unique minimizer of the cost functional $\mathsf{Cost}(.)$ among the semicouplings of λ^ω and $1_A \mu^\omega$.
(iii) $\mathfrak{Cost}(Q_A) = \int_\Omega \mathsf{Cost}(q_A^\omega)\, \mathbb{P}(d\omega)$.

For a bounded Borel set $A \subset M$, the *transportation cost on A* is given by the random variable $\mathsf{C}_A : \Omega \to [0, \infty]$ as

$$\mathsf{C}_A(\omega) := \mathsf{Cost}(q_A^\omega) = \inf\{\mathsf{Cost}(q^\omega) : \; q^\omega \text{ semicoupling of } \lambda^\omega \text{ and } 1_A\, \mu^\omega\}.$$

Lemma 2.16. *(i) If $A_1, \ldots, A_n$ are disjoint then $\forall \omega \in \Omega$*

$$\mathsf{C}_{\bigcup_{i=1}^n A_i}(\omega) \quad \geq \quad \sum_{i=1}^n \mathsf{C}_{A_i}(\omega)$$

(ii) If $A_1 = gA_2$ for some $g \in G$, then C_{A_1} and C_{A_2} are identically distributed.

Proof. Property (ii) follows directly from the joint invariance of $\lambda^\bullet$ and $\mu^\bullet$. The intuitive argument for (i) is, that minimizing the cost on $\bigcup_i A_i$ is more restrictive than doing it separately on each of the A_i. The more detailed argument is the following. Given any semicoupling q^ω of λ^ω and $1_{\bigcup_i A_i}\mu^\omega$ then for each i the measure $q_i^\omega := 1_{M \times A_i} q^\omega$ is a semicoupling of λ^ω and $1_{A_i}\mu^\omega$. Choosing q^ω as the minimizer of $\mathsf{C}_{\bigcup_{i=1}^n A_i}(\omega)$ yields

$$\mathsf{C}_{\bigcup_i A_i}(\omega) = \mathsf{Cost}(q^\omega) = \sum_i \mathsf{Cost}(q_i^\omega) \geq \sum_i \mathsf{C}_{A_i}(\omega).$$

□

2.6 Standard tessellations

In this section, we construct the fundamental region B_0 and thereby a tessellation or a tiling of M. We will call this tessellation a standard tessellation. The specific choice of fundamental domain is not really important for us. However, we will choose one to fix ideas.
We now define the Dirichlet region. To this end let $p \in M$ be arbitrary. Due to the assumption of freeness, the stabilizer of p is trivial. Construct the Voronoi tessellation with respect to Gp, the orbit of p. The cell containing p is the Dirichlet region.

Definition 2.17 (Dirichlet region)**.** *Let $p \in M$ be arbitrary. The Dirichlet region of G centered at p is defined by*

$$D_p(G) \;=\; \{x \in M \;:\; d(x,p) \leq d(x, gp) \;\forall g \in G\}.$$

From now on we will fix p and write for simplicity of notation $D = D_p(G)$. We want to construct a fundamental domain from D. For every $x \in \overset{\circ}{D}$ we have $d(x, p) < d(gx, p)$ for every $id \neq g \in G$, that is $|Gx \cap D| = 1$, where $|H|$ denotes the cardinality of H. However, if $x \in \partial D$ we have

$x \in D \cap gD \neq \emptyset$ for some $g \in G$. This implies that $|Gx \cap D| \geq 2$. Yet, for the fundamental region, B_0, we need exactly one representative from every orbit. Hence, we need to chose from any orbit Gx intersecting the boundary of D exactly one representative $z \in Gx \cap \partial D$. Let V be a measurable selection of these and finally define $B_0 = \overset{\circ}{D} \cup V$. By definition, B_0 is a fundamental region. There exists indeed a measurable selection by the following theorem (Theorem 17 and the following Corollary in [Del75]):

Theorem 2.18 (Dellacherie)**.** *A surjective Borel map f between Polish spaces such that the fibers $f^{-1}(y)$ are all compact, admits a Borel right inverse.*

By assumption on the group action, D is compact and the orbit of any $x \in M, Gx$, is discrete. Consider the map $f : D \to M/G, x \mapsto [x]$. The inverse of $[x] \in M/G$ under f is $f^{-1}([x]) = Gx \cap D$. Hence, it is a closed subset of a compact set and therefore compact. Apply the theorem to construct a Borel right inverse and therefore the desired selection V.

Example 2.19. Considering $\mathbb{R}^d$ with group action translations by $\mathbb{Z}^d$ a choice for the fundamental region would be $B_0 = [0,1)^d$. If we consider $M = \mathbb{H}^2$ the two dimensional hyperbolic space we can take for G a Fuchsian group acting cocompactly and freely, that is, with no elliptic elements. The Dirichlet region then becomes a hyperbolic polygon (see [Kat92]).

2.7 Optimality

The standard notion of optimality – minimizers of Cost or $\mathfrak{Cost}$ – is not well adapted to our setting. For example for any semicoupling $q^\bullet$ between the Lebesgue measure and a Poisson point process of intensity $\beta \leq 1$ we have $\mathfrak{Cost}(q^\bullet) = \infty$. Hence, we need to introduce a different notion which we explain in this section.

The collection of admissible sets is defined as $\mathrm{Adm}(M) = \{B \in \mathcal{B}(M) : \exists I \subset G, 1 \leq |I| < \infty, F \text{ fundamental region} : B = \bigcup_{g \in I} gF\}$

For a semicoupling $q^\bullet$ between $\lambda^\bullet$ and $\mu^\bullet$ the *mean transportation cost* of $q^\bullet$ is defined by

$$\mathfrak{C}(q^\bullet) \ := \sup_{B \in \mathrm{Adm}(M)} \frac{1}{m(B)} \mathbb{E}\left[\int_{M \times B} c(x,y)\, q^\bullet(dx,dy)\right].$$

Definition 2.20. *A semicoupling $q^\bullet$ between $\lambda^\bullet$ and $\mu^\bullet$ is called*

i) asymptotically optimal *iff*

$$\mathfrak{C}(q^\bullet) = \inf_{\tilde{q}^\bullet \in \Pi_s(\lambda^\bullet, \mu^\bullet)} \mathfrak{C}(\tilde{q}^\bullet) \ =: \ \mathfrak{c}_\infty.$$

ii) optimal *iff $q^\bullet$ is equivariant and asymptotically optimal.*

Note that the set of optimal semicouplings is convex. This will be useful for the proof of uniqueness.

Remark 2.21. Equivariant semicouplings $q^\bullet$ are invariant. Hence, they are asymptotically optimal iff

$$\mathfrak{C}(q^\bullet) = \mathbb{E}\left[\int_{M \times B_0} c(x,y) q^\bullet(dx,dy)\right] = \mathfrak{c}_\infty.$$

Because of the invariance, the supremum does not play any role. Moreover, for two different fundamental regions B_0 and $\tilde{B}_0$ define

$$f(g,h) \ = \ \mathbb{E}[\mathsf{Cost}(1_{M \times (gB_0 \cap h\tilde{B}_0)} q^\bullet)].$$

Then, for $k \in G$ and equivariant $q^\bullet$ we have $f(g,h) = f(kg,kh)$. Hence, we can apply the mass transport principle to get

$$\mathbb{E}\left[\int_{M\times B_0} c(x,y)q^\bullet(dx,dy)\right] = \sum_{h\in G} f(id,h) = \sum_{g\in G} f(g,id) = \mathbb{E}\left[\int_{M\times \tilde{B}_0} c(x,y)q^\bullet(dx,dy)\right].$$

Thus, the specific choice of fundamental region is not important for the cost functional $\mathfrak{C}(\cdot)$ if we restrict to equivariant semicouplings.

Remark 2.22. The notion of optimality explains why we restrict to stationary probability measures and equivariant random measures. If $\lambda^\bullet$ and $\mu^\bullet$ are just invariant, there does not have to be any invariant semicoupling between them. Indeed, take $\lambda^\bullet$ a Poisson point process of unit intensity in $\mathbb{R}^d$. It can be written as $\mu^\omega = \sum_{\xi\in\Xi(\omega)}\delta_\xi$. Define $\lambda^\omega := \sum_{\xi\in\Xi(\omega)}\delta_{-\xi}$ to be the Poisson process that we get if we reflect the first one at the origin. Then $\lambda^\bullet$ and $\mu^\bullet$ are invariant but not jointly invariant, e.g. consider the set $[0,1)^d \times [-1,0)^d$, and not both of them can be equivariant.
Assume there is an invariant coupling $q^\bullet$ of $\lambda^\bullet$ and $\mu^\bullet$. Then, for any measurable $f: \mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$ we have for any $z \in \mathbb{R}^d$ $\int f(x,y)q^\bullet(dx,dy) \overset{d}{=} \int f(x+z,y+z)q^\bullet(dx,dy)$ if one of the two random variables is finite. Now take any measurable bounded $g: \mathbb{R}^d \to \mathbb{R}_+$ with $\mathrm{supp}(g) \subset [0,1)^d$ and define $h(x) = g(-x)$ such that $\mathrm{supp}(h) \subset [-1,0)^d$. Take $f(x,y) = g(x) + h(y)$. Then, we have

$$\mathbb{E}\left[\left(\int f(x,y)q^\bullet(dx,dy)\right)^2\right] = \mathbb{E}\left[\left(2\int g(x)\mu^\bullet(dx)\right)^2\right] =: 4\mathbb{E}[X^2]$$

for $X = \int g(x)\mu^\bullet(dx)$. On the other hand, writing $\mathbf{1} = (1,\ldots,1)$ this has to equal (by invariance)

$$\mathbb{E}\left[\left(\int f(x+\mathbf{1},y+\mathbf{1})q^\bullet(dx,dy)\right)^2\right] =: \mathbb{E}[(X+Y)^2],$$

with $Y \overset{d}{=} X$ but independent of X as the shifted functions $g(\cdot+\mathbf{1})$ and $h(\cdot+\mathbf{1})$ have support in $A_1 = [-1,0)^d$ and $A_2 = [-2,-1)^d$ respectively and $\mu^\bullet(A_1)$ and $\lambda^\bullet(A_2)$ are independent and identically distributed by definition of the Poisson point process. But this leads to a contradiction by e.g. choosing $g(x) = 1_{[0,1)^d}(x)$.

2.8 An abstract existence result

In this section we want to show that the existence of an optimal coupling given that the mean transportation cost is finite can be shown by an abstract compactness result. A similar reasoning is used to prove Corollary 11 in [Hol09].

Proposition 2.23. *Let $\lambda^\bullet$ and $\mu^\bullet$ be two equivariant random measures on M with intensities 1 and $\beta \le 1$ respectively. Assume that $\inf_{q^\bullet\in\Pi_{is}(\lambda^\bullet,\mu^\bullet)}\mathfrak{C}(q^\bullet) = \mathfrak{c}_{i,\infty} < \infty$, then there exists some equivariant semicoupling $q^\bullet$ between $\lambda^\bullet$ and $\mu^\bullet$ with $\mathfrak{C}(q^\bullet) = \mathfrak{c}_{i,\infty}$.*

Proof. As $\mathfrak{c}_{i,\infty} < \infty$ there is a sequence $q_n^\bullet \in \Pi_{is}(\lambda^\bullet,\mu^\bullet)$ such that $\mathfrak{C}(q_n^\bullet) = c_n \searrow \mathfrak{c}_{i,\infty}$. Moreover, we can assume that the transportation cost is uniformly bounded by $c_n \le 2\mathfrak{c}_{i,\infty} =: c$ for all n. We claim that there is a subsequence $(q_{n_k}^\bullet)_{k\in\mathbb{N}}$ of $(q_n^\bullet)_{n\in\mathbb{N}}$ converging to some $q^\bullet \in \Pi_{is}(\lambda^\bullet,\mu^\bullet)$ with $\mathfrak{C}(q^\bullet) = \mathfrak{c}_{i,\infty}$. We prove this in four steps:

i) The functional $\mathfrak{C}(\cdot)$ is lower semicontinuous:
It is sufficient to prove that the functional $\mathsf{Cost}(\cdot)$ is lower semicontinuous. Let $(\rho_n)_{n\in\mathbb{N}}$ be any sequence of couplings between finite measures converging to some measure ρ in the vague

topology. If $\mathsf{Cost}(\rho_n) = \infty$ for almost all n we are done. Hence, we can assume, that the transportation cost are bounded. Let $(B_0)_r$ denote the r-neighbourhood of B_0. For $k \in \mathbb{R}$ let $\phi_k : M \times M \to [0,1]$ be nice cut off functions with $\phi_k(x,y) = 1$ on $(B_0)_k \times (B_0)_k$ and $\phi_k(x,y) = 0$ if $x \in \complement((B_0)_{k+1})$ or $y \in \complement((B_0)_{k+1})$. Then, we have using continuity of the cost function c(x,y) and by the definition of vague convergence

$$\begin{aligned}
\liminf_{n\to\infty} \mathsf{Cost}(\rho_n) &= \liminf_{n\to\infty} \int_{M\times M} c(x,y)\rho_n(dx,dy) \\
&= \liminf_{n\to\infty} \sup_{k\in\mathbb{N}} \int_{M\times M} \phi_k(x,y)\ c(x,y)\rho_n(dx,dy) \\
&\geq \sup_k \liminf_{n\to\infty} \int_{M\times M} \phi_k(x,y)\ c(x,y)\rho_n(dx,dy) \\
&= \sup_k \int_{M\times M} \phi_k(x,y)\ c(x,y)\rho(dx,dy) = \mathsf{Cost}(\rho).
\end{aligned}$$

Applying this to $1_{M\times B_0} q_n^\bullet$ shows the lower semicontinuity of $\mathfrak{C}(\cdot)$.
ii) The sequence $(q_n^\bullet)_{n\in\mathbb{N}}$ is tight in $\mathcal{M}(M \times M \times \Omega)$:
Put $f \in C_c(M \times M \times \Omega)$. According to Lemma 2.8 we have to show $\sup_{n\in\mathbb{N}} q_n^\bullet\mathbb{P}(f) \leq M_f < \infty$ for some constant M_f. To this end let $A \subset M$ compact be such that $\mathrm{supp}(f) \subset A \times M \times \Omega$. We estimate

$$\begin{aligned}
\int_{M\times M\times\Omega} f(x,y,\omega) q_n^\omega(dx,dy)\mathbb{P}(d\omega) &\leq \|f\|_\infty\ \lambda^\bullet\mathbb{P}(A\times\Omega) \\
&\leq \|f\|_\infty\ m(A) =: M_f.
\end{aligned}$$

Hence, there is some measure $q^\bullet$ and a subsequence $q_{n_k}^\bullet$ with $q_{n_k}^\bullet \to q^\bullet$ in vague topology on $\mathcal{M}(M \times M \times \Omega)$. By lower semicontinuity, we have $\mathfrak{C}(q^\bullet) \leq \liminf \mathfrak{C}(q_{n_k}^\bullet) = \mathfrak{c}_{i,\infty}$. Now we have a candidate. We still need to show that it is admissible.

iii) $q^\bullet$ is equivariant:
Take any continuous compactly supported $f \in C_c(M \times M \times \Omega)$. By definition of vague convergence

$$\int f(x,y,\omega) q_{n_k}^\omega(dx,dy)\mathbb{P}(d\omega) \to \int f(x,y,\omega) q^\omega(dx,dy)\mathbb{P}(d\omega).$$

As all the $q_{n_k}^\bullet$ are equivariant, we have for any $g \in G$

$$\begin{aligned}
\int f(x,y,\omega) q_{n_k}^\omega(dx,dy)\mathbb{P}(\omega) &= \int f(g^{-1}x, g^{-1}y, \theta_g\omega) q_{n_k}^{\theta_g\omega}(dx,dy)\mathbb{P}(d\omega) \\
&\to \int f(g^{-1}x, g^{-1}y, \theta_g\omega) q^{\theta_g\omega}(dx,dy)\mathbb{P}(d\omega).
\end{aligned}$$

Putting this together, we have for any $g \in G$

$$\int f(x,y,\omega) q^\omega(dx,dy)\mathbb{P}(d\omega) = \int f(g^{-1}x, g^{-1}y, \theta_g\omega) q^{\theta_g\omega}(dx,dy)\mathbb{P}(d\omega).$$

Hence, $q^\bullet$ is equivariant.

iv) $q^\bullet$ is a semicoupling of $\lambda^\bullet$ and $\mu^\bullet$:
Fix $h \in C_c(M \times \Omega)$. Put $A \subset M$ compact such that $\mathrm{supp}(h) \subset A \times \Omega$ and $A \in Adm(M)$. Denote the $R-$neighbourhood of A by A_R. By the uniform bound on transportation cost we have

$$q_n^\bullet\mathbb{P}(\complement(A_R), A, \Omega) \leq m(A)\frac{c}{\vartheta(R)}, \tag{2.1}$$

uniformly in n. Let $f_R : M \to [0,1]$ be a continuous compactly supported function such that $f_R(x) = 1$ for $x \in A_R$ and $f_R(x) = 0$ for $x \in \complement A_{R+1}$. As $q_n^\bullet \mathbb{P}$ is a semicoupling of $\lambda^\bullet$ and $\mu^\bullet$ we have due to monotone convergence

$$\begin{aligned}\int_{M\times\Omega} h(y,\omega)\mu^\omega(dy)\mathbb{P}(d\omega) &= \int_{M\times M\times\Omega} h(y,\omega)q_n^\omega(dx,dy)\mathbb{P}(d\omega)\\ &= \lim_{R\to\infty}\int_{M\times M\times\Omega} f_R(x)h(y,\omega)q_n^\omega(dx,dy)\mathbb{P}(d\omega).\end{aligned}$$

Because of the uniform bound (2.1) we have

$$\left|\int_{M\times\Omega} h(x,\omega)\mu^\omega(dx)\mathbb{P}(d\omega) - \int_{M\times M\times\Omega} f_R(x)h(y,\omega)q_{n_k}^\omega(dx,dy)\mathbb{P}(d\omega)\right| \le m(A)\frac{c\cdot\|h\|_\infty}{\vartheta(R)}.$$

Taking first the limit of $n_k \to \infty$ and then the limit of $R \to \infty$ we conclude using vague convergence and monotone convergence that

$$\begin{aligned}0 &= \lim_{R\to\infty}\lim_{k\to\infty}\left|\int_{M\times\Omega} h(y,\omega)\mu^\omega(dy)\mathbb{P}(d\omega) - \int_{M\times M\times\Omega} f_R(x)h(y,\omega)q_{n_k}^\omega(dx,dy)\mathbb{P}(d\omega)\right|\\ &= \lim_{R\to\infty}\left|\int_{M\times\Omega} h(y,\omega)\mu^\omega(dy)\mathbb{P}(d\omega) - \int_{M\times M\times\Omega} f_R(x)h(y,\omega)q^\omega(dx,dy)\mathbb{P}(d\omega)\right|\\ &= \left|\int_{M\times\Omega} h(y,\omega)\mu^\omega(dy)\mathbb{P}(d\omega) - \int_{M\times M\times\Omega} h(y,\omega)q^\omega(dx,dy)\mathbb{P}(d\omega)\right|\end{aligned}$$

This shows that the second marginal equals $\mu^\bullet$. For the first marginal we have for any $k \in C_c(M\times\Omega)$

$$\int_{M\times\Omega} k(x,\omega)q_{n_k}^\omega(dx,dy)\mathbb{P}(d\omega) \le \int_{M\times\Omega} k(x,\omega)\lambda^\omega(dx)\mathbb{P}(d\omega).$$

In particular, using the function f_R from above we have,

$$\int_{M\times\Omega} f_R(y)\; k(x,\omega)q_{n_k}^\omega(dx,dy)\mathbb{P}(d\omega) \le \int_{M\times\Omega} k(x,\omega)\lambda^\omega(dx)\mathbb{P}(d\omega).$$

Taking the limit $n_k \to \infty$ yields by vague convergence

$$\int_{M\times\Omega} f_R(y)\; k(x,\omega)q^\omega(dx,dy)\mathbb{P}(d\omega) \le \int_{M\times\Omega} k(x,\omega)\lambda^\omega(dx)\mathbb{P}(d\omega).$$

Finally taking the supremum over R shows that $q^\bullet$ is indeed a semicoupling of $\lambda^\bullet$ and $\mu^\bullet$. $\square$

Remark 2.24. i) This coupling need not be a factor coupling. We do not know if it is in general true or not that $\mathfrak{c}_\infty = \mathfrak{c}_{i,\infty}$, that is, if minimizing the functional $\mathfrak{C}(\cdot)$ over all semicouplings is the same as minimizing over all equivariant semicouplings. However, in the case that the balls $\Lambda_r \subset G$ are Følner sets, we can show equality (see Corollary 5.5 and Remark 5.6).

ii) The same proof shows the existence of optimal semicouplings between $\lambda^\bullet$ and $\mu^\bullet$ with intensities 1 and $\beta \ge 1$ respectively. In this case the "semi" is on the side of $\mu^\bullet$ (see also chapter 6).

Lemma 2.25. *Let $q^\bullet$ be an invariant semicoupling of two random measures $\lambda^\bullet$ and $\mu^\bullet$ with intensities 1 and $\beta \le 1$ respectively. Then, $q^\bullet$ is a coupling iff $\beta = 1$.*

Proof. This is another application of the mass transport principle. Let B_0 be a fundamental region and define $f(g,h) = \mathbb{E}[q^\bullet(gB_0, hB_0)]$. By invariance of $q^\bullet$, we have $f(g,h) = f(kg,kh)$ for any $k \in G$. Hence, we get

$$1 = \mathbb{E}[\lambda^\bullet(B_0)] \geq \mathbb{E}[q^\bullet(B_0, M)] = \sum_{g\in G} f(id, g) = \sum_{h \in G} f(h, id) = \mathbb{E}[q^\bullet(M, B_0)] = \beta.$$

We have equality iff $\beta = 1$. By definition of semicoupling, we also have $q^\omega(A, M) \leq \lambda^\omega(A)$ for any $A \subset M$. Hence, in the case of equality we must have $q^\omega(A, M) = \lambda^\omega(A)$ for $\mathbb{P}-$almost all ω. □

Remark 2.26. The above remark applies again. Considering the case of intensity $\beta \geq 1$ gives that $q^\bullet$ is a coupling iff $\beta = 1$.

2.9 Assumptions

Let us summarize the setting and assumptions we work with in the rest of the thesis.

- M will be a smooth connected non-compact Riemannian manifold with Riemannian volume measure m, such that there is a group G of isometries acting properly discontinuous, cocompactly and freely on M.
- B_0 will denote the chosen fundamental region.
- $c(x,y) = \vartheta(d(x,y))$ with $\vartheta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\vartheta(0) = 0$ and $\lim_{r\to\infty} \vartheta(r) = \infty$. Given two compactly supported probability measures on M, $\lambda \ll m$ and μ arbitrary, we will assume that the optimal transportation problem admits a unique solution which is induced by a measurable map T, i.e. $q = (id, T)_*\lambda$.
- $(\Omega, \mathfrak{A}, \mathbb{P})$ will be a probability space admitting a measurable flow $(\theta_g)_{g\in G}$. $\mathbb{P}$ is assumed to be stationary and Ω is assumed to be a compact metric space.
- $\lambda^\bullet$ and $\mu^\bullet$ will be equivariant measure of intensities one respectively $\beta \in (0, \infty)$. Moreover, we assume that $\lambda^\bullet$ is absolutely continuous.

Chapter 3

Semicouplings on bounded sets

The goal of this chapter is to prove Theorem 2.15, the crucial existence and uniqueness result for optimal semicouplings between $\lambda^\bullet$ and $\mu^\bullet$ restricted to a bounded set. The strategy will be to first prove existence and uniqueness of optimal semicouplings $q = q^\omega$ for deterministic measures $\lambda = \lambda^\omega$ and $\mu = \mu^\omega$. Secondly, we will show that the map $\omega \mapsto q^\omega$ is measurable, which will allow us to deduce Theorem 2.15.

The theory of optimal semicouplings is a concept of independent interest. Optimal semicouplings are solutions of a twofold optimization problem: the optimal choice of a density $\rho \leq 1$ of the first marginal λ and subsequently the optimal choice of a coupling between $\rho\lambda$ and μ. This twofold optimization problem can also be interpreted as a transport problem with free boundary values. Throughout this chapter, we fix the cost function $c(x, y) = \vartheta(d(x, y))$ with ϑ – as before – being a strictly increasing, continuous function from $\mathbb{R}_+$ to $\mathbb{R}_+$ with $\vartheta(0) = 0$ and $\lim\limits_{r\to\infty} \vartheta(r) = \infty$. As already mentioned, we additionally assume that the optimal transportation problem between two compactly supported probability measures λ and μ such that $\lambda \ll m$ has a unique solution given by a transportation map, e.g. the optimal coupling is given by $q = (id, T)_*\lambda$. There are very general results on the uniqueness of the solution to the Monge problem for which we refer to chapters 9 and 10 of [Vil09]. To be more concrete we state a uniqueness result for compact manifolds due to McCann [McC01] and will prove an uniqueness result in the simple but for us very interesting case that the measure μ is discrete.

Theorem 3.1 (McCann)**.** *Let N be a compact manifold, $\lambda \ll m$ and μ be probability measures and $c(x, y) = \int_0^{d(x,y)} \tau(s)ds$ with $\tau : \mathbb{R}_+ \to \mathbb{R}$ continuously increasing and $\tau(0) = 0$. Then, there is a measurable map $T : M \to M \cup \{\eth\}$ such that the unique optimal coupling between λ and μ is given by $q = (id, T)_*\lambda$.*

The "cemetery" $\eth$ in the statement is not really important. This is the place where all points outside of the support of λ are sent. We just include it to make some notations easier.
If we assume μ to be discrete, we can actually take ϑ to be any continuous strictly increasing function.

Lemma 3.2. *Given a finite set $\Xi = \{\xi_1, \ldots, \xi_k\} \subset M$, positive numbers $(a_i)_{1\leq i\leq k}$ summing to one and a probability density $\rho \in L^1(M, m)$. Consider the cost function $c(x, y) = \vartheta(d(x, y))$ for some continuous strictly increasing function $\vartheta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\vartheta(0) = 0$ and $\lim_{r\to\infty} \vartheta(r) = \infty$. If $dim(M) = 1$ we exclude the case $\vartheta(r) = r$.*
i) There exists a unique coupling q of $\rho \cdot m$ and $\sigma = \sum_{i=1}^k a_i\delta_{\xi_i}$ which minimizes the cost functional $\mathsf{Cost}(\cdot)$.
ii) There exists a (m-a.e. unique) map $T : \{\rho > 0\} \to \Xi$ with $T_*(\rho \cdot m) = \sigma$ which minimizes $\int c(x, T(x))\rho(x)\, m(dx)$.
iii) There exists a (m-a.e. unique) map $T : \{\rho > 0\} \to \Xi$ with $T_*(\rho\cdot m) = \sigma$ which is c-monotone (in the sense that the closure of $\{(x, T(x)) : \ \rho(x) > 0\}$ is a c-cyclically monotone set).

iv) The minimizers in (i), (ii) and (iii) are related by $q = (Id, T)_*(\rho \cdot m)$ *or, in other words,*

$$q(dx, dy) \;=\; \delta_{T(x)}(dy)\, \rho(x)\, m(dx).$$

Proof. We prove the lemma in three steps.
a) By compactness of $\Pi(\rho \cdot m, \sigma)$ w.r.t. weak convergence and continuity of $c(\cdot, \cdot)$ there is a coupling q minimizing the cost function $\mathsf{Cost}(\cdot)$ (see also [Vil09], Theorem 4.1).
b) Write $\rho \cdot m =: \lambda = \sum_{i=1}^{k} \lambda_i$ where $\lambda_i(.) := q(. \times \{\xi_i\})$ for each $i = 1, \dots, k$. We claim that the measures $(\lambda_i)_i$ are mutually singular. Assuming that there is a Borel set N such that for some $i \neq j$ we have $\lambda_i(N) = \alpha > 0$ and $\lambda_j(N) = \beta > 0$ we will redistribute the mass on N being transported to ξ_i and ξ_j in a cheaper way. This will show that the measures $(\lambda_i)_i$ are mutually singular. In particular, the proof implies the existence of a measurable c-monotone map T such that $q = (Id, T)_*(\rho \cdot m)$.
W.l.o.g. we may assume that $(\rho \cdot m)(N) = \alpha + \beta$. Otherwise write $\rho = \rho_1 + \rho_2$ such that on N $\lambda_i(dx) + \lambda_j(dx) = (\rho_1 m)(dx)$ and just work with the density ρ_1.
Put $f(x) := c(x, \xi_i) - c(x, \xi_j)$. As $c(\cdot, \cdot)$ is continuous, f is continuous. The function $c(x, y)$ is a strictly increasing function of the distance $d(x, y)$. Thus, the level sets $\{f \equiv b\}$ define (locally) $(d - 1)$ dimensional submanifolds (e.g. use implicit function theorem for non smooth functions, see Corollary 10.52 in [Vil09]) changing continuously with b. Choose b_0 such that $\rho \cdot m(\{f < b_0\} \cap N) = \alpha$ (which implies $\rho \cdot m(\{f > b_0\} \cap N) = \beta$) and set $N_i := \{f < b_0\} \cap N$ and $N_j := \{f \geq b_0\} \cap N$.
For $l = i, j$

$$\tilde{\lambda}_l(dx) := \lambda_l(dx) - 1_N(x)\lambda_l(dx) + 1_{N_l}(x)(\rho \cdot m)(dx).$$

For $l \neq i, j$ set $\tilde{\lambda}_l = \lambda_l$. By construction, $\tilde{q} = \sum_{l=1}^{k} \tilde{\lambda}_l \otimes \delta_{\xi_l}$ is a coupling of $\rho \cdot m$ and σ. Moreover, $\tilde{q}$ is c-cyclically monotone on N, that is $\forall x_i \in N_i, x_j \in N_j$ we have

$$c(x_i, \xi_i) + c(x_j, \xi_j) \leq c(x_j, \xi_i) + c(x_i, \xi_j).$$

Furthermore, the set where equality holds is a null set because $c(x, y)$ is a strictly increasing function of the distance. Then, we have

$$\begin{aligned} &\mathsf{Cost}(q) - \mathsf{Cost}(\tilde{q}) \\ = &\int_N c(x, \xi_i)\lambda_i(dx) + c(x, \xi_j)\lambda_j(dx) - \int_{N_i} c(x, \xi_i)\tilde{\lambda}_i(dx) - \int_{N_j} c(x, \xi_j)\tilde{\lambda}_j(dx) > 0, \end{aligned}$$

by cyclical monotonicity. This proves that λ_i and λ_j are singular to each other.
Hence, the family $(\lambda_i)_{i=1,\dots,k}$ is mutually singular which in turn implies that there exist Borel sets $S_i \subset M$ with $\dot{\bigcup}_i S_i = M$ and $\lambda_i(S_j) = 0$ for all $i \neq j$. Define the map $T : M \to \Xi$ by $T(x) := \xi_i$ for all $x \in S_i$. Then $q = (Id, T)_*(\rho \cdot m)$.
c) Assume there are two minimizers of the cost function Cost, say q_1 and q_2. Then $q_3 := \frac{1}{2}(q_1 + q_2)$ is a minimizer as well. By step (b) we have $q_i = (Id, T_i)_*\rho \cdot m$ for $i = 1, 2, 3$. This implies

$$\begin{aligned} \delta_{T_3(x)}(dy)\, \rho \cdot m(dx) \;&=\; q_3(dx, dy) \;=\; \left(\frac{1}{2} q_1(dx, dy) + \frac{1}{2} q_2(dx, dy)\right) \\ &=\; \left(\frac{1}{2}\delta_{T_1(x)}(dy) + \frac{1}{2}\delta_{T_2(x)}(dy)\right)\, \rho \cdot m(dx) \end{aligned}$$

This, however, implies $T_1(x) = T_2(x)$ for $\rho \cdot m$ a.e. $x \in M$ and thus $q_1 = q_2$. □

Remark 3.3. In the case that $dim(M) = 1$ and cost function $c(x, y) = d(x, y)$ the optimal coupling between an absolutely continuous measure and a discrete measure need not be unique. In higher dimensions this is the case, as we get strict inequalities in the triangle inequalities. A counterexample for one dimension is the following. Take λ to be the Lebesgue measure on $[0,1]$ and put $\mu = \frac{1}{3}\delta_0 + \frac{2}{3}\delta_{1/16}$. Then, for any $a \in [1/16, 1/3]$

$$q_a(dx,dy) \;=\; 1_{[0,a)}(x)\delta_0(dy)\lambda(dx) + 1_{[a,2/3+a)}(x)\delta_{1/16}(dy)\lambda(dx) + 1_{[a+2/3,1]}(x)\delta_0(dy)\lambda(dx)$$

is an optimal coupling of λ and μ with $\mathsf{Cost}(q_a) = 11/24$.

Remark 3.4. In the case of cost function $c(x,y) = \frac{1}{p}d^p(x,y)$ the optimal transportation map is given by

$$T(x) \;=\; \exp\left(d(x,\xi_j)\frac{\nabla\Phi_j(x)}{|\nabla\Phi_j(x)|}\right)$$

for functions $\Phi_i(z) = -\frac{1}{p}d^p(z,\xi_i) + b_i$ with constants b_i and j such that $\Phi_j(x) = \max_{1\le i\le k}\Phi_i(x)$ (e.g. see [McC01]).

Given two deterministic measures $\lambda = f\cdot m$ for some compactly supported density f (in particular $\lambda \ll m$) and an arbitrary finite measure μ with $\mathrm{supp}(\mu) \subset A$ for some compact set A such that $\mu(M) \le \lambda(M) < \infty$. We are looking for minimizers of

$$\mathsf{Cost}(q) = \int c(x,y)q(dx,dy)$$

under all semicouplings q of λ and μ. The key step is a nice observation by Figalli, namely Proposition 2.4 in [Fig10]. The version we state here is adapted to our setting.

Proposition 3.5 (Figalli)**.** *Let q be a Cost minimizing semicoupling between λ and μ. Write $f_q\cdot m = (\pi_1)_*q$. Consider the Monge-Kantorovich problem:*

$$\textit{minimize } C(\gamma) \;=\; \int_{M\times M} c(x,y)\gamma(dx,dy)$$

among all γ which have λ and $\mu + (f - f_q)\cdot m$ as first and second marginals, respectively. Then, the unique minimizer is given by

$$q + (id\times id)_*(f - f_q)\cdot m.$$

This allows us to show that all minimizers of Cost are concentrated on the same graph which also gives us uniqueness:

Proposition 3.6. *There is a unique Cost minimizing semicoupling between λ and μ. It is given as $q = (id,T)_*(\rho\cdot\lambda)$ for some measurable map $T: M\to M\cup\{\eth\}$ and density ρ.*

Proof. (i) The functional $\mathsf{Cost}(\cdot)$ is lower semicontinuous on $\mathcal{M}(M\times M)$ wrt weak convergence of measures. Indeed, take a sequence of measures $(q_n)_{n\in\mathbb{N}}$ converging weakly to some q. Then we have by continuity of the cost function $c(\cdot,\cdot)$:

$$\begin{aligned}\int c(x,y)\; q(dx,dy) \;&=\; \sup_{k\in\mathbb{N}}\int c(x,y)\wedge k\; q(dx,dy)\\ &=\; \sup_{k\in\mathbb{N}}\lim_{n\to\infty}\int c(x,y)\wedge k\; q_n(dx,dy)\\ &\le\; \liminf_{n\to\infty}\int c(x,y)\; q_n(dx,dy).\end{aligned}$$

(ii) Let $\mathcal{O}$ denote the set of all semicouplings of λ and μ and $\mathcal{O}_1$ denote the set of all semicouplings q satisfying $\mathsf{Cost}(q) \le 2\inf_{q\in\mathcal{O}}\mathsf{Cost}(q) =: 2c$. Then $\mathcal{O}_1$ is relatively compact wrt weak topology. Indeed, $q(M\times \complement A) = 0$ for all $q\in\mathcal{O}_1$ and

$$q(\complement(A_r)\times A) \le \frac{1}{\vartheta\,(r)}\cdot\mathsf{Cost}(q) \le \frac{2}{\vartheta\,(r)}c$$

for each $r > 0$ where A_r denotes the closed r-neighborhood of A in M. Thus, for any $\epsilon > 0$ there exists a compact set $K = A_r \times A$ in $M \times M$ such that $q(\complement K) \leq \epsilon$ uniformly in $q \in \mathcal{O}_1$.
(iii) The set $\mathcal{O}$ is closed wrt weak topology. Indeed, if $q_n \to q$ then $(\pi_1)_* q_n \to (\pi_1)_* q$ and $(\pi_2)_* q_n \to (\pi_2)_* q$. Hence $\mathcal{O}_1$ is compact and Cost attains its minimum on $\mathcal{O}$. Let q denote one such minimizer. Its first marginal is absolutely continuous to m. By Theorem 3.1 and Lemma 3.2, there is a measurable map $T : M \to M \cup \{\eth\}$ and densities $\tilde{f}_q, f_q$ such that $q = (id, T)_*(\tilde{f}_q \cdot \lambda) = (id, T)_*(f_q \cdot m)$.
(iv) Given a minimizer of Cost, say q. By Proposition 3.5, $\tilde{q} := q + (id, id)_*(f - f_q) \cdot m$ solves

$$\min C(\gamma) = \int c(x, y)\gamma(dx, dy)$$

under all γ which have λ and $\mu + (f - f_q)m$ as first respectively second marginals, where $f_q \cdot m = (\pi_1)_* q$ as above. By Theorem 3.1 and Lemma 3.2, there is a measurable map S such that $\tilde{q} = (id, S)_* \lambda$. That is, $\tilde{q}$ and in particular q are concentrated on the graph of S. By definition $\tilde{q} = q + (id, id)_*(f - f_q) \cdot m$ and, therefore, we must have $S(x) = x$ on $\{f > f_q\}$.
(v) This finally allows us to deduce uniqueness. By the previous step, we know that any convex combination of optimal semicouplings is concentrated on a graph. This implies that all optimal semicouplings are concentrated on the *same* graph. Moreover, Proposition 3.5 implies that if we do not transport all the λ mass in one point we leave it where it is. Hence, all optimal semicouplings choose the same density ρ of λ and therefore coincide.
Assume there are two optimal semicouplings q_1 and q_2. Then $q_3 := \frac{1}{2}(q_1 + q_2)$ is optimal as well. By the previous step for any $i \in \{1, 2, 3\}$, we get maps S_i such that q_i is concentrated on the graph of S_i. Moreover, we have $S_3(x) = x$ on the set $\{f > f_{q_3}\} = \{f > f_{q_1}\} \cup \{f > f_{q_2}\}$, where again $f_{q_i} \cdot m = (\pi_1)_* q_i$. As q_3 is concentrated on the graph of S_3, q_1 and q_2 must be concentrated on the same graph. Hence, we have $S_3 = S_i$ on $\{f_{q_i} > 0\}$ for $i = 1, 2$. We also know from the previous step that $S_i(x) = x$ on $\{f > f_{q_i}\} \subset \{f > f_{q_3}\}$. This gives, that $S_3 = S_1 = S_2$ on $\{f > 0\}$.
We still need to show that $\{f_{q_1} > 0\} = \{f_{q_2} > 0\}$. Put $A_1 := \{f_{q_1} > f_{q_2}\}$ and $A_2 := \{f_{q_2} > f_{q_1}\}$ and assume $m(A_1) > 0$. As $A_1 \subset \{f > f_{q_2}\}$ we know that $S_3(x) = x$ on A_1 and similarly $S_3(x) = x$ on A_2. Now consider

$$A \ := \ S_3^{-1}(A_1) \ = \ (A \cap \{f_{q_1} = f_{q_2}\}) \cup (A \cap A_1) \cup (A \cap A_2).$$

As $S_3(A_2) \subset A_2$ and $A_1 \cap A_2 = \emptyset$ we have $A \cap A_2 = \emptyset$. Therefore, we can conclude

$$\begin{aligned}
\mu(A_1) \ &= \ (S_3)_* f_{q_1} m(A_1) \ = \ f_{q_1} m(A_1) + f_{q_1} m(A \cap \{f_{q_1} = f_{q_2}\}) \\
&> \ f_{q_2} m(A_1) + f_{q_2} m(A \cap \{f_{q_1} = f_{q_2}\}) \\
&= \ (S_3)_* f_{q_2} m(A_1) \ = \ \mu(A_1),
\end{aligned}$$

which is a contradiction, proving $q_1 = q_2$. □

Remark 3.7. Let $q = (id, T)_*(\rho\lambda)$ be the optimal semicoupling of λ and μ. If μ happens to be discrete, we have $\rho(x) \in \{0, 1\}$ m almost everywhere. Indeed, assume the contrary. Then, there is $\xi \in \operatorname{supp}(\mu)$ such that on $U := T^{-1}(\xi)$ we have $\rho \in (0, 1)$ on some set of positive λ measure. Let R be such that $\lambda(U \cap B(\xi, R)) = \mu(\{\xi\})$, where $B(\xi, R)$ denotes the ball of radius R around ξ. Put $V = U \cap B(\xi, R)$ and

$$\tilde{q}(dx, dy) \ = \ q(dx, dy) - 1_U(x)\rho(x)\delta_\xi(dy)\lambda(dx) + 1_V(x)\delta_\xi(dy)\lambda(dx).$$

This means, we take the same transportation map, but use the λ mass more efficiently. $\tilde{q}$ leaves some λ mass far out and instead uses the same amount of λ mass which is closer to the target ξ. By construction, we have $\mathsf{Cost}(q) > \mathsf{Cost}(\tilde{q})$ contradicting optimality of q.

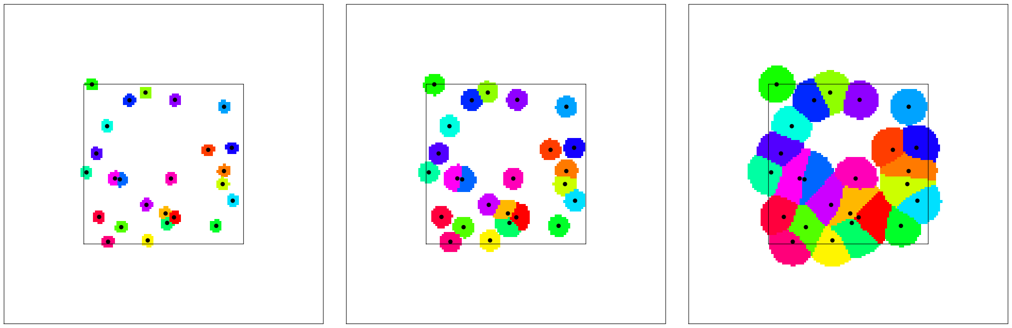

Figure 3.1: Optimal semicoupling of Lebesgue and 25 points in the cube with $c(x,y) = |x-y|$ where each point gets mass $1/9, 1/3, 1$ respectively.

We showed the existence and uniqueness of optimal semicouplings between deterministic measures. The next step in the proof of Theorem 2.15 is to show the measurability of the mapping $\omega \mapsto \Phi(\lambda^\omega, 1_A\mu^\omega) = q_A^\omega$ the unique optimal semicoupling between λ^ω and $1_A\mu^\omega$. The mapping $\omega \mapsto (\lambda^\omega, 1_A\mu^\omega)$ is measurable by definition. Hence, we have to show that $(\lambda^\omega, 1_A\mu^\omega) \mapsto \Phi(\lambda^\omega, 1_A\mu^\omega)$ is measurable. We will show a bit more, namely that this mapping is actually continuous. We start with a simple but important observation about optimal semicouplings.

Denote the one-point compactification of M by $M \cup \{\eth\}$ and let $\tilde{\vartheta}(r)$ be such that it is equal to $\vartheta(r)$ on a very large box, say $[0,K]$ and then tends continuously to zero such that $\tilde{c}(x,\eth) = \tilde{\vartheta}(d(x,\eth)) = \lim_{r\to\infty} \tilde{\vartheta}(r) = 0$ for any $x \in M$. By a slight abuse of notation, we also write $\eth : M \to \{\eth\}$ for the map $x \mapsto \eth$.

Lemma 3.8. *Let two measures λ and μ on M be given such that $\infty > \lambda(M) = N \geq \mu(M) = \alpha$ and assume there is a ball $B(x, K/2)$ such that supp(λ), supp$(\mu) \subset B(x, K/2)$. Then, q is an optimal semicoupling between λ and μ wrt to the cost function $c(\cdot,\cdot)$ iff $\tilde{q} = q + (id, \eth)_*(1-f_q)\cdot\lambda$ is an optimal* coupling *between λ and $\tilde{\mu} = \mu + (N-\alpha)\delta_\eth$ wrt the cost function $\tilde{c}(\cdot,\cdot)$, where $(\pi_1)_*q = f_q\lambda$.*

Proof. Let q be any semicoupling between λ and μ. Then $\tilde{q} = q + (id,\eth)_*(1-f_q)\cdot\lambda$ defines a coupling between λ and $\tilde{\mu}$. Moreover, the transportation cost of the semicoupling and the one of the coupling are exactly the same, that is $\mathsf{Cost}(q) = \mathsf{Cost}(\tilde{q})$. Hence, q is optimal iff $\tilde{q}$ is optimal. □

This allows to deduce the continuity of Φ from the classical theory of optimal transportation.

Lemma 3.9. *Given a sequence of measures $(\lambda_n)_{n\in\mathbb{N}}$ converging vaguely to some λ, all absolutely continuous to m with $\lambda_n(M) = \lambda(M) = \infty$. Moreover, let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of finite measures converging weakly to some finite measure μ, all concentrated on some bounded set $A \subset M$. Let q_n be the optimal semicoupling between λ_n and μ_n and q be the optimal semicoupling between λ and μ. Then, q_n converges weakly to q. In particular, the map $(\lambda,\mu) \mapsto \Phi(\lambda,\mu) = q$ is continuous.*

Proof. i) As $(\mu_n)_{n\in\mathbb{N}}$ converge to μ and μ is finite, we can assume that $\sup_n \mu_n(M), \mu(M) \leq \alpha < \infty$. As λ_n and λ have infinite mass for any $x \in M$ and $k \in \mathbb{R}$ there is a radius $R(x,k) < \infty$ such that $\lambda(B(x, R(x,k))) \geq k$, where $B(x,R)$ denotes the closed ball around x of radius R. Fix an arbitrary $x \in A$ and set $R_1 = R(x,\alpha) + \mathrm{diam}(A)$ and $R_2 = R(x, 2\alpha) + \mathrm{diam}(A)$. Because $(\lambda_n)_{n\in\mathbb{N}}$ converge to λ we can assume that for any n $\lambda_n(B(x,R_2)) \geq \lambda(B(x,R_1) = N > \alpha$.

ii) Optimality of q_n and q implies that $\mathrm{supp}(q_n) \subset B(x,R_2) \times A$ and $\mathrm{supp}(q) \subset B(x,R_1) \times A$. Because otherwise there is still some mass lying closer to the target than the mass which is transported into the target. For any n let $r_n \leq R_2$ be such that $\lambda_n(B(x,r_n)) = N$. Such choices exist as $\lambda_n \ll m$ for all n. Then, we even know that $\mathrm{supp}(q_n) \subset B(x, r_n) \times A$. Set $\tilde{\lambda}_n = 1_{B(x,r_n)}\lambda_n$ and $\tilde{\lambda} = 1_{B(x,R_1)}\lambda$. Then the optimal semicoupling between λ_n and μ_n is the same as the optimal semicoupling between $\tilde{\lambda}_n$ and μ_n and similarly the optimal semicoupling between λ and μ is the same as the optimal semicoupling between $\tilde{\lambda}$ and μ. Moreover, because for any n $\tilde{\lambda}_n$ is compactly supported with total mass N the vague convergence $\lambda_n \to \lambda$ implies weak convergence of $\tilde{\lambda}_n \to \tilde{\lambda}$.

iii) Now we are in a setting where we can apply the previous Lemma. Set $K = 2R_2$ and define $\tilde{\vartheta}, \tilde{\mu}_n, \tilde{\mu}$ as above. Then $\tilde{q}_n$ and $\tilde{q}$ are optimal couplings between $\tilde{\lambda}_n$ and $\tilde{\mu}_n$ and $\tilde{\lambda}$ and $\tilde{\mu}$ respectively wrt to the cost function $\tilde{c}(\cdot,\cdot)$. The cost function $\tilde{c}$ is continuous and M and $M \cup \{\eth\}$ are Polish spaces. Hence, we can apply the stability result of the classical optimal transportation theory (e.g. Theorem 5.20 in [Vil09] or Proposition 1.3) to conclude that $\tilde{q}_n \to \tilde{q}$ weakly and therefore $q_n \to q$ weakly. $\square$

Take a pair of equivariant random measure $(\lambda^\bullet, \mu^\bullet)$ with $\lambda^\omega \ll m$ as usual. For a given $\omega \in \Omega$ we want to apply the results of the previous Lemma to a fixed realization $(\lambda^\omega, \mu^\omega)$. Then, for any bounded Borel set $A \subset M$, there is a unique optimal semicoupling q_A^ω between λ^ω and $1_A\mu^\omega$, that is, a unique minimizer of the cost function Cost among all semicouplings of λ^ω and $1_A\mu^\omega$.

Lemma 3.10. *For each bounded Borel set $A \subset M$ the map $\omega \mapsto q_A^\omega$ is measurable.*

Proof. We saw that the map $\Phi : (\lambda^\omega, 1_A\mu^\omega) = q_A^\omega$ is continuous. By definition of random measures the map $\omega \mapsto (\lambda^\omega, 1_A\mu^\omega)$ is measurable. Hence, the map

$$\omega \mapsto \Phi(\lambda^\omega, 1_A\mu^\omega) = q_A^\omega$$

is measurable. $\square$

The uniqueness and measurably of q_A^ω allows us to finally deduce

Theorem 3.11. *(i) For each bounded Borel set $A \subset M$ there exists a unique semicoupling Q_A of $\lambda^\bullet\mathbb{P}$ and $(1_A\mu^\bullet)\mathbb{P}$ which minimizes the mean cost functional $\mathfrak{Cost}(.)$.*
(ii) The measure Q_A can be disintegrated as $Q_A(dx, dy, d\omega) := q_A^\omega(dx, dy)\,\mathbb{P}(d\omega)$ where for $\mathbb{P}$-a.e. ω the measure q_A^ω is the unique minimizer of the cost functional $\mathsf{Cost}(.)$ among the semicouplings of λ^ω and $1_A\mu^\omega$.
(iii) $\mathfrak{Cost}(Q_A) = \int_\Omega \mathsf{Cost}(q_A^\omega)\,\mathbb{P}(d\omega)$.

Proof. The existence of a minimizer is proven along the same lines as in the previous proposition: We choose an approximating sequences Q_n in $\mathcal{M}(M \times M \times \Omega)$ – instead of a sequence q_n in $\mathcal{M}(M \times M)$ – minimizing the lower semicontinuous functional $\mathfrak{Cost}(\cdot)$. Existence of a limit follows as before from tightness of the set of all semicouplings Q with $\mathfrak{Cost}(Q) \le 2\inf_{\tilde{Q}} \mathfrak{Cost}(\tilde{Q})$. For each semicoupling Q of $\lambda^\bullet$ and $1_A\mu^\bullet$ with disintegration as $q^\bullet\mathbb{P}$ we obviously have

$$\mathfrak{Cost}(Q) = \int_\Omega \mathsf{Cost}(q^\omega)\,d\mathbb{P}(\omega).$$

Hence, Q is a minimizer of the functional $\mathfrak{Cost}(\cdot)$ (among all semicouplings of $\lambda^\bullet$ and $1_A\mu^\bullet$) if and only if for $\mathbb{P}$-a.e. $\omega \in \Omega$ the measure q^ω is a minimizer of the functional $\mathsf{Cost}(.)$ (among all semicouplings of λ^ω and $1_A\mu^\omega$).
Uniqueness of the minimizer of $\mathsf{Cost}(\cdot)$ therefore implies uniqueness of the minimizer of $\mathfrak{Cost}(\cdot)$. $\square$

Corollary 3.12. *The optimal semicouplings $Q_A = q_A^\bullet\mathbb{P}$ are equivariant in the sense that*

$$Q_{gA}(gC, gD, \theta_g\omega) = Q_A(C, D, \omega),$$

for any $g \in G$ and $C, D \in \mathcal{B}(M)$.

Proof. This is a consequence of the equivariance of $\lambda^\bullet$ and $\mu^\bullet$ and the fact that q_A^ω is a deterministic function of λ^ω and $1_A\mu^\omega$. $\square$

Chapter 4

Uniqueness

The aim of this chapter is to prove Theorem 1.4, the uniqueness of optimal semicouplings. Moreover, the representation of optimal semicouplings, that we get as a byproduct of the uniqueness statement, allows to draw several conclusions about the geometry of the cells of the induced allocations.
Throughout this chapter we fix two equivariant random measures $\lambda^\bullet$ and $\mu^\bullet$ of unit resp. subunit intensities on M with finite mean transportation cost $\mathfrak{c}_\infty$. Moreover, we assume that $\lambda^\bullet$ is absolutely continuous. The whole chapter follows closely the respective section in [HS10].

Proposition 4.1. *Given a semicoupling q^ω of λ^ω and μ^ω for fixed $\omega \in \Omega$, then the following properties are equivalent.*

(i) For each bounded Borel set $A \subset M$, the measure $1_{M\times A}q^\omega$ is the unique optimal coupling of the measures $\lambda^\omega_A(\cdot) := q^\omega(\cdot, A)$ and $1_A\mu^\omega$.

(ii) The support of q^ω is c-cyclically monotone, more precisely,

$$\sum_{i=1}^{N} c(x_i, y_i) \leq \sum_{i=1}^{N} c(x_i, y_{i+1})$$

for any $N \in \mathbb{N}$ and any choice of points $(x_1, y_1), \dots, (x_N, y_N)$ in $\operatorname{supp}(q^\omega)$ with the convention $y_{N+1} = y_1$.

(iii) There exists a nonnegative density ρ^ω and a c-cyclically monotone map $T^\omega : \{\rho^\omega > 0\} \to M$ such that

$$q^\omega = (Id, T^\omega)_* (\rho^\omega \lambda^\omega). \tag{4.1}$$

Recall that, by definition, a map T is c-cyclically monotone iff the closure of its graph $\{(x, T(x)) : x \in \{\rho^\omega > 0\}\}$ is a c-cyclically monotone set.

Proof. $(iii) \Rightarrow (ii) \Rightarrow (i)$ follows from Theorem 3.1 and Lemma 3.2.

$(i) \Rightarrow (iii)$: Take a nested sequence of convex sets $(K_n)_n$ such that $K_n \nearrow M$. By assumption $1_{M\times K_n}q^\omega$ is the unique optimal coupling between $\lambda^\omega_{K_n} \ll m$ and $1_{K_n}\mu^\omega$. By Proposition 3.6 or Theorem 3.1, there exists a density ρ^ω_n and a map $T^\omega_n : \{\rho^\omega_n > 0\} \to M$ such that $1_{M\times K_n}q^\omega = (id, T^\omega_n)_*(\rho^\omega_n \lambda^\omega_{K_n})$. Set $A^\omega_n = \{\rho^\omega_n > 0\}$. As $K_n \subset K_{n+1}$ we have $A^\omega_n \subset A^\omega_{n+1}$. Subtransports of optimal transports are optimal again. Therefore, we have

$$T^\omega_{n+1} = T^\omega_n \quad \text{on } A^\omega_n$$

implying $\rho^\omega_{n+1} = \rho^\omega_n$ on A^ω_n. Hence, the limits

$$T^\omega = \lim_n T^\omega_n, \; A^\omega = \lim_n A^\omega_n \;\text{ and }\; \rho^\omega = \lim_n \rho^\omega_n$$

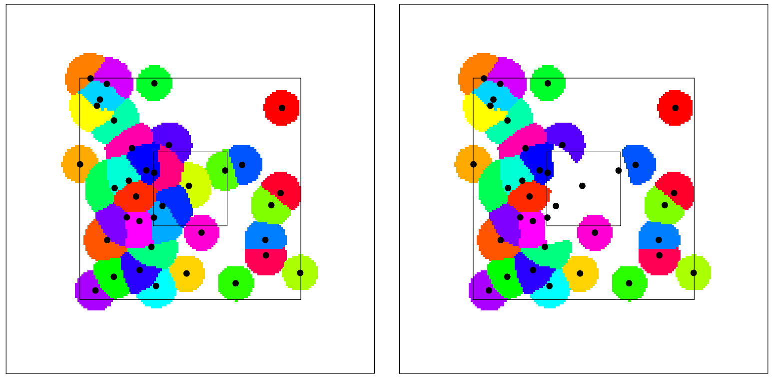

Figure 4.1: The left picture is a semicoupling of Lebesgue and 36 points with cost function $c(x, y) = |x - y|^4$. In the right picture, the five points within the small cube can choose new partners from the mass that was transported to them in the left picture (corresponding to the measure λ^ω_A). If the semicoupling on the left hand side is locally optimal, then the points in the small cube on the right hand side will choose exactly the partners they have in the left picture.

exist and define a c-cyclically monotone map $T^\omega : A^\omega \to M$ such that on $A^\omega \times M$:

$$q^\omega = (id, T^\omega)_*(\rho^\omega \lambda^\omega).$$

$\square$

Remark 4.2. In the sequel, any *transport map* $T^\omega : A^\omega \to M$ as above will be extended to a map $T^\omega : M \to M \cup \{\eth\}$ by putting $T^\omega(x) := \eth$ for all $x \in M \setminus A^\omega$ where $\eth$ denotes an isolated point added to M ('point at infinity', 'cemetery'). Then (4.1) reads

$$q^\omega = (Id, T^\omega)_* \, \rho^\omega \lambda^\omega \quad \text{on } M \times M. \tag{4.2}$$

Moreover, we put $c(x, T^\omega(x)) = c(x, \eth) := 0$ for $x \in M \setminus A^\omega$. If we know a priori that $\rho^\omega(x) \in \{0, 1\}$ almost surely (4.2) simplifies to

$$q^\omega = (Id, T^\omega)_* \, \lambda^\omega \quad \text{on } M \times M. \tag{4.3}$$

Definition 4.3. *A semicoupling $Q = q^\bullet \mathbb{P}$ of $\lambda^\bullet$ and $\mu^\bullet$ is called* locally optimal *iff some (hence every) property of the previous proposition is satisfied for $\mathbb{P}$-a.e. $\omega \in \Omega$.*

Remark 4.4. (i) Asymptotic optimality is not sufficient for uniqueness and it does not imply local optimality: Consider the Lebesgue measure $\lambda^\bullet = \mathcal{L} = \lambda$ and a Poisson point process $\mu^\bullet$ of unit intensity on $\mathbb{R}^d$. Let us fix the cost function $c(x, y) = |x-y|^2$. Lemma 2.16 shows that $\mathfrak{c}_0$ the optimal mean transportation cost on the unit cube, that is the cost of the optimal semicoupling of λ and $1_{[0,1)^d}\mu^\bullet$, is strictly less than the optimal mean transportation cost on a big cube, say $[0, 10^{10})^d$. Moreover, for any semicoupling $q^\bullet$ between λ and $\mu^\bullet$ Lemma 2.16 implies that

$$\mathfrak{C}(q^\bullet) = \liminf_{n\to\infty} \frac{1}{\lambda(B_n)} \mathbb{E}\left[\mathsf{Cost}(1_{\mathbb{R}^d \times B_n} q^\bullet)\right],$$

for $B_n = [-2^{n-1}, 2^{n-1})^d$. In other words, it is more costly to transport in one big cube than in many small cubes separately. For all $n \in \mathbb{N}$, let $\rho_n : \Omega \times \mathbb{R}^d \to [0, 1]$ be the unique optimal density for the transport problem between λ and $1_{B_n}\mu^\bullet$, that is the optimal semicoupling is given by $q^\bullet_n = (id, T^\bullet_n)_*(\rho^\bullet_n \lambda)$. Let $\kappa^\bullet_n$ be the following semicoupling between λ and $1_{B_n}\mu^\bullet$

$$\kappa^\omega(dx, dy) \;=\; q^\omega_0(dx, dy) + \sigma^\omega_n(dx, dy),$$

where σ_n^ω is the unique optimal *coupling* between $(\rho_n^\omega - \rho_0^\omega)\cdot\lambda$ and $1_{B_n\setminus B_0}\mu^\omega$. Let $f_n:\Omega\times\mathbb{R}^d\to[0,1]$ be such that $1_{R^d\times(B_n\setminus B_0)}q_n^\omega = (id, T_n^\omega)_*(f_n^\omega\lambda)$. Denote by $\mathbb{W}_2$ the expectation of the usual L^2- Wasserstein distance. Then, we can estimate using the triangle inequality

$$\mathfrak{Cost}(\kappa_n^\bullet) \;=\; \mathfrak{Cost}(q_0^\bullet)+\mathbb{W}_2((\rho_n-\rho_0)\cdot\lambda, 1_{(B_n\setminus B_0)}\mu^\bullet) \;\le\; \mathfrak{Cost}(q_0^\bullet)+\mathfrak{Cost}(q_n^\bullet)+\mathbb{W}_2((\rho_n-\rho_0)\cdot\lambda, f_n\cdot\lambda).$$

Set $Z_l=\mu^\bullet(B_l)$. Note that $(\rho_n^\omega-\rho_0^\omega)$ and f_n^ω coincide on a set of Lebesgue measure of mass at least $Z_n^\omega - Z_0^\omega$. This allows to estimate $\mathbb{W}_2((\rho_n-\rho_0)\cdot\lambda, f_n\cdot\lambda)$ very roughly from above (for similar less rough and much more detailed estimates we refer to chapter 7). We have to transport mass of amount at most Z_0 at most a distance $R_1 = 2h\cdot Z_n^{1/d}+2\sqrt{d}2^n$ for some constant h, e.g. $h=(\Gamma(\frac{d}{2}+1))^{1/d}$ would do. Indeed, ρ_n^ω must be supported in a $h\cdot Z^{1/d}$ neighbourhood of B_n because we could otherwise produce a cheaper semicoupling (see Lemma 3.9). This gives using the estimates on Poisson moments of Lemma 7.12

$$\begin{aligned}
&\mathbb{W}_2((\rho_n-\rho_0)\cdot\lambda, f_n\cdot\lambda)\le\mathbb{E}\left[R_1^2\cdot Z_0\right]\\
\le\;& C_1\;\left(\mathbb{E}[Z_0^2]^{1/2}\mathbb{E}[Z_n^{4/d}]^{1/2}+\lambda(B_n)^{2/d}+2\cdot\mathbb{E}[Z_0^2]^{1/2}\lambda(B_n)^{1/d}\mathbb{E}[Z_n^{2/d}]^{1/2}\right)\\
\le\;& C_2\;\lambda(B_n)^{2/d},
\end{aligned}$$

for some constants C_1 and C_2. In particular, if we take $d\ge 3$ this shows that

$$\liminf_{n\to\infty}\frac{1}{\lambda(B_n)}\mathsf{Cost}(\kappa_n^\bullet)=\liminf_{n\to\infty}\frac{1}{\lambda(B_n)}\mathsf{Cost}(q_n^\bullet).$$

Hence, as in Proposition 2.23 we can show that $\kappa_n^\bullet$ converges along a subsequence to some semicoupling $\kappa^\bullet$ between λ and $\mu^\bullet$ which is asymptotically optimal but not locally optimal.

(ii) Local optimality does not imply asymptotic optimality and it is not sufficient for uniqueness: For instance in the case $M=\mathbb{R}^d, c(x,y)=|x-y|^2$, given any coupling $q^\bullet$ of $\mathcal{L}$ and a Poisson point process $\mu^\bullet$ and $z\in\mathbb{R}^d\setminus\{0\}$ then

$$\tilde{q}^\omega(dx,dy):=q^\omega(d(x+z),dy)$$

defines another locally optimal coupling of $\mathcal{L}$ and $\mu^\bullet$. Not all of them can be asymptotically optimal.

(iii) The name *local optimality* might be misleading in the context of semicouplings. Consider a Poisson process $\mu^\bullet$ of intensity $1/2$ and let $q^\bullet$ be an optimal coupling between $1/2\mathcal{L}$ and $\mu^\bullet$. Then, it is locally optimal (see Theorem 4.6) according to this definition. However, as we left half of the Lebesgue measure laying around we can everywhere locally produce a coupling with less cost. In short, the optimality does not refer to the choice of density only to the use of the chosen density.

(iv) Note that local optimality — in contrast to asymptotic optimality and equivariance — is not preserved under convex combinations. It is an open question if local optimality and asymptotic optimality imply uniqueness.

Given two random measures $\gamma^\bullet,\eta^\bullet:\Omega\to\mathcal{M}(M)$ with $\gamma^\omega(M)=\eta^\omega(M)<\infty$ for all $\omega\in\Omega$ we define the *optimal mean transportation cost* by

$$\mathfrak{Cost}(\gamma^\bullet,\eta^\bullet)\;:=\;\inf\left\{\mathfrak{Cost}(q^\bullet):\;q^\omega\in\Pi(\gamma^\omega,\eta^\omega)\text{ for a.e. }\omega\in\Omega\right\}.$$

Given a semicoupling $q^\bullet$ of $\lambda^\bullet$ and $\mu^\bullet$ and a bounded Borel set $A\subset M$ recall the definition of $\lambda_A^\bullet$ from Prop. 4.1. We define the *efficiency of the semicoupling* $q^\bullet$ *on the set* A by

$$\mathfrak{e}_A(q^\bullet):=\frac{\mathfrak{Cost}(\lambda_A^\bullet\,,\,1_A\mu^\bullet)}{\mathfrak{Cost}(1_{M\times A}q^\bullet)}.$$

It is a number in $(0,1]$. The semicoupling $q^\bullet$ is said to be efficient on A iff $\mathfrak{e}_A(q^\bullet)=1$. Otherwise, it is inefficient on A. As noted in the remark above in the case of true *semi*couplings this notion might mislead the intuition.

Lemma 4.5. *(i) $q^\bullet$ is locally optimal if and only if $\mathfrak{e}_A(q^\bullet) = 1$ for all bounded Borel sets $A \subset M$. (ii) $\mathfrak{e}_A(q^\bullet) = 1$ for some $A \subset M$ implies $\mathfrak{e}_{A'}(q^\bullet) = 1$ for all $A' \subset A$, where we set $0/0 = 1$.*

Proof. (i) Let A be given and $\omega \in \Omega$ be fixed. Then $1_{M\times A} q^\omega$ is the optimal semicoupling of the measures λ_A^ω and $1_A \mu^\omega$ if and only if

$$\mathsf{Cost}(1_{M\times A} q^\omega) = \mathsf{Cost}\left(\lambda_A^\omega\,,\, 1_A \mu^\omega\right). \tag{4.4}$$

On the other hand, $\mathfrak{e}_A(q^\bullet) = 1$ is equivalent to

$$\mathbb{E}\left[\mathsf{Cost}(1_{M\times A} q^\bullet)\right] = \mathbb{E}\left[\mathsf{Cost}\left(\lambda_A^\bullet\,,\, 1_A \mu^\bullet\right)\right].$$

The latter, in turn, is equivalent to (4.4) for $\mathbb{P}$-a.e. $\omega \in \Omega$.
(ii) If the transport $q^\bullet$ restricted to $M \times A$ is optimal then also each of its sub-transports. $\square$

Remember that due to the stationarity of $\mathbb{P}$ equivariance of $q^\bullet$ translates into invariance of its distribution. The next Theorem is a key step in establishing uniqueness because it shows that every optimal semicoupling is induced by a map.

Theorem 4.6. *Every optimal semicoupling between $\lambda^\bullet$ and $\mu^\bullet$ is locally optimal.*

Proof. Assume we are given a semicoupling $q^\bullet$ of $\lambda^\bullet$ and $\mu^\bullet$ that is equivariant but not locally optimal. According to the previous lemma, the latter implies that there is $g \in G$ and $r \in \mathbb{N}$ such that $q^\bullet$ is not efficient on gB_r, i.e.

$$\eta = \mathfrak{e}_{gB_r}(q^\bullet) < 1.$$

By invariance, this implies that $\eta = \mathfrak{e}_{hB_r}(q^\bullet) < 1$ for all $h \in G$. Hence, for any $h \in G$ there is a coupling $\tilde{q}^\bullet_{hB_r}$ of $\lambda^\bullet_{hB_r}$ and $1_{hB_r}\mu^\bullet$, the unique optimal coupling, which is more efficient than $1_{M\times hB_r} q^\bullet$, i.e. such that

$$\mathbb{E}[\mathsf{Cost}(\tilde{q}^\bullet_{hB_r})] \leq \eta \cdot \mathbb{E}[\mathsf{Cost}(1_{M\times hB_r} q^\bullet)].$$

Moreover, because of the equivariance of $q^\bullet$ we have $\tilde{q}^\omega_{hB_r}(dx, dy) = \tilde{q}^{\theta_g\omega}_{ghB_r}(d(gx), d(gy))$ (see also Corollary 3.12). Hence, all convex combinations of the measures $\tilde{q}^\omega_{hB_r}$ will have similar equivariance properties.
We would like to have the estimate above also for the restriction of $\tilde{q}^\bullet_{hB_r}$ to $M \times hB_0$ in order to produce a semicoupling with less transportation cost than $q^\bullet$. This is not directly possible as we cannot control the contribution of hB_0 to the cost of $\tilde{q}^\bullet_{hB_r}$. However, we can use a trick which we will also use for the construction in the next chapter that will give us the desired result. Remember that Λ_r denotes the 2^r neighbourhood of the identity in the Cayley graph of G. Set

$$\bar{q}^\bullet_{hB_0} = \frac{1}{|\Lambda_r|} \sum_{g \in h\Lambda_r} 1_{M\times hB_0} \tilde{q}^\bullet_{gB_r}.$$

Then we have

$$\begin{aligned}
\mathfrak{Cost}(\bar{q}^\bullet_{hB_0}) &= \frac{1}{|\Lambda_r|} \sum_{g \in h\Lambda_r} \mathbb{E}\left[\int_{M\times hB_0} c(x,y)\tilde{q}^\bullet_{gB_r}(dx,dy)\right] \\
&= \frac{1}{|\Lambda_r|} \cdot \mathbb{E}\left[\int_{M\times hB_r} c(x,y)\tilde{q}^\bullet_{hB_r}(dx,dy)\right] \\
&= \frac{1}{|\Lambda_r|} \cdot \mathbb{E}\left[\mathsf{Cost}(\tilde{q}^\bullet_{hB_r})\right].
\end{aligned}$$

The second equality holds because fixing hB_0 and summing over all gB_r containing hB_0 is, due to the invariance of $\tilde{q}^\bullet_{gB_r}$, the same as fixing hB_r and summing over all gB_0 contained in

hB_r. Put $\bar{q}^\bullet = \sum_{h\in G} \bar{q}^\bullet_{hB_0}$. By construction, $\bar{q}^\bullet$ is an equivariant semicoupling of $\lambda^\bullet$ and $\mu^\bullet$. Furthermore, for any $g \in G$ we have

$$\mathbb{E}[\mathsf{Cost}(\bar{q}^\bullet_{gB_0})] \leq \eta \cdot \mathbb{E}[1_{M\times gB_0} q^\bullet].$$

This means, that $q^\bullet$ is not asymptotically optimal.

$\square$

Remark 4.7. We really need to consider $r > 1$ in the above proof as it can happen that $q^\bullet$ is efficient on every fundamental region but not locally optimal. Indeed, consider $\mu = \sum_{z\in\mathbb{Z}^2} \delta_z$ and let $q^\bullet$ denote the coupling transporting one quarter of the Lebesgue measure of the square of edge length 2 centered at z to z. This is efficient on every fundamental region, which contains exactly one $z \in \mathbb{Z}^2$, but not efficient on say $[0, 5)^2$.

Theorem 4.8. *Assume that $\mu^\bullet$ has intensity one, then there is a unique optimal coupling between $\lambda^\bullet$ and $\mu^\bullet$.*

Proof. Assume we are given two optimal couplings $q_1^\bullet$ and $q_2^\bullet$. Then also $q^\bullet := \frac{1}{2}q_1^\bullet + \frac{1}{2}q_2^\bullet$ is an optimal coupling because asymptotic optimality and equivariance are stable under convex combination. Hence, by the previous theorem all three couplings – $q_1^\bullet$, $q_2^\bullet$ and $q^\bullet$ – are locally optimal. Thus, for a.e. ω by the results of Proposition 4.1 there exist maps $T_1^\omega, T_2^\omega, T^\omega$ such that

$$\begin{aligned}\delta_{T^\omega(x)}(dy)\ \lambda^\omega(dx) &= q^\omega(dx, dy)\\ &= \left(\frac{1}{2}\delta_{T_1^\omega(x)}(dy) + \frac{1}{2}\delta_{T_2^\omega(x)}(dy)\right)\ \lambda^\omega(dx)\end{aligned}$$

This, however, implies $T_1^\omega(x) = T_2^\omega(x)$ for a.e. $x \in M$. Thus $q_1^\omega = q_2^\omega$. (By Lemma 2.25 we know that every invariant semicoupling between $\lambda^\bullet$ and $\mu^\bullet$ has to be a coupling.) $\square$

Before we can prove the uniqueness of optimal semicouplings we have to translate Proposition 3.5 to this setting.

Proposition 4.9. *Assume $\mu^\bullet$ has intensity $\beta \leq 1$ and let $q^\bullet$ be an optimal semicoupling between $\lambda^\bullet$ and $\mu^\bullet$. Let $(\pi_1)_* q^\bullet = \rho \cdot \lambda^\bullet$ for some density $\rho : \Omega \times M \to [0, 1]$. Then,*

$$\tilde{q}^\bullet \;=\; q^\bullet + (id \times id)_*((1-\rho)\cdot\lambda^\bullet)$$

is the unique optimal coupling between $\lambda^\bullet$ and $\hat{\mu}^\bullet := \mu^\bullet + (1-\rho)\cdot\lambda^\bullet$.

Proof. Because $q^\bullet$ is equivariant by assumption also $\rho\lambda^\bullet(\cdot) = q^\bullet(\cdot, M)$ is equivariant. But then $\hat{\mu}^\bullet = \mu^\bullet + (1-\rho)\cdot\lambda^\bullet$ is equivariant. Moreover, by assumption we have $\mathfrak{C}(\tilde{q}) = \mathfrak{C}(q) < \infty$ which implies

$$\inf_{\kappa^\bullet\in\Pi(\lambda^\bullet,\hat{\mu}^\bullet)} \mathfrak{C}(\kappa^\bullet) < \infty.$$

By the previous theorem, there is a unique optimal coupling $\kappa^\bullet$ between $\lambda^\bullet$ and $\hat{\mu}^\bullet$ given by $\kappa^\bullet = (id, S)_*\lambda^\bullet$. Moreover,

$$\mathfrak{C}(\kappa^\bullet) \;\leq\; \mathfrak{C}(\tilde{q}^\bullet) \;=\; \mathfrak{C}(q^\bullet).$$

Because $S_*\lambda^\bullet = \hat{\mu}^\bullet$ there is a density f such that $S_*(f\cdot\lambda^\bullet) = (1-\rho)\cdot\lambda^\bullet$. Indeed, for any $g \in G$ we can disintegrate

$$1_{M\times gB_0}\kappa^\omega(dx, dy) \;=\; \kappa_y^{\omega,g}(dx)(\mu^\omega(dy) + (1-\rho^\omega(y))\lambda^\omega(dy)).$$

The measure $\sum_{g\in G}\kappa_y^{\omega,g}(dx)((1-\rho^\omega(y))\lambda^\omega(dy))$ does the job. In particular this implies that

$$\tilde{\kappa}^\bullet \;=\; (id \times S)_*((1-f)\cdot\lambda^\bullet)$$

is a semicoupling between $\lambda^\bullet$ and $\mu^\bullet$. The mean transportation cost of $\tilde{\kappa}^\bullet$ are bounded above by the mean transportation cost of $\kappa^\bullet$ as we just transport less mass. Hence, we have

$$\mathfrak{C}(\tilde{\kappa}^\bullet) \leq \mathfrak{C}(\kappa^\bullet) \leq \mathfrak{C}(\tilde{q}^\bullet) = \mathfrak{C}(q^\bullet).$$

As $q^\bullet$ was assumed to be optimal, hence asymptotically optimal, we must have equality everywhere. By uniqueness of optimal couplings this implies that $\tilde{q}^\bullet = \kappa^\bullet$ almost surely. □

Lemma 4.10. *Assume $\mu^\bullet$ has intensity $\beta \leq 1$ and let $q^\bullet = (id, T)_*(\rho \cdot \lambda^\bullet)$ be an optimal semicoupling between $\lambda^\bullet$ and $\mu^\bullet$. Then, on the set $\{0 < \rho^\omega < 1\}$ we have $T^\omega(x) = x$.*

Proof. Just as in the previous proposition consider $\tilde{q}^\bullet = (id, S)_*\lambda^\bullet$ the optimal coupling between $\lambda^\bullet$ and $\hat{\mu}^\bullet$. $\tilde{q}^\bullet$ is concentrated on the graph of S and therefore also $q^\bullet$ has to be concentrated on the graph of S. In particular, this shows that $S = T$ almost everywhere almost surely (we can safely extend T by S on $\{\rho = 0\}$). But on $\{\rho < 1\}$ we have $S(x) = x$. Hence, we also have $T(x) = x$ on $\{0 < \rho < 1\}$. □

This finally enables us to prove uniqueness of optimal semicouplings.

Theorem 4.11. *There exists at most one optimal semicoupling of $\lambda^\bullet$ and $\mu^\bullet$.*

Proof. Assume we are given two optimal semicouplings $q_1^\bullet$ and $q_2^\bullet$. Then also $q^\bullet := \frac{1}{2}q_1^\bullet + \frac{1}{2}q_2^\bullet$ is an optimal semicoupling. Hence, by Theorem 4.6 all three couplings – $q_1^\bullet$, $q_2^\bullet$ and $q^\bullet$ – are locally optimal. Thus, for a.e. ω by the results of Proposition 4.1 there exist maps $T_1^\omega, T_2^\omega, T^\omega$ and densities $\rho_1^\omega, \rho_2^\omega, \rho^\omega$ such that

$$\begin{aligned}
\delta_{T^\omega(x)}(dy)\,\rho^\omega(x)\,\lambda^\omega(dx) &= q^\omega(dx, dy) \\
&= \left(\frac{1}{2}\delta_{T_1^\omega(x)}(dy)\rho_1^\omega(x) + \frac{1}{2}\delta_{T_2^\omega(x)}(dy)\rho_2^\omega(x)\right)\lambda^\omega(dx)
\end{aligned}$$

This, however, implies $T_1^\omega(x) = T_2^\omega(x)$ for a.e. $x \in \{\rho_1^\omega > 0\} \cap \{\rho_2^\omega > 0\}$. In particular, all optimal semicouplings are concentrated on the *same* graph. To show uniqueness, we have to show that $\rho_1^\omega = \rho_2^\omega$ almost everywhere almost surely. To this end, put $A_1^\omega = \{\rho_1^\omega > \rho_2^\omega\}$. Assume $\lambda^\omega(A_1^\omega) > 0$. On A_1^ω we have $\rho^\omega < 1$. Hence, by the previous Lemma we have $T^\omega(x) = T_1^\omega(x) = T_2^\omega(x) = x$. Similarly, on $A_2^\omega = \{\rho_2^\omega > \rho_1^\omega\}$ we have $T^\omega(x) = x$. Hence, we have

$$(T^\omega)^{-1}(A_1^\omega) \cap A_2^\omega = \emptyset,$$

because $A_i \subset (T^\omega)^{-1}(A_i)$ for $i = 1, 2$. As q_1^ω and q_2^ω are semicouplings, we must have $\mu^\omega(A) = \rho_i^\omega \cdot \lambda^\omega((T^\omega)^{-1}(A))$ for $i = 1, 2$ and any Borel set A. Putting this together gives

$$\begin{aligned}
\mu^\omega(A_1^\omega) &= \rho_1^\omega \cdot \lambda^\omega((T^\omega)^{-1}(A_1^\omega)) \\
&= \rho_1^\omega \cdot \lambda^\omega((T^\omega)^{-1}(A_1^\omega) \cap A_1^\omega) + \rho_1^\omega \cdot \lambda^\omega((T^\omega)^{-1}(A_1^\omega) \cap \{\rho_1^\omega = \rho_2^\omega\}) \\
&> \rho_2^\omega \cdot \lambda^\omega((T^\omega)^{-1}(A_1^\omega) \cap A_1^\omega) + \rho_2^\omega \cdot \lambda^\omega((T^\omega)^{-1}(A_1^\omega) \cap \{\rho_1^\omega = \rho_2^\omega\}) \\
&= \mu^\omega(A_1^\omega).
\end{aligned}$$

This is a contradiction and therefore proving $\lambda^\omega(A_1^\omega) = 0$. Thus, $\rho_1^\omega = \rho_2^\omega$ almost everywhere almost surely and $q_1^\bullet = q_2^\bullet$. □

4.1 Geometry of tessellations induced by fair allocations

The fact that any optimal semicoupling is locally optimal allows us to say something about the geometries of the cells of fair allocations to point processes. The following result was already shown for probability measures in section 4 of [Stu09] and also in [AHA92]. We will use the representation of optimal transportation maps recalled in Remark 3.4.

Corollary 4.12. *In the case $\vartheta(r) = r^2$, given an optimal coupling $q^\bullet$ of Lebesgue measure $\mathcal{L}$ and a point process $\mu^\bullet$ of unit intensity in $M = \mathbb{R}^d$ (for a Poisson point process this implies $d \geq 3$ as otherwise the mean transportation cost will be infinite, see Theorem 1.7) then for a.e. $\omega \in \Omega$ there exists a convex function $\varphi^\omega : \mathbb{R}^d \to \mathbb{R}$ (unique up to additive constants) such that*

$$q^\omega = (Id, \nabla\varphi^\omega)_* \mathcal{L}.$$

In particular, a 'fair allocation rule' is given by the monotone map $T^\omega = \nabla\varphi^\omega$.
Moreover, for a.e. ω and any center $\xi \in \Xi(\omega) := \mathrm{supp}(\mu^\omega)$, the associated cell

$$S^\omega(\xi) \;=\; (T^\omega)^{-1}(\{\xi\})$$

is a convex polytope of volume $\mu^\omega(\xi) \in \mathbb{N}$. If the point process is simple then all these cells have volume 1.

Proof. By Proposition 4.1 we know that $T^\omega = \lim_{n\to\infty} T_n^\omega$, where T_n^ω is an optimal transportation map from some set A_n^ω to K_n. From the classical theory (see [Bre91, GM96]) we know that, $T_n^\omega = \nabla\varphi_n^\omega$ for some convex function φ_n^ω. More precisely,

$$\varphi_n^\omega(x) = \max_{\xi\in\Xi(\omega)\cap K_n} (x^2 - |x-\xi|^2/2 + b_\xi)$$

for some constants b_ξ. Moreover, we know that $T_{n+k}^\omega = T_n^\omega$ on A_n^ω for all $k \in \mathbb{N}$. Fix any $\xi_0 \in \Xi(\omega)$. Then, there is $n \in \mathbb{N}$ such that $\xi_0 \in K_n$. Then, $(T_{n+k}^\omega)^{-1}(\xi_0) = (T_n^\omega)^{-1}(\xi_0)$ for any $k \in \mathbb{N}$. Furthermore,

$$T_n^\omega(x) = \xi_0 \quad \Leftrightarrow \quad -|x-\xi_0|^2/2 + b_{\xi_0} > -|x-\xi|^2/2 + b_\xi \quad \forall \xi \in \Xi(\omega)\cap K_n,\ \xi \neq \xi_0.$$

For fixed $\xi \neq \xi_0$ this equation describes two halfspaces separated by a hyperplane (defined by equality in the equation above). The set $S^\omega(\xi_0)$ is then given as the intersection of all these halfspaces defined by ξ_0 and $\xi \in \Xi(\omega)\cap K_n$. Hence, it is a convex polytope. □

Corollary 4.13. *In the case $\vartheta(r) = r$, given an optimal coupling $q^\bullet$ of m and a point process $\mu^\bullet$ of unit intensity on M with $dim(M) \geq 2$, there exists an allocation rule T such that the optimal coupling is given by*

$$q^\omega = (Id, T^\omega)_* m.$$

Moreover, for a.e. ω and any center $\xi \in \Xi(\omega) := \mathrm{supp}(\mu^\omega)$, the associated cell

$$S^\omega(\xi) \;=\; (T^\omega)^{-1}(\{\xi\})$$

is starlike with respect to ξ.

Proof. We argue as in the last proof. However, the defining equation for the cells now becomes

$$T_n^\omega(x) = \xi_0 \quad \Leftrightarrow \quad -d(x,\xi_0) + b_{\xi_0} > -d(x,\xi) + b_\xi \quad \forall \xi \in \Xi(\omega)\cap K_n,\ \xi \neq \xi_0.$$

Hence, the cell can again be written as the intersection of "halfspaces" $H_j^0 := \{x : -d(x,\xi_0) + b_{\xi_0} > -d(x,\xi_j) + b_{\xi_j}\}$. Therefore, it is sufficient to show that for any $z \in H_j^0$ the whole geodesic from z to ξ_0 lies inside H_j^0. For convenience we write $\Phi_0(x) = -d(x,\xi_0) + b_{\xi_0}$ and $\Phi_j(x) = -d(x,\xi_j) + b_{\xi_j}$.
Assume $\xi_0 \in \partial H_j^0$ and w.l.o.g. $b_{\xi_0} = 0$. Then, we have

$$\Phi_0(\xi_0) = 0 = \Phi_j(\xi_0) \Rightarrow b_{\xi_j} = d(\xi_j, \xi_0).$$

The set $N = \{z \in M : d(\xi_j, z) = d(\xi_j, \xi_0) + d(\xi_0, z)\}$ is a m-null set. For all $z \notin N$ we have

$$\Phi_j(z) = -d(\xi_j, z) + b_{\xi_j} > -d(\xi_j, \xi_0) + b_{\xi_j} - d(\xi_0, z) = \Phi_0(z)$$

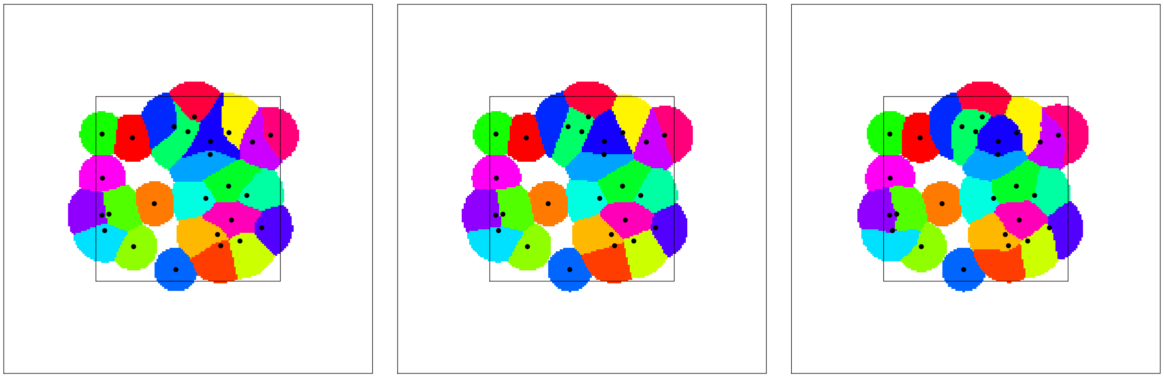

Figure 4.2: Optimal semicoupling of Lebesgue and 25 points in the cube with cost function $c(x,y)=|x-y|^p$ and (from left to right) p=1, 2, 4 respectively.

This implies that $m(T_n^{-1}(\xi_i))=0$ contradicting the assumption of T being an allocation. Thus, $\xi_0 \notin \partial H_j^0$ and in particular $T(\xi_0)\in\Xi=\mathrm{supp}(\mu)$.
Assume $T(\xi_0)\neq\xi_0$. Then, there is a $\xi_j\neq\xi_0$ such that $T(\xi_0)=\xi_j$, i.e. $\Phi_j(\xi_0)=-d(\xi_0,\xi_j)+b_{\xi_j}>b_{\xi_0}=\Phi_0(\xi_0)$. Then, we have for any $p\in M, p\neq\xi_0$

$$-d(p,\xi_j)+b_{\xi_j}\geq -d(p,\xi_0)-d(\xi_0,\xi_j)+b_{\xi_j}>-d(p,\xi_0)+b_{\xi_0}.$$

This implies, that $m(T^{-1}(\xi_0))=0$ contradicting the assumption of T being an allocation. Thus, $T(\xi_0)=T_n(\xi_0)=\xi_0$.
Take any $w\in T_n^{-1}(\xi_0)$ (hence, $\Phi_0(w)>\Phi_j(w)$ for all $j\neq 0$) and $p\in M$ such that $d(\xi_0,w)=d(\xi_0,p)+d(p,w)$, i.e. p lies on the minimizing geodesic from ξ_0 to w. Then, we have for any $j\neq 0$ by using the triangle inequality once more

$$\begin{aligned}-d(p,\xi_0)+b_{\xi_0} &= -d(\xi_0,w)+d(p,w)+b_{\xi_0}\\ &\geq -d(\xi_0,w)+b_{\xi_0}+d(w,\xi_j)-d(p,\xi_j)\\ &> -d(p,\xi_j)+b_{\xi_j},\end{aligned}$$

which means that $\Phi_0(p)>\Phi_j(p)$ for all $j\neq 0$. Hence, $p\in H_j^0$ proving the claim. □

Remark 4.14. i) Questions on the geometry of the cells of fair allocations are highly connected to the very difficult problem of the regularity of optimal transportation maps (see [MTW05, Loe09, KM07]). The link is of course the cyclical monotonicity. The geometry of the cells of the "optimal fair allocation" is dictated by the cyclical monotonicity and the optimal choice of cyclical monotone map to get an asymptotic optimal coupling.
Consider the classical transport problem between two probability measures one being absolutely continuous to the volume measure on M with full support on a convex set and the other one being a convex combination of N Dirac masses. Assume that the cell being transported to one of the N points is not connected. Then, it is not difficult to imagine that it is possible to smear out the Dirac masses slightly to get two absolutely continuous probability measures (even with very nice densities) but a discontinuous transportation map.

ii) Considering L^p cost on $\mathbb{R}^d$ with $p\notin\{1,2\}$, the cell structure is much more irregular than in the two cases considered above. The cells do not even have to be connected. Indeed, just as in the proof of the two Corollaries above it holds also for general p that $T^\omega(x)=\xi_0$ iff $\Phi_0(x)>\Phi_i(x)$ for all $i\neq 0$ where $\Phi_i(x)=-|x-\xi_i|^p+b_i$ for some constants b_i (see also Example 1.6 in [GM96]). By considering the sets $\Phi_i\equiv\Phi_0$ it is not difficult to cook up examples of probability measures such that the cells do not have to be connected.

In the case that $p \in (0,1)$ similar to the case that $p = 1$ we always have that the center of the cells lies in the cell, that is $T(\xi_i) = \xi_i$ for all $\xi_i \in \mathrm{supp}(\mu^\bullet)$ because the cost function defines a metric (see [GM96]).

iii) As was shown by Loeper in section 8.1 of [Loe09] the cells induced by the optimal transportation problem in the hyperbolic space between an absolutely continuous measure and a discrete measure with respect to the cost function $c(x, y) = d^2(x, y)$ do not have to be connected. In the same article he shows that for the same problem on the sphere the cells have to be connected. In [vN09] von Nessi studies more general cost functions on the sphere, including the L^p cost function $c(x, y) = d^p(x, y)$. He shows that in general for $p \neq 2$ the cells do not have to connected. This suggests that on a general manifold the cell structure will probably be rather irregular.

Chapter 5

Construction of optimal semicouplings

We fix again a pair of equivariant random measures $\lambda^\bullet$ and $\mu^\bullet$ of unit resp. subunit intensity with finite mean transportation cost $\mathfrak{c}_\infty$ such that $\lambda^\bullet$ is absolutely continuous. Additionally, we assume that G satisfies some strong form of amenability. Recall that the 2^r neighbourhood of the identity element in the Cayley graph $\Delta(G,S)$ of G is denoted by Λ_r and the range of its action on the fundamental region by B_r, i.e. $B_r = \bigcup_{g\in\Lambda_r} gB_0$. Then, we assume that for any $g \in G$

$$\frac{|\Lambda_r \triangle g\Lambda_r|}{|\Lambda_r|} \to 0 \text{ as } r \to \infty, \tag{5.1}$$

where $|A|$ denotes the cardinality of A. In other words, we assume that the "balls" Λ_r are Følner sets. This of course implies for any $g \in G$

$$\frac{m(B_r \triangle gB_r)}{m(B_r)} \to 0 \text{ as } r \to \infty.$$

The aim of this chapter is to construct the optimal semicoupling and thereby proving Theorem 1.5. The construction is based on approximation by semicouplings on bounded sets. We will also show a nice convergence result of these approximations, proving Theorem 1.6.
This chapter follows closely the respective section in [HS10].

5.1 Symmetrization and Annealed Limits

The crucial step in our construction of optimal semicouplings between $\lambda^\bullet$ and $\mu^\bullet$ is the introduction of a *symmetrization* or *second randomization*. We want to construct the optimal semicoupling by approximation of optimal semicouplings on bounded sets. The difficulty in this approximation lies in the estimation of the contribution of the fundamental regions gB_0 to the transportation cost, i.e. what does it cost to transport mass into gB_0? How can the cost be bounded in order for us to be able to conclude that the limiting measure still transports the right amount of mass into gB_0? The solution is to mix several optimal semicouplings and thereby get a symmetry which will be very useful (see proof of Lemma 5.1 (i)). One can also think of the mixing as an expectation of the random choice of increasing sequences of sets hB_r exhausting M.
For each $g \in G$ and $r \in \mathbb{N}$, recall that Q_{gB_r} denotes the minimizer of $\mathfrak{Cost}$ among the semicouplings of $\lambda^\bullet$ and $1_{gB_r}\mu^\bullet$ as constructed in Theorem 2.15. It inherits the equivariance from $\lambda^\bullet$ and $\mu^\bullet$, namely $Q_{gA}(g\cdot, g\cdot, \theta_g\omega) = Q_A(\cdot,\cdot,\omega)$ (see Corollary 3.12). In particular, the stationarity of $\mathbb{P}$ implies $(\tau_h)_* Q_{gBr} \stackrel{d}{=} Q_{hgB_r}$. Put

$$Q_g^r(dx, dy, d\omega) \ := \ 1_{gB_0}(y)\frac{1}{|\Lambda_r|}\sum_{h\in g\Lambda_r} Q_{hB_r}(dx, dy, d\omega).$$

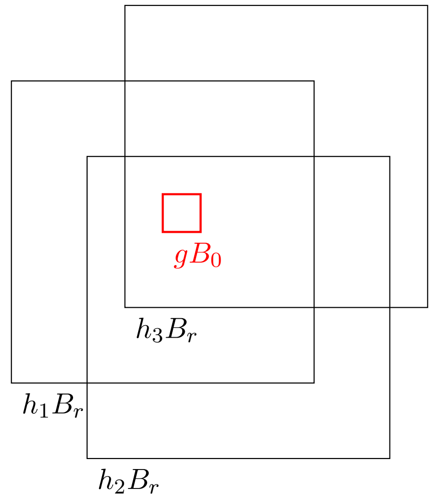

Figure 5.1: Schematic picture of the mixing procedure.

The measure Q_g^r defines a semicoupling between $\lambda^\bullet$ and $1_{gB_0}\mu^\bullet$. It is a deterministic, fractional allocation in the following sense:

- for any ω it is a deterministic function of λ^ω and μ^ω and does not depend on any additional randomness,
- for any ω the first marginal is absolutely continuous with respect to λ^ω with density ≤ 1.

The last fact implies that the semicoupling Q_g^r is *not* optimal in general, e.g. if one transports the Lebesgue measure to a point process. The first fact implies that all the objects derived from Q_g^r in the sequel – like Q_g^∞ and Q^∞ – are also deterministic. Moreover, Q_g^r shares the equivariance properties of the measures Q_{hB_r}.

Lemma 5.1. *(i) For each $r \in \mathbb{N}$ and $g \in G$*

$$\int_{M\times gB_0\times\Omega} c(x,y)Q_g^r(dx,dy,d\omega) \leq \mathfrak{c}_\infty.$$

(ii) The family $(Q_g^r)_{r\in\mathbb{N}}$ of probability measures on $M \times M \times \Omega$ is relatively compact in the weak topology.

(iii) There exist probability measures Q_g^∞ and a subsequence $(r_l)_{l\in\mathbb{N}}$ such that for all $g \in G$:

$$Q_g^{r_l} \quad \longrightarrow \quad Q_g^\infty \qquad \textit{weakly as } l \to \infty.$$

Proof. (i) Let us fix $g \in G$ and start with the important observation: *For given $r \in \mathbb{N}$ and $g \in G$ averaging over all $h\Lambda_r$ with $h \in g\Lambda_r$ has the effect that "g attains each possible position inside Λ_r with equal probability"* (see also the proof of Theorem 4.6).

Hence, together with the invariance of Q_{kB_r} we obtain

$$\begin{aligned}
&\int_{M\times gB_0\times\Omega} c(x,y)Q_g^r(dx,dy,d\omega)\\
=&\ \frac{1}{|\Lambda_r|}\sum_{h\in g\Lambda_r}\int_{M\times gB_0\times\Omega} c(x,y)Q_{hB_r}(dx,dy,d\omega)\\
=&\ \frac{1}{|\Lambda_r|}\int_{M\times gB_r\times\Omega} c(x,y)Q_{gB_r}(dx,dy,d\omega)\\
=&\ \frac{1}{|\Lambda_r|}\mathsf{C}_{gB_r} =: \mathfrak{c}_r \le \mathfrak{c}_\infty,
\end{aligned}$$

by definition of $\mathfrak{c}_\infty$.

(ii) In order to prove tightness of $(Q_g^r)_{r\in\mathbb{N}}$, let $(gB_0)_l$ denote the closed l–neighborhood of gB_0 in M. Then,

$$\begin{aligned}
Q_g^r(\complement(gB_0)_l, gB_0,\Omega) &\le \frac{1}{\vartheta(l)}\int_{M\times gB_0\times\Omega} c(x,y)Q_g^r(dx,dy,d\omega)\\
&\le \frac{1}{\vartheta(l)}\mathfrak{c}_\infty.
\end{aligned}$$

Since $\vartheta(l)\to\infty$ as $l\to\infty$ this proves tightness of the family $(Q_g^r)_{r\in\mathbb{N}}$ on $M\times M\times\Omega$. (Recall that Ω was assumed to be compact from the very beginning.)

(iii) Tightness yields the existence of Q_g^∞ and of a converging subsequence for each $g\in G$. A standard argument ('diagonal sequence') then gives convergence for all $g\in G$ along a common subsequence (G is countable as it is finitely generated). $\square$

Note that the measures Q_g^∞ inherit as weak limits the property $Q_{hg}^\infty(h\cdot,h\cdot,\theta_h\cdot)=Q_g^\infty(\cdot,\cdot,\cdot)$ from the measures Q_g^r (see also the proof of the equivariance property in Proposition 2.23). The next Lemma allows to control the difference in the first marginals of Q_g^∞ and Q_h^∞ for $g\neq h$. This is the first point where we use amenability.

Lemma 5.2. *i) For all $l>0$ there exists numbers $\epsilon_r(l)$ with $\epsilon_r(l)\to 0$ as $r\to\infty$ s.t. for all $g,g'\in G$ and all $r\in\mathbb{N}$*

$$\frac{1}{|\Lambda_r|}\sum_{h\in g'\Lambda_r} Q_{hB_r}(A)\le\frac{1}{|\Lambda_r|}\sum_{h\in g\Lambda_r} Q_{hB_r}(A)+\epsilon_r(d_\Delta(g,g'))\cdot\sup_{h\in g'\Lambda_r} Q_{hB_r}(A)$$

for any Borel set $A\subset M\times M\times\Omega$.

ii) For all $g_1,\dots,g_n\in G$, all $r\in\mathbb{N}$ and all Borel sets $A\subset M, D\subset\Omega$

$$\sum_{i=1}^n Q_{g_i}^r(A,M,D)\le\left(1+\sum_{i=1}^n\epsilon_r(d_\Delta(g_1,g_i))\right)\cdot\lambda(D,A),$$

where $\lambda(D,A):=\int_D\int_A\lambda^\omega(dx)\mathbb{P}(d\omega)$.

Proof. (i) First note that for all $g,g'\in G$ and $r\in\mathbb{N}$ we have

$$g'\in g\Lambda_r \quad\Leftrightarrow\quad g\in g'\Lambda_r.$$

In this case, for $h\in g\Lambda_r$ with $g'\in h\Lambda_r$ we also have $h\in g'\Lambda_r$ and $g\in h\Lambda_r$. Moreover,

$$\frac{|\{h\in g\Lambda_r : g'\notin h\Lambda_r\}|}{|\Lambda_r|}\le\epsilon_r(d_\Delta(g,g')),$$

for some $\epsilon_r(l)$ with $\epsilon_r(l) \to 0$ as $r \to \infty$. One possible choice for ϵ_r is

$$\epsilon_r(d_\Delta(id, g)) = \frac{|\Lambda_r \triangle g\Lambda_r|}{|\Lambda_r|},$$

which tends to zero as r tends to infinity for any $g \neq id$ by assumption. This implies that for each pair $g, g' \in G$ and each $r \in \mathbb{N}$

$$\frac{|\{h \in g\Lambda_r : g' \in h\Lambda_r\}|}{|\Lambda_r|} \geq 1 - \epsilon_r(d_\Delta(g, g')).$$

Therefore, for each Borel set $A \subset M \times M \times \Omega$

$$\frac{1}{|\Lambda_r|} \sum_{h \in g'\Lambda_r} Q_{hB_r}(A) \leq \frac{1}{|\Lambda_r|} \sum_{h \in g\Lambda_r} Q_{hB_r}(A) + \epsilon_r(d_\Delta(g, g')) \cdot \sup_{h \in g'\Lambda_r} Q_{hB_r}(A).$$

(ii) According to the previous part (i), for each Borel sets $A \subset M, D \subset \Omega$

$$\begin{aligned}
&\sum_{i=1}^n Q^r_{g_i}(A, M, D) \\
= \; &\sum_{i=1}^n \frac{1}{|\Lambda_r|} \sum_{h \in g_i\Lambda_r} Q_{hB_r}(A, g_iB_0, D) \\
\leq \; &\sum_{i=1}^n \left(\frac{1}{|\Lambda_r|} \sum_{h \in g_1\Lambda_r} Q_{hB_r}(A, g_iB_0, D) + \epsilon_r(d_\Delta(g_1, g_i)) \cdot \sup_{h \in g_i\Lambda_r} Q_{hB_r}(A, g_iB_0, D) \right) \\
\leq \; &\left(1 + \sum_{i=1}^n \epsilon_r(d_\Delta(g_1, g_i)) \right) \lambda(D, A)
\end{aligned}$$

$\square$

Theorem 5.3. *The measure $Q^\infty := \sum_{g \in G} Q^\infty_g$ is an optimal semicoupling of $\lambda^\bullet$ and $\mu^\bullet$.*

Proof. (i) Second/third marginal: This is a direct consequence of the construction. For any $f \in \mathcal{C}_b^+(M \times \Omega)$ we have due to Lemma 5.1

$$\begin{aligned}
&\int_{M \times \Omega} f(y, \omega) Q^\infty(dx, dy, d\omega) \\
= \; &\sum_{g \in G} \int_{M \times \Omega} f(y, \omega) Q^\infty_g(dx, dy, d\omega) \\
= \; &\sum_{g \in G} \lim_{l \to \infty} \int_{M \times \Omega} f(y, \omega) Q^{k_l}_g(dx, dy, d\omega) \\
= \; &\sum_{g \in G} \int_{M \times \Omega} f(y, \omega) 1_{gB_0}(y) (\mu^\bullet \mathbb{P})(dy, d\omega) \\
= \; &\int_{M \times \Omega} f(y, \omega) (\mu^\bullet \mathbb{P})(dy, d\omega).
\end{aligned}$$

(ii) First/third marginal: Let an arbitrary bounded open set $A \subset M$ and an arbitrary Borel set $D \subset \Omega$ be given and let $(g_i)_{i \in \mathbb{N}}$ be an enumeration of G. According to the previous Lemma 5.2, for any $n \in \mathbb{N}$ and any $r \in \mathbb{N}$

$$\sum_{i=1}^n Q^r_{g_i}(A \times M \times D) \leq \left(1 + \sum_{i=1}^n \epsilon_r(d_\Delta(g_1, g_i)) \right) \cdot \lambda(D, A).$$

Letting first r tend to ∞ yields

$$\sum_{i=1}^{n} Q_{g_i}^{\infty}(A\times M\times D)\le\lambda(D,A).$$

Then with $n\to\infty$ we obtain

$$Q^{\infty}(A\times M\times D)\le\lambda(D,A)$$

which proves that $(\pi_{1,3})_*Q^\infty\le\lambda^\bullet\mathbb{P}$.

(iii) Optimality: By construction, Q^∞ is equivariant. Due to the resulting invariance, the asymptotic cost is given by

$$\begin{aligned}\int_{M\times B_0\times\Omega} c(x,y)\,Q^\infty(dx,dy,d\omega) &= \sum_{g\in G}\int_{M\times B_0\times\Omega} c(x,y)\,Q_g^\infty(dx,dy,d\omega)\\ &= \int_{M\times B_0\times\Omega} c(x,y)\,Q_{id}^\infty(dx,dy,d\omega)\ \le\ \mathfrak{c}_\infty.\end{aligned}$$

Here the final *inequality* is due to Lemma 5.1, property (i) (which remains true in the limit $r=\infty$), and the last *equality* comes from the fact that

$$\int_{M\times gB_0\times\Omega} c(x,y)\,Q_h^r(dx,dy,d\omega)=0$$

for all $g\ne h$ and for all $r\in\mathbb{N}$ (which also remains true in the limit $r=\infty$). □

Corollary 5.4. *(i) For $r\to\infty$, the sequence of measures $Q^r := \sum_{g\in G} Q_g^r$, $r\in\mathbb{N}$, converges vaguely to the unique optimal semicoupling Q^∞.*
(ii) For each $g\in G$ and $r\in\mathbb{N}$ put

$$\tilde Q_g^r(dx,dy,d\omega)\ :=\ \frac{1}{|\Lambda_r|}\sum_{h\in g\Lambda_r} Q_{hB_r}(dx,dy,d\omega).$$

The sequence $(\tilde Q_g^r)_{r\in\mathbb{N}}$ converges vaguely to the unique optimal semicoupling Q^∞.

Proof. (i) A slight extension of the previous Lemma 5.1(iii) + Theorem 5.3 yields that each subsequence $(Q^{r_n})_n$ of the above sequence $(Q^r)_r$ will have a sub-subsequence converging vaguely to an optimal coupling of $\lambda^\bullet$ and $\mu^\bullet$. Since the optimal coupling is unique, all these limit points coincide. Hence, the whole sequence $(Q^r)_r$ converges to this limit point (see e.g. [Dud02], Prop. 9.3.1).
(ii) Lemma 5.2 (i) implies that for $g,g',h\in G$ and every measurable $A\subset M\times M\times\Omega$

$$\begin{aligned}&|\tilde Q_g^r(A\cap(M\times hB_0\times\Omega))-\tilde Q_{g'}^r(A\cap(M\times hB_0\times\Omega))|\\ &\le\ \epsilon_r(d_\Delta(g,g'))\cdot\sup_{k\in G}\tilde Q_{kB_r}(A\cap(M\times hB_0\times\Omega))\\ &\le\ \varepsilon_r(d_\Delta(g,g'))\to 0\end{aligned}$$

as $r\to\infty$. Hence, for each $f\in\mathcal{C}_c(M\times M\times\Omega)$ and each $g'\in M$

$$\left|\sum_{g\in G}\int f(x,y,\omega)\,1_{gB_0}(y)\,d\tilde Q_g^r-\int f(x,y,\omega)\,d\tilde Q_{g'}^r\right|\to 0.$$

That is, $|\int f\,dQ^r-\int f\,d\tilde Q_{g'}^r|\to 0$ as $r\to\infty$. □

Corollary 5.5. *Denote the set of all semicoupling of $\lambda^\bullet$ and $\mu^\bullet$ by Π_s. Then it holds*

$$\begin{aligned}&\inf_{q^\bullet\in\Pi_s}\liminf_{r\to\infty}\frac{1}{m(B_r)}\mathbb{E}\left[\int_{M\times B_r}c(x,y)q^\bullet(dx,dy)\right]\\ &= \liminf_{r\to\infty}\inf_{q^\bullet\in\Pi_s}\frac{1}{m(B_r)}\mathbb{E}\left[\int_{M\times B_r}c(x,y)q^\bullet(dx,dy)\right].\end{aligned}$$

In particular, we have

$$\mathfrak{c}_\infty \;=\; \inf_{q^\bullet\in\Pi_s}\mathfrak{C}(q^\bullet) \;=\; \inf_{q^\bullet\in\Pi_{is}}\mathfrak{C}(q^\bullet) \;=\; \mathfrak{c}_{i,\infty}.$$

Proof. For any semicoupling $q^\bullet$ we have due to the supremum in the definition of $\mathfrak{C}(\cdot)$ that

$$\liminf_{r\to\infty}\frac{1}{m(B_r)}\mathbb{E}\left[\int_{M\times B_r}c(x,y)q^\bullet(dx,dy)\right] \;\le\; \mathfrak{C}(q^\bullet).$$

Hence, the left hand side is bounded from above by $\inf_{q^\bullet\in\Pi_s}\mathfrak{C}(q^\bullet)$. However, we just constructed a semicoupling, the unique optimal semicoupling Q^∞ which attains equality, i.e. with $Q^\infty = q^\bullet\mathbb{P}$

$$\liminf_{r\to\infty}\frac{1}{m(B_r)}\mathbb{E}\left[\int_{M\times B_r}c(x,y)q^\bullet(dx,dy)\right] \;=\; \mathfrak{C}(Q^\infty).$$

Hence, the left hand side equals $\inf_{q^\bullet\in\Pi_s}\mathfrak{C}(q^\bullet)$.
The right hand side equals $\liminf_{r\to\infty}\mathfrak{c}_r$ which is bounded by $\mathfrak{c}_\infty = \inf_{q^\bullet\in\Pi_s}\mathfrak{C}(q^\bullet)$ by Lemma 2.16. By our construction, the asymptotic transportation cost of Q^∞ are bounded by the right hand side, i.e.

$$\mathfrak{C}(Q^\infty)\le\liminf_{r\to\infty}\mathfrak{c}_r$$

by Lemma 5.1. Hence, also the right hand side equals $\inf_{q^\bullet\in\Pi_s}\mathfrak{C}(q^\bullet)$. Thus, we have equality. □

Remark 5.6. i) Because of the uniqueness of the optimal semicoupling the limit of the sequence Q^r does not depend on the choice of fundamental region. The approximating sequence $(Q^r)_{r\in\mathbb{N}}$ does of course depend on B_0 and also the choice of generating set S that defines the Cayley graph.

ii) In the construction of the semicoupling Q^∞ we only used finite transportation cost, invariance of Q_A in sense that $(\tau_h)_*Q_A \overset{d}{=} Q_{hA}$ and the amenability assumption on G. The only specific property of $\lambda^\bullet$ and $\mu^\bullet$ that we used is the uniqueness of the semicoupling on bounded sets which makes is easy to choose a good optimal semicoupling Q_{gB_r}. Hence, we can use the same algorithm to construct an optimal coupling between two arbitrary random measures. In particular this shows, that $\mathfrak{c}_\infty = \mathfrak{c}_{i,\infty}$ (see also Proposition 2.23).
Indeed, given two arbitrary equivariant measures $\nu^\bullet$ and $\mu^\bullet$ of unit respectively subunit intensity. For any $r\in\mathbb{N}$ let $Q_{B_r} = q^\bullet_{B_r}\mathbb{P}$ be an optimal semicoupling between $\nu^\bullet$ and $1_A\mu^\bullet$. In particular, we made some measurable choice of optimal semicoupling for each ω (they do not have to be unique), e.g. like in Corollary 5.22 of [Vil09]. *Define* Q_{gB_r} via $q^{\theta_g\omega}_{gB_r}(d(gx),d(gy)) := q^\omega_{B_r}(dx,dy)$. Due to equivariance, this is again a measurable choice of optimal semicouplings. Stationarity of $\mathbb{P}$ implies $(\tau_h)_*Q_{B_r} \overset{d}{=} Q_{hB_r}$. Hence, by the same construction there is some optimal semicoupling Q^∞ of $\nu^\bullet$ and $\mu^\bullet$ with cost bounded by $\mathfrak{c}_\infty$.

5.2 Quenched Limits

According to chapter 4, the unique optimal semicoupling between $\lambda^\bullet$ and $\mu^\bullet$ can be represented on $M\times M\times\Omega$ as

$$Q^\infty(dx,dy,d\omega) = \delta_{T(x,\omega)}(dy)\,\rho^\omega(x)\lambda^\omega(dx)\,\mathbb{P}(d\omega)$$

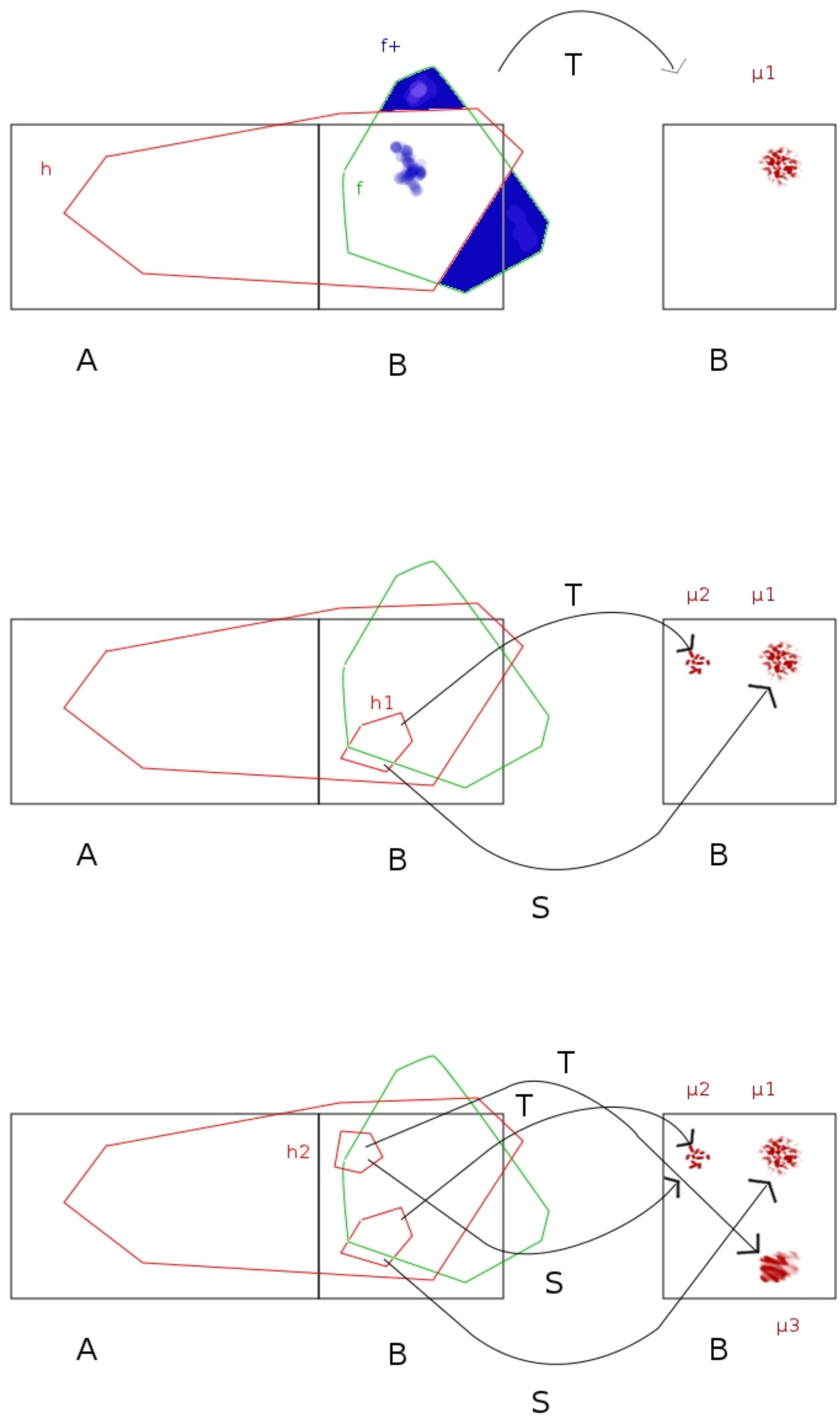

Figure 5.2: This little comic is to capture the iterative idea of the proof of Lemma 5.8. Note that the support of f_+ need not be disjoint of the support of h nor does any density have to be 0–1 valued.

by means of a measurable map
$$T : M \times \Omega \to M \cup \{\eth\},$$
defined uniquely almost everywhere and a density ρ^ω. Similarly, for each $g \in G$ and $r \in \mathbb{N}$ there exists a measurable map
$$T_{g,r} : M \times \Omega \to M \cup \{\eth\}$$
and a density $\rho^\omega_{g,r}$ such that the measure
$$Q_{gB_r}(dx, dy, d\omega) = \delta_{T_{g,r}(x,\omega)}(dy)\, \rho^\omega_{g,r} \lambda^\omega(dx)\, \mathbb{P}(d\omega)$$
on $M \times M \times \Omega$ is the unique optimal semicoupling of $\lambda^\bullet$ and $1_{gB_r}\mu^\bullet$.

Proposition 5.7. *For every $g \in G$*
$$T_{g,r}(x,\omega) \quad \to \quad T(x,\omega) \quad \textit{as} \quad r \to \infty \quad \textit{in } \lambda^\bullet \otimes \mathbb{P}\textit{-measure.}$$

The claim relies on the following two Lemmas. For the first one we use our amenability assumption once more. The second one is a slight modification (and extension) of a result in [Amb03].

Lemma 5.8. *i) Fix $\omega \in \Omega$ and take two disjoint bounded Borel sets $A, B \subset M$. Let $q^\omega_A = (id, T^\omega_A)_*(\rho^\omega_A \lambda^\omega)$ be the optimal semicoupling between λ^ω and $1_A\mu^\omega$. Similarly, let q^ω_B and $q^\omega_{A\cup B}$ be the unique optimal semicouplings between λ^ω and $1_B\mu^\omega$ respectively $1_{A\cup B}\mu^\omega$ with transport maps T^ω_B and $T^\omega_{A\cup B}$ and densities ρ^ω_B and $\rho^\omega_{A\cup B}$. Then, it holds that*
$$\rho^\omega_{A\cup B}(x) \geq \max\{\rho^\omega_A(x), \rho^\omega_B(x)\} \quad \lambda^\omega a.s..$$

ii) For any $g \in G$ and $r \in \mathbb{N}$ we have $\rho^\omega_{g,r}(x) \leq \rho^\omega(x) \quad (\lambda^\bullet \otimes \mathbb{P})$ a.s..

iii) For any $g \in G$ we have $\lim_{r\to\infty} \rho^\omega_{g,r}(x) \nearrow \rho^\omega(x) \quad (\lambda^\bullet \otimes \mathbb{P})$ a.s..

Proof. i) Firstly, note that if $\{\rho^\omega_A > 0\} \cap \{\rho^\omega_B > 0\} = \emptyset$ we have $\rho^\omega_{A\cup B} = \rho^\omega_A + \rho^\omega_B$. Because of the symmetry in A and B it is sufficient to prove that $\rho^\omega_{A\cup B} \geq \rho^\omega_B$. The proof is rather technical and involves an iterative choice of possibly different densities. The idea behind this iteration is sketched in figure 5.2.

For simplicity of notation we will suppress ω and write $f = \rho_B$ and $h = \rho_{A\cup B}$ and $T = T_B, S = T_{A\cup B}$. We will show the claim by contradiction. Assume there is a set D of positive λ measure such that $f(x) > h(x)$ on D. Put $f_+ := (f - h)_+$ and $\mu_1 := T_*(f_+\lambda)$. Let $h_1 \leq h$ be such that $S_*(h_1\lambda) = \mu_1$, that is h_1 is a subdensity of h such that $T_*(f_+\lambda) = S_*(h_1\lambda)$ (for finding this density we can use disintegration as in the proof of Proposition 4.9).

If $1_{\{h_1>0\}}h > f$ on some set D_1 of positive λ measure, we are done. Indeed, as f is the unique Cost minimizing choice for the semicoupling between λ and $1_B\mu$ the transport $S_*(1_{D_1}h_1\lambda) =: \tilde{\mu}_1$ must be more expensive than the respective transport $T_*(1_{\tilde{D}_1}f_+\lambda) = \tilde{\mu}_1$ for some suitable set $\tilde{D}_1$. Hence, $q_{A\cup B}$ cannot be minimizing and therefore not optimal, a contradiction.

If $1_{\{h_1>0\}}h \leq f$ we can assume wlog that $T_*(h_1\lambda) = \mu_2$ and μ_1 are singular to each other. Indeed, if they are not singular we can choose a different h_1 because $1_B\mu$ has to get its mass from somewhere. To be more precise, if $\tilde{h} \leq h_1$ is such that $T_*(\tilde{h}\lambda) \leq \mu_1$ we have $T_*((f_+ + \tilde{h})\lambda) > \mu_1$. Therefore, there must be some density h' such that $h' + h_1 \leq h$ and $S_*((h' + h_1)\lambda) = T_*((f_+ + \tilde{h})\lambda)$. Because, $f_+ > 0$ on some set of positive measure and $T_*(f\lambda) \leq S_*(h\lambda)$, there must be such an h_1 as claimed.

Take a density $h_2 \leq h$ such that $S_*(h_2\lambda) = \mu_2$. If $1_{\{h_2>0\}}h > f$ on some set D_2 of positive λ measure, we are done. Indeed, the optimality of q_B implies that the choice of f_+ and h_1 is cheaper than the choice of h_1 and h_2 for the transport into $\mu_1 + \mu_2$ (or maybe subdensities of these).

If $1_{\{h_2>0\}}h \le f$ and $\{h_2 > 0\} \cap \{f_+ > 0\}$ has positive λ measure, we get a contradiction of optimality of $q_{A\cup B}$ by cyclical monotonicity. Otherwise, we can again assume that $T_*(h_2\lambda) =: \mu_3$ and μ_2 are singular to each other. Hence, we can take a density $h_3 \le h$ such that $S_*(h_3\lambda) = \mu_3$. Proceeding in this manner, because $f_+\lambda(M) = h_i\lambda(M) > 0$ for all i and the finiteness of $q_B(M,M)$ one of the following two alternatives must happen

- there is j such that $1_{\{h_j>0\}}h > f$ on some set of positive λ measure.
- there are $j \neq i$ such that $\{h_j > 0\} \cap \{h_i > 0\}$ on some set of positive λ measure with $f_+ = h_0$.

Both cases lead to a contradiction by using the optimality of q_B, either by producing a cheaper semicoupling (in the first case) or by arguing via cyclical monotonicity (in the second case).

ii) Fix ω, g and r. Denote the density of the first marginal of $\tilde{Q}^l_f$ by $\zeta^\omega_{f,l}$. It is a convex combination of $\rho^\omega_{h,l}$ with $h \in f\Lambda_l$. For $h \in G$ with $d(g,h) \le n$ we have $g\Lambda_r \subset h\Lambda_{r+n}$. Hence, we have $\rho^\omega_{g,r} \le \rho^\omega_{h,r+n}$ by the first part of the Lemma. Therefore, the contribution of $\rho^\omega_{g,r}(x)$ to $\zeta^\omega_{g,r+n}(x)$ is at least the number of $h \in G$ such that $d(g,h) \le n$ divided by $|\Lambda_{r+n}|$. Hence,

$$\frac{|\Lambda_n|}{|\Lambda_{r+n}|}\rho^\omega_{g,r}(x) \le \zeta^\omega_{g,r+n}(x).$$

By the assumption (5.1) we have

$$\lim_{r\to\infty} \frac{|K\Lambda_r \triangle \Lambda_r|}{|\Lambda_r|} = 0,$$

for any finite $K \subset G$. If we take $K = \{h : d(h, id) = r\}$ we can conclude

$$\frac{|\Lambda_{n+r}|}{|\Lambda_n|} \le 1 + \frac{|K\Lambda_n \triangle \Lambda_n|}{|\Lambda_n|} \to 1 \text{ as } n \to \infty.$$

Fix $\epsilon > 0$. If $\rho^\omega_{g,r} > \epsilon + \rho^\omega$ on some positive $(\lambda^\bullet \otimes \mathbb{P})-$ set, we have that $\zeta^\omega_{g,r+n}(x) > \rho^\omega(x) + \epsilon/2$ on some positive $(\lambda^\bullet \otimes \mathbb{P})-$ set for all n such that $\frac{|\Lambda_n|}{|\Lambda_{n+r}|} \ge 1 - \epsilon/2$, because $\rho^\omega_{g,r} \le 1$ and thus $\rho^\omega \le 1-\epsilon$. Denote this set by A, so $A \subset M \times \Omega$. Then, we have $\tilde{Q}^{r+n}_g(A\times M) > Q^\infty(A\times M)+\epsilon/2$ for all n big enough. However, this is a contradiction to the vague convergence of $\tilde{Q}^r_g$ to Q^∞ which was shown in the last section.

iii) The last part allows to interpret $\rho^\omega_{g,r}$ as a density of $(\rho^\omega\lambda^\omega)$ instead of as a density of λ^ω. We will adopt this point of view and show that $\rho^\omega_{g,r}$ converges to 1 $(\lambda^\bullet \otimes \mathbb{P})$ a.s..
Assume that $\rho^\omega_{g,r}(x) \le \gamma < 1$ for all $r \in \mathbb{N}$. Moreover, assume that there is $k \in G$ and $s \in \mathbb{N}$ such that $\rho^\omega_{k,s}(x) > \gamma$. Then there is a $t \in \mathbb{N}$ such that $g\Lambda_t \supset k\Lambda_s$. The first part of the Lemma then implies that $\rho^\omega_{g,t}(x) \ge \rho^\omega_{k,s}(x) > \gamma$ which contradicts the assumption of $\rho^\omega_{g,r}(x) \le \gamma$. Hence, if we have $\rho^\omega_{g,r}(x) \le \gamma < 1$ for all $r \in \mathbb{N}$ on a set of positive $(\lambda^\bullet \otimes \mathbb{P})$ measure we must have $\rho^\omega_{k,s}(x) \le \gamma$ for all $k \in G$ ans $s \in \mathbb{N}$ on this set. Denote this set again by A, $A \subset M \times \Omega$. As $\zeta^\omega_{g,r}$ is a convex combination of the densities $\rho^\omega_{h,r}$ it must also be bounded away from 1 by γ on the set A. However, this is again a contradiction to the vague convergence of $\tilde{Q}^r_g$ to Q^∞.

□

Lemma 5.9. *Let X, Y be locally compact separable spaces, θ a Radon measure on X and ρ a metric on Y compatible with the topology.*
(i) For all $n \in \mathbb{N}$ let $T_n, T : X \to Y$ be Borel measurable maps. Put $Q_n(dx, dy) := \delta_{T_n(x)}(dy)\theta(dx)$ and $Q(dx, dy) := \delta_{T(x)}(dy)\theta(dx)$. Then,

$$T_n \to T \quad \text{in measure on } X \quad \Longleftrightarrow \quad Q_n \to Q \text{ vaguely in } \mathcal{M}(X \times Y).$$

(ii) More generally, let T and Q be as before whereas

$$Q_n(dx,dy) := \int_{X'} \delta_{T_n(x,x')}(dy)\,\theta'(dx')\,\theta(dx)$$

for some probability space $(X', \mathfrak{A}', \theta')$ and suitable measurable maps $T_n : X \times X' \to Y$. Then

$$Q_n \to Q \text{ vaguely in } \mathcal{M}(X \times Y) \quad \Longrightarrow \quad T_n(x,x') \to T(x) \quad \text{in measure on } X \times X'.$$

Proof. (i) Assume $T_n \to T$ in θ-measure. Then also $f \circ (Id, T_n) \to f \circ (Id, T)$ in θ-measure for any $f \in C_c(X \times Y)$. Then, every subsequence of $(f \circ (Id, T_n))_{n \in \mathbb{N}}$ has a further subsequence converging to $f \circ (Id, T)$ θ a.e.. As f has compact support and in particular is bounded this implies by the dominated convergence theorem

$$\int f(x,y)dQ_n = \int f(x, T_n(x))d\theta \to \int f(x, T(x))d\theta = \int f(x,y)dQ.$$

This proves the vague convergence of Q_n towards Q.
For the opposite direction, fix $\tilde{K} \subset X$ compact and $\varepsilon > 0$ and $\delta > 0$. By Lusin's theorem there is a compact set $K \subset \tilde{K}$ such that $T|_K$ is continuous and $\theta(\tilde{K} \backslash K) < \delta$. Put $\eta : \mathbb{R}_+ \to \mathbb{R}_+, t \mapsto 1 \wedge |t| / \varepsilon$. The function

$$\phi(x,y) = 1_K(x)\eta(\rho(y, T(x)))$$

is upper semicontinuous, nonnegative and compactly supported. Hence, there exist $\phi_l \in C_c(X \times Y)$ with $\phi_l \searrow \phi$. By assumption, we have for each l

$$\int \phi(x,y)Q_n(dx,dy) \leq \int \phi_l(x,y)Q_n(dx,dy) \overset{n\to\infty}{\to} \int \phi_l(x,y)Q(dx,dy).$$

Moreover,

$$\int \phi_l(x,y)Q(dx,dy) \overset{l\to\infty}{\to} \int \phi(x,y)Q(dx,dy) = 0.$$

Therefore, $\lim_{n\to\infty} \int \phi(x,y)Q_n(dx,dy) = 0$. In other words,

$$\lim_{n\to\infty} \int 1_K(x)\eta(\rho(T_n(x), T(x)))\theta(dx) = 0.$$

This implies $\lim_{n\to\infty} \theta(\{x \in K : \rho(T_n(x), T(x)) \geq \varepsilon\}) = 0$ and then in turn

$$\lim_{n\to\infty} \theta(\{x \in \tilde{K} : \rho(T_n(x), T(x)) \geq \varepsilon\}) \leq \delta.$$

As $\delta > 0$ was arbitrary, we can conclude

$$\lim_{n\to\infty} \theta(\{x \in \tilde{K} : \rho(T_n(x), T(x)) \geq \varepsilon\}) = 0.$$

(ii) Given any compact $\tilde{K} \subset X$ and any $\varepsilon > 0$, choose ϕ as before. Then vague convergence again implies $\lim_{n\to\infty} \int \phi(x,y)Q_n(dx,dy) = 0$. This, in other words, now reads as

$$\lim_{n\to\infty} \int_X \int_{X'} 1_K(x)\eta(\rho(T_n(x,x'), T(x)))\,\theta'(dx')\,\theta(dx) = 0.$$

Therefore,

$$\lim_{n\to\infty} (\theta \otimes \theta') \left(\left\{ (x,x') \in \tilde{K} \times X' : \; \rho(T_n(x,x'), T(x)) \geq \varepsilon \right\} \right) = 0.$$

This is the claim. $\square$

Proof of the Proposition. Firstly, we will show that the Proposition holds for 'sufficiently many' $g \in G$. We want to apply the previous Lemma. Recall that

$$\tilde{Q}_g^r \to Q^\infty \quad \text{vaguely on } M \times M \times \Omega,$$

where

$$Q^\infty(dx, dy, d\omega) = \delta_{T(x,\omega)}(dy)\rho^\omega(x)\lambda^\omega(dx)\mathbb{P}(d\omega)$$

and

$$\tilde{Q}_g^r(dx, dy, d\omega) = \frac{1}{|\Lambda_r|}\sum_{h \in gB_r} Q_{hB_r}(dx, dy, d\omega) = \frac{1}{|\Lambda_r|}\sum_{h \in gB_r} \delta_{T_{h,r}(x)}(dy)\rho_{h,r}^\omega(x)\lambda^\omega(dx)\mathbb{P}(d\omega),$$

with transport maps $T, T_{h,r} : M \times \Omega \to M \cup \{\eth\}$ and densities $\rho, \rho_{h,r} : M \times \Omega \to \mathbb{R}_+$. The Lemma above allows to interpret $\rho_{h,r}$ as density of the measure $\rho\lambda^\bullet$. Fix $k \in G$ and let θ_r' be the uniform measure on $k\Lambda_r$. Take $\theta = \rho\lambda^\bullet \otimes \mathbb{P}$, $X = M \times \Omega$ and $Y = M \cup \{\eth\}$. Apply the same reasoning as in the proof of the second assertion in the last lemma, however, now with changing θ', to get

$$\lim_{r\to\infty} (\theta \otimes \theta_r') \left(\left\{(x, \omega, h) \in \tilde{K} \times G \ : \ \rho_{h,r}^\omega(x) \cdot d(T_{h,r}(x, \omega), T(x, \omega)) \geq \epsilon\right\}\right) = 0. \tag{5.2}$$

Let $H \subset G$ be those h for which

$$\lim_{r\to\infty} \theta \left(\left\{(x, \omega) \in \tilde{K} \ : \ \rho_{h,r}^\omega(x) d(T_{h,r}(x, \omega), T(x, \omega)) \geq \epsilon\right\}\right) > 0.$$

Because we know that (5.2) holds, we must have $\lim_{r\to\infty} \theta_r'(H) = 0$. Hence, there are countably many $g \in G$ such that

$$\lim_{r\to\infty} \theta \left(\left\{(x, \omega) \in \tilde{K} \ : \ d(T_{g,r}(x, \omega), T(x, \omega)) \geq \epsilon\right\}\right) = 0,$$

where we used that $\rho_{g,r}^\omega \nearrow 1$ for $(\lambda^\bullet \otimes \mathbb{P})$ a.e. (x, ω), according to the Lemma above. This shows that the Proposition holds for those g.
Pick one such $g \in G$. Then the first part of the previous lemma implies

$$Q_{gB_r} \to Q^\infty \quad \text{vaguely on } M \times M \times \Omega.$$

This in turn implies that for any $h \in G$ $(\tau_h)_* Q_{gB_r} \to (\tau_h)_* Q^\infty \stackrel{(d)}{=} Q^\infty$ by invariance of Q^∞. Moreover, by Corollary 3.12 we have $(\tau_h)_* Q_{gB_r} \stackrel{(d)}{=} Q_{hgB_r}$. This means, that for any $h \in G$ we have

$$Q_{hgB_r} \to Q^\infty \quad \text{vaguely on } M \times M \times \Omega.$$

Applying once more the first part of the previous Lemma proves the Proposition. □

Corollary 5.10. *There is a measurable map $\Psi : \mathcal{M}(M) \times \mathcal{M}(M) \to \mathcal{M}(M \times M)$ s.t. $q^\omega := \Psi(\lambda^\omega, \mu^\omega)$ denotes the unique optimal semicoupling between λ^ω and μ^ω. In particular the optimal semicoupling is a factor.*

Proof. We showed that the optimal semicoupling Q^∞ can be constructed as the unique limit point of a sequence of deterministic functions of $\lambda^\bullet$ and $\mu^\bullet$. Hence, the map $\omega \mapsto q^\omega$ is measurable with respect to the sigma algebra generated by $\lambda^\bullet$ and $\mu^\bullet$. Thus, there is a measurable map Ψ such that $q^\bullet = \Psi(\lambda^\bullet, \mu^\bullet)$. □

5.2.1 Semicouplings of $\lambda^\bullet$ and a point process.

If $\mu^\bullet$ is known to be a point process the above convergence result can be significantly improved.

Theorem 5.11. *For any $g \in G$ and every bounded Borel set $A \subset M$*

$$\lim_{r\to\infty} (\lambda^\bullet \otimes \mathbb{P}) (\{(x,\omega) \in A \times \Omega \ : \ T_{g,r}(x,\omega) \neq T(x,\omega)\}) \ = \ 0.$$

Proof. Let A be as above and $\varepsilon > 0$ be given. Finiteness of the asymptotic mean transportation cost implies that there exists a bounded set $A' \subset M$ such that

$$(\lambda^\bullet \otimes \mathbb{P}) \left(\left\{(x,\omega) \in A \times \Omega : \ T(x,\omega) \notin A'\right\}\right) \leq \varepsilon.$$

Given the bounded set A' there exists $\delta > 0$ such that the probability to find two distinct particles of the point process at distance $< \delta$, at least one of them within A', is less than ε, i.e.

$$\mathbb{P}\left(\left\{\omega : \exists (y,y') \in A' \times M : \ 0 < d(y,y') < \delta, \ \mu^\omega(\{y\}) > 0, \ \mu^\omega(\{y'\}) > 0\right\}\right) \leq \varepsilon.$$

On the other hand, Proposition 5.7 states that with high probability the maps T and $T_{g,r}$ have distance less than δ. More precisely, for each $\delta > 0$ there exists r_0 such that for all $r \geq r_0$

$$(\lambda^\bullet \otimes \mathbb{P}) (\{(x,\omega) \in M \times \Omega : \ d(T_{g,r}(x,\omega), T(x,\omega)) \geq \delta\}) \leq \varepsilon.$$

Since all the maps T and $T_{g,r}$ take values in the support of the point process (plus the point $\eth$) it follows that

$$(\lambda^\bullet \otimes \mathbb{P}) (\{(x,\omega) \in M \times \Omega : \ T_{g,r}(x,\omega) \ \neq \ T(x,\omega)\}) \leq 3\varepsilon$$

for all $r \geq r_0$. $\square$

Corollary 5.12. *There exists a subsequence $(r_l)_l$ such that*

$$T_{g,r_l}(x,\omega) \quad \to \quad T(x,\omega) \qquad \text{as} \quad l \to \infty$$

for almost every $x \in M$, $\omega \in \Omega$ and every $g \in G$. Indeed, the sequence $(T_{g,r_l})_l$ is finally stationary. That is, there exists a random variable $l_g : M \times \Omega \to \mathbb{N}$ such that almost surely

$$T_{g,r_l}(x,\omega) \quad = \quad T(x,\omega) \qquad \text{for all} \ \ l \geq l_g(x,\omega).$$

Chapter 6

The other semicouplings

In the previous chapters we studied semicouplings between two equivariant random measures $\lambda^\bullet$ and $\mu^\bullet$ with intensities 1 and $\beta \leq 1$ respectively. In this chapter we want to study the case that $\mu^\bullet$ has intensity $\beta > 1$. Then, $q^\bullet$ is a semicoupling between $\lambda^\bullet$ and $\mu^\bullet$ iff for all $\omega \in \Omega$

$$(\pi_1)_* q^\omega = \lambda^\omega \quad \text{and} \quad (\pi_2)_* q^\omega \leq \mu^\omega.$$

This will complete the picture of semicouplings with one marginal being absolutely continuous. In the terminology of section 2.2 we should better talk about semicouplings between $\mu^\bullet$ and $\lambda^\bullet$. However, we prefer to keep $\lambda^\bullet$ as first marginal as it better suits our intuition of transporting a continuous quantity somewhere. We will not repeat all proofs but mostly just stress the parts where something changes. In general, it will get easier because we do not have to worry about densities.

6.1 Semicouplings on bounded sets

Lemma 6.1. *Let $\rho \in L^1(M, m)$ be a nonnegative density. Let μ be an arbitrary measure on M with $\mu(M) \geq (\rho \cdot m)(M)$. Then, there is a unique semicoupling q between $(\rho \cdot m)$ and μ minimizing $\mathsf{Cost}(\cdot)$. Moreover, $q = (id, T)_*(\rho \cdot m)$ for some measurable cyclically monotone map T.*

Proof. The existence of one Cost minimizing semicoupling q goes along the same lines as for example in Lemma 3.2. Let q_1 be one such minimizer. As q_1 is minimizing it has to be an optimal coupling between its marginals. Therefore, it is induced by a map, that is $q_1 = (id, T_1)_*(\rho \cdot m)$. Let $q_2 = (id, T_2)_*(\rho \cdot m)$ be another minimizer. Then, $q_3 = \frac{1}{2}(q_1 + q_2)$ is minimizing as well. Hence, $q_3 = (id, T_3)_*(\rho \cdot m)$. However, just as in the proof of Lemma 3.2 this implies $T_1 = T_2$ (ρm) almost everywhere and therefore $q_1 = q_2$. $\square$

Given a pair of random measures $(\lambda^\bullet, \mu^\bullet)$. We can apply this result to a fixed realization $(\lambda^\omega, \mu^\omega)$. For any bounded Borel set $A \subset M$, there is a unique semicoupling q_A^ω between $1_A \lambda^\omega$ and μ^ω minimizing $\mathsf{Cost}(\cdot)$. We can argue as in Lemma 3.10 to conclude

Lemma 6.2. *For each bounded Borel set $A \subset M$, the map $\omega \mapsto q_A^\omega$ is measurable.*

This allows to deduce, just as before:

Theorem 6.3. *(i) For each bounded Borel set $A \subset M$ there exists a unique semicoupling Q_A of $(1_A \lambda^\bullet)\mathbb{P}$ and $\mu^\bullet \mathbb{P}$ which minimizes the mean cost functional $\mathfrak{Cost}(.)$.*
(ii) The measure Q_A can be disintegrated as $Q_A(dx, dy, d\omega) := q_A^\omega(dx, dy)\, \mathbb{P}(d\omega)$ where for $\mathbb{P}$-a.e. ω the measure q_A^ω is the unique minimizer of the cost functional $\mathsf{Cost}(.)$ among the semicouplings of $1_A \lambda^\omega$ and μ^ω.
(iii) $\mathfrak{Cost}(Q_A) = \int_\Omega \mathsf{Cost}(q_A^\omega)\, \mathbb{P}(d\omega)$.

6.2 Uniqueness

We will again prove uniqueness by showing that every optimal semicoupling is induced by a transport map.

Lemma 6.4. *Given a semicoupling q^ω of λ^ω and μ^ω for fixed $\omega \in \Omega$, then the following properties are equivalent.*

(i) For each bounded Borel set $A \subset M$, the measure $1_{A\times M} q^\omega$ is the unique optimal coupling of the measures $1_A \lambda^\omega$ and $\mu^\omega_A(\cdot) := q^\omega(A, \cdot)$.

(ii) The support of q^ω is c-cyclically monotone.

(iii) There exists a c-cyclically monotone map $T^\omega : M \to M$ such that

$$q^\omega = (Id, T^\omega)_* \lambda^\omega.$$

Proof. The proof goes along the same lines as the proof of Proposition 4.1. The part (i) implies (iii) will even get easier as only indicator functions appear and not general densities, e.g. $1_{K_n\times M} q^\omega$ is the unique optimal coupling between $1_{K_n}\lambda^\omega \ll m$ and $\mu^\omega_{K_n}$. We omit the details. □

Definition 6.5. *A semicoupling $q^\bullet \mathbb{P}$ between $\lambda^\bullet$ and $\mu^\bullet$ is called*

i) locally optimal iff some (hence every) property of the previous proposition is satisfied for $\mathbb{P}$-a.e. $\omega \in \Omega$.

ii) efficient on $A \subset M$ bounded Borel iff

$$\mathfrak{e}_A(q^\bullet) := \frac{\mathfrak{Cost}(1_A\lambda^\bullet,\, \mu^\bullet_A)}{\mathfrak{Cost}(1_{A\times M} q^\bullet)} = 1.$$

iii) optimal iff it is asymptotically optimal and equivariant.

Copying the proofs basically line to line we get

Proposition 6.6. *i) $q^\bullet$ is locally optimal iff it is efficient on all bounded Borel sets $A \subset M$.*

ii) Every optimal semicoupling between $\lambda^\bullet$ and $\mu^\bullet$ is locally optimal.

iii) There exists at most one optimal semicoupling between $\lambda^\bullet$ and $\mu^\bullet$.

Proof of part (iii). Take two optimal semicouplings $q_1^\bullet$ and $q_2^\bullet$. Then, $q^\bullet = \frac{1}{2}q_1^\bullet + \frac{1}{2}q_2^\bullet$ is optimal as well. Hence, by part (ii) all three semicouplings – $q_1^\bullet$, $q_2^\bullet$ and $q^\bullet$ – are locally optimal. Thus, for a.e. ω by the results of the Lemma above there exist maps $T_1^\omega, T_2^\omega, T^\omega$ such that

$$\begin{aligned} \delta_{T^\omega(x)}(dy)\ \lambda^\omega(dx) &= q^\omega(dx, dy) \\ &= \left(\frac{1}{2}\delta_{T_1^\omega(x)}(dy) + \frac{1}{2}\delta_{T_2^\omega(x)}(dy)\right) \lambda^\omega(dx) \end{aligned}$$

This, however, implies $T_1^\omega(x) = T_2^\omega(x)$ for a.e. $x \in M$. Thus $q_1^\omega = q_2^\omega$. □

Remark 6.7 (Geometry of tessellation induced by semicouplings)**.** The results about the cells remain true, as the transport maps are still c-cyclically monotone. In particular, considering the cost function $|x-y|^2$ on $\mathbb{R}^d$ an optimal semicoupling between the Lebesgue measure $\mathcal{L}$ and a Poisson point process of intensity 42 will induce a tessellation of $\mathbb{R}^d$ into convex polytopes. However, the different cells will have mass *at most* one.

6.3 Construction

Just as in chapter 5 we will assume that the balls Λ_r are Følner sets. The construction is very similar. We only have to switch the roles of $\lambda^\bullet$ and $\mu^\bullet$. Apart from this the proofs remain unchanged as we did not use any specific properties of $\lambda^\bullet$ or $\mu^\bullet$.
Put

$$Q_g^r(dx, dy, d\omega) = 1_{gB_0}(x)\frac{1}{|\Lambda_r|}\sum_{h\in g\Lambda_r} Q_{hB_r}(dx, dy, d\omega).$$

The very same proof as before (Lemma 5.1) yields

Lemma 6.8. *(i) For each $r \in \mathbb{N}$ and $g \in G$*

$$\int_{gB_0\times M\times\Omega} c(x, y)Q_g^r(dx, dy, d\omega) \le \mathfrak{c}_\infty.$$

(ii) The family $(Q_g^r)_{r\in\mathbb{N}}$ of probability measures on $M \times M \times \Omega$ is relatively compact in the weak topology.

(iii) There exist probability measures Q_g^∞ and a subsequence $(r_l)_{l\in\mathbb{N}}$ such that for all $g \in G$:

$$Q_g^{r_l} \quad\longrightarrow\quad Q_g^\infty \qquad \textit{weakly as } l \to \infty.$$

Arguing as in the proof for Theorem 5.3 yields (interchanging the roles of $\lambda^\bullet$ and $\mu^\bullet$)

Theorem 6.9. *The measure $Q^\infty := \sum_{g\in G} Q_g^\infty$ is an optimal semicoupling of $\lambda^\bullet$ and $\mu^\bullet$.*

As before we can lift the restriction in the definition of Q_g^r to get

Corollary 6.10. *(i) For $r \to \infty$, the sequence of measures $Q^r := \sum_{g\in G} Q_g^r$, $r \in \mathbb{N}$, converges vaguely to the unique optimal semicoupling Q^∞.*
(ii) For each $g \in G$ and $r \in \mathbb{N}$ put

$$\tilde{Q}_g^r := \frac{1}{|\Lambda_r|}\sum_{h\in g\Lambda_r} Q_{hB_r}(dx, dy, d\omega).$$

The sequence $(\tilde{Q}_g^r)_{r\in\mathbb{N}}$ converges vaguely to the unique optimal semicoupling Q^∞.

The proof for the quenched results is even slightly easier than in the $\beta \le 1$ case, as we know that all the mass of λ^ω will be transported. Hence, we can directly apply Lemma 5.9:
The unique optimal semicoupling between $\lambda^\bullet$ and $\mu^\bullet$ can be represented on $M \times M \times \Omega$ as

$$Q^\infty(dx, dy, d\omega) = \delta_{T(x,\omega)}(dy)\,\lambda^\omega(dx)\,\mathbb{P}(d\omega)$$

by means of a measurable map

$$T : M \times \Omega \to M,$$

defined uniquely almost everywhere. Similarly, for each $g \in G$ and $r \in \mathbb{N}$ there exists a measurable map

$$T_{g,r} : M \times \Omega \to M$$

such that the measure

$$Q_{gB_r}(dx, dy, d\omega) = \delta_{T_{g,r}(x,\omega)}(dy)\,1_{gB_r}\lambda^\omega(dx)\,\mathbb{P}(d\omega)$$

on $M \times M \times \Omega$ is the unique optimal semicoupling of $1_{gB_r}\lambda^\bullet$ and $\mu^\bullet$.

Proposition 6.11. *For every $g \in G$*

$$T_{g,r}(x,\omega) \quad \to \quad T(x,\omega) \quad \text{as} \quad r \to \infty \quad \text{in } \lambda^\bullet \otimes \mathbb{P}\text{-measure.}$$

Proof. We want to use the second part of Lemma 5.9 with $X = M \times \Omega, Y = M, \theta = \lambda^\bullet \otimes \mathbb{P}$ and θ'_r the uniform measure on $k\Lambda_r$ for some fixed $k \in G$. By the same reasoning as in the proof of Proposition 5.7 or as in the proof of the second assertion of Lemma 5.9 we can deduce for any compact $K \subset M \times \Omega$:

$$\lim_{r\to\infty} \theta\left(\{(x,\omega) \in K \ : \ d(T_{g,r}(x,\omega), T(x,\omega)) \geq \epsilon\}\right) = 0.$$

□

In particular this implies the following Corollaries, whose proofs can be copied line to line from respective results in chapter 5.

Corollary 6.12. *The optimal semicoupling between $\lambda^\bullet$ and $\mu^\bullet$ is a factor.*

Corollary 6.13. *If $\mu^\bullet$ is a point process,*

i) for any $g \in G$ and every bounded Borel set $A \subset M$

$$\lim_{r\to\infty} (\lambda^\bullet \otimes \mathbb{P})\left(\{(x,\omega) \in A \times \Omega \ : \ T_{g,r}(x,\omega) \neq T(x,\omega)\}\right) \ = \ 0.$$

ii) there exists a subsequence $(r_l)_l$ such that

$$T_{g,r_l}(x,\omega) \quad \to \quad T(x,\omega) \qquad \text{as} \quad l \to \infty$$

for almost every $x \in M$, $\omega \in \Omega$ and every $g \in G$. Indeed, the sequence $(T_{g,r_l})_l$ is finally stationary. That is, there exists a random variable $l_g : M \times \Omega \to \mathbb{N}$ such that almost surely

$$T_{g,r_l}(x,\omega) \quad = \quad T(x,\omega) \qquad \text{for all} \ \ l \geq l_g(x,\omega).$$

Remark 6.14. Given a pair of equivariant random measure $(\lambda^\bullet, \mu^\bullet)$ such that $\lambda^\bullet$ is absolutely continuous as usual and $\mu^\bullet$ is a simple point process of intensity $\beta > 1$. One could define a semicoupling $q^\bullet$ between $\lambda^\bullet$ and $\mu^\bullet$ also by requiring that for $\mathbb{P}$ a.e. $\omega \in \Omega$

$$q^\omega(\cdot, M) = \lambda^\omega(\cdot) \quad \text{ and } \quad q^\omega(M,\cdot) = \sum_{\xi \in \text{supp}(\mu^\omega)} \rho^\omega(\xi)\delta_\xi,$$

with some "density" $\rho : \text{supp}(\mu^\omega) \to \{0,1\}$. This means, we are looking for the cheapest way to build a sub-point process of intensity one from $\mu^\bullet$. If β is very large this thinning should (if it exists) give something which is close to a lattice. One can interpret this semicoupling as a way to distribute resources to consumers or production sites with the constraint that the resource is just useful for the recipient if she/he gets it in a complete unit, e.g. if you want to produce a car you need a certain amount of steel. If you only get half of what you need, you cannot build the car.

Chapter 7

Cost estimate

In this chapter we show the estimates of the transportation cost between the Lebesgue measure $\lambda = \mathcal{L}$ on $\mathbb{R}^d$ and an equivariant Poisson point process $\mu^\bullet$ of intensity $\beta \in (0,\infty)$, i.e. we prove the missing parts of Theorem 1.7 and Corollary 1.8. Furthermore, we will show transportation cost estimates between $\mathcal{L}$ and a compound Poisson process. Hence, we take $M = \mathbb{R}^d, G = \mathbb{Z}^d$ acting on $\mathbb{R}^d$ by translation and as a fundamental region we can choose $B_0 = [0,1)^d$. We will often call invariant measures translation invariant. The presentation of the Poisson estimates follows closely the respective section in [HS10].

The asymptotic mean transportation cost for $\mu^\bullet$ will be denoted by

$$\mathfrak{c}_\infty = \mathfrak{c}_\infty(\vartheta, d, \beta)$$

or, if $\vartheta(r) = r^p$, by $\mathfrak{c}_\infty(p,d,\beta)$. We will present sufficient as well as necessary conditions for finiteness of $\mathfrak{c}_\infty$. These criteria will be quite sharp. Moreover, in the case of L^p-cost, we also present explicit sharp estimates for $\mathfrak{c}_\infty$. Note that, because of Lemma 2.16 (see also Corollary 5.5) we have

$$\mathfrak{c}_\infty = \liminf_{r\to\infty} \inf_{q\in\Pi_s} \frac{1}{2^{dr}} \mathbb{E}\left[\int_{\mathbb{R}^d\times[0,2^r)^d} c(x,y)q^\bullet(dx,dy)\right].$$

To begin with, let us summarize some elementary monotonicity properties of $\mathfrak{c}_\infty(\vartheta,d,\beta)$.

Lemma 7.1. *i) $\vartheta \le \overline{\vartheta}$ implies $\mathfrak{c}_\infty(\vartheta,d,\beta) \le \mathfrak{c}_\infty(\overline{\vartheta},d,\beta)$.*
More generally, $\limsup_{r\to\infty} \frac{\overline{\vartheta}(r)}{\vartheta(r)} < \infty$ and $\mathfrak{c}_\infty(\vartheta,d,\beta) < \infty$ imply $\mathfrak{c}_\infty(\overline{\vartheta},d,\beta) < \infty$.

ii) If $\overline{\vartheta} = \varphi \circ \vartheta$ for some convex increasing $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ then $\varphi\left(\beta^{-1}\mathfrak{c}_\infty(\vartheta,d,\beta)\right) \le \beta^{-1}\mathfrak{c}_\infty(\overline{\vartheta},d,\beta)$.

iii) $\beta \le \overline{\beta} \le 1$ implies $\mathfrak{c}_\infty(\vartheta,d,\beta) \le \mathfrak{c}_\infty(\vartheta,d,\overline{\beta})$.

iv) $\beta \ge \overline{\beta} \ge 1$ implies $\mathfrak{c}_\infty(\vartheta,d,\beta) \le \mathfrak{c}_\infty(\vartheta,d,\overline{\beta})$.

Proof. i) is obvious. ii) If $\overline{q}$ denotes the optimal semicoupling for $\overline{\vartheta}$ then Jensen's inequality implies

$$\begin{aligned}\beta^{-1}\mathfrak{c}_\infty(\overline{\vartheta},d,\beta) &= \beta^{-1}\mathbb{E}\int_{\mathbb{R}^d\times[0,1)^d} \varphi\left(\vartheta(|x-y|)\right)\overline{q}(dx,dy)\\ &\ge \varphi\left(\beta^{-1}\mathbb{E}\int_{\mathbb{R}^d\times[0,1)^d} \vartheta(|x-y|)\,\overline{q}(dx,dy)\right) \ge \varphi(\beta^{-1}\mathfrak{c}_\infty(\vartheta,d,\beta)).\end{aligned}$$

iii) Given a realization $\overline{\mu}^\omega$ of a Poisson point process with intensity $\overline{\beta}$. Delete each point $\xi \in \operatorname{supp}[\overline{\mu}^\omega]$ with probability $1-\beta/\overline{\beta}$, independently of each other. Then the remaining point

process μ^ω is a Poisson point process with intensity β (easy computation or see chapter 11.3 in [DVJ07]). Hence, each semicoupling $\overline{q}^\omega$ between $\mathcal{L}$ and $\overline{\mu}^\omega$ leads to a semicoupling q^ω between $\mathcal{L}$ and μ^ω with less or equal transportation cost: the centers which survive are coupled with the same cells as before.)
iv) Given a Poisson point process $\overline{\mu}^\bullet$ with intensity $\overline{\beta}$. Let $\overline{q}^\bullet$ be the optimal semicoupling between $\mathcal{L}$ and $\overline{\mu}^\bullet$. Take another Poisson point process $\tilde{\mu}^\bullet$ of intensity $(\beta - \overline{\beta})$ independent of $\overline{\mu}^\bullet$. Then, $\overline{q}^\bullet$ is also a semicoupling between $\mathcal{L}$ and $\mu^\bullet := \overline{\mu}^\bullet + \tilde{\mu}^\bullet$ which is a Poisson point process of intensity β. However, $\overline{q}^\bullet$ need not be optimal. $\square$

Remark 7.2. The same results hold for compound Poisson processes.

7.1 Lower Estimates

Theorem 7.3 ([HL01])**.** *Assume $\beta = 1$ and $d \leq 2$. Then for all equivariant couplings of Lebesgue and Poisson*

$$\mathbb{E}\left[\int_{\mathbb{R}^d \times [0,1)^d} |x-y|^{d/2}\, q^\bullet(dx,dy)\right] = \infty.$$

Theorem 7.4. *For all $\beta \in (0,\infty)$ and $d \geq 1$ there exists a constant $\kappa' = \kappa'(d,\beta)$ such that for all translation invariant semicouplings of Lebesgue and Poisson*

$$\mathbb{E}\left[\int_{\mathbb{R}^d \times [0,1)^d} \exp\left(\kappa' |x-y|^d\right)\, q^\bullet(dx,dy)\right] = \infty.$$

The result is well-known in the case $\beta = 1$. In this case, it is based on a lower bound for the event "no Poisson particle in the cube $[-r,r)^d$" and on a lower estimate for the cost of transporting the Lebesgue measure in $[-r/2,r/2)^d$ to some distribution on $\mathbb{R}^d \setminus [-r,r)^d$:

$$\mathfrak{c}_\infty \geq \exp\left(-(2r)^d\right) \cdot \vartheta\left(\frac{r}{2}\right) \cdot 2^{-d}.$$

Hence, $\mathfrak{c}_\infty \to \infty$ as $r \to \infty$ if $\vartheta(r) = \exp(\kappa' r^d)$ with $\kappa' > 2^{2d}$.
However, this argument breaks down in the case $\beta < 1$ because some of the Lebesgue mass will not be used. We will present a different argument which works for all $\beta \in (0,\infty)$.

Proof. $\beta \leq 1$ Consider the event "more than $(3r)^d$ Poisson particles in the box $[-r/2,r/2)^d$" or, formally,

$$\Omega(r) = \left\{\mu^\bullet\left([-r/2,r/2)^d\right) \geq (3r)^d\right\}.$$

Note that $\mathbb{E}\mu^\bullet\left([-r/2,r/2)^d\right) = \beta r^d$ with $\beta \leq 1$. For $\omega \in \Omega(r)$, the cost of a semicoupling between $\mathcal{L}$ and $1_{[-r/2,r/2)^d}\mu^\omega$ is bounded from below by

$$\vartheta(r/2) \cdot r^d$$

(since r^d Poisson points – or more – must be transported at least a distance $r/2$). The large deviation result formulated in the next lemma allows to estimate

$$\mathbb{P}(\Omega(r_n)) \geq e^{-k \cdot {r_n}^d}$$

for any $k > I_\beta(3^d)$ and suitable $r_n \to \infty$. Hence, if $\vartheta(r) \geq \exp(\kappa' r^d)$ with $\kappa' > 2^d \cdot I_\beta(3^d)$ then (normalizing with r^d) for suitable k

$$\mathfrak{c}_\infty \geq \mathbb{P}(\Omega(r)) \cdot \vartheta(r/2) \geq \exp((\kappa' 2^{-d} - k) r^d) \to \infty$$

as $r \to \infty$.

$\beta > 1$: We argue similarly as in the case $\beta \leq 1$ (Now we could also use the already known argument as all the Lebesgue mass will be transported). Consider the event "less than $r^d/3$ Poisson points in $A_r := [-r/2, r/2)^d$ *and* less than $r^d/3$ Poisson points in $A'_r := [-2^{1/d}r/2, 2^{1/d}r/2)^d \setminus [-r/2, r/2)^d$", formally

$$\Phi(r) = \{\mu^\bullet(A_r) \leq \frac{r^d}{3}, \mu^\bullet(A'_r) \leq \frac{r^d}{3}\}.$$

As $\mathcal{L}(A_r) = \mathcal{L}(A'_r)$ and $A_r \cap A'_r = \emptyset$, the independence of the Poisson point process on A_r and A'_r implies that $\mathbb{P}[\Phi(r)] = \mathbb{P}[\{\mu^\bullet(A_r) \leq \frac{r^d}{3}\}]^2$. For $\omega \in \Phi(r)$ the cost of a semicoupling between $\mathcal{L}$ and $\mu^\bullet$ is bounded from below by

$$\vartheta\left(\frac{2^{1/d}-1}{2}r\right) \cdot \frac{r^d}{3},$$

because Lebesgue measure of mass at least $r^d/3$ has to be transported at least a distance $\frac{2^{1/d}-1}{2}r$. The large deviation result formulated in the next Lemma allows to estimate

$$\mathbb{P}(\Phi(r_n)) \geq e^{-k \cdot r_n{}^d}$$

for any $k > 2 \cdot I_\beta(1/3)$ and suitable $r_n \to \infty$. Hence, if $\vartheta(r) \geq \exp(\kappa' r^d)$ with $\kappa' > (2^{d+1})/(2^{1/d} - 1)^d \cdot I_\beta(1/3)$ then (normalizing with r^d) for suitable k

$$\mathfrak{c}_\infty \geq \mathbb{P}(\Phi(r)) \cdot \vartheta\left(\frac{2^{1/d}-1}{2}r\right) \geq \exp((\kappa' \frac{(2^{1/d}-1)^d}{2^d} - k)r^d) \to \infty$$

as $r \to \infty$. $\square$

Lemma 7.5. *Given any nested sequence of boxes $B_n \subset \mathbb{R}^d$ of the form $B_n = B_n(z_n) = z_n + [0, 2^n)^d$ with $z_n \in \mathbb{Z}^d$. Put $I_\beta(t) = t\log(t/\beta) - t + \beta$.*

i) if $t \geq \beta$ we have

$$\lim_{n\to\infty} \frac{-1}{\mathcal{L}(B_n)} \log \mathbb{P}\left[\frac{1}{\mathcal{L}(B_n)}\mu^\bullet(B_n) \geq t\right] = I_\beta(t).$$

ii) if $t \leq \beta$ we have

$$\lim_{n\to\infty} \frac{-1}{\mathcal{L}(B_n)} \log \mathbb{P}\left[\frac{1}{\mathcal{L}(B_n)}\mu^\bullet(B_n) \leq t\right] = I_\beta(t).$$

Proof. For a fixed sequence B_n, $n \in \mathbb{N}$, consider the sequence of random variables $Z_n(.) = \mu^\bullet(B_n)$. For each $n \in \mathbb{N}$

$$Z_n = \sum_{i \in B_n \cap \mathbb{Z}^d} X_i$$

with $X_i = \mu^\bullet(B_0(i))$. The X_i are iid Poisson random variables with mean β. Hence, Cramér's Theorem states that for all $t \geq \beta$

$$\liminf_{n\to\infty} \frac{-1}{\mathcal{L}(B_n)} \log \mathbb{P}\left[\frac{1}{\mathcal{L}(B_n)} Z_n \geq t\right] \geq I_\beta(t)$$

with

$$I_\beta(t) = \sup_x [tx - \log\hat{\mu}(x)] = t\log(t/\beta) - t + \beta.$$

The second part is similar. $\square$

7.2 Upper Estimates for Concave Cost

In this section we treat the case of a concave scale function ϑ. In particular this implies that the cost function $c(x,y)=\vartheta(|x-y|)$ defines a metric on $\mathbb{R}^d$. The results of this section will be mainly of interest in the case $d\le 2$; in particular, they will prove assertion (ii) of Theorem 1.7. It suffices to consider the case $\beta=1$. Similar to the early work of Ajtai, Komlós and Tusnády [AKT84], our approach will be based on iterated transports between cuboids of doubled edge length.

We put

$$\Theta(r):=\int_0^r \vartheta(s)\mathrm{d}s \quad \text{and} \quad \varepsilon(r):=\sup_{s\ge r}\frac{\vartheta(s)}{s^{d/2}}.$$

7.2.1 Modified Cost

In order to prove the finiteness of the asymptotic mean transportation cost, we will estimate the cost of a semicoupling between $\mathcal{L}$ and $1_A\mu^\bullet$ from above in terms of the cost of another, related coupling.

Given two measure valued random variables $\nu_1^\bullet,\nu_2^\bullet:\Omega\to\mathcal{M}(\mathbb{R}^d)$ with $\nu_1^\omega(\mathbb{R}^d)=\nu_2^\omega(\mathbb{R}^d)$ for a.e. $\omega\in\Omega$ we define their transportation distance by

$$\mathbb{W}_\vartheta(\nu_1^\bullet,\nu_2^\bullet):=\int_\Omega W_\vartheta(\nu_1^\omega,\nu_2^\omega)\,\mathbb{P}(d\omega)$$

where

$$W_\vartheta(\eta_1,\eta_2)=\inf\left\{\int_{\mathbb{R}^d\times\mathbb{R}^d}\vartheta(|x-y|)\,q(dx,dy):\ q\text{ is a coupling of }\eta_1,\eta_2\right\}$$

denotes the usual L^1-Wasserstein distance – w.r.t. the distance $\vartheta(|x-y|)$ – between (not necessarily normalized) measures $\eta_1,\eta_2\in\mathcal{M}(\mathbb{R}^d)$ of equal total mass.

Lemma 7.6. *(i) For any triple of measure-valued random variables $\nu_1^\bullet,\nu_2^\bullet,\nu_3^\bullet:\Omega\to\mathcal{M}(\mathbb{R}^d)$ with $\nu_1^\omega(\mathbb{R}^d)=\nu_2^\omega(\mathbb{R}^d)=\nu_3^\omega(\mathbb{R}^d)$ for a.e. $\omega\in\Omega$ we have the triangle inequality*

$$\mathbb{W}_\vartheta(\nu_1^\bullet,\nu_3^\bullet)\le\mathbb{W}_\vartheta(\nu_1^\bullet,\nu_2^\bullet)+\mathbb{W}_\vartheta(\nu_2^\bullet,\nu_3^\bullet).$$

(ii) For each countable families of pairs of measure-valued random variables $\nu_{1,k}^\bullet,\nu_{2,k}^\bullet:\Omega\to\mathcal{M}(\mathbb{R}^d)$ with $\nu_{1,k}^\omega(\mathbb{R}^d)=\nu_{2,k}^\omega(\mathbb{R}^d)$ for a.e. $\omega\in\Omega$ and all k we have

$$\mathbb{W}_\vartheta\left(\sum_k\nu_{1,k}^\bullet\,,\,\sum_k\nu_{2,k}^\bullet\right)\le\sum_k\mathbb{W}_\vartheta\left(\nu_{1,k}^\bullet\,,\,\nu_{2,k}^\bullet\right).$$

Proof. Gluing lemma (cf. [Dud02] or [Vil09], chapter 1) plus Minkowski inequality yield (i); (ii) is obvious. □

For each bounded measurable $A\subset\mathbb{R}^d$ let us now define a random measure $\nu_A^\bullet:\Omega\to\mathcal{M}(\mathbb{R}^d)$ by

$$\nu_A^\omega:=\frac{\mu^\omega(A)}{\mathcal{L}(A)}\cdot 1_A\,\mathcal{L}.$$

Note that – by construction – the measures ν_A^ω and $1_A\,\mu^\omega$ have the same total mass. The *modified transportation cost* is defined as

$$\widehat{\mathsf{C}}_A(\omega)\ =\ \inf\left\{\int c(x,y)\widehat{q}(dx,dy)\ :\ \widehat{q}\text{ is a coupling of }\nu_A^\omega\text{ and }1_A\,\mu^\omega\right\}\ =\ W_\vartheta(\nu_A^\omega,1_A\mu^\omega).$$

Put

$$\widehat{\mathfrak{c}}_n\ =\ 2^{-nd}\cdot\mathbb{E}\left[\widehat{\mathsf{C}}_{B_n}\right]$$

with $B_n=[0,2^n)^d$.

The general idea for estimating the transportation cost is sketched in figure 7.1.

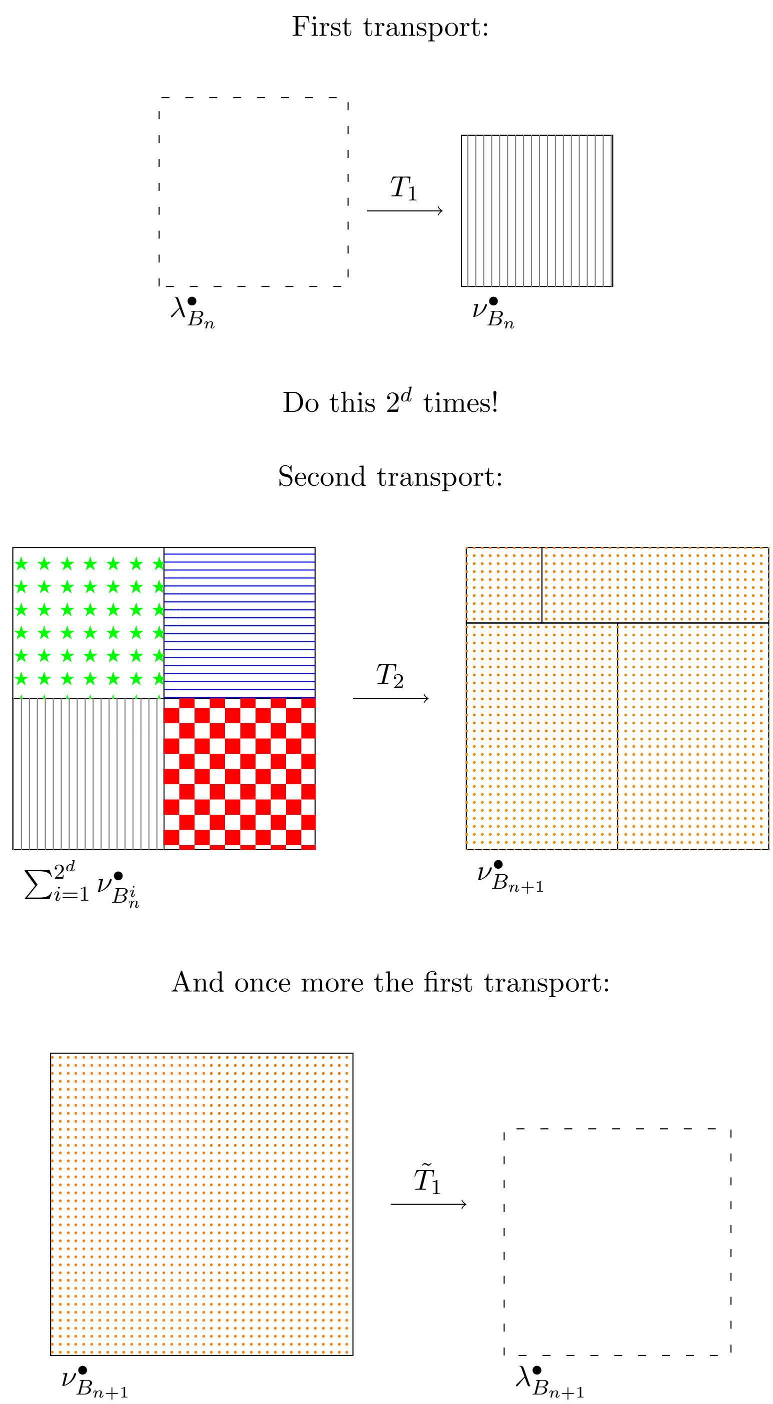

Figure 7.1: Sketch of transportation cost estimate.

7.2.2 Semi-Subadditivity of Modified Cost

The crucial advantage of this modified cost function $\widehat{\mathsf{C}}_A$ is that it is semi-subadditive (i.e. subadditive up to correction terms) on suitable classes of *cuboids* which we are going to introduce now. For $n \in \mathbb{N}_0, k \in \{1,\ldots,d\}$ and $i \in \{0,1\}^k$ put

$$B^i_{n+1} := [0,2^n)^k \times [0,2^{n+1})^{d-k} + 2^n \cdot (i_1,\ldots,i_k,0,\ldots,0)$$

These cuboids can be constructed by iterated subdivision of the standard cube B_{n+1} as follows: We start with $B_{n+1} = [0,2^{n+1})^d$ and subdivide it (along the first coordinate) into two disjoint congruent pieces $B^{(0)}_{n+1} = [0,2^n) \times [0,2^{n+1})^{d-1}$ and $B^{(1)}_{n+1} = B^{(0)}_{n+1} + 2^n \cdot (1,0,\ldots,0)$. In the k-th step, we subdivide each of the $B^i_{n+1} = B^{(i_1,\ldots,i_{k-1})}_{n+1}$ for $i \in \{0,1\}^{k-1}$ along the k-th coordinate into two disjoint congruent pieces $B^{(i_1,\ldots,i_{k-1},0)}_{n+1}$ and $B^{(i_1,\ldots,i_{k-1},1)}_{n+1}$. After d steps we are done. Each of the B^i_{n+1} for $i \in \{0,1\}^d$ is a copy of the standard cube $B_n = [0,2^n)^d$, more precisely,

$$B^i_{n+1} = B_n + 2^n \cdot i.$$

Lemma 7.7. *Given $n \in \mathbb{N}_0, k \in \{1,\ldots,d\}$ and $i \in \{0,1\}^k$ put $D_0 = B^{(i_1,\ldots,i_{k-1},0)}_{n+1}, D_1 = B^{(i_1,\ldots,i_{k-1},1)}_{n+1}$ and $D = D_0 \cup D_1 = B^{(i_1,\ldots,i_{k-1})}_{n+1}$. Then*

$$\mathbb{W}_\vartheta\left(\nu^\bullet_{D_0} + \nu^\bullet_{D_1}, \nu^\bullet_D\right) \le 2^{-(n+1)}\Theta(2^{n+1})2^{d/2(n+1)-k/2}.$$

Proof. Put $Z_j(\omega) := \mu^\omega(D_j)$ for $j \in \{0,1\}$. Then Z_0, Z_1 are independent Poisson random variables with parameter $\alpha_0 = \alpha_1 = \mathcal{L}(D_j) = 2^{d(n+1)-k}$ and $Z := \mu^\bullet(D) = Z_0 + Z_1$ is a Poisson random variable with parameter $\alpha = 2^{d(n+1)-k+1}$.
The measure $\nu^\bullet_D$ has density $\frac{Z}{\alpha}$ on D whereas the measure $\tilde\nu^\bullet_D := \nu^\bullet_{D_0} + \nu^\bullet_{D_1}$ has density $\frac{2Z_0}{\alpha}$ on the part $D_0 \subset D$ and it has density $\frac{2Z_1}{\alpha}$ on the remaining part $D_1 \subset D$. If $Z = 0$ nothing has to be transported since $\tilde\nu^\bullet_D$ already coincides with $\nu^\bullet_D$. Hence, for the sequel we may assume $Z > 0$. Assume that $Z_0 > Z_1$. Then a total amount of mass $\frac{Z_0 - Z_1}{2}$, uniformly distributed over D_0, will be transported with the map

$$T:\ (x_1,\ldots,x_{k-1},x_k,x_{k+1},\ldots,x_d) \mapsto (x_1,\ldots,x_{k-1},2^{n+1}-x_k,x_{k+1},\ldots,x_d)$$

from D_0 to D_1. The rest of the mass remains where it is. Hence, the cost of this transport is

$$\frac{|Z_0 - Z_1|}{2} \cdot 2^{-n} \int_0^{2^n} \vartheta(2^{n+1} - 2x_k)\,\mathrm{d}x_k = 2^{-(n+2)}\Theta(2^{n+1}) \cdot |Z_0 - Z_1|.$$

Hence, we get

$$\begin{aligned}
\mathbb{W}_\vartheta\left(\tilde\nu^\bullet_D, \nu^\bullet_D\right) &= 2^{-(n+2)}\Theta(2^{n+1}) \cdot \mathbb{E}\left[|Z_0 - Z_1|\right] \\
&\le 2^{-(n+1)}\Theta(2^{n+1}) \cdot \mathbb{E}\left[|Z_0 - \alpha_0|\right] \\
&\le 2^{-(n+1)}\Theta(2^{n+1}) \cdot \alpha_0^{1/2} = 2^{-(n+1)}\Theta(2^{n+1})2^{d/2(n+1)-k/2}.
\end{aligned}$$

□

Proposition 7.8. *For all $n \in \mathbb{N}$ and arbitrary dimension d it holds that*

$$\widehat{\mathfrak{c}}_{n+1} \quad \le \quad \widehat{\mathfrak{c}}_n + 2^{d/2+1} \cdot 2^{-(n+1)(d/2+1)}\Theta(2^{n+1}).$$

Proof. Let us begin with the trivial observations

$$\mathbb{W}_\vartheta\left(1_{B_{n+1}}\mu^\bullet, \nu^\bullet_{B_{n+1}}\right) = 2^{d(n+1)} \cdot \widehat{\mathfrak{c}}_{n+1}$$

and

$$\mathbb{W}_\vartheta\left(1_{B_{n+1}}\mu^\bullet\,,\,\sum_{i\in\{0,1\}^d}\nu^\bullet_{B^i_n}\right)\leq\sum_{i\in\{0,1\}^d}\mathbb{W}_\vartheta\left(1_{B^i_n}\mu^\bullet,\nu^\bullet_{B^i_n}\right)=2^d\cdot\mathbb{W}_\vartheta\left(1_{B_n}\mu^\bullet,\nu^\bullet_{B_n}\right)=2^{d(n+1)}\cdot\widehat{\mathfrak{c}}_n.$$

Hence, by the triangle inequality for $\mathbb{W}_\vartheta$ an upper estimate for $\widehat{\mathfrak{c}}_{n+1}\;-\;\widehat{\mathfrak{c}}_n$ will follow from an upper bound for $\mathbb{W}_\vartheta\left(\sum_{i\in\{0,1\}^d}\nu^\bullet_{B^i_n}\,,\,\nu^\bullet_{B_{n+1}}\right)$.
In order to estimate the cost of transportation from $\nu^\bullet_{(d)}:=\sum_{i\in\{0,1\}^d}\nu^\bullet_{B^i_n}$ to $\nu^\bullet_{(0)}:=\nu^\bullet_{B_{n+1}}$ for fixed $n\in\mathbb{N}_0$, we introduce $(d-1)$ further ('intermediate') measures

$$\nu^\bullet_{(k)}=\sum_{i\in\{0,1\}^k}\nu^\bullet_{B^i_{n+1}}$$

and estimate the cost of transportation from $\nu^\bullet_{(k)}$ to $\nu^\bullet_{(k-1)}$ for $k\in\{1,\dots,d\}$. For each k, these cost arise from *merging* 2^{k-1} *pairs of cuboids* into 2^{k-1} cuboids of twice the size. More precisely, from moving mass within pairs of adjacent cuboids in order to obtain equilibrium in the unified cuboid of twice the size. These cost – for each of the 2^{k-1} pairs involved – have been estimated in the previous lemma:

$$\mathbb{W}_\vartheta\left(\nu^\bullet_{(k)},\nu^\bullet_{(k-1)}\right)\leq 2^{k-1}\cdot\mathbb{W}_\vartheta\left(\nu^\bullet_{B^{i,0}_{n+1}}+\nu^\bullet_{B^{i,1}_{n+1}}\,,\,\nu^\bullet_{B^i_{n+1}}\right)\leq 2^{k-1}\cdot 2^{-(n+1)}\Theta(2^{n+1})2^{d/2(n+1)-k/2}$$

for $k\in\{1,\dots,d\}$ (and arbitrary $i\in\{0,1\}^{k-1}$). Thus

$$\begin{aligned}
2^{d(n+1)}\cdot[\widehat{\mathfrak{c}}_{n+1}-\widehat{\mathfrak{c}}_n] &\leq \mathbb{W}_\vartheta\left(1_{B_{n+1}}\mu^\bullet\,,\,\nu^\bullet_{(0)}\right)-\mathbb{W}_\vartheta\left(1_{B_{n+1}}\mu^\bullet\,,\,\nu^\bullet_{(d)}\right)\\
&\leq \sum_{k=1}^d\mathbb{W}_\vartheta\left(\nu^\bullet_{(k-1)}\,,\,\nu^\bullet_{(k)}\right)\\
&\leq \sum_{k=1}^d 2^{k/2}\cdot 2^{-(n+2)}\Theta(2^{n+1})2^{d/2(n+1)}\\
&\leq 4\cdot 2^{(n+2)(d/2-1)}\cdot\Theta(2^{n+1})
\end{aligned}$$

which yields the claim. □

Corollary 7.9. *If* $\sum_{n\geq 1}2^{-(n+1)(d/2+1)}\Theta(2^{n+1})<\infty$, *we have*

$$\widehat{\mathfrak{c}}_\infty\;:=\;\lim_{n\to\infty}\widehat{\mathfrak{c}}_n$$

exists and is finite.

Proof. According to the previous theorem

$$\lim_{n\to\infty}\widehat{\mathfrak{c}}_n\leq\widehat{\mathfrak{c}}_N+\sum_{m\geq N}2^{-(m+1)(d/2+1)}\Theta(2^{m+1}),\tag{7.1}$$

for each $N\in\mathbb{N}$. As the sum was assumed to converge the claim follows. □

7.2.3 Comparison of Costs

Proposition 7.10. *For all* $d\in\mathbb{N}$ *and for all* $n\in\mathbb{N}_0$

$$\mathfrak{c}_n\quad\leq\quad\widehat{\mathfrak{c}}_n+\sqrt{2d}\cdot\varepsilon(2^n).$$

Proof. Let a box $B = B_n = [0, 2^n)^d$ for some fixed $n \in \mathbb{N}_0$ be given. We define a measure-valued random variable $\lambda_B^\bullet : \Omega \to \mathcal{M}(\mathbb{R}^d)$ by

$$\lambda_B^\omega = 1_{\widehat{B}(\omega)} \cdot \mathcal{L}$$

with a randomly scaled box $\widehat{B}(\omega) = [0, Z(\omega)^{1/d})^d \subset \mathbb{R}^d$ and $Z(\omega) = \mu^\omega(B)$. Recall that Z is a Poisson random variable with parameter $\alpha = 2^{nd}$. Moreover, note that

$$\lambda_B^\omega(\mathbb{R}^d) = \mu^\omega(B) = \nu_B^\omega(\mathbb{R}^d)$$

and that $\lambda_B^\omega \leq \mathcal{L}$ for each $\omega \in \Omega$. Each coupling of λ_B^ω and $1_B\mu^\omega$, therefore, is also a semicoupling of $\mathcal{L}$ and $1_B\mu^\omega$. Hence,

$$2^{nd} \cdot \mathfrak{c}_n \leq \mathbb{W}_\vartheta(\lambda_B^\bullet, 1_B\mu^\bullet).$$

On the other hand, obviously,

$$2^{nd} \cdot \widehat{\mathfrak{c}}_n = \mathbb{W}_\vartheta(\nu_B^\bullet, 1_B\mu^\bullet)$$

and thus

$$2^{nd} \cdot (\mathfrak{c}_n - \widehat{\mathfrak{c}}_n) \leq \mathbb{W}_\vartheta(\nu_B^\bullet, \lambda_B^\bullet).$$

If $Z > \alpha$ a transport $T_*\nu_B = \lambda_B$ can be constructed as follows: at each point of B the portion $\frac{\alpha}{Z}$ of ν_B remains where it is; the rest is transported from B into $\widehat{B} \setminus B$. The maximal transportation distance is $\sqrt{d} \cdot Z^{1/d}$. Hence, the cost can be estimated by

$$\vartheta\left(\sqrt{d} \cdot Z^{1/d}\right) \cdot (Z - \alpha).$$

On the other hand, if $Z < \alpha$ in a similar manner a transport $T'_*\lambda_B = \nu_B$ can be constructed with cost bounded from above by

$$\vartheta\left(\sqrt{d} \cdot \alpha^{1/d}\right) \cdot (\alpha - Z).$$

Therefore, by definition of the function $\varepsilon(.)$

$$\begin{aligned}
\mathbb{W}_\vartheta(\nu_B^\bullet, \lambda_B^\bullet) &\leq \mathbb{E}\left[\vartheta\left(\sqrt{d}(Z \vee \alpha)^{1/d}\right) \cdot |Z - \alpha|\right] \\
&\leq \varepsilon\left(\alpha^{1/d}\right) \cdot \sqrt{d} \cdot \mathbb{E}\left[(Z \vee \alpha)^{1/2} \cdot |Z - \alpha|\right] \\
&\leq \varepsilon\left(\alpha^{1/d}\right) \cdot \sqrt{d} \cdot \mathbb{E}\left[Z + \alpha\right]^{1/2} \cdot \mathbb{E}\left[|Z - \alpha|^2\right]^{1/2} \\
&= \varepsilon\left(2^n\right) \cdot \sqrt{d} \cdot \left[2 \cdot 2^{nd} \cdot 2^{nd}\right]^{1/2}.
\end{aligned}$$

This finally yields

$$\mathfrak{c}_n - \widehat{\mathfrak{c}}_n \leq 2^{-nd} \cdot \mathbb{W}_\vartheta(\nu_B^\bullet, \lambda_B^\bullet) \leq \varepsilon(2^n) \cdot \sqrt{2d}.$$

□

Theorem 7.11. *Assume that*

$$\int_1^\infty \frac{\vartheta(r)}{r^{1+d/2}} dr < \infty \tag{7.2}$$

then

$$\mathfrak{c}_\infty \leq \widehat{\mathfrak{c}}_\infty < \infty.$$

Proof. Since

$$\int_1^\infty \frac{\vartheta(r)}{r^{1+d/2}}dr < \infty \quad \Longleftrightarrow \quad \sum_{n=1}^\infty \frac{\Theta(2^n)}{2^{n(1+d/2)}} < \infty,$$

Corollary 7.9 applies and yields $\widehat{\mathfrak{c}}_\infty < \infty$. Moreover, since ϑ is increasing, the integrability condition (7.2) implies that

$$\varepsilon(r) = \sup_{s\geq r} \frac{\vartheta(s)}{s^{d/2}} \ \to \ 0$$

as $r \to \infty$. Hence, $\mathfrak{c}_\infty \ \leq \ \widehat{\mathfrak{c}}_\infty$ by Proposition 7.10. $\square$

The previous Theorem essentially says that $\mathfrak{c}_\infty < \infty$ if ϑ grows 'slightly' slower than $r^{d/2}$. This criterion is quite sharp in dimensions 1 and 2. Indeed, according to Theorem 7.3 in these two cases we also know that $\mathfrak{c}_\infty = \infty$ if ϑ grows like $r^{d/2}$ or faster.

7.3 Estimates for L^p-Cost

The results of the previous section in particular apply to L^p-cost for $p < d/2$ in $d \leq 2$ and to L^p-cost for $p \leq 1$ in $d \geq 3$. A slight modification of these arguments will allow us to deduce cost estimates for L^p cost for arbitrary $p \geq 1$ in the case $d \geq 3$.
In this case, the finiteness of $\mathfrak{c}_\infty$ will also be covered by the more general results of [HP05] and [MT11], see Theorem 1.7 (i). However, using the idea of modified cost we get reasonably well quantitative estimates on $\mathfrak{c}_\infty$. Throughout this section we assume $\beta = 1$.

7.3.1 Some Moment Estimates for Poisson Random Variables

For $p \in \mathbb{R}$ let us denote by $\lceil p \rceil$ the smallest integer $\geq p$.

Lemma 7.12. *For each $p \in (0,\infty)$ there exist constants $C_1(p), C_2(p), C_3(p)$ such that for every Poisson random variable Z with parameter $\alpha \geq 1$:*

(i) $\mathbb{E}[Z^p] \leq C_1(p)\cdot\alpha^p$, *e.g.* $C_1(1) = 1,\ C_1(2) = 4$.
For general p one may choose $C_1(p) = \lceil p \rceil^p$ or $C_1(p) = 2^{p-1}\cdot(\lceil p \rceil - 1)!$.

(ii) $\mathbb{E}\left[Z^{-p}\cdot 1_{\{Z>0\}}\right] \leq C_2(p)\cdot\alpha^{-p}$.
For general p one may choose $C_2(p) = (\lceil p \rceil + 1)!$.

(iii) $\mathbb{E}[(Z-\alpha)^p] \leq C_3(p)\cdot\alpha^{p/2}$, *e.g.* $C_3(2) = 1,\ C_1(4) = 2$.
For general p one may choose $C_3 = 2^{p-1}\cdot(2\lceil \frac{p}{2} \rceil - 1)!$.

Proof. In all cases, by Hölder's inequality it suffices to prove the claim for integer $p \in \mathbb{N}$.

(i) The moment generating function of Z is $M(t) := \mathbb{E}[e^{tZ}] = \exp\big(\alpha(e^t - 1)\big)$. For integer p, the p-th moment of Z is given by the p-th derivative of M at the point $t = 0$, i.e. $\mathbb{E}[Z^p] = M^{(p)}(0)$. As a function of α, the p-th derivative of M is a polynomial of order p (with coefficients depending on t). As $\alpha \geq 1$ we are done.
To get quantitative estimates for C_1, observe that differentiating $M(t)$ p times yields at most 2^{p-1} terms, each of them having a coefficient $\leq (p-1)!$ (if we do not merge terms of the same order). Thus, we can take $C_1 = 2^{p-1}\cdot(p-1)!$.

Alternatively, we may use the recursive formula

$$T_{n+1}(\alpha) = \alpha \sum_{k=0}^n \binom{n}{k} T_k(\alpha)$$

for the *Touchard polynomials* $T_n(\alpha) := \mathbb{E}[Z^n]$, see e.g. [Tou56]. Assuming that $T_k(\alpha) \le (k\alpha)^k$ for all $k = 1, \dots, n$ leads to the corresponding estimate for $k = n+1$.

(iii) Put $p = 2k$ with integer k. The moment generating function of $(Z - \alpha)$ is

$$\begin{aligned} N(t) &:= \exp\big(\alpha(e^t - 1 - t)\big) \\ &= \exp\Big(\frac{\alpha}{2} t^2 h(t)\Big) = 1 + \frac{\alpha}{2} t^2 h(t) + \frac{1}{2}\Big(\frac{\alpha}{2}\Big)^2 t^4 h^2(t) + \frac{1}{6}\Big(\frac{\alpha}{2}\Big)^3 t^6 h^3(t) + \dots \end{aligned}$$

with $h(t) = \frac{2}{t^2}(e^t - 1 - t)$. Hence, the $2k$-th derivative of N at the point $t = 0$ is a polynomial of order k in α. Since $\alpha \ge 1$ by assumption, $\mathbb{E}[(Z-\alpha)^{2k}] = N^{(2k)}(0) \le C_3 \cdot \alpha^k$ for some C_3. To estimate C_3, again observe that differentiating $N(t)$ (2k) times yields at most 2^{2k-1} terms. Each of these terms has a coefficient $\le (2k-1)!$ (if we do not merge terms). Hence we can take $C_3(2k) = 2^{2k-1} \cdot (2k-1)!$.

(ii) The result follows from the inequality

$$\frac{1}{x^k} \le \frac{(k+1)!x!}{(k+x)!}$$

for positive integers k and x. The inequality is equivalent to

$$\binom{x+k}{x-1} \le x^{k+1}.$$

For fixed k the latter inequality holds for $x = 1$. If x increases from x to $x+1$ the right hand side grows by a factor of $\big(\frac{x+1}{x}\big)^{k+1}$ and the l.h.s. by a factor of $\frac{x+k+1}{x}$. As $(x+k+1)x^k \le (x+1)^{k+1}$, the inequality holds. Then, we can estimate

$$\begin{aligned} \mathbb{E}\left[\frac{1}{Z^k} \cdot 1_{Z>0}\right] &\le \mathbb{E}\left[\frac{(k+1)!}{(Z+1)\cdots(Z+k)} \cdot 1_{Z>0}\right] \\ &= e^{-\alpha} \cdot \sum_{j=1}^{\infty} \frac{\alpha^j}{j!} \cdot \frac{(k+1)!}{(j+1)\cdots(j+k)} = \frac{(k+1)!}{\alpha^k} \cdot e^{-\alpha} \cdot \sum_{j=1}^{\infty} \frac{\alpha^{j+k}}{(j+k)!} \le \frac{(k+1)!}{\alpha^k}. \end{aligned}$$

If we choose $k = \lceil p \rceil$ this yields the claim. □

7.3.2 L^p-Cost for $p \ge 1$ in $d \ge 3$

Given two measure valued random variables $\nu_1^\bullet, \nu_2^\bullet : \Omega \to \mathcal{M}(\mathbb{R}^d)$ with $\nu_1^\omega(\mathbb{R}^d) = \nu_2^\omega(\mathbb{R}^d)$ for a.e. $\omega \in \Omega$ we define their L^p-transportation distance by

$$\mathbb{W}_p(\nu_1^\bullet, \nu_2^\bullet) := \left[\int_\Omega W_p^p(\nu_1^\omega, \nu_2^\omega)\, \mathbb{P}(d\omega)\right]^{1/p}$$

where

$$W_p(\eta_1, \eta_2) = \inf\left\{\left[\int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p\, \theta(dx, dy)\right]^{1/p} : \ \theta \text{ is a coupling of } \eta_1, \eta_2\right\}$$

denotes the usual L^p-Wasserstein distance between measures $\eta_1, \eta_2 \in \mathcal{M}(\mathbb{R}^d)$ of equal total mass (they do not have to be normalized). Similar to the concave case the triangle inequality holds and we define the *modified transportation cost* as

$$\widehat{\mathsf{C}}_A(\omega) \ = \ \inf\left\{\int |x-y|^p \widehat{q}^\omega(dx, dy) \ : \ \widehat{q}^\omega \text{ is a coupling of } \nu_A^\omega \text{ and } 1_A\, \mu^\omega\right\} \ = \ W_p^p(\nu_A^\omega, 1_A \mu^\omega).$$

Put

$$\widehat{\mathfrak{c}}_n \ = \ 2^{-nd} \cdot \mathbb{E}\left[\widehat{\mathsf{C}}_{B_n}\right] = \mathbb{W}_p^p(\nu_{B_n}^\bullet, 1_{B_n}\mu^\bullet)$$

with $B_n = [0, 2^n)^d$ as usual.

Lemma 7.13. *Given $n \in \mathbb{N}_0, k \in \{1,\dots,d\}$ and $i \in \{0,1\}^k$ put $D_0 = B_{n+1}^{(i_1,\dots,i_{k-1},0)}, D_1 = B_{n+1}^{(i_1,\dots,i_{k-1},1)}$ and $D = D_0 \cup D_1 = B_{n+1}^{(i_1,\dots,i_{k-1})}$. Then for some constant κ_1 depending only on p:*

$$\mathbb{W}_p^p\left(\nu_{D_0}^\bullet + \nu_{D_1}^\bullet,\, \nu_D^\bullet\right) \le \kappa_1 \cdot 2^{(n+1)(p+d-pd/2)} \cdot 2^{k(p/2-1)+1}.$$

One may choose $\kappa_1(p) = \frac{1}{p+1}2^{-p} \cdot C_3(2p) \cdot C_2(2(p-1))$.

Proof. The proof will be a modification of the proof of Lemma 7.7. An optimal transport map $T : D \to D$ with $T_*\tilde{\nu}_D^\bullet = \nu_D^\bullet$ is now given by

$$T:\ (x_1,\dots,x_{k-1},x_k,x_{k+1},\dots,x_d) \mapsto (x_1,\dots,x_{k-1},\frac{2Z_0}{Z}\cdot x_k, x_{k+1},\dots,x_d)$$

on D_0 and

$$T:\ (x_1,\dots,x_{k-1},x_k,x_{k+1},\dots,x_d) \mapsto (x_1,\dots,x_{k-1},2^{n+1}-(2^{n+1}-x_k)\cdot\frac{2Z_1}{Z}, x_{k+1},\dots,x_d)$$

on D_1. (If $p > 1$ this is indeed the only optimal transport map.) The cost of this transport can easily be calculated:

$$\int_{D_0} |T(x)-x|^p\, d\tilde{\nu}^\bullet(x) = Z_0 \cdot 2^{-n}\int_0^{2^n}\left|\frac{2Z_0}{Z}\cdot x_k - x_k\right|^p dx_k = \frac{2^{np}}{p+1}\cdot Z_0 \cdot \left|\frac{Z_0-Z_1}{Z}\right|^p$$

and analogously

$$\int_{D_1} |T(x)-x|^p\, d\tilde{\nu}^\bullet(x) = \frac{2^{np}}{p+1}\cdot Z_1 \cdot \left|\frac{Z_0-Z_1}{Z}\right|^p.$$

Hence, together with the estimates from Lemma 7.12 this yields

$$\begin{aligned}
\mathbb{W}_p^p\left(\tilde{\nu}_D^\bullet, \nu_D^\bullet\right) &= \frac{2^{np}}{p+1}\cdot\mathbb{E}\left[\frac{|Z_0-Z_1|^p}{Z^{p-1}}\cdot 1_{\{Z>0\}}\right]\\
&\le \frac{2^{np}}{p+1}\cdot\mathbb{E}\left[|Z_0-Z_1|^{2p}\right]^{1/2}\cdot\mathbb{E}\left[Z^{-2(p-1)}\cdot 1_{\{Z>0\}}\right]^{1/2}\\
&\le \frac{2^{(n+1)p}}{p+1}\cdot\mathbb{E}\left[|Z_0-\alpha_0|^{2p}\right]^{1/2}\cdot\mathbb{E}\left[Z^{-2(p-1)}\cdot 1_{\{Z>0\}}\right]^{1/2}\\
&\le \frac{2^{(n+1)p}}{p+1}\cdot C_3\cdot\alpha_0^{p/2}\cdot C_2\cdot\alpha^{1-p}\\
&\le \kappa_1\cdot 2^{(n+1)(p+d-pd/2)}\cdot 2^{k(p/2-1)+1}
\end{aligned}$$

which is the claim. □

With the very same proof as before (Proposition 7.8), just insert different results, we get

Proposition 7.14. *For all $d \ge 3$ and all $p \ge 1$ there is a constant $\kappa_2 = \kappa_2(p,d)$ such that for all $n \in \mathbb{N}_0$*

$$\hat{\mathfrak{c}}_{n+1}^{1/p} \quad\le\quad \hat{\mathfrak{c}}_n^{1/p} + \kappa_2\cdot 2^{(n+1)(1-d/2)}.$$

In particular,

$$\hat{\mathfrak{c}}_\infty^{1/p} \quad\le\quad \hat{\mathfrak{c}}_n^{1/p} + \kappa_2\cdot\frac{2^{-(n+1)(d/2-1)}}{1-2^{-(d/2-1)}}.$$

One may choose $\kappa_2(p,d) = \kappa_1(p)^{1/p}\cdot\sum_{k=1}^d 2^{k/2} \le \kappa_1(p)^{1/p}\cdot 2^{d/2+2}$.

Corollary 7.15. *For all $d \geq 3$ and all $p \geq 1$*

$$\widehat{\mathfrak{c}}_\infty \quad := \quad \lim_{n\to\infty} \widehat{\mathfrak{c}}_n \quad < \quad \infty.$$

More precisely,

$$\widehat{\mathfrak{c}}_\infty^{1/p} \quad \leq \quad \widehat{\mathfrak{c}}_0^{1/p} + \frac{8\kappa_1(p)^{1/p}}{1-2^{-(d/2-1)}}.$$

Comparison of cost $\widehat{\mathfrak{c}}_n$ and $\mathfrak{c}_n$ now yields

Proposition 7.16. *For all $d \geq 3$ and all $p \geq 1$ there is a constant κ_3 such that for all $n \in \mathbb{N}_0$*

$$\mathfrak{c}_n^{1/p} \quad \leq \quad \widehat{\mathfrak{c}}_n^{1/p} + \kappa_3 \cdot 2^{n(1-d/2)}.$$

Proof. It is a modification of the proof of Proposition 7.10. This time, the map $T : B \mapsto \widehat{B}$

$$T: \; x \mapsto \left(\frac{Z}{\alpha}\right)^{1/d} \cdot x$$

defines an optimal transport $T_*\nu_B^\bullet = \lambda_B^\bullet$. Put $\tau' = \tau'(d,p) = \int_{[0,1)^d} |x|^p \, dx$. (This can easily be estimated, e.g. by $\tau' \leq \frac{1}{p+1} d^{p/2}$ if $p \geq 2$.) The cost of the transport T is

$$\begin{aligned}
\int_B |T(x) - x|^p \, d\nu_B^\bullet(x) \quad &= \quad \tau' \cdot 2^{np} \cdot Z \cdot \left| \left(\frac{Z}{\alpha}\right)^{1/d} - 1 \right|^p \\
&\leq \quad \tau' \cdot 2^{np} \cdot Z \cdot \left| \frac{Z}{\alpha} - 1 \right|^p
\end{aligned}$$

The inequality in the above estimation follows from the fact that $|t-1| \leq |t-1| \cdot (t^{d-1} + \ldots + t + 1) = |t^d - 1|$ for each real $t > 0$. The previous cost estimates hold true for each fixed ω (which for simplicity we had suppressed in the notation). Integrating w.r.t. $\mathbb{P}(d\omega)$ yields

$$\begin{aligned}
\mathbb{W}_p^p(\nu_B^\bullet, \lambda_B^\bullet) \quad &\leq \quad \tau' \cdot 2^{np} \cdot \mathbb{E}\left[Z \cdot \left| \frac{Z}{\alpha} - 1 \right|^p \right] \\
&\leq \quad \tau' \cdot 2^{np} \cdot \alpha^{-p} \cdot \mathbb{E}\left[Z^2\right]^{1/2} \cdot \mathbb{E}\left[|Z-\alpha|^{2p}\right]^{1/2} \\
&\leq \quad \tau' \cdot 2^{np} \cdot \alpha^{-p} \cdot \alpha \cdot C_3 \cdot \alpha^{p/2} \quad = \quad \kappa_3^p \cdot 2^{n(d+p-dp/2)}
\end{aligned}$$

and thus

$$\mathfrak{c}_n^{1/p} \; - \; \widehat{\mathfrak{c}}_n^{1/p} \; \leq \; \kappa_3 \cdot 2^{n(1-d/2)}.$$

□

Corollary 7.17. *For all $d \geq 3$ and all $p \geq 1$*

$$\mathfrak{c}_\infty \quad \leq \quad \widehat{\mathfrak{c}}_\infty \quad < \quad \infty.$$

Remark 7.18. It is not clear if the two quantities $\mathfrak{c}_\infty$ and $\hat{\mathfrak{c}}_\infty$ are really different or might actually coincide (see also Remark 8.16).

7.3.3 Quantitative Estimates

Throughout this section, we assume that $\vartheta(r) = r^p$ with $p < \overline{p}(d)$ where

$$p < \overline{p}(d) := \begin{cases} \infty, & \text{for } d \geq 3 \\ 1, & \text{for } d = 2 \\ \frac{1}{2}, & \text{for } d = 1. \end{cases}$$

Proposition 7.19. *Put* $\tau(p,d)=\frac{d}{d+p}\cdot\left(\Gamma(\frac{d}{2}+1)^{1/d}\cdot\pi^{-1/2}\right)^p$. *Then*

$$\mathfrak{c}_\infty\geq\mathfrak{c}_0\geq\tau(p,d).$$

Proof. The number τ as defined above is the minimal cost of a semicoupling between $\mathcal{L}$ and a single Dirac mass, say δ_0. Indeed, this Dirac mass will be transported onto the d-dimensional ball $K_r=\{x\in\mathbb{R}^d:|x|<r\}$ of unit volume, i.e. with radius r chosen s.t. $\mathcal{L}(K_r)=1$. The cost of this transport is $\int_{K_r}|x|^p\,dx=\frac{d}{d+p}r^p=\tau$.
For each integer $Z\geq 2$, the minimal cost of a semicoupling between $\mathcal{L}$ and a sum of Z Dirac masses will be $\geq Z\cdot\tau$. Hence, if Z is Poisson distributed with parameter 1

$$\mathfrak{c}_0\geq\mathbb{E}[Z]\cdot\tau=\tau.$$

□

Remark 7.20. Explicit calculations yield

$$\tau(p,1)=\frac{1}{1+p}\cdot 2^{-p},\qquad \tau(p,2)=\frac{2}{2+p}\cdot\pi^{-p/2},\qquad \tau(p,3)=\frac{3}{3+p}\cdot\left(\frac{3}{4\pi}\right)^{p/3}$$

whereas Stirling's formula yields a uniform lower bound, valid for all $d\in\mathbb{N}$ (which indeed is a quite good approximation for large d)

$$\tau(p,d)\geq\frac{d}{d+p}\cdot\left(\frac{d}{2\pi e}\right)^{p/2}.$$

Proposition 7.21. *Put* $\widehat{\tau}=\widehat{\tau}(d,p)=\int_{[0,1)^d}\int_{[0,1)^d}|x-y|^p\,dydx$. *Then*

$$e^{-1}\cdot\widehat{\tau}\quad\leq\quad\widehat{\mathfrak{c}}_0\quad\leq\quad\widehat{\tau}.$$

Moreover, $\widehat{\tau}\leq\frac{1}{(1+p)(1+p/2)}\cdot d^{p/2}$ *for all* $p\geq 2$ *and* $\widehat{\tau}\leq\left(\frac{d}{6}\right)^{p/2}$ *for all* $0<p\leq 2$

Proof. If there is exactly one Poisson particle in the cube $B_0=[0,1)^d$ – which then is uniformly distributed– the transportation cost are exactly $\widehat{\tau}(d,p)$. In general, $\widehat{\tau}$ still is an upper bound for the cost per particle. The number of particles will be Poisson distributed with parameter 1. The lower estimate for the cost follows from the fact that with probability e^{-1} there is exactly one Poisson particle in $B_0=[0,1)^d$.

Using the inequality $(x_1^2+\ldots+x_d^2)^{p/2}\leq d^{p/2-1}\cdot(x_1^p+\ldots+x_d^p)$ – valid for all $p\geq 2$ – the upper estimate for $\widehat{\tau}$ can be derived as follows

$$\begin{aligned}\int_{[0,1)^d}\int_{[0,1)^d}|x-y|^p\,dydx &\leq d^{p/2-1}\sum_{i=1}^d\int_{[0,1]^d}\int_{[0,1]^d}|x_i-y_i|^p\,dydx\\ &= d^{p/2}\int_0^1\int_0^1|s-t|^p\,dsdt\\ &= \frac{1}{(1+p)(1+p/2)}\cdot d^{p/2}.\end{aligned}$$

Applying Hölder's inequality to the inequality for $p=2$ yields the claim for all $p\leq 2$. □

Theorem 7.22. *For all* $p\leq 1$ *and* $d>2p$

$$\frac{d}{d+p}\cdot\left(\frac{d}{2\pi e}\right)^{p/2}\quad\leq\quad\mathfrak{c}_\infty\quad\leq\quad\left(\frac{d}{6}\right)^{p/2}+\frac{1}{(p+1)(2^{d/2-p}-1)}$$

whereas for all $p\geq 1$ *and* $d\geq 3$

$$\left(\frac{d}{d+p}\right)^{1/p}\cdot\left(\frac{d}{2\pi e}\right)^{1/2}\quad\leq\quad\mathfrak{c}_\infty^{1/p}\quad\leq\quad\frac{d^{1/2}}{6^{1/2}\wedge[(1+p)(1+p/2)]^{1/p}}+28\cdot\kappa_1^{1/p}.$$

Proof. Proposition 7.19 and the subsequent remark yield the lower bound

$$\frac{d}{d+p} \cdot \left(\frac{d}{2\pi e}\right)^{p/2} \quad \leq \quad \tau \quad \leq \quad \mathfrak{c}_\infty,$$

valid for all d and p. In the case $p \geq 1$ the upper bound follows from Proposition 7.21 and Corollary 7.15 by

$$\mathfrak{c}_\infty^{1/p} \quad \leq \quad \widehat{\tau}^{1/p} + \frac{4\kappa_1^{1/p}}{1 - 2^{-(d/2-1)}} \leq \frac{d^{1/2}}{6^{1/2} \wedge [(1+p)(1+p/2)]^{1/p}} + 28 \cdot \kappa_1^{1/p}.$$

In the case $p \leq 1$, estimate (7.1) with $\Theta(r) = \frac{1}{p+1} r^{p+1}$ yields

$$\widehat{\mathfrak{c}}_\infty \leq \widehat{\mathfrak{c}}_0 + \sum_{m=0}^{\infty} 2^{-(m+1)(d/2+1)} \cdot \frac{1}{p+1} 2^{(m+1)(p+1)} = \widehat{\mathfrak{c}}_0 + \frac{1}{(p+1)(2^{d/2-p} - 1)}.$$

provided $p < d/2$. Together with Proposition 7.10 this yields the claim. □

Corollary 7.23. *(i) For all $p \in (0, \infty)$*

$$\frac{1}{\sqrt{2\pi e}} \quad \leq \quad \liminf_{d\to\infty} \frac{\mathfrak{c}_\infty^{1/p}}{d^{1/2}} \quad \leq \quad \limsup_{d\to\infty} \frac{\mathfrak{c}_\infty^{1/p}}{d^{1/2}} \quad \leq \quad \frac{1}{\sqrt{6} \wedge [(1+p)(1+p/2)]^{1/p}}.$$

Note that the ratio of right and left hand sides is less than 5, – and for $p \leq 2$ even less than 2.
(ii) For all $p \in (0, \infty)$ there exist constants k, k' such that for all $d > 2(p \wedge 1)$

$$k \cdot d^{p/2} \ \leq \ \mathfrak{c}_\infty \ \leq \ k' \cdot d^{p/2}.$$

7.4 An allocation of optimal tail

In this section we want to sketch a construction of Marko and Timar [MT11] giving an equivariant allocation with finite mean transportation cost for $\vartheta(r) = \exp(\kappa r^d)$ for some positive $\kappa \in \mathbb{R}_+$. It is based on the algorithm of Ajtai, Komlós and Tusnády. This will prove the missing part of Theorem 1.7 (i). Moreover, we will adapt this argument to prove a similar result for the case of a compound Poisson process with exponential weights in the following section. We will call realizations ω or μ^ω of the Poisson process configurations. An atom of μ^ω will be called configuration point.

The idea is to use an infinite version of the already mentioned Ajtai, Komlós and Tusnády (AKT) scheme. A very short version of the Markó Timar construction goes as follows:
Fix a configuration ω such that 0 is a configuration point, consider the lattice $v + 2^n\mathbb{Z}^d$ and partition each of the cubes $v + u + [0, 2^n)^d$, $u \in 2^n\mathbb{Z}^d$ into dyadic subcubes such that in each of the subcubes there is at most one point of the configuration. Then run the infinite AKT scheme up to stage n, that is until you reach the cube of side length 2^n. Let $f_{v,n}$ be the indicator function of the cell associated to the point at zero. If n tends to ∞ the volume of the cell tends to one. This gives an allocation. However, this allocation will depend on v. To get an equivariant allocation, we mix over all such allocations, that is, consider $f_n = \int_{v\in[0,2^n)^d} f_{v,n} dv$, a function from $\mathbb{R}^d$ to $[0, 1]$, and take limits. As the transportation scheme is very explicit it is possible to follow the movements of the different cells and control this limit. This will result in a *fractional allocation*, that is an allocation in which one Lebesgue point might be divided between different Poisson points. However, because of the good control, one can modify this fractional allocation rule into a real allocation rule.

Let us start the more detailed description by recalling the classical AKT scheme. We start with a cube $[0,2^n)^d$ and a couple of points inside this cube. Then, we take an horizontal hyperplane orthogonal to the first coordinate axis and shift this axis such that the volume ratio of the two halves equals the ratio of the number of points in the respective halves. In the next step, consider in each of the two cuboids the hyperplane orthogonal to the second coordinate axis dividing the cuboid in two halves and shift this hyperplane as in the first step. Proceed until in every cube there is just one point left (see also the the picture in the middle of figure 7.1).

For the infinite version of the AKT scheme we do the very same just the other way around as we cannot divide the whole space in two halves and move the hyperplane to adjust the volume ratio to the ratio of the configuration in the two halves. Fix a realization and start again with a cube $[0,2^n)^d$. Divide the cube in subcubes of the form $z+[0,1)^d$ with $z \in \mathbb{Z}^d$. Assume that there is at most one point in each such cube. Otherwise subdivide each cube dyadically further until you reach a level, in which any point has his own cube. In the first step consider all the cuboids of the form $\{(k_1,\ldots,k_d)+[0,1)^{d-1}\times[0,2) : 0 \le k_i \le 2^n-1 \text{ for } 1 \le i \le d-1 \text{ and } 0 \le k_d \le 2^{n-1}-1\}$ and take the bisector hyperplane orthogonal to the d-th coordinate axis of each of these cuboids. Move them along the d-th axis until the volume ratio of the two halves gets equal to the ratio of the number of configuration points in them. Transform the interior of the cell affinely.

In the second step consider the cuboids $\{(k_1,\ldots,k_d)+[0,1)^{d-2}\times[0,2)^2 : 0 \le k_i \le 2^n-1 \text{ for } 1 \le i \le d-2 \text{ and } 0 \le k_{d-1},k_d \le 2^{n-1}-1\}$, take their bisector orthogonal to the (d-1)-th coordinate axis and move it as before. Continue this procedure until the starting cube $[0,2^n)^d$ is reached. We end up with a partition of $[0,2^n)^d$ into $\mu^\omega([0,2^n)^d)$ cuboids of equal volume. Each transformation is called a *step*. The d-tuple of the first d steps, the d-tuple of the second d steps and so on are called *stages*.
For any $v \in \mathbb{R}^d$ we can run this algorithm for any $2^n \times 2^n \times \ldots \times 2^n$ cube of the partition $v+2^n\mathbb{Z}^d$ of $\mathbb{R}^d$. We call this the AKT(v) scheme up to stage n. In each stage there is exactly on step in every coordinate direction. The movements in the different directions are independent. Hence, if we want to get an upper bound on the total transformation of the box, we can treat the different directions separately.
We want to condition on $0 \in \omega$, that is on 0 being a configuration point. Hence, we consider the Palm version (see Introduction) of the Poisson point process. Taking expectation and computing probabilities wrt to the Palm measure will be indicated by a $'$, e.g. $\mathbb{P}'$. Let $C_{v,n}$ be the cell assigned to 0 by AKT(v) run up to stage n. Denote the indicator function of $C_{v,n}$ by $f_{v,n}$. The task will be to control $\operatorname{diam}(C_{v,n} \cup \{0\})$. We can think of $C_{v,n}$ as the transformed initial cube containing 0 after nd steps (or (n+k)d steps if 0 was in the cube with more than one point in the beginning and we further subdivided this cube k-times.). For bounding the diameter it is therefore sufficient to bound the shift of 0 after n stages together with the length of the edges of $C_{v,n}$. The key part in proving the theorem are the following two lemmas. The first is a standard property of Poisson random variables.

Lemma 7.24. *Let X be a random variable with Poisson distribution of mean α. If $0 \le \rho \le 2$ then*
$$\mathbb{P}[|X-\alpha| \ge \alpha\rho] < 2\exp\left(-\frac{\alpha\rho^2}{4}\right).$$

Lemma 7.25. *There exist $c, C > 0$ such that for any $R > 0$ there exist an event E_R, such that*

i) $1-\mathbb{P}'[E_R] < C\exp(-cR^d)$.

ii) On E_R, for every n, the total movement of 0 in the AKT(v) run up to stage n is at most R, for all $v \in \mathbb{R}^d$.

iii) On E_R, the total length of the shifts of 0 in the i-th stage is at most $c'_i = c'_i(R)$, and $(c'_i)_{i\in\mathbb{N}}$ is absolutely summable.

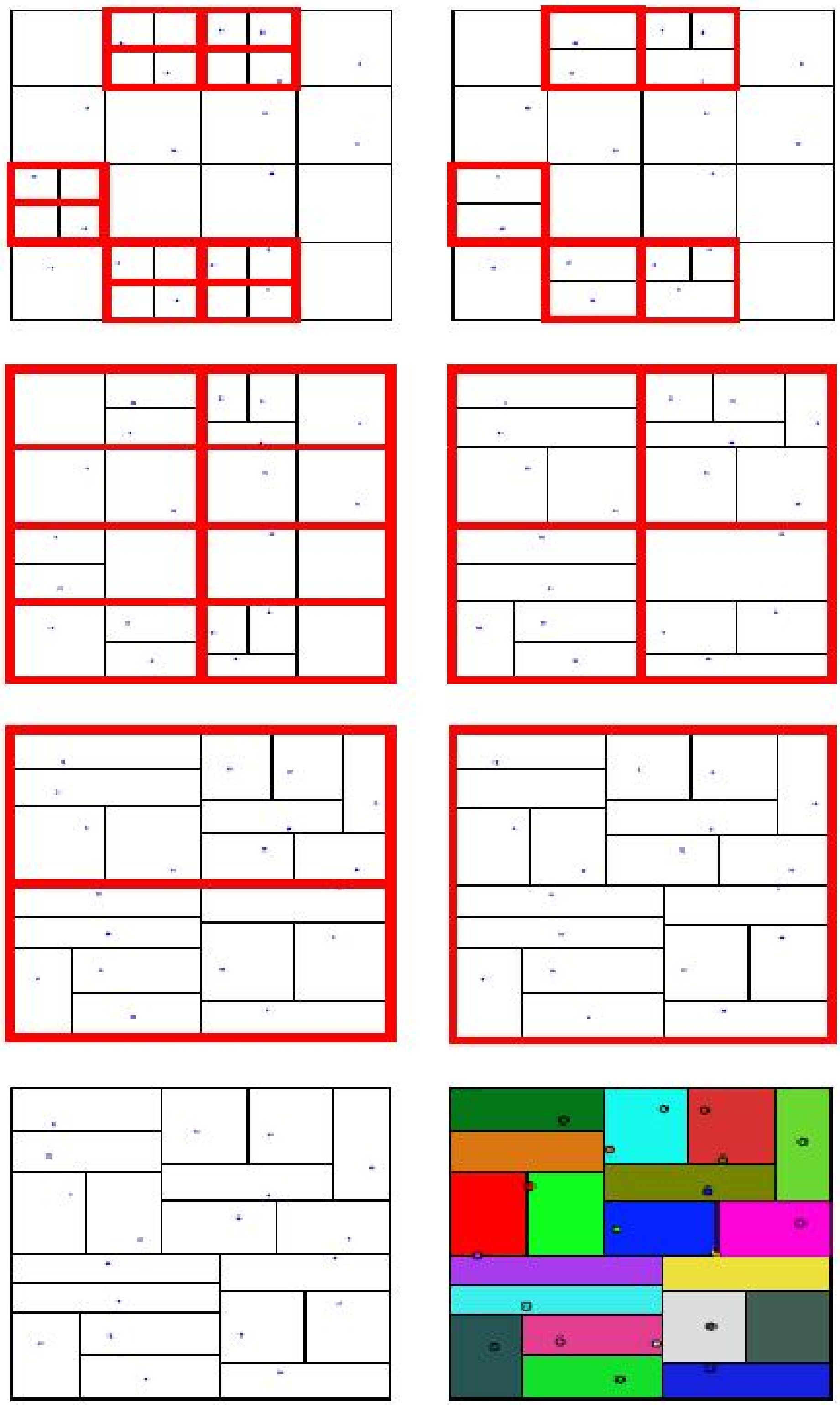

Figure 7.2: The AKT(v) scheme on a 4x4 box with 19 points. Picture from [MT11].

Sketch of the proof: As we do not want to get optimal constants we will just prove this Lemma for $R > R_0$. Fix an arbitrary $v \in \mathbb{R}^d$. As already mentioned the movements in the different directions in one stage of the AKT(v) scheme are independent and the sizes are comparable. Hence, it is sufficient to consider first just the movement in one direction, treat the other directions similarly and conclude by using the triangle inequality.
The idea is to use the fact that the number of Poisson points in a very big box is very close to the volume of this box. Therefore, in the AKT(v) scheme after sufficiently many stages, the movements will become very small. The exponential bound in the previous Lemma allows to control this "smallness" and show that it is actually summable. Therefore, after N stages the different cells do merely move any more and we get the desired bounds.
Let $r_0 \in \mathbb{N}$ be such that $R/4 \leq 2^{r_0} < R/2$ and put $\rho_n = 2^{-5n/4} R^{5/4} 10^{-2}$. Note that $\rho_n < 1$ if $n > r_0$. We want to consider an event giving bounds on the number of configuration points in *all* cubes of side length 2^n containing the origin. To this end, we define an event $A_R = A_{R,1}$ giving bounds on the number of configuration points in the two halves of the finitely many cubes for each $n > r_0$ with the properties that the cube has corners in $G_n(v) = v + 2^{-n}\mathbb{Z}^d$, side length $2^n \pm 2^{-n}$ and contains the origin.
For $n = r_0 + 1, \ldots, m_1 = 1, \ldots, (2^{2^n} - 1)^d$ and $m_2 = 1, \ldots, (2^{2^d} + 1)^d$ let $\underline{B_{n,m_1}}$ denote the $m_1 - th$ cube with vertices in $G_n(v)$ of side length $2^n - 2^{-n}$ containing the origin (the cubes are numbered in some arbitrary way), $\underline{U_{n,m_1}}$ its "left" side and $\underline{V_{n,m_1}}$ its "right" side. Moreover, let $\overline{B_{n,m_2}}$ denote the $m_2 - th$ cube of side length $2^n + 2^{-n}$ with vertices in $G_n(v)$ containing the origin, $\overline{U_{n,m_2}}$ and $\overline{V_{n,m_2}}$ its left and right side.
Let A_R be the event such that all the following holds:

$$\begin{aligned} \underline{A_R^{n,m_1}}: \quad & \left| \mu(\underline{U_{n,m_1}} \setminus \{0\}) - \frac{(2^n - 2^{-n})^d}{2} \right| < \frac{(2^n - 2^{-n})^d}{2} \rho_n, \quad \text{and} \\ & \left| \mu(\underline{V_{n,m_1}} \setminus \{0\}) - \frac{(2^n - 2^{-n})^d}{2} \right| < \frac{(2^n - 2^{-n})^d}{2} \rho_n, \end{aligned}$$

and

$$\begin{aligned} \overline{A_R^{n,m_2}}: \quad & \left| \mu(\overline{U_{n,m_2}} \setminus \{0\}) - \frac{(2^n + 2^{-n})^d}{2} \right| < \frac{(2^n + 2^{-n})^d}{2} \rho_n, \quad \text{and} \\ & \left| \mu(\overline{V_{n,m_2}} \setminus \{0\}) - \frac{(2^n + 2^{-n})^d}{2} \right| < \frac{(2^n + 2^{-n})^d}{2} \rho_n, \end{aligned}$$

for all $n = r_0 + 1, \ldots, m_1 = 1, \ldots, (2^{2^n} - 1)^d$ and $m_2 = 1, \ldots, (2^{2^d} + 1)^d$.

Using the previous lemma we can estimate

$$\begin{aligned} 1 - \mathbb{P}'[A_R] \quad & \leq \quad \sum_{n=r_0+1}^{\infty} (2^{2^n} - 1)^d \cdot (1 - \mathbb{P}'[\underline{A_R^{n,1}}]) + \sum_{n=r_0+1}^{\infty} (2^{2^n} + 1)^d \cdot (1 - \mathbb{P}'[\overline{A_R^{n,1}}]) \\ & \leq \quad \sum_{n=r_0+1}^{\infty} 2^{2(n+1)d} 8 \cdot \exp\left(-\frac{\rho_n^2 2^{(n-1)d}}{4} \right) \\ & \leq \quad \sum_{n=r_0+1}^{\infty} 2^{2(n+1)d+3} \cdot \exp\left(-2^{(n+1)(d-5/2)} \frac{R^{5/2} 10^{-4}}{2^{3d}} \right). \end{aligned}$$

By the definition of r_0 the first term in this series is at most $R^{2d} 2^{2d+3} \exp(-R^d 10^{-4}/2^{3d})$. Hence, taking $R > R_0(d)$ large enough this series can be dominated by a converging geometric series. This implies the estimate

$$1 - \mathbb{P}'[A_R] < C' \exp(-c' R^d).$$

Now we want to remove the condition that the cubes have to have vertices in the set $G_n(v)$. Let B be an arbitrary cube of side length 2^n containing the origin and U and V its left and right half. Then, we can find a cube $\underline{B_{n,m_1}}$ of side length $2^n - 2^{-n}$ with vertices in $G_n(v)$ such that $V \subset \underline{V_{n,m_1}}$ and a cube $\overline{B_{n,m_2}}$ of side length $2^n + 2^{-n}$ with vertices in $G_n(v)$ such that $U \subset \overline{U_{n,m_2}}$. This allows to conclude after some computation that conditioned on A_R and 0 being a configuration point

$$\left|\frac{\mu(U)-\mu(V)}{\mu(U)+\mu(V)}\right| < 8\rho_n.$$

Hence, we can estimate the shift of the cell containing 0.
Let $v' \in \mathbb{R}^d$ be arbitrary and consider the $v'-$partition sequence $(v'+\mathbb{Z}^d, v'+2\mathbb{Z}^d, \ldots, v'+2^i\mathbb{Z}^d, \ldots)$. Let u be an arbitrary point inside the initial cube containing 0. Denote its signed shift along the first axis in the n-th stage of AKT(v') by $D_n = D_n(u, v')$. Until the $n-th$ stage every displacement takes place inside a cube of side length 2^n. Hence, we have (by the choice of r_0):

$$\left|\sum_{n=-k}^{r_0} D_n\right| \leq 2^{r_0} < \frac{R}{2},$$

if we start the algorithm from the level of cubes of side length 2^{-k}. As in the previous sections (cf Lemma 7.7 and Lemma 7.13) we can compute the shift of u by

$$D_n = (1 - C_n(u))2^{n-1}\frac{\mu(U_n)-\mu(V_n)}{\mu(U_n)+\mu(V_n)},$$

where U_n and V_n are the left and right half of the cuboid B_n in the $v' + 2^n\mathbb{Z}^d$ partition grid containing 0 and $C_n(u)$ is the relative distance of u to the bisector of B_n orthogonal to the first axis before the n-th stage. Conditioning on A_R and 0 being a configuration point we can estimate:

$$\begin{aligned}
\left|\sum_{n=r_0+1}^{\infty} D_n\right| &\leq \sum_{n=r_0+1}^{\infty} \left|(1-C_n(u))2^{n-1}\frac{\mu(U_n)-\mu(V_n)}{\mu(U_n)+\mu(V_n)}\right| \\
&\leq \sum_{n=r_0+1}^{\infty} 2^{2-n/4}R^{5/4}10^{-2} < \frac{R}{2}.
\end{aligned}$$

This shows, that conditioned on A_R and 0 being a configuration point, for any v' and any u in the initial cube containing 0 in the v'-partition sequence, the total shift of u along the first axis is at most R. Moreover, the shift along the first axis in the n-th stage is at most $c_n'(R) := 2^{2-n/4}R^{5/4}10^{-2}$ for any $n > r_0(R)$. Defining $c_n'(R) = 2^{n-1}$ for $n \leq r_0$, which is the maximal possible shift in the n-th stage, defines a summable sequence.

For $1 < i \leq d$ one can define the events $A_{R,i}$ similarly and get with the same ρ_n sequence the respective bounds on the probability and shifts along the i-th axis. Hence, we can define the event

$$E_R := \bigcap_{i=1}^{d} A_{R/d,i},$$

satisfying all the assertions in the Lemma and thereby proving the claim. □

The Lemma and especially its proof allow to draw a couple of conclusions. Firstly, as the number of points in a big box concentrate around the volume of that box we have that

$$\lim_{n\to\infty} \mathrm{Vol}(C_{v,n}) = \lim_{n\to\infty} \mathcal{L}(C_{v,n}) = \lim_{n\to\infty} \int_{\mathbb{R}^d} f_{v,n}(x)\mathcal{L}(dx) = 1,$$

uniformly in v. The uniformity is not directly clear but can be achieved in the same way as in the last proof by using some approximating cubes belonging to a countable family.
Let $l(C_{v,n}) \in \mathbb{R}^d$ denote the d-dimensional vector whose i-th component denotes the length of the i-th side of $C_{v,n}$. Let $|\cdot|_{max}$ denote the maximum norm on $\mathbb{R}^d$. In the proof of the last Lemma we saw that each side length can change in one stage at most by the factor $(1 + 8\rho_n)$. Hence, we get

$$|l(C_{v,n+1})|_{max} \leq |l(C_{v,n})|_{max}(1 + 8\rho_{n+1}).$$

Proposition 7.26. *Conditioned on E_R and 0 being a configuration point, there is a constant $K > 0$ such that $|l(C_{v,n})|_{max} \leq KR$ uniformly in v. Moreover,*

$$|l(C_{v,n+1}) - l(C_{v,n})|_{max} \leq e_n(R)$$

for $n > r_0$, where $(e_n(R))_{n\in\mathbb{N}}$ is an absolutely summable sequence with sum at most cR.

Putting all of this together gives

Lemma 7.27. *There exist $c, C > 0$ such that for all $R > 0$ there exists an event E_R such that*

i) $1 - \mathbb{P}'[E_R] < C\exp(-cR^d)$.

ii) Conditioned on E_R, the diameter of the cell $C_{v,n}$ of 0 in the AKT(v) ran up to stage n is at most $c'R$, for all $v \in \mathbb{R}^d$.

iii) On E_R we have $\|f_{v,n} - f_{v,n+1}\|_1 = \mathcal{L}(C_{v,n} \triangle C_{v,n+1}) < c_n(R)$, for $n > r_0(R)$, where the series $c_n(R)$ is absolutely summable. The constants do not depend on v.

We could try to consider the limit of the AKT(v) algorithm. However, this would not yield an equivariant allocation as it depends on the v-partition sequence. The idea is to average over all AKT(v) transport schemes and then take the limit. Because of the uniformity and summability of the bounds this will give the desired estimate. Put

$$f_n = f_n^\omega := \frac{1}{2^{nd}} \int_{[0,2^n]^d} f_{v,n} dv.$$

As for $\mathbb{P}'$ almost every ω $f_{v,n}$ is L^1-continuous in v except for possibly the union of finitely many hyperplanes, f_n is well defined and is a function from $\mathbb{R}^d$ to $[0, 1]$. The difficult part is to establish the following

Proposition 7.28. *With probability one, the L^1 limit f of f_n exists. It is a function with values in $[0, 1]$, integral one and with*

$$\mathbb{P}'[diam((\{0\} \cup supp(f)) > r] \leq C\exp(-cr^d).$$

As E_R is an increasing sequence of events exhausting a set of measure one, it is sufficient to prove the Proposition for $\omega \in E_R$ and show that the diameter of supp(f^ω) is bounded by cR, where c is independent of R. This follows by some computation using the last Lemma, especially part iii).
With the function f we can define a fractional allocation using equivariance. Let $\xi \in \omega$ be a configuration point. Then we allocate to ξ the cell defined by the nonnegative L^1 function $\Psi(\xi) = f^{\omega-\xi} + \xi$. In other words we think of ξ as being the origin and do the whole construction for ξ. By the construction it is clear that each Poisson point gets mass exactly one, that the allocation is measurable, equivariant and satisfies the desired bound on the diameter of the cells. Moreover, from the construction it is also clear that this allocation distributes exactly the

Lebesgue measure, that is we have for $\mathcal{L}$ a.e. $x: \sum_{\xi\in\omega}\Psi(\xi)(x) = 1$. This is the case as this is already true for $\Psi_n(\xi) = f_n^{\omega-\xi} + \xi$. Indeed,

$$\begin{aligned}\sum_{\xi\in\omega}\Psi_n(\xi)(x) &= \sum_{\xi\in\omega} 2^{-nd}\left(\int_{[0,2^n)^d}(f_{v,n}^{\omega-\xi}+\xi)\ dv\right)(x)\\ &= 2^{-nd}\int_{[0,2^n)^d}\sum_{\xi\in\omega}(f_{v,n}^{\omega-\xi}+\xi)(x)dv\\ &= 2^{-nd}\int_{[0,2^n)^d} 1\ dv = 1.\end{aligned}$$

Therefore, $\Psi(\xi)$ is a fractional allocation. The last and remaining step is to construct out of the fractional allocation a real allocation, so that almost every Lebesgue point is allocated to exactly one Poisson or configuration point. The observation which allows to do this is that for $x \in \mathbb{R}^d$, $\mathbb{P}$ almost surely there is an $\epsilon > 0$ such that the epsilon ball around x is contained in just finitely many cells, that is $B(x,\epsilon) \subset \cap_{i=1}^k \mathrm{supp}(f^{\omega-\xi}+\xi)$ and $B(x,\epsilon)\cap\mathrm{supp}(f^{\omega-\xi_j}+\xi_j) = \emptyset$ for any $\xi_j \notin \{\xi_1,\dots,\xi_k\}$. In other words, every x has $\mathbb{P}$ a.s. a neighbourhood which belongs to only finitely many cells. The reason for this are the good diameter and probability bounds, which allow to use the Borel Cantelli Lemma.
The final step is easy. Take a set belonging to k different cells and redistribute it in an equivariant way, respecting the different volume ratios of the cells in this set. This can be done e.g. by optimal transport or the stable marriage algorithm (see [HHP06]). Note, that this will at most decrease the diameter as the center of the cell is by definition included in the cell.

7.5 Estimates on compound Poisson processes

In this section we turn to $\gamma-$compound Poisson processes. Using the results of the previous sections we show how one can get good estimates on the transportation cost for $\gamma-$compound Poisson processes. We still assume $M = \mathbb{R}^d$ and $\lambda = \mathcal{L}$ being the Lebesgue measure. Let $\mu^\bullet$ be an equivariant $\gamma-$compound Poisson process of intensity $\beta \in (0,\infty)$ with compound measure γ having nonnegative support and mean 1. In other words, we consider a "standard" Poisson process where each atom gets the random weight X_i (instead of the constant 1). The different weights $(X_i)_{i\in\mathbb{N}}$ are assumed to be nonnegative iid random variables with mean 1 and distribution γ. The transportation cost estimates will depend on the integrability of the weight X_1. If X_1 does not have all moments, the transportation cost will be infinite for L^p cost with high p. If X_1 has exponential moments, the transportation cost behave as in the Poisson case.

7.5.1 Lower estimates

Let us first consider the case that X_1 does not have finite moments of order p. In that case we have

Proposition 7.29. *Let $\beta \le 1$ and $p > 1$ be such that $\mathbb{E}[X_1^p] = \infty$. Let $\vartheta(r) \ge r^{(p-1)d}$, then $\mathfrak{c}_\infty = \infty$.*

Proof. $\mathfrak{c}_\infty$ can easily be bounded from below by the cost of transporting mass X_1 optimally into a single point. This transportation cost behaves for fixed $X_1 = X_1(\omega)$ like

$$\int_0^{cX_1^{1/d}} \vartheta(r)r^{d-1}dr \ \gtrsim\ X_1^p.$$

Taking expectation wrt X_1 yields the desired result. $\square$

Remark 7.30. We do not expect this result to hold for $\beta > 1$ because the higher Poisson density should have the effect of cutting off the high X_i values resulting in effectively bounded weights. However, we do not know how to prove this.

If we assume that X_1 does not have all exponential moments, that is $\mathbb{E}[\exp(\kappa X_1)] = \infty$ for some positive κ, we get by the same reasoning $\mathfrak{c}_\infty = \infty$ for $\vartheta(r) \geq \exp(\kappa' r^d)$ with $\kappa' = \kappa/c^{1/d}$. However, this is already (up to constants) the most expensive cost function producing finite transportation cost. Indeed, one direction is the following Theorem, the other one Theorem 7.36.

Theorem 7.31. *For all $\beta \in (0,\infty)$ and $d \geq 1$ there exists a constant $\kappa = \kappa(d,\beta,X_1)$ such that for any invariant semicoupling $q^\bullet$ of $\mathcal{L}$ and $\mu^\bullet$*

$$\mathbb{E}\left[\int_{\mathbb{R}^d\times[0,1)^d} \exp\left(\kappa|x-y|^d\right) q^\bullet(dx,dy)\right] = \infty.$$

Proof. The proof is the very same as before. Just use the large deviation principle for the iid sequence $Z_i = \mu^\bullet(i+[0,1)^d) = \sum_{j=1}^N X_j$ with $i \in \mathbb{Z}^d$ and N a Poisson random variable with mean β. □

7.5.2 Upper estimates

L^p estimates

In this section we want to apply the techniques and estimates on L^p cost for the Poisson process to the case of the $\gamma-$compound Poisson process with iid weights $(X_i)_{i\in\mathbb{N}}$. To this end, we need to find good estimates on the $p-th$ (central) moments and $p-th$ inverse moments of a $\gamma-$compound Poisson process. However, a general estimate of the $p-th$ inverse moment is not so easy and even the existence is in general not given, if X_1 is not bounded away from zero. In fact, if X_1 has a density $f(x)$ with respect to the Lebesgue measure, the existence of the $p-th$ inverse moment of X_1 is equivalent to the existence of

$$\lim_{x\searrow 0} \frac{f(x)}{x^q}$$

for some $q > p-1$ (see [DHT09]). In order to avoid these technicalities at this stage, we will assume for simplicity that $X_1 > \chi$ almost surely for some $1 > \chi > 0$. This trivially implies the existence of all inverse moments.
Let N be a Poisson random variable with mean α independent of $(X_i)_{i\in\mathbb{N}}$. Put $Z = \sum_{i=1}^N X_i$.

Proposition 7.32. *Let $p = 2q > 0$ be given and assume that $\mathbb{E}[X_1^p] < \infty$.*

i) There is a finite constant $C_1 = C_1(p,X_1)$ such that $\mathbb{E}[Z^p] \leq C_1\alpha^p$.

ii) There is a finite constant $C_2 = C_2(p,X_1)$ such that $\mathbb{E}[(Z-\alpha)^{2\lfloor q\rfloor}] \leq C_2\alpha^{\lfloor q\rfloor}$.

iii) Assume that $p \geq 2$. If α is sufficiently large, there is a finite constant $C_3 = C_3(p,X_1)$ such that $\mathbb{E}[Z^{-p}\cdot 1_{\{Z>0\}}] \leq C_3\alpha^{-p}$.

Proof. In all cases, by Hölder's inequality, it suffices to prove the inequalities for integer $p \in \mathbb{N}$.

i) For $p \geq 1$ this is a direct consequence of Hölder's inequality and Lemma 7.12.

$$\begin{aligned}
\mathbb{E}[Z^p] &= \sum_{n=1}^{\infty} \mathbb{P}[N=n]\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^p\right] \\
&\leq \sum_{n=1}^{\infty} \mathbb{P}[N=n] n^{p-1}\mathbb{E}\left[\sum_{i=1}^{n} X_i^p\right] \\
&= \sum_{n=1}^{\infty} \mathbb{P}[N=n] n^{p}\mathbb{E}[X_1^p] \\
&= \mathbb{E}[N^p]\mathbb{E}[X_1^p] \leq C_1\alpha^p
\end{aligned}$$

ii) The argument is very similar to the one in the third part of Lemma 7.12. However, in this case, the moment generating function need not exist as the moment generating function of X_1 might not exist. Instead of the moment generating function, we can work with the characteristic function of $Y = Z - \alpha$. It is given by $\phi_Y(t) = \mathbb{E}\left[\exp(it(Z-\alpha))\right] = \exp(\alpha(\phi_{X_1}(t) - 1 - it))$. Let q be an integer such that $2q \leq p$. Then, the 2q-th moment of $(Z-\alpha)$ is given by $i^{2q}\phi_Y^{(2q)}(0)$. However, the $(2q)-th$ derivative of ϕ_Y at the point $t=0$ is a polynomial of order q in α because $\phi'_{X_1}(0) = i\mathbb{E}[X_1] = i$.

iii) The existence of the inverse moment is clear as for any $k > 0$ we have $0 \leq \mathbb{E}[Z^{-k} \cdot 1_{\{Z>0\}}] \leq \chi^{-k}$. The estimate above follows from the following Theorem.

Theorem 7.33 ([GP01])**.** *Let $(V_n)_{n\in\mathbb{N}}$ be a sequence of nonnegative random variables with $V_n > \gamma > 0$ a.s.. Let η_n and σ_n be such that*

i) $\frac{V_n - \eta_n}{\sigma_n} \xrightarrow{d} N(0,1)$, *as* $n \to \infty$

ii) $\lim_{n\to\infty} \eta_n = \infty$

iii) there is $\epsilon < 1$ such that $\eta_n^{\epsilon}/\sigma_n > 1$, for n sufficiently large.

Then

$$\frac{\mathbb{E}\left[\frac{1}{(V_n)^k}\right]}{\frac{1}{\mathbb{E}[V_n]^k}} \to 1$$

as $n \to \infty$ for any $k > 0$.

To conclude, we can take $V_n = \sum_{i=1}^{N_n} X_i$ with N_n a Poisson random variable with mean α_n such that $\alpha_n \to \infty$. Then $\mathbb{E}[V_n] = \alpha_n = \eta_n$ and $Var(V_n) = (Var(X_1)+1)\alpha_n = \sigma_n^2$. □

Looking back at the estimates for the transportation cost between the Lebesgue measure and a Poisson point process the estimates above give us directly the following result.

Theorem 7.34. *i) Assume $d \leq 2$ and $\mathbb{E}[X_1^2] < \infty$. Then for any concave $\hat{\vartheta} : [1,\infty) \to \mathbb{R}$ dominating ϑ*

$$\int_1^{\infty} \frac{\hat{\vartheta}(r)}{r^{1+d/2}} dr < \infty \quad \Longrightarrow \quad \mathfrak{c}_\infty < \infty.$$

ii) Assume $d \geq 3$ and $p < p_0 - 1$ with $2 < p_0 = \sup\{q : \mathbb{E}[X_1^q] < \infty\} < \infty$. Then for $\vartheta(r) = r^p$ we have $\mathfrak{c}_\infty < \infty$.

iii) Assume $d \geq 3$ and $\mathbb{E}[X_1^p] < \infty$ for all $p > 0$. Then for $\vartheta(r) = r^p$ we have $\mathfrak{c}_\infty < \infty$ for any $p > 0$.

Proof. Thanks to the estimates in the Proposition above, we can use the very same estimates as in the Poisson case. Only in the second case, we need to change the variables for the Hölder inequality in two estimates because X_1 does not have arbitrary moments. We do not repeat the whole argument. We only show the very little pieces that change, namely the moment estimates in the calculation of the transportation cost.
Let $\epsilon > 0$ be such that $1/(p_0 - 1) < \epsilon < (p_0 - p)/p$. This condition ensures that $(1+\epsilon)/\epsilon < p_0$ and $p(1+\epsilon) < p_0$. The condition $p < p_0 - 1$ ensures that $1/(p_0 - 1) < (p_0 - p)/p$. The first estimate which is needed for the convergence of $\hat{\mathfrak{c}}_n$ becomes

$$\begin{aligned} &\mathbb{E}\left[\frac{|Z_0 - Z_1|^p}{Z^{p-1}} \cdot 1_{\{Z>0\}}\right] \\ \leq\ & \mathbb{E}\left[|Z_0 - Z_1|^{(1+\epsilon)p}\right]^{1/(1+\epsilon)} \cdot \mathbb{E}\left[Z^{-(p-1)(1+\epsilon)/\epsilon}\right]^{\epsilon/(1+\epsilon)} \\ \leq\ & C \cdot \alpha^{p/2}\alpha^{-(p-1)}. \end{aligned}$$

This already gives the finiteness of $\hat{\mathfrak{c}}_\infty$. To be able to use this estimate to show that $\mathfrak{c}_\infty < \infty$ we need to compute

$$\begin{aligned} &\mathbb{E}\left[Z \cdot \left|\frac{Z}{\alpha} - 1\right|^p\right] \\ \leq\ & \mathbb{E}\left[Z^{(1+\epsilon)/\epsilon}\right]^{\epsilon/(1+\epsilon)} \mathbb{E}\left[\left|\frac{Z}{\alpha} - 1\right|^{p(1+\epsilon)}\right]^{1/(1+\epsilon)} \\ \leq\ & C\alpha \cdot \alpha^{-p} \cdot \alpha^{p/2}, \end{aligned}$$

which is exactly what we need. □

Remark 7.35. i) In the first part of the theorem we require X_1 to have finite second moment, as otherwise our methods do not apply.

ii) In the second part of the last theorem, the requirement that $p < p_0 - 1$ instead of the more natural requirement $p < p_0$ fits well with the lower estimate in dimension 1.

iii) We do not know which estimate is responsible for the gap between the lower and upper estimate for the case that $\mathbb{E}[X_1^p] = \infty$ for some finite p in dimensions $d \geq 2$.

Compounding with optimal tail

We want to show using the construction of Markó and Timar (see [MT11] and section 7.4) that choosing X_1 according to an exponential law with mean 1 has optimal tail behavior in the sense that there is a constant κ such that the mean transportation cost $\mathfrak{c}_\infty < \infty$ for $\vartheta(r) \leq \exp(\kappa r^d)$. The aim of this choice is twofold. Firstly, we want to present a non-trivial example (trivial would be $X_1 = 1$ a.s.) of a $\gamma-$compound Poisson process having finite transportation cost for $\vartheta(r) = \exp(\kappa r^d)$. Secondly, this covers an example to which not only the last section is not applicable (X_1 can have values arbitrary close to zero) but also an example to which the machinery developed in the Poisson case is not directly applicable as X_1 does not have any inverse moment (see [DHT09] and the beginning of the last section). Hence, let $\mu^\bullet$ be a $\gamma-$compound Poisson process of unit intensity with compounding measure γ an exponential law with mean 1.

Theorem 7.36. *There is $\kappa > 0$ such that mean transportation cost between $\mathcal{L}$ and $\mu^\bullet$ for the cost function $c(x,y) = \vartheta(|x-y|)$ with $\vartheta(r) = \exp(\kappa r^d)$ are finite, that is $\mathfrak{c}_\infty < \infty$.*

Before we prove this theorem we need the following concentration estimate for $\gamma-$compound Poisson random variables. Let $Z = \sum_{i=1}^N X_i$ with N a Poisson random variable with mean α and $(X_i)_{i\in\mathbb{N}}$ a sequence of iid exponentially distributed random variables with mean 1 independent of N.

Lemma 7.37. *For any $0 < \rho < 1$ it holds that*

$$\mathbb{P}[|Z - \alpha| > \alpha\rho] \leq 2 \cdot \exp(-\alpha(2 + \rho - 2\sqrt{1 + \rho})) \leq 2 \cdot \exp\left(-\alpha\left(\frac{\rho^2}{4} - \frac{\rho^3}{8}\right)\right).$$

Proof. The moment generating function for X_1 is $M_X(t) = \frac{1}{1-t}$ defined for $t < 1$ and the moment generating function for Z is $M_Z(t) = \exp(\alpha(M_X(t) - 1))$. Then, we can estimate using the Markov inequality with $t < 1$

$$\mathbb{P}[Z - \alpha > \alpha\rho] \leq \exp(-t\alpha(\rho + 1))\mathbb{E}[\exp(tZ)] = \exp\left(\alpha\left(\frac{t}{1-t} - t(\rho + 1)\right)\right).$$

This expression is minimized by $t = 1 - \sqrt{\frac{1}{1+\rho}}$ yielding

$$\mathbb{P}[Z - \alpha > \alpha\rho] \leq \exp(-\alpha(2 + \rho - 2\sqrt{\rho + 1})).$$

Similarly, we can estimate for $t > -1$

$$\mathbb{P}[\alpha - Z > \alpha\rho] \leq \exp(-\alpha t(\rho - 1))M_Z(-t) = \exp\left(\alpha\left(\frac{-t}{1+t} + t(1 - \rho)\right)\right).$$

This expression is minimized by $t = \sqrt{\frac{1}{1-\rho}} - 1 > 0$ yielding

$$\mathbb{P}[\alpha - Z > \alpha\rho] \leq \exp(-\alpha(2 - \rho - 2\sqrt{1 - \rho})).$$

Because

$$-\rho - 2\sqrt{1 - \rho} > \rho - 2\sqrt{1 + \rho}$$

we can conclude

$$\mathbb{P}[|Z - \alpha| > \alpha\rho] \leq 2 \cdot \exp(-\alpha(2 + \rho - 2\sqrt{1 + \rho})).$$

Using the Taylor expansion of $\sqrt{1 + x}$ for $|x| < 1$ we can estimate

$$\sqrt{1 + \rho} \leq 1 + \frac{\rho}{2} - \frac{\rho^2}{8} + \frac{\rho^3}{16}.$$

This finally gives

$$\mathbb{P}[|Z - \alpha| > \alpha\rho] \leq 2 \cdot \exp\left(-\alpha\left(\frac{\rho^2}{4} - \frac{\rho^3}{8}\right)\right).$$

$\square$

Proof of the Theorem: The proof goes by explicitly constructing an allocation using the approach of Markó and Timar. The key point is to note that the properties of the Poisson point process entering their proof are independence in disjoint sets and exponential concentration around the mean in big boxes. The compound Poisson process inherits the independence between disjoint sets from the Poisson point process. The concentration follows from the Lemma above. Hence, we only need to choose an appropriate sequence $(\rho_n)_{n\in\mathbb{N}}$ to define events A_R and E_R analogously to the Poisson case. Having established this we can also copy the rest of their proof line to line (which we will not do explicitly). Hence it remains to show

There exist $c, C > 0$ such that for any $R > 0$ there exist an event E_R, such that

i) $1 - \mathbb{P}'[E_R] < C\exp(-cR^d)$.

ii) On E_R, for every n, the total movement of 0 in the AKT(v) run up to stage n is at most R, for all $v \in \mathbb{R}^d$.

iii) On E_R, the total length of the shifts of 0 in the i-th stage is at most $c'_i = c'_i(R)$, and $(c'_i)_{i\in\mathbb{N}}$ is absolutely summable.

Here $\mathbb{P}'$ denotes the Palm measure of the $\gamma-$compound Poisson process. Define the event A_R as disjoint union of the events $\underline{A_R^{n,m_1}}$ and $\overline{A_R^{n,m_2}}$ and r_0, R_0 and ρ_n as in section 7.4, that is $R/4 \le 2^{r_0} < R/2$ and $\rho_n = 2^{-5n/4}R^{5/4}10^{-2}$. In particular, we have $2^{r_0} = \kappa R$ with $1/4 \le \kappa < 2$. Hence, we can estimate (as above or as in [MT11])

$$\begin{aligned}
1 - \mathbb{P}'[A_R] &\le \sum_{n=r_0+1}^{\infty} (2^{2^n} - 1)^d \cdot (1 - \mathbb{P}'[\underline{A_R^{n,1}}]) + \sum_{n=r_0+1}^{\infty} (2^{2^n} + 1)^d \cdot (1 - \mathbb{P}'[\overline{A_R^{n,1}}]) \\
&\le \sum_{n=r_0+1}^{\infty} 2^{2(n+1)d} 8 \cdot \exp\left(-2^{(n-1)d}\left(\frac{\rho_n^2}{4} - \frac{\rho_n^3}{8}\right)\right) \\
&\le \sum_{n=r_0+1}^{\infty} 2^{2(n+1)d+3} \cdot \exp\left(-2^{(n-1)d} \cdot b_n\right).
\end{aligned}$$

with

$$\begin{aligned}
b_n &= \frac{1}{4}\left(2^{-5n/2}R^{5/2}10^{-4} - \frac{1}{2}2^{-15n/4}R^{15/4}10^{-6}\right) \\
&= \frac{1}{4}\left(2^{-5j/2}(\kappa R)^{-5/2}R^{5/2}10^{-4} - \frac{1}{2}2^{-15j/4}(\kappa R)^{-15/4}R^{15/4}10^{-6}\right) \\
&= \frac{1}{4}\left(2^{-5j/2}\kappa^{-5/2}10^{-4} - \frac{1}{2}2^{-15j/4}\kappa^{-15/4}10^{-6}\right),
\end{aligned}$$

where we wrote $n = r_0 + j$. This gives as exponent in exp

$$\begin{aligned}
-2^{(n-1)d} \cdot b_n &= -(\kappa \cdot R)^d 2^{(j-1)d} \cdot \frac{1}{4}\left(2^{-5j/2}\kappa^{-5/2}10^{-4} - \frac{1}{2}2^{-15j/4}\kappa^{-15/4}10^{-6}\right) \\
&= -(\kappa \cdot R)^d \left(C_1 2^{j(d-5/2)} - C_2 2^{j(d-15/4)}\right).
\end{aligned}$$

As $(d - 5/2) > (d - 15/4)$ the whole sum can be dominated by a geometric series. Hence, we can bound the sum by a constant times its first summand. The first summand is bounded by

$$k^{2d}R^{2d}2^{4d+3}\exp\left(-\kappa^d R^d \frac{1}{4}\left(2^{-5/2}\kappa^{-5/2}10^{-4} - \frac{1}{2}2^{-15/4}\kappa^{-15/4}10^{-6}\right)\right).$$

Putting this together yields

$$1 - \mathbb{P}'[A_R] < C' \exp(-c' R^d).$$

As we defined the sequence of ρ_n in the same way as Markó and Timar we also get the same bounds on the total shifts of points in the initial cube of 0 conditioned on A_R and 0 being a configuration point. Hence, we can define the event E_R in the same way by

$$E_R = \bigcap_{i=1}^{d} A_{R,i},$$

proving the claim and thereby the theorem. □

Remark 7.38. We could also have taken a different law for X_1. The important feature we need is the exponential concentration around the mean.

Chapter 8

Couplings on the regular k-tree

The aim of this chapter is twofold. Firstly, we want to show how it is possible to change the presented construction of an optimal coupling to be able to construct an equivariant coupling with finite cost on a non-amenable space. It is not clear, if the constructed coupling is actually optimal (see Remark 8.16). Secondly, we want to give an example of a non-smooth space to which this toolbox can be applied in order to produce equivariant couplings of equivariant random measures.

To keep things as easy as possible we chose the regular k-tree $\Upsilon_k = \Upsilon$, that is the tree in which every vertex has degree k. It has the advantage, that it can be well controlled because of its symmetry and simple geometry. Each edge will be identified with the unit interval and given an arbitrary orientation. $\lambda^\bullet = \lambda$ will be the Lebesgue measure on the edges and $\mu^\bullet$ will denote a Poisson point process of intensity 1. We write $V(\Upsilon)$ for the vertices of Υ and $E(\Upsilon)$ for the edges of Υ. A point $p \in \Upsilon$ will be written as $p = (e, x)$ with $e \in E(\Upsilon)$ and $x \in [0, 1]$, where the orientation of e is from 0 to 1. This representation is unique iff $x \in (0, 1)$.

G will be the group of all automorphisms of Υ, that is of all bijections $\alpha : V(\Upsilon) \to V(\Upsilon)$ such that $\alpha(v_1)$ and $\alpha(v_2)$ are joint by an edge iff v_1 and v_2 are joined by an edge. This action is not free but this does not cause any harm in this case because we do not explicitly use the group in the construction. Clearly, λ is G-equivariant. As usual we assume that $\mu^\bullet$ is equivariant as well and modeled on a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ admitting a measurable flow. $\mathbb{P}$ is assumed to be stationary. As Υ_2 can be identified with $\mathbb{R}$ we will assume from now on $k \geq 3$. Let d denote the natural distance on Υ. We will consider the L^p cost functions $c(x, y) = d(x, y)^p$ for $p \geq 1$. As it will turn out in section 8.3 this will imply that the mean asymptotic transportation cost are finite. Moreover, just as in Lemma 3.2 we exclude the case of $p = 1$ due to non-uniqueness of the solution to the optimal transportation problem.

We restrict ourselves to this rather special situation due to two reasons. Firstly, the general transport problem on Υ_k is very difficult to solve and estimates on the transportation cost are difficult to obtain. Secondly, Υ_k resembles many features of hyperbolic space $\mathbb{H}$. Hence, one might be able to use this ansatz to construct (maybe optimal) couplings with finite cost on $\mathbb{H}$.

8.1 Basic Results

We want to adapt the results of the preceding sections to this setting. We say that a finite subset $\Xi = \{\xi_1, \ldots, \xi_k\} \in \Upsilon$ has the non-equidistance property (NE) iff for all $1 \leq i \leq k$ $\xi_i = (e_i, x_i)$ we have $x_i \in (0, 1)$ and $1 - x_j \neq x_i \neq x_j$ for $j \neq i$. The second requirement says that the sets $E_{ij} = \{z \in \Upsilon \ : \ d(z, x_i) = d(z, x_j)\}$ should have zero λ measure.

Lemma 8.1. *Given a finite set $\Xi = \{\xi_1, \ldots, \xi_k\}$ with property (NE), a probability density $\rho \in L^1(\Upsilon, \lambda)$ and positive numbers $a_1, \ldots, a_k$ summing to one.*

(i) There exists a unique coupling q of $\rho\lambda$ and $\sigma = \sum_{j=1}^k a_j \delta_{\xi_j}$ which minimizes the cost functional $\mathsf{Cost}(\cdot)$.
(ii) There exists a (λ-a.e. unique) map $T : \{\rho > 0\} \to \Xi$ with $T_*(\rho\lambda) = \sigma$ which minimizes $\int c(x, T(x))\rho(x)\,\lambda(dx)$.
(iii) There exists a (λ-a.e. unique) map $T : \{\rho > 0\} \to \Xi$ with $T_*(\rho\lambda) = \sigma$ which is c-monotone (in the sense that the closure of $\{(x, T(x)) : \ \rho(x) > 0\}$ is a c-cyclically monotone set).
(iv) The minimizers in (i), (ii) and (iii) are related by $q = (Id, T)_*(\rho\lambda)$ or, in other words,

$$q(dx, dy) \ = \ \delta_{T(x)}(dy)\,\rho(x)\,\lambda(dx).$$

Proof. The proof goes along the same lines as the proof of Lemma 3.2. Write again $\rho\lambda = \sum_{i=1}^k \lambda_i$ with $\lambda_i := q(\cdot, \{\xi_i\})$. To show that the measures λ_i are mutually singular it is sufficient to prove this for every edge e. However, looking at the transport from a fixed edge to two points this is just a usual transport on $\mathbb{R}$. Because of property (NE) we get a unique solution which is given by a transport map. □

Remark 8.2. Identifying each edge with an higher dimensional set, say $[0, 1]^2$, instead of the unit interval allows to avoid the (NE) assumption as the sets E_{ij} are λ null sets.

Lemma 8.3. *The support of almost every realization μ^ω of the Poisson point process has the (NE) property.*

Proof. By the definition of a Poisson point process the random variables $\mu^\bullet(e)$ and $\mu^\bullet(f)$ are independent for different edges e and f. Moreover, given that $\mu^\omega(e) = n$ the positions of the n points in e are independent and uniformly distributed. Hence, it is sufficient to prove that countably many independent uniformly distributed points on e, say $(X_i)_{i\in\mathbb{N}}$, do not meet almost surely.
Consider the event $A_{i,j} = \{X_i = X_j\} \cup \{X_i = 1 - X_j\}$, an event which occurs if particle i and particle j meet. However, due to the diffusiveness of λ this event has zero probability. Taking the union over all such events yields the probability that two particles meet. This is a countable union of a null set and therefore a null set. □

In this setting the mean transportation cost of a semicoupling $q^\bullet$ between λ and $\mu^\bullet$ is defined to be

$$\mathfrak{C}(q^\bullet) \ = \ \sup_{B\in \mathrm{Adm}(\Upsilon)} \frac{1}{\lambda(B)} \mathbb{E}\left[\int_{\Upsilon\times B} c(x, y)\ q^\bullet(dx, dy)\right],$$

where the collection of admissible sets is defined by $\mathrm{Adm}(M) = \{B \in \mathcal{B}(M) : \exists I \subset E(\Upsilon), 1 \leq |I| < \infty, B = \bigcup_{e\in I} e\}$, that is B is admissible iff it is a finite union of edges. For an equivariant semicoupling $q^\bullet$ the mean transportation cost reduce to

$$\mathfrak{C}(q^\bullet) \ = \ \mathbb{E}\left[\int_{\Upsilon\times e} c(x, y)\ q^\bullet(dx, dy)\right],$$

for some arbitrary edge e. The optimal transport cost are again defined by

$$\mathfrak{c}_{i,\infty} = \inf_{q^\bullet\in\Pi_{is}(\lambda,\mu^\bullet)} \mathfrak{C}(q^\bullet).$$

Definition 8.4. *A semicoupling $q^\bullet$ between λ and $\mu^\bullet$ is called*

i) asymptotically optimal *iff*

$$\mathfrak{C}(q^\bullet) \leq \mathfrak{c}_{i,\infty}.$$

ii) optimal *iff $q^\bullet$ is equivariant and asymptotically optimal.*

Remark 8.5. It is clear that $\mathfrak{c}_{i,\infty} \geq \mathfrak{c}_\infty$. We do not know if they are actually equal or not. Because we are only interested in equivariant semicouplings we define asymptotic optimality via $\mathfrak{c}_{i,\infty}$.

We can copy the proof of Proposition 4.1 more or less line to line to get

Proposition 8.6. *Given a coupling q^ω of λ and μ^ω for fixed $\omega \in \Omega$, then the following properties are equivalent:*

(i) For each bounded Borel set $A \subset \Upsilon$, the measure $1_{A\times\Upsilon} q^\omega$ is the unique optimal coupling of the measures $1_A\lambda$ and $\mu_A^\omega := q^\omega(A, \cdot)$.

(ii) The support of q^ω is c-cyclically monotone.

(iii) There exists a c-cyclically monotone map $T^\omega : \Upsilon \to \Upsilon$ such that

$$q^\omega = (Id, T^\omega)_* \lambda. \tag{8.1}$$

Following the argumentation in chapter 4 or chapter 6 we get

Theorem 8.7. *There exists at most one optimal semicoupling of λ and $\mu^\bullet$.*

Remark 8.8. These results also hold for the case of $\mu^\bullet$ having intensity $\beta \neq 1$. However, the semicoupling we will construct will just have the chance of being optimal for $\beta = 1$.

For an edge e denote the r neighbourhood of e, the set of points with distance at most r to e, by $B_e(r)$. Similarly for a vertex x denote its r neighbourhood by $B_x(r)$. For two edges e and f define their distance by

$$d(e, f) \ := \ \inf\{m \in \mathbb{N} : e \in B_m(f)\},$$

which is nothing but the Hausdorff distance between e and f. We hope that no confusion is caused by using the same notation for the distance on Υ and the Hausdorff distance on Υ. The setting should be clear from the context. For the construction we need to count some edges.

Lemma 8.9. *Fix an edge e. Then the following holds.*

(i) The number of edges of distance j to e is $2 \cdot (k-1)^j$.

(ii) $\lambda(B_e(n)) = 1 + \sum_{j=1}^n 2 \cdot (k-1)^j = \frac{2(k-1)^{n+1}-k}{k-2}$.

(iii) The number of edges f such that we have $e \in B_f(n)$ is $\lambda(B_e(n))$.

Proof. (i) is a consequence of the tree structure of Υ. (ii) follows from (i). For (iii) note that we have $e \in B_f(n)$ iff $d(e, f) \leq n$ iff $f \in B_e(n)$ and each edge has volume one. □

We will use the mass transport principle again. However, this time we will use a slightly different version than before.

Lemma 8.10 (mass transport principle)**.** *Let $\gamma : E(\Upsilon) \times E(\Upsilon) \to \mathbb{R}_+$ be invariant under the diagonal action of G, i.e. $\gamma(e, f) = \gamma(ge, gf)$ for all $e, f \in E(\Upsilon)$ and $g \in G$. Then,*

$$\sum_{f\in E(\Upsilon)} \gamma(e, f) = \sum_{e\in E(\Upsilon)} \gamma(e, f).$$

Proof. Fix two edges e_0 and f_0. Then, we have

$$\sum_{f\in E(\Upsilon)} \gamma(e_0, f) \ = \ \sum_{n=0}^{\infty} \sum_{f:d(e_0,f)=n} \gamma(e_0, f) \ = \ \sum_{n=0}^{\infty} \sum_{f:d(e,f_0)=n} \gamma(e, f_0) \ = \ \sum_{e\in E(\Upsilon)} \gamma(e, f_0).$$

□

Just as before this allows us to resolve the asymmetry in the definition of transportation cost for equivariant semicouplings.

Lemma 8.11. *Let $q^\bullet$ be an equivariant semicoupling between λ and $\mu^\bullet$. Then,*

$$\mathbb{E}\left[\int_{e\times\Upsilon} c(x,y)\ q^\bullet(dx,dy)\right] \ = \ \mathbb{E}\left[\int_{\Upsilon\times e} c(x,y)\ q^\bullet(dx,dy)\right].$$

Proof. Use the mass transport principle with $\gamma(e,f) = \mathbb{E}\left[\int_{e\times f} c(x,y)\ q^\bullet(dx,dy)\right]$. □

We will use the following version of the law of large numbers due to Pruitt [Pru66]:

Theorem 8.12. *Let $(Z_j)_{j\in\mathbb{N}}$ be a sequence of iid L^1 random variables with mean α. Let $(a_{nk})_{n,k\in\mathbb{N}}$ be an array of numbers such that*

$$\lim_{n\to\infty} a_{nk} = 0 \quad \text{for any } k, \ \lim_{n\to\infty}\sum_{k=1}^{\infty} a_{nk} = 1 \ \ \text{and} \ \sum_{k=1}^{\infty} |a_{nk}| \le A < \infty \ \ \text{for all } n.$$

If $\max_{k\in\mathbb{N}} |a_{nk}| = O(n^{-1})$ *then* $\mathbb{E}[Z_1^2] < \infty$ *implies that*

$$Y_n = \sum_{k=1}^{\infty} a_{nk} Z_k \to \alpha \quad \text{almost surely.}$$

8.2 Construction

We want to construct a coupling by approximation very similar to the case of amenable spaces. However, instead of semicouplings between λ and $1_B\mu$ for some bounded set B, we consider the "classical" optimal transportation problem between the two random measures $\nu_B^\bullet := \frac{\mu^\bullet(B)}{\lambda(B)} 1_B\lambda$ and $1_B\mu^\bullet$. $\nu_B^\bullet$ inherits some equivariance properties from $\mu^\bullet$, i.e. $\nu_{gB}^{\theta_g\omega}(gC) = \nu_B^\omega(C)$ for all Borel sets C and $\omega\in\Omega$. To get a uniform cost bound we mix as in the amenable case. The advantage of this transportation problem is that it is easier to control the amount of the first marginal being transported somewhere by a careful analysis (or counting) in the mixing procedure because we do not allow mass being transported over the boundary of B. It is not clear if this just gives a coupling with finite cost or even the optimal coupling (see Remark 8.16).

For an edge e let $q_{B_e(m)}^\bullet$ be the unique optimal coupling between $\nu_{B_e(m)}^\bullet$ and $1_{B_e(m)}\mu^\bullet$. By uniqueness, it inherits the same equivariance properties as $\nu_{B_e(m)}^\bullet$. Put

$$Q_e^n(dx,dy,d\omega) \ := \ \frac{1}{\lambda(B_e(2^n))} \sum_{f\in B_e(2^n)} q_{B_f(2^n)}^\bullet(dx,dy)\mathbb{P}(d\omega),$$

where the sum is over all edges $f\in B_e(n)$. We do not restrict the second marginal of Q_e^n to e, as we do not know from where e gets its mass. In the amenable case this missing information was not necessary, because we could use amenability. However, we do not have amenability here, therefore we do not restrict to e. Put

$$Q^n \ := \ \frac{1}{\alpha_n}\sum_{e\in E(\Upsilon)} Q_e^n,$$

where α_n is chosen such that $(\pi_{2,3})_* Q^n = \mu^\bullet\mathbb{P}$, that is

$$\alpha_n \ = \ \sum_{f\in E(\Upsilon)} \frac{\lambda(B_e(2^n)\cap B_f(2^n))}{\lambda(B_e(2^n))}.$$

Indeed, fix an edge e. Mass will be transported to $1_e\mu^\bullet$ under Q^n_f iff there is an edge h such that $e,f \in B_h(2^n)$ that is iff $h \in B_e(2^n) \cap B_f(2^n)$. If there is such an edge h, $q^\bullet_{B_h(2^n)}$ transports the mass $\mu^\bullet(e)$ to the edge e. Thus, Q^n_f transports the mass $\frac{\lambda(B_e(2^n)\cap B_f(2^n))}{\lambda(B_e(2^n))} \cdot \mu^\bullet(e)$ to e. Summing over all edges yields this α_n.

Q^n is equivariant by construction. The first marginal of Q^n restricted to one edge, say e, can be written as

$$\lambda^n_e = 1_e \cdot (\pi_1)_* Q^n = \sum_{f \in E(\Upsilon)} a_{nf} Z_f,$$

where $Z_f = \mu^\bullet(f)$ are iid Poisson random variables with mean 1. The a_{nf} count the contribution of the Poisson points Z_f on the edge f to the first marginal in the mixing procedure. To be more precise, let $f \in E(\Upsilon)$ be such that $e \in B_f(2^n)$. The first marginal of $q^\bullet_{B_f(2^n)}$ is by definition $\sum_{h \in B_f(2^n)} Z_h/\lambda(B_f(2^n))$. Hence, every time the 2^n-ball around f comes up in the mixing of Q^n each edge $h \in B_f(2^n) \cap B_e(2^n)$ contributes to λ^n_e with $Z_h/\lambda(B_f(2^n))$. Considering all these contributions together with the normalization yields the numbers a_{nf}. Without having to work too much we can say quite a bit about these numbers.

Lemma 8.13. *i)* $0 \le a_{nf} \le \frac{1}{\lambda(B_e(2^n))}$

ii) $\sum_{f \in E(\Upsilon)} a_{nf} = 1.$

Proof. Just as $\lambda^n_e = \sum_{f\in E(\Upsilon)} a_{nf} Z_f$ we also have $\lambda^g_e = 1_e \cdot (\pi_1)_* Q^n_g = \sum_{f \in E(\Upsilon)} b_{nf} Z_f$. To get an expression for b_{nf} we can argue similarly as above. For Z_f to contribute to λ^g_e we need to find an $h \in B_g(2^n)$ such that $e,f \in B_h(2^n)$ that is $e,f,g \in B_h(2^n)$. There are $|B_e(2^n) \cap B_g(2^n) \cap B_f(2^n)|$ such h each contributing with $Z_f/\lambda(B_e(2^n))$. This gives

$$b_{nf} = \frac{1}{(\lambda(B_e(2^n)))^2} \lambda(B_e(2^n) \cap B_g(2^n) \cap B_f(2^n)).$$

This yields directly an expression for a_{nf}. We just need to add up all the contributions from the different edges.

$$a_{nf} = \frac{1}{(\lambda(B_e(2^n)))^2} \frac{1}{\alpha_n} \sum_{g \in E(\Upsilon)} \lambda(B_e(2^n) \cap B_g(2^n) \cap B_f(2^n)).$$

Clearly a_{nf} is nonnegative for any f and maximal if $f = e$. In that case

$$a_{ne} = \frac{1}{\lambda(B_e(2^n))},$$

proving the first part of the Lemma. For the second part put $F(e,f) = \mathbb{E}[Q^n(e,f)]$ for two edges e and f. By invariance of Q^n we have $F(e,f) = F(ge,gf)$ for any $g \in G$ ($F(e,f)$ actually just depends on the distance between e and f). Therefore we can use the mass transport principle

$$\sum_{f \in E(\Upsilon)} F(e,f) = \mathbb{E}[Q^n(e,\Upsilon)] = \mathbb{E}[\lambda^n_e(e)] = \sum_{f\in E(\Upsilon)} a_{nf}\mathbb{E}[Z_f] = \sum_{f \in E(\Upsilon)} a_{nf}.$$

On the other hand this equals

$$\sum_{e \in E(\Upsilon)} F(e,f) = \mathbb{E}[Q^n(\Upsilon,f)] = \mathbb{E}[Z_f] = 1.$$

□

Setting $\hat{\mathfrak{c}}_n := \frac{1}{\lambda(B_f(2^n))} \mathfrak{Cost}(q^\bullet_{B_f(2^n)})$ for some arbitrary edge f we have

Proposition 8.14. *(i)* $\limsup_{n\to\infty} \widehat{\mathfrak{c}}_n =: \widehat{\mathfrak{c}}_\infty < \infty$.

(ii) For all edges e and $n \in \mathbb{N}$

$$\int_{\Upsilon\times e\times\Omega} c(x,y)Q^n(dx,dy,d\omega) \le \widehat{\mathfrak{c}}_\infty.$$

(iii) The sequence $(Q^n)_n$ is tight and there is a subsequence converging to some Q^∞ which has cost bounded by $\widehat{\mathfrak{c}}_\infty$.

Proof. We postpone the proof of (i) to subsection 8.3. For (ii) note that only those $q^\bullet_{B_f(2^n)}$ contribute to the integral for which we have $e \in B_f(2^n)$. Secondly, each of the balls $B_f(2^n)$ that do contribute to the integral appear in the mixing procedure exactly the same number of times because all balls appear in the mixing the same number of times. Hence, we can think of the balls contributing to the integral as appearing just once. Thirdly, just as in the amenable case, summing over all balls $B_f(2^n)$ containing e has the effect that e appears at each possible position within the ball of radius 2^n with equal probability. Together with invariance this leads to the following estimate.

$$\begin{aligned}
&\int_{\Upsilon\times e\times\Omega} c(x,y)Q^n(dx,dy,d\omega)\\
=\ & \frac{1}{\lambda(B_e(2^n))}\sum_{f\in B_e(2^n)}\int_{\Upsilon\times e\times\Omega} c(x,y)\cdot q^\bullet_{B_f(2^n)}(dx,dy)\mathbb{P}(d\omega)\\
=\ & \frac{1}{\lambda(B_e(2^n))}\int_{\Upsilon\times B_e(2^n)\times\Omega} c(x,y)\cdot q^\bullet_{B_e(2^n)}(dx,dy)\mathbb{P}(d\omega)\\
=\ & \hat{\mathfrak{c}}_n \le \widehat{\mathfrak{c}}_\infty,
\end{aligned}$$

uniformly in n as $\limsup_{n\to\infty} \widehat{\mathfrak{c}}_n = \widehat{\mathfrak{c}}_\infty$. Because of the invariance this is independent of the specific edge e.
For (iii) note that the uniform cost bound directly gives tightness of the measure $Q^n_e = 1_{\Upsilon\times e}Q^n$ and therefore the convergence of a subsequence to some limit Q^∞_e exactly as in Lemma 5.1. Put $Q^\infty = \sum_{e\in E(\Upsilon)} Q^\infty_e$. By invariance we have $\mathfrak{C}(Q^n) \le \hat{\mathfrak{c}}_n$. Hence, we have $\mathfrak{C}(Q^\infty) \le \hat{\mathfrak{c}}_\infty$. □

Theorem 8.15. *Q^∞ is an equivariant coupling between λ and $\mu^\bullet$.*

Proof. (i) Second/third marginal: For any n the second and third marginal of Q^n is $(\pi_{2,3})_*Q^n = \mu^\bullet\mathbb{P}$, by construction. Therefore, we can argue as in the proof of Theorem 5.3. Take a function $\phi \in C_b^+(\Upsilon\times\Omega)$. Then, we have due to the previous Lemma

$$\begin{aligned}
&\int_{\Upsilon\times\Omega} \phi(y,\omega)Q^\infty(dx,dy,d\omega)\\
=\ & \sum_{e\in E(\Upsilon)}\int_{\Upsilon\times\Omega} \phi(y,\omega)Q^\infty_e(dx,dy,d\omega)\\
=\ & \sum_{e\in E(\Upsilon)}\lim_{l\to\infty}\int_{\Upsilon\times\Omega} \phi(y,\omega)Q^{k_l}_e(dx,dy,d\omega)\\
=\ & \sum_{e\in E(\Upsilon)}\int_{\Upsilon\times\Omega} \phi(y,\omega)1_e(y)(\mu^\bullet\mathbb{P})(dy,d\omega)\\
=\ & \int_{\Upsilon\times\Omega} \phi(y,\omega)(\mu^\bullet\mathbb{P})(dy,d\omega).
\end{aligned}$$

(ii) First marginal: We saw that the first marginal of Q^n restricted to an edge e has density $\sum_{f\in E(\Upsilon)} a_{nf} Z_f$. Moreover, due to Lemma 8.13 the array $(a_{nf})_{n\in\mathbb{N}, f\in E(\Upsilon)}$ satisfies all the conditions of Theorem 8.12. Hence, the density converges to 1. Thus, λ^n converges to λ.

(iii) Equivariance: Q^∞ inherits the equivariance from Q^n as in the previous chapters, e.g. see proof of Proposition 2.23.

□

Remark 8.16. In order to show that this coupling is actually the unique optimal coupling one would have to show that $\mathfrak{c}_\infty = \hat{\mathfrak{c}}_\infty$. To this end, one needs to be able to control the mass being transported over the boundary by the optimal semicoupling between λ and $1_{B_e(n)}\mu^\bullet$, i.e. the transportation cost between ν_B and the optimal choice of density ρ_B for the semicoupling between λ and $1_B\mu^\bullet$. However, this is a rather difficult problem which is not solved (see Remark 2 in [DSS11], comments after Theorem 2 in [BB11] and also Remark 7.18).

8.3 Cost estimates

In this section we want to prove part (i) and (ii) of Proposition 8.14. To this end let $r\Upsilon$ be the rooted k-tree, that is a tree with a distinguished vertex r, the root, whose degree is $(k-1)$ and all other vertices having degree k. Moreover, denote the j-neighborhood of r by rB_j. Recall the L^p-transportation distance between two measure valued random variables $\eta_1^\bullet, \eta_2^\bullet$ from section 7.3.2

$$\mathbb{W}_p(\eta_1^\bullet, \eta_2^\bullet) = \inf\left\{ \left[\mathbb{E}\left[\int_{\Upsilon\times\Upsilon} d(x,y)^p\, \theta^\bullet(dx,dy)\right]\right]^{1/p} \;:\; \theta^\bullet \text{ is a coupling of } \eta_1^\bullet, \eta_2^\bullet \right\}.$$

We will use the triangle inequality for $\mathbb{W}_p$ several times. First, we establish an transportation cost estimate on $r\Upsilon$. This will allow us to deduce the claimed bound. Put

$$\tilde{c}_n \;=\; \frac{1}{\lambda(rB_n)} \mathbb{W}_p^p\left(\nu_{rB_n}, 1_{rB_n}\mu^\bullet\right).$$

Lemma 8.17. *For $n\in\mathbb{N}$ it holds that*

$$\tilde{c}_{2n}^{1/p} \;\le \tilde{c}_n^{1/p} + 8n\cdot\lambda(rB_n)^{-1/2p}.$$

Proof. rB_{2n} can be decomposed into $v = ((k-1)^n + 1)$ disjoint copies of rB_n. This gives, using the triangle inequality

$$\begin{aligned}
\mathbb{W}_p\left(\nu^\bullet_{rB_{2n}}, 1_{rB_{2n}}\mu^\bullet\right) \;&\le\; \mathbb{W}_p\left(\sum_{i=1}^v \nu^\bullet_{rB_n^{(i)}}, 1_{rB_{2n}}\mu^\bullet\right) + \mathbb{W}_p\left(\sum_{i=1}^v \nu^\bullet_{rB_n^{(i)}}, \nu^\bullet_{rB_{2n}}\right)\\
&\le\; v^{1/p}\mathbb{W}_p\left(\nu^\bullet_{rB_n}, 1_{rB_n}\mu^\bullet\right) + \mathbb{W}_p\left(\sum_{i=1}^v \nu^\bullet_{rB_n^{(i)}}, \nu^\bullet_{rB_{2n}}\right).
\end{aligned}$$

We need to get a bound on the second term. Put $Z_i = \mu^\bullet(rB_n^{(i)})$. We estimate rather roughly. For each i we have to transport at most mass $\left|Z_i - \frac{\sum_{j=1}^v Z_j}{v}\right|$ at most distance $4n$. Put $\alpha = \lambda(rB_n)$.

$$\begin{aligned}
\mathbb{W}_p^p\left(\nu^\bullet_{rB_{2n}}, \sum_{i=1}^v \nu^\bullet_{rB_n^{(i)}}\right) \;&\le\; v\cdot(4n)^p\cdot\mathbb{E}\left[\left|Z_i - \frac{\sum_{j=1}^q Z_j}{v}\right|\right]\\
&\le\; v\cdot(4n)^p\cdot\mathbb{E}\left[|Z_i-\alpha| + \left|\alpha - \frac{\sum_{j=1}^v Z_j}{v}\right|\right]\\
&\le\; v\cdot(4n)^p\cdot\left(\sqrt{\alpha}+\sqrt{\frac{\alpha}{v}}\right).
\end{aligned}$$

Dividing by $(v \cdot \alpha)^{1/p}$ yields the result. $\square$

The error term $8n \cdot \lambda(rB_n)^{-1/2p}$ is clearly summable in n. Hence, we have

Corollary 8.18. *There is a finite constant $\tilde{c}_\infty$ such that*

$$\lim_{n\to\infty} \tilde{c}_n \;=\; \tilde{c}_\infty.$$

The next step is to show that $\limsup_{n\to\infty} \hat{\mathfrak{c}}_n =: \hat{\mathfrak{c}}_\infty$ is bounded. The key is to decompose $B_e(2^n)$ into two rooted trees and one edge, $rB^1_{2^n}, rB^2_{2^n}$ and e. Note that $\lambda(B_e(n)) \cdot \hat{\mathfrak{c}}_n = \mathbb{W}_p^p\left(\nu^\bullet_{B_e(n)}, 1_{B_e(n)}\mu^\bullet\right)$.

Lemma 8.19. $\hat{\mathfrak{c}}_n^{1/p} \;\le\; \tilde{c}_n^{1/p} + \mathbb{W}_p(\nu^\bullet_e, 1_e\mu^\bullet) \cdot \lambda(B_e(n))^{-1/p} + (2n+1)\cdot 5^{1/p} \cdot \lambda(B_e(n))^{-1/2p}$.

Proof. We need to estimate the transportation cost between $\nu^\bullet_{B_e(n)}$ and $\nu^\bullet_{rB^1_n} + \nu^\bullet_{rB^2_n} + \nu^\bullet_e$. Indeed, by the triangle inequality we have

$$\begin{aligned}
&\mathbb{W}_p\left(\nu^\bullet_{B_e(n)}, 1_{B_e(n)}\mu^\bullet\right)\\
&\le\; \mathbb{W}_p\left(\nu^\bullet_{B_e(n)}, \nu^\bullet_{rB^1_n} + \nu^\bullet_{rB^2_n} + \nu^\bullet_e\right) + 2^{1/p}\mathbb{W}_p\left(\nu_{rB_n}, 1_{rB_n}\mu^\bullet\right) + \mathbb{W}_p\left(\nu_e, 1_e\mu^\bullet\right).
\end{aligned}$$

As before we estimate very roughly. Put $Z_e = \mu^\bullet(e)$ and $Z_i = \mu^\bullet(rB^i_n)$ for $i = 1, 2$. The transportation distance is at most $(2n+1)$ and we need to transport at most mass

$$\left|Z_e - \frac{Z_1+Z_2+Z_e}{\lambda(B_e(n))}\right| + \left|Z_1 - \frac{(Z_1+Z_2+Z_e)\cdot\lambda(rB_n)}{\lambda(B_e(n))}\right| + \left|Z_2 - \frac{(Z_1+Z_2+Z_e)\cdot\lambda(rB_n)}{\lambda(B_e(n))}\right|.$$

The expectation can be bounded as before, e.g. for the middle term we get

$$\begin{aligned}
&\mathbb{E}\left[\left|Z_1 - \frac{(Z_1+Z_2+Z_e)\cdot\lambda(rB_n)}{\lambda(B_e(n))}\right|\right]\\
&\le\; \mathbb{E}\left[|Z_1 - \lambda(rB_n)|\right] + \mathbb{E}\left[\left|\lambda(rB_n) - \frac{(Z_1+Z_2+Z_e)\cdot\lambda(rB_n)}{\lambda(B_e(n))}\right|\right]\\
&\le\; \sqrt{\lambda(rB_n)} + \frac{\lambda(rB_n)}{\sqrt{\lambda(B_e(n))}}.
\end{aligned}$$

The other terms are similar. Therefore, we get

$$\mathbb{W}_p^p\left(\nu^\bullet_{B_e(n)}, \nu^\bullet_{rB^1_n} + \nu^\bullet_{rB^2_n} + \nu^\bullet_e\right) \le (2n+1)^p \cdot 5 \cdot \sqrt{\lambda(rB_n)}.$$

Taking p-th root and normalizing everything with $\lambda(B_e(n))^{1/p}$ yields the claim. $\square$

Corollary 8.20.

$$\limsup_{n\to\infty} \hat{\mathfrak{c}}_n \;\le\; \infty.$$

Proof. The last Lemma together with the previous Corollary implies

$$\limsup_n \hat{\mathfrak{c}}_n \le \limsup_n \tilde{c}_n = \tilde{c}_\infty < \infty.$$

$\square$

In the previous section we constructed an equivariant coupling between λ and $\mu^\bullet$ with cost bounded by $\hat{\mathfrak{c}}_\infty < \infty$. Therefore, we can conclude.

Corollary 8.21. $\mathfrak{c}_{i,\infty} < \infty$.

Chapter 9

Stability

As an application of the previous results, especially the existence and uniqueness results, we want to study stability properties of the optimal coupling between two random measures. Moreover, we will show some metric properties of the mean transportation cost.

Given sequences of random measures $(\lambda_n^\bullet)_{n\in\mathbb{N}}, (\mu_n^\bullet)_{n\in\mathbb{N}}$ and their optimal couplings $(q_n^\bullet)_{n\in\mathbb{N}}$ we want to understand which kind of convergence $\lambda_n^\bullet \to \lambda^\bullet, \mu_n^\bullet \to \mu^\bullet$ implies the convergence $q_n^\bullet \to q^\bullet$, where $q^\bullet$ denotes the/an optimal coupling between $\lambda^\bullet$ and $\mu^\bullet$. If for all n $\lambda_n, \lambda, \mu_n, \mu$ are probability measures q_n the optimal coupling between λ_n and μ_n (all transportation cost involved bounded by some constant) and $\lambda_n \to \lambda, \mu_n \to \mu$ weakly, then, by the classical theory (see Theorem 5.20 in [Vil09] or Proposition 1.3), also along a subsequence $q_n \to q$ weakly, where q is an optimal coupling between λ and μ.

A naive approach to our problem would be to ask for $\lambda_n^\bullet \xrightarrow{d} \lambda^\bullet$ and $\mu_n^\bullet \xrightarrow{d} \mu^\bullet$. However, in this case let $\lambda^\bullet$ and $\mu^\bullet$ be two independent Poisson point process and set $\lambda_n^\bullet \equiv \mu_n^\bullet \equiv \lambda^\bullet$. Then, we indeed have $\lambda_n^\bullet \xrightarrow{d} \lambda^\bullet$ and $\mu_n^\bullet \xrightarrow{d} \mu^\bullet$. Yet, the optimal couplings $(q_n^\bullet)_{n\in\mathbb{N}}$, which are just $q_n^\omega(dx,dy) = \delta_x(dy)\lambda_n^\omega(dx)$, do not converge to any coupling between $\lambda^\bullet$ and $\mu^\bullet$ in any reasonable sense. Moreover, the couplings $(q_n^\bullet)_{n\in\mathbb{N}}$ do 'converge' (they are all the same) to some coupling $\tilde{q}^\bullet$ with marginals being $\lambda^\bullet$ and $\lambda^\bullet$ having the same *distribution* as $\lambda^\bullet$ and $\mu^\bullet$.

The next best guess, instead of vague convergence in distribution is vague convergence on $M \times M \times \Omega$. Together with some integrability condition this will be the answer if the cost of the couplings converge.

For two random measure $\lambda^\bullet, \mu^\bullet$ with intensity one and $c(x,y) = d^p(x,y)$ with $p \in [1,\infty)$ write

$$\mathbb{W}_p^p(\lambda^\bullet,\mu^\bullet) = \inf_{q^\bullet\in\Pi_{is}(\lambda^\bullet,\mu^\bullet)} \mathfrak{C}(q^\bullet) = \inf_{q^\bullet\in\Pi_{is}(\lambda^\bullet,\mu^\bullet)} \mathbb{E}\left[\int_{M\times B_0} d^p(x,y)\ q^\bullet(dx,dy)\right].$$

We want to establish a triangle inequality for $\mathbb{W}_p$ and therefore restrict to L^p cost functions. We could also extend this to more general cost functions by using Orlicz type norms as developed in [Stu11]. However, to keep notations simple we stick to this case.

In this chapter, we will assume that all pairs of random measures considered will be equivariant and modeled on the same probability space $(\Omega, \mathfrak{A}, \mathbb{P})$. As usual $\mathbb{P}$ is assumed to be stationary. Moreover, we will always assume without explicitly mentioning it that the mean transportation cost is finite.

Recall the disintegration Theorem 2.9. This will allow us to use the gluing lemma.

Proposition 9.1. *Let $\mu^\bullet, \lambda^\bullet, \xi^\bullet$ be three equivariant random measures of unit intensity.*

i) $\mathbb{W}_p(\lambda^\bullet,\mu^\bullet) = 0 \quad \Leftrightarrow \quad \lambda^\omega = \mu^\omega \quad \mathbb{P} - a.s..$

ii) $\mathbb{W}_p(\lambda^\bullet,\mu^\bullet) = \mathbb{W}_p(\mu^\bullet,\lambda^\bullet).$

iii) $\mathbb{W}_p(\lambda^\bullet, \mu^\bullet) \le \mathbb{W}_p(\lambda^\bullet, \xi^\bullet) + \mathbb{W}_p(\xi^\bullet, \mu^\bullet)$.

Proof. i) $\mathbb{W}_p(\lambda^\bullet, \mu^\bullet) = 0$ iff there is a coupling of $\lambda^\bullet$ and $\mu^\bullet$ which is entirely concentrated on the diagonal almost surely, that is iff $\lambda^\omega = \mu^\omega$ $\mathbb{P}-$ almost surely.

ii) Let $q^\bullet$ be an optimal coupling between $\lambda^\bullet$ and $\mu^\bullet$. By definition

$$\mathbb{W}_p^p(\lambda^\bullet, \mu^\bullet) = \mathbb{E}\left[\int_{M\times B_0} d^p(x,y)q^\bullet(dx,dy)\right].$$

For $g, h \in G$ put

$$f(g,h) = \mathbb{E}\left[\int_{gB_0\times hB_0} d^p(x,y)q^\bullet(dx,dy)\right].$$

By equivariance and stationarity, we have $f(g,h) = f(kg,kh)$ for all $k \in G$. Hence, we can apply the mass transport principle.

$$\sum_{h\in G} f(g,h) = \mathbb{E}\left[\int_{gB_0\times M} d^p(x,y)q^\bullet(dx,dy)\right] = \sum_{g\in G} f(g,h) = \mathbb{E}\left[\int_{M\times hB_0} d^p(x,y)q^\bullet(dx,dy)\right].$$

This proves the symmetry.

iii) The random measures are random variables on some Polish space. Therefore, we can use the gluing Lemma (cf. [Dud02] or [Vil09], chapter 1) to construct an equivariant random measure $q^\bullet$ on $M \times M \times M$ such that

$$(\pi_{1,2})_* q^\bullet \in \Pi_{opt}(\lambda^\bullet, \mu^\bullet) \quad \text{and} \quad (\pi_{2,3})_* q^\bullet \in \Pi_{opt}(\mu^\bullet, \xi^\bullet),$$

where $\Pi_{opt}(\lambda^\bullet, \mu^\bullet)$ denotes the set of all optimal couplings between $\lambda^\bullet$ and $\mu^\bullet$. $q^\bullet$ is equivariant as the optimal couplings are equivariant and $q^\bullet$ is glued together along the common marginal of these two couplings.
We can indeed use the gluing lemma. Let $q_1^\bullet \in \Pi_{opt}(\lambda^\bullet, \mu^\bullet)$ and $q_2^\bullet \in \Pi_{opt}(\mu^\bullet, \xi^\bullet)$. Then, consider $1_{M\times gB_0\times\Omega} q_1^\bullet$ and $1_{gB_0\times M\times\Omega} q_2^\bullet$ to produce with the usual gluing Lemma a measure $q_g^\bullet$ on $M\times M\times M\times\Omega$ with the desired marginals on $M\times gB_0\times M\times\Omega$. As all these sets are disjoint we can add up the different $q_g^\bullet$ yielding $q^\bullet = \sum_{g\in G} q_g^\bullet$ a measure with the desired properties.
For $g, h \in G$ put

$$e(g,h) = \mathbb{E}\left[\int_{M\times gB_0\times hB_0} d^p(x,z)q^\bullet(dx,dy,dz)\right].$$

By equivariance of $q^\bullet$, we have $e(kg,kh) = e(g,h)$ for all $k \in G$. By the mass transport principle this implies

$$\mathbb{E}\left[\int_{M\times B_0\times M} d^p(x,z)q^\bullet(dx,dy,dz)\right] = \mathbb{E}\left[\int_{M\times M\times B_0} d^p(x,z)q^\bullet(dx,dy,dz)\right].$$

Then we can conclude, using the Minkowski inequality

$$\begin{aligned}
&\mathbb{W}_p(\lambda^\bullet, \xi^\bullet)\\
\le\ & \mathbb{E}\left[\int_{M\times M\times B_0} d^p(x,z)q^\bullet(dx,dy,dz)\right]^{1/p}\\
=\ & \mathbb{E}\left[\int_{M\times B_0\times M} d^p(x,z)q^\bullet(dx,dy,dz)\right]^{1/p}\\
\le\ & \mathbb{E}\left[\int_{M\times B_0\times M} d^p(x,y)q^\bullet(dx,dy,dz)\right]^{1/p} + \mathbb{E}\left[\int_{M\times B_0\times M} d^p(y,z)q^\bullet(dx,dy,dz)\right]^{1/p}\\
=\ & \mathbb{W}_p(\lambda^\bullet, \mu^\bullet) + \mathbb{W}_p(\mu^\bullet, \xi^\bullet).
\end{aligned}$$

In the last step we used the symmetry shown in part ii).

□

Remark 9.2. Note that the first two properties also hold for general cost functions and general semicouplings. The assumption of equal intensity is not needed for these statements.

Let $\mathcal{P}_p = \{\text{equivariant random measures } \mu^\bullet : \mathbb{W}_p(m, \mu^\bullet) < \infty\}$.

Proposition 9.3. *Let $(\mu_n^\bullet)_{n\in\mathbb{N}}, \mu^\bullet \in \mathcal{P}_p$ be random measures of intensity one. Let $q_n^\bullet$ denote the optimal coupling between m and $\mu_n^\bullet$ and $q^\bullet$ the optimal coupling between m and $\mu^\bullet$. Consider the following statements.*

i) $\mathbb{W}_p(\mu_n^\bullet, \mu^\bullet) \to 0$ *as* $n \to \infty$.

ii) $\mu_n^\bullet \mathbb{P} \to \mu^\bullet \mathbb{P}$ *vaguely and* $\mathbb{W}_p(\mu_n^\bullet, m) \to \mathbb{W}_p(\mu^\bullet, m)$ *as* $n \to \infty$.

iii) $q_n^\bullet \mathbb{P} \to q^\bullet \mathbb{P}$ *vaguely and* $\mathbb{W}_p(\mu_n^\bullet, m) \to \mathbb{W}_p(\mu^\bullet, m)$ *as* $n \to \infty$.

iv) $q_n^\bullet \mathbb{P} \to q^\bullet \mathbb{P}$ *vaguely and*

$$\lim_{R\to\infty} \limsup_{n\to\infty} \mathbb{E}\left[\int_{(\complement(B_0)_R)\times B_0} d^p(x,y) q_n^\bullet(dx,dy)\right] = 0,$$

where $(B_0)_R$ denotes the $R-$neighbourhood of B_0.

Then i) implies ii). iii) and iv) are equivalent and either of them implies i).

Proof. *i)* $\Rightarrow$ *ii)* : For any $f \in C_c(M \times \Omega)$ we have to show that $\lim_{n\to\infty} \mathbb{E}[\mu_n(f) - \mu(f)] = 0$. To this end, fix $f \in C_c(M \times \Omega)$ such that $\text{supp}(f) \subset K \times \Omega$ for some compact set K. f is uniformly continuous. Let $\eta > 0$ be arbitrary and set $\epsilon = \eta/(2m(K))$. Then, there is δ such that $d(x,y) \leq \delta$ implies $d(f(x,\omega), f(y,\omega)) \leq \epsilon$. Put $A = \{(x,y) : d(x,y) \geq \delta\}$ and denote by $q_n^\bullet$ an optimal coupling between $\mu_n^\bullet$ and $\mu^\bullet$. By assumption, there is $N \in \mathbb{N}$ such that for all $n > N$ we have $\mathbb{W}_p^p(\mu_n^\bullet, \mu^\bullet) \leq \frac{\eta\delta^p}{4\|f\|_\infty}$. Then, we can estimate for $n > N$

$$\begin{aligned}
|\mathbb{E}[\mu_n^\omega(f) - \mu^\omega(f)]| &\leq \left|\mathbb{E}\left[\int_{M\times M} (f(x,\omega) - f(y,\omega)) q_n^\omega(dx,dy)\right]\right| \\
&\leq \epsilon \cdot m(K) + \left|\mathbb{E}\left[\int_A (f(x,\omega) - f(y,\omega)) q_n^\omega(dx,dy)\right]\right| \\
&\leq \frac{\eta}{2} + 2\|f\|_\infty \mathbb{E}\left[q_n^\bullet(A)\right] \\
&\leq \frac{\eta}{2} + 2\|f\|_\infty \frac{1}{\delta^p} \mathbb{W}_p^p(\mu_n^\bullet, \mu^\bullet) \\
&\leq \frac{\eta}{2} + \frac{\eta}{2} = \eta,
\end{aligned}$$

The second assertion in *ii)* is a direct consequence of the triangle inequality:

$$\mathbb{W}_p(\mu_n^\bullet, m) \leq \mathbb{W}_p(\mu_n^\bullet, \mu^\bullet) + \mathbb{W}_p(\mu^\bullet, m)$$

and

$$\mathbb{W}_p(\mu^\bullet, m) \leq \mathbb{W}_p(\mu_n^\bullet, \mu^\bullet) + \mathbb{W}_p(\mu_n^\bullet, m).$$

Taking limits yields the claim.

iii) $\Leftrightarrow$ *iv)* : By the existence and uniqueness result we know that $q_n^\omega(dx,dy) = \delta_{T_n^\omega(x)}(dy) m(dx)$ and $q^\omega(dx,dy) = \delta_{T^\omega(x)}(dy) m(dx)$. In particular, we have $\mu_n^\omega(dx)\mathbb{P}(d\omega) = (T_n^\omega)_* m(dx)\mathbb{P}(d\omega)$. By Lemma 5.9 we know that the vague convergence of $q_n^\bullet \mathbb{P} \to q^\bullet \mathbb{P}$ implies that $T_n \to T$ in $m \otimes \mathbb{P}$ measure. This in turn implies the convergence of $f \circ (id, T_n) \to f \circ (id, T)$ in $m \otimes \mathbb{P}$ measure for any continuous and compactly supported function $f : M \times M \to \mathbb{R}$. Then, it follows as in the proof of Lemma 5.9 that

$$\mathbb{E}\int f(x, T_n(x))m(dx) \to \mathbb{E}\int f(x, T(x))m(dx).$$

Let $c_k(x, y)$ be a continuous compactly supported function such that for any $(x, y) \in (B_0)_{k-1} \times B_0$ we have $d^p(x, y) = c_k(x, y)$, for any $x \in \complement(B_0)_k$ we have $c_k(x, y) = 0$ and $c_k(x, y) \le d^p(x, y)$ for all $(x, y) \in M \times M$. Then, we have

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbb{E}\left[\int_{\complement((B_0)_R)\times B_0} d^p(x, y)q_n^\bullet(dx, dy)\right] \\
\le\ & \limsup_{n\to\infty} \left(\mathbb{E}\left[\int_{M\times B_0} d^p(x, y)q_n^\bullet(dx, dy)\right] - \mathbb{E}\left[\int_{M\times B_0} c_R(x, y)q_n^\bullet(dx, dy)\right]\right) \\
=\ & \mathbb{E}\left[\int_{M\times B_0} d^p(x, y)q^\bullet(dx, dy)\right] - \mathbb{E}\left[\int_{M\times B_0} c_R(x, y)q^\bullet(dx, dy)\right] \\
\le\ & \mathbb{E}\left[\int_{\complement((B_0)_{R-1}\times B_0} d^p(x, y)q^\bullet(dx, dy)\right].
\end{aligned}$$

Taking the limit of $R \to \infty$ proves the implication $iii) \Rightarrow iv)$. The other direction is similar.

$iv) \Rightarrow i)$: We will show that $\mathbb{W}_p(\mu_n^\bullet, \mu^\bullet) \to 0$ by constructing a not optimal coupling between $\mu_n^\bullet$ and $\mu^\bullet$ whose transportation cost converges to zero. Let T_n, T be the transportation maps from the previous steps. Put $Q_n(dx, dy) := (T_n, T)_* m$. This is an equivariant coupling of $\mu_n^\bullet$ and $\mu^\bullet$ because the maps T_n, T are equivariant in the sense that (see also Example 2.13)

$$T^{\theta_g\omega}(x) = gT^\omega(g^{-1}x).$$

The transportation cost are given by

$$\mathfrak{C}(Q_n) = \mathbb{E}\left[\int_{B_0\times M} d^p(x, y)Q_n(dx, dy)\right] = \mathbb{E}\left[\int_{B_0} d^p(T_n(x), T(x))m(dx)\right].$$

We want to divide the integral into four parts. Put $A^R = \{x : d(T(x), x) \ge R\}$ and similarly $A_n^R = \{x : d(T_n(x), x) \ge R\}$. The four parts will be the integrals over $B_0 \cap \complement^a A_n^R \cap \complement^b A^R$ with $a, b \in \{0, 1\}$ and $\complement^0 A = A$. We estimate the different integrals separately.

$$\mathbb{E}\left[\int_{B_0\cap\complement A_n^R\cap\complement A^R} d^p(T_n(x), T(x))m(dx)\right] \to 0,$$

by a similar argument as in the previous step due to the convergence of $T_n \to T$ in $m \otimes \mathbb{P}$ measure and the boundedness of the integrand.

$$\begin{aligned}
&\mathbb{E}\left[\int_{B_0\cap A_n^R\cap A^R} d^p(T_n(x), T(x))m(dx)\right] \\
\le\ & 2^p\ \mathbb{E}\left[\int_{B_0\cap A^R} d^p(x, T(x))m(dx)\right] + 2^p\ \mathbb{E}\left[\int_{B_0\cap A_n^R} d^p(x, T_n(x))m(dx)\right].
\end{aligned}$$

If $d(x, y) \le R, d(x, z) \ge R$ and $d(y, z) \le d(x, z) + R + a$ for some constant a$(= \mathrm{diam}(B_0))$, there is a constant C_1, e.g. $C_1 = 2+\mathrm{diam}(B_0)$, such that $d(y, z) \le C_1 d(x, z)$ (because $d(x, z)+R+a \le (2 + a)d(x, z)$). This allows to estimate with $(x = x, T(x) = z, T_n(x) = y)$

$$\mathbb{E}\left[\int_{B_0\cap\complement A_n^R\cap A^R} d^p(T_n(x), T(x))m(dx)\right] \le C_1^p\ \mathbb{E}\left[\int_{B_0\cap A^R} d^p(x, T(x))m(dx)\right].$$

Similarly

$$\mathbb{E}\left[\int_{B_0\cap A_n^R\cap \complement A^R} d^p(T_n(x),T(x))m(dx)\right] \le C_1^p\,\mathbb{E}\left[\int_{B_0\cap A_n^R} d^p(x,T_n(x))m(dx)\right].$$

This finally gives

$$\begin{aligned}
&\limsup_{n\to\infty}\mathbb{E}\left[\int_{B_0\times M} d^p(x,y)Q_n(dx,dy)\right]\\
\le\ & \lim_{R\to\infty}\limsup_{n\to\infty}\Bigg(2^p\,\mathbb{E}\left[\int_{B_0\cap A^R} d^p(x,T(x))m(dx)\right]+2^p\,\mathbb{E}\left[\int_{B_0\cap A_n^R} d^p(x,T_n(x))m(dx)\right]\\
&+C_1^p\,\mathbb{E}\left[\int_{B_0\cap A^R} d^p(x,T(x))m(dx)\right]+C_1^p\,\mathbb{E}\left[\int_{B_0\cap A_n^R} d^p(x,T_n(x))m(dx)\right]\Bigg)\\
=\ & 0,
\end{aligned}$$

by assumption. □

Remark 9.4. For an equivalence of all statements we would need that ii) implies iii). In the classical theory this is precisely the stability result (Theorem 5.20 in [Vil09]). This result is proven by using the characterization of optimal transports by cyclical monotone supports. However, as mentioned in the discussion on local optimality (see Remark 4.4) a cyclical monotone support is not sufficient for optimality in our case.

We do not have real stability in general but we get at least close to it.

Proposition 9.5. *Let $(\lambda_n^\bullet)_{n\in\mathbb{N}}$ and $(\mu_n^\bullet)_{n\in\mathbb{N}}$ be two sequences of equivariant random measures. Let $q_n^\bullet$ be the unique optimal coupling between $\lambda_n^\bullet$ and $\mu_n^\bullet$. Assume that $\lambda_n^\bullet\mathbb{P}\to\lambda^\bullet\mathbb{P}$ vaguely, $\mu_n^\bullet\mathbb{P}\to\mu^\bullet\mathbb{P}$ vaguely and $\sup_n\mathfrak{C}(q_n^\bullet)\le c<\infty$. Then, there is an equivariant coupling $q^\bullet$ of $\lambda^\bullet$ and $\mu^\bullet$ and a subsequence $(q_{n_k}^\bullet)_{k\in\mathbb{N}}$ such that $q_{n_k}^\bullet\mathbb{P}\to q^\bullet\mathbb{P}$ vaguely, the support of $q^\bullet$ is cyclically monotone and*

$$\mathfrak{C}(q^\bullet)\le\liminf_{n\to\infty}\mathfrak{C}(q_n^\bullet).$$

In particular, if

$$\lim_{n\to\infty}\mathfrak{C}(q_n^\bullet)=\inf_{\tilde q^\bullet\in\Pi_{is}(\lambda^\bullet,\mu^\bullet)}\mathfrak{C}(\tilde q^\bullet)$$

$q^\bullet$ is the/an optimal coupling between $\lambda^\bullet$ and $\mu^\bullet$ and $q_n^\bullet\mathbb{P}\to q^\bullet\mathbb{P}$ vaguely.

Proof. The proof is basically the same as for Proposition 2.23:
i) $(q_n^\bullet\mathbb{P})_{n\in\mathbb{N}}$ is relatively compact in the vague topology on $\mathcal{M}(M\times M\times\Omega)$:
Fix $f\in C_c(M\times M\times\Omega)$ and set $h(x,\omega)=\sup_y f(x,y,\omega)$. Then, due to the uniform continuity of f we have $h\in C_c(M\times\Omega)$. This yields

$$\int_{M\times M\times\Omega} f(x,y,\omega)q_n^\omega(dx,dy)\mathbb{P}(d\omega)\le\int_{M\times\Omega} h(x,\omega)\lambda_n^\omega\mathbb{P}(d\omega)\le\kappa<\infty,$$

uniformly in n because $\lambda_n^\bullet\mathbb{P}\to\lambda^\bullet\mathbb{P}$ vaguely. Hence, $(q_n^\bullet\mathbb{P})_{n\in\mathbb{N}}$ is relatively compact in the vague topology on $\mathcal{M}(M\times M\times\Omega)$ and there is a subsequence $(q_{n_k}^\bullet\mathbb{P})_{k\in\mathbb{N}}$ and a measure $q^\bullet\mathbb{P}\in\mathcal{M}(M\times M\times\Omega)$ such that $q_{n_k}^\bullet\mathbb{P}\to q^\bullet\mathbb{P}$ vaguely.

ii) $q^\bullet\mathbb{P}$ is a coupling of $\lambda^\bullet$ and $\mu^\bullet$:
Fix $g\in C_c(M\times\Omega)$. Put $A\subset M$ compact such that $\mathrm{supp}(g)\subset A\times\Omega$ and $A\in Adm(M)$. Denote the $R-$neighbourhood of A by A_R. By the uniform bound on transportation cost we have

$$q_n^\bullet\mathbb{P}(A,\complement(A_R),\Omega)\le m(A)\frac{c}{\vartheta(R)},\tag{9.1}$$

uniformly in n. Let $f_R : M \to [0,1]$ be a continuous compactly supported function such that $f_R(y) = 1$ for $y \in A_R$ and $f_R(y) = 0$ for $y \in \complement A_{R+1}$. As $q_n^\bullet \mathbb{P}$ is a coupling of $\lambda_n^\bullet$ and $\mu_n^\bullet$ we have due to monotone convergence

$$\begin{aligned}\int_{M\times\Omega} g(x,\omega)\lambda_n^\omega(dx)\mathbb{P}(d\omega) &= \int_{M\times M\times\Omega} g(x,\omega)q_n^\omega(dx,dy)\mathbb{P}(d\omega)\\ &= \lim_{R\to\infty}\int_{M\times M\times\Omega} g(x,\omega)f_R(y)q_n^\omega(dx,dy)\mathbb{P}(d\omega).\end{aligned}$$

Because of the uniform bound (9.1) we have

$$\left|\int_{M\times\Omega} g(x,\omega)\lambda_{n_k}^\omega(dx)\mathbb{P}(d\omega) - \int_{M\times M\times\Omega} g(x,\omega)f_R(y)q_{n_k}^\omega(dx,dy)\mathbb{P}(d\omega)\right| \le m(A)\frac{c\cdot\|g\|_\infty}{\vartheta(R)}.$$

Taking first the limit of $n_k \to \infty$ and then the limit of $R \to \infty$ we conclude using vague convergence and monotone convergence that

$$\begin{aligned}0 &= \lim_{R\to\infty}\lim_{k\to\infty}\left|\int_{M\times\Omega} g(x,\omega)\lambda_{n_k}^\omega(dx)\mathbb{P}(d\omega) - \int_{M\times M\times\Omega} g(x,\omega)f_R(y)q_{n_k}^\omega(dx,dy)\mathbb{P}(d\omega)\right|\\ &= \lim_{R\to\infty}\left|\int_{M\times\Omega} g(x,\omega)\lambda^\omega(dx)\mathbb{P}(d\omega) - \int_{M\times M\times\Omega} g(x,\omega)f_R(y)q^\omega(dx,dy)\mathbb{P}(d\omega)\right|\\ &= \left|\int_{M\times\Omega} g(x,\omega)\lambda^\omega(dx)\mathbb{P}(d\omega) - \int_{M\times M\times\Omega} g(x,\omega)q^\omega(dx,dy)\mathbb{P}(d\omega)\right|.\end{aligned}$$

This shows that the first marginal is $\lambda^\bullet$. Similarly, we can show that the second marginal is $\mu^\bullet$. The equivariance is inherited from the equivariance of $q_n^\bullet$. Just consider $f(x,y,\omega)$ and $\tilde{f}(x,y,\omega) = f(g^{-1}x, g^{-1}y, \theta_g\omega)$ as in the proof of Proposition 2.23.

iii) The support of $q^\bullet\mathbb{P}$ is cyclically monotone and $\mathfrak{C}(q^\bullet) \le \liminf \mathfrak{C}(q_{n_k}^\bullet)$:
The second assertion is a consequence of the lower semicontinuity of $\mathfrak{Cost}(\cdot)$ with respect to vague convergence. For the first assertion note that due to local optimality for any n $(q_n^\bullet)^{\otimes N}$ is concentrated on the set of $((x_1,y_1),\ldots,(x_N,y_N))$ such that

$$\sum_{i=1}^N c(x_i,y_i) \le \sum_{i=1}^N c(x_i,y_{i+1}),$$

with $y_{N+1} = y_1$. As the cost function c is continuous this is a closed set. If we restrict to points $(x_i,y_i) \in K$ for some compact set $K \subset M\times M$, such that $q^\bullet\mathbb{P}(\partial K) = 0$, then, for all n, $(1_K q_n^\bullet)^{\otimes N}$ is concentrated on this *compact* set. By vague convergence, this implies that $(1_K q^\bullet)^{\otimes N}$ is also concentrated on this set. Thus, the support of $q^\bullet$ is cyclically monotone. □

Remark 9.6. The last proposition also holds if we consider semicouplings instead of couplings. In that case we have to show one inequality and one equality for the marginals of $q^\bullet$. This is again the same as in the proof of Proposition 2.23.

Example 9.7 (Wiener mosaic)**.** Let $\mu_0^\bullet$ be a Poisson point process of intensity one on $\mathbb{R}^3$. Let each atom of μ_0 evolve according to independent Brownian motions for some time t. The resulting discrete random measure is again a Poisson point process, denoted by $\mu_t^\bullet$ (e.g. see page 404 of [Doo53]). Consider the transport problem between the Lebesgue measure $\mathcal{L}$ and $\mu_t^\bullet$ with cost function $c(x,y) = |x-y|^2$. Let $q_t^\bullet$ be the unique optimal coupling between $\mathcal{L}$ and $\mu_t^\bullet$. Then, $\mathfrak{C}(q_t^\bullet) = \mathbb{W}_2(\mathcal{L},\mu_t^\bullet) = \mathbb{W}_2(\mathcal{L},\mu_s^\bullet)$ for any $s \in \mathbb{R}$ as $\mu_s^\bullet$ and $\mu_t^\bullet$ are both Poisson point processes of intensity one. Moreover, we clearly have $\mu_s^\bullet\mathbb{P} \to \mu_t^\bullet\mathbb{P}$ vaguely as $s \to t$ and therefore $q_s^\bullet\mathbb{P} \to q_t^\bullet\mathbb{P}$ vaguely. By Lemma 5.9, this implies the convergence of the transport maps $T_s \to T_t$ in $\mathcal{L}\otimes\mathbb{P}$ measure. In particular, we get a continuously moving mosaic.

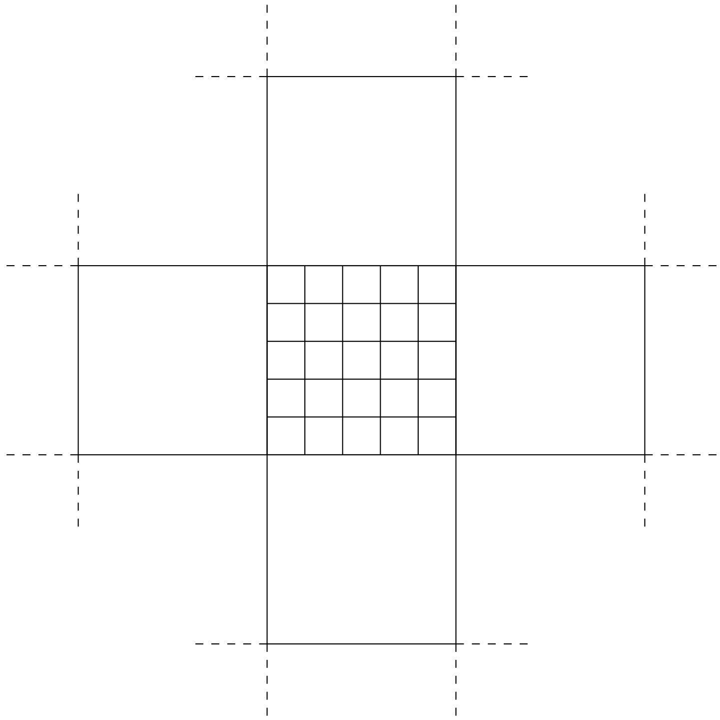

Figure 9.1: This is a schematic picture of the cubes used in the definition of $\Omega^a(\beta')$ and $\Omega(a,\rho)$. The cube in the center is $[-a/2, a/2)^d$.

Example 9.8 (Voronoi tessellation)**.** Let $\mu^\bullet$ be a simple point process of unit intensity. Put $\mu^\bullet_\beta = \beta \cdot \mu^\bullet$. We want to consider semicouplings between the Lebesgue measure $\mathcal{L}$ and $\mu^\bullet_\beta$ for $\beta > 1$. By the results of chapter 6, there is a unique optimal semicoupling $q^\bullet_\beta$ between $\mathcal{L}$ and $\mu^\bullet_\beta$. Moreover, $q^\omega_\beta = (id, T^\omega_\beta)_*\mathcal{L}$ for some measurable map $T_\beta : \mathbb{R}^d \times \Omega \to \mathbb{R}^d$. It is clear that q^ω_∞ induces the Voronoi tessellation with respect to the support of μ^ω_1 no matter which cost function ϑ we consider. We want to show that $q^\bullet_\beta \to q^\bullet_\gamma$ vaguely as $\beta \to \gamma$ for large γ. For this it is sufficient to show that $C : \beta \mapsto \mathfrak{C}(q^\bullet_\beta)$ is continuous. For $\gamma > \beta > 1$ note that $q^\bullet_\beta$ is also a semicoupling of $\mathcal{L}$ and $\mu^\bullet_\gamma$. Hence, $C(\beta)$ is monotonously decreasing in β. Moreover, from the previous Lemma we know that

$$\mathfrak{C}(q^\bullet_\gamma) \leq \liminf_{\beta \to \gamma} \mathfrak{C}(q^\bullet_\beta).$$

Therefore, $C(\beta)$ is a monotonously decreasing lower semicontinuous function. This implies that it is right continuous. Hence, we need to show that $C(\cdot)$ is left continuous.

We show the left continuity of $C(\cdot)$ in the case that $\mu^\bullet$ is a Poisson point process for large $\beta < \infty$. For the case $\beta = \infty$ we show the vague convergence of $q^\bullet_\beta \to q^\bullet_\infty$ directly. By scaling $\tau : x \mapsto \beta^{1/d} \cdot x$ this is the same as showing left continuity of $\tilde{C} : \beta \mapsto \mathfrak{C}(\tilde{q}^\bullet_\beta)$, with $\tilde{q}^\bullet_\beta$ the optimal semicoupling between $\mathcal{L}$ and $\eta^\bullet_\beta$, where $\eta^\bullet_\beta$ is a Poisson point process of intensity $\beta > 1$. Indeed, $\tau_*\eta^\bullet_\beta$ is a Poisson process of intensity one. If $\tilde{q}^\bullet_\beta$ is a semicoupling between $\mathcal{L}$ and $\eta^\bullet_\beta$, then $q^\bullet_\beta := \beta \cdot (\tau, \tau)_*\tilde{q}^\bullet_\beta$ is a semicoupling between $\mathcal{L}$ and $\mu^\bullet_\beta$. Moreover, we have

$$\begin{aligned}
\mathfrak{C}(q^\bullet_\beta) &= \mathbb{E}\left[\int_{\mathbb{R}^d \times [0,1)^d} |x-y|^p \; q^\bullet_\beta(dx,dy)\right] \\
&= \beta \cdot \mathbb{E}\left[\int_{\mathbb{R}^d \times [0,\beta^{-1/d})^d} \beta^{p/d}|x-y|^p \; \tilde{q}^\bullet_\beta(dx,dy)\right] \\
&= \beta^{p/d} \cdot \mathfrak{C}(\tilde{q}^\bullet_\beta).
\end{aligned}$$

Hence, $q_\beta^\bullet$ is optimal iff $\tilde q_\beta^\bullet$ is optimal and $C(\cdot)$ is continuous iff $\tilde C(\cdot)$ is continuous.

For $0<\rho<2$ and $a,\beta>1$ let $\Omega^a(\beta)$ be the event such that

$$\left|\eta_\beta^\omega\left(z+[-a/2,a/2)^d\right)-\beta a^d\right| \;\le\; \rho\cdot\beta a^d$$

$\forall\ z\in a\cdot\left[\mathbb{Z}^d\cap[-2,2]^d\right]$ and

$$\left|\eta_\beta^\omega\left(z+[-5^k\cdot a/2,5^k\cdot a/2)^d\right)-\beta 5^{kd}a^d\right| \;\le\; \rho\cdot\beta\cdot 5^{kd}a^d$$

$\forall\ z\in 5^k a\cdot\left[\mathbb{Z}^d\cap[-2,2]^d\setminus\{0\}\right],\ \forall k\in\mathbb{N}$. Using Lemma 7.24 we can estimate

$$\begin{aligned}
\mathbb{P}[\complement\Omega^a(\beta)] &\le \sum_{k=0}^{\infty}2\cdot 5^d\exp\left(-\beta a^d 5^{kd}\cdot\frac{\rho^2}{4}\right)\\
&\le 2\cdot 5^d\left(\sum_{k=1}^{\infty}\exp\left(-k\cdot\beta a^d 5^d\cdot\frac{\rho^2}{4}\right)+\exp\left(-\beta a^d\cdot\frac{\rho^2}{4}\right)\right)\\
&\le 6\cdot 5^d\cdot\exp\left(-\beta a^d\cdot\frac{\rho^2}{4}\right)\\
&=: C\cdot\exp\left(-\beta a^d\cdot\frac{\rho^2}{4}\right),
\end{aligned}$$

for β large enough such that the geometric series becomes bounded by 2, e.g. $\beta>42$ would do. Fix a large β and a slightly smaller β'. We can write $\eta_\beta^\bullet=\eta_{\beta'}^\bullet+\eta_{\beta-\beta'}^\bullet$ with independent $\eta_{\beta'}^\bullet$ and $\eta_{\beta-\beta'}^\bullet$. Put $A_{\beta'}^\omega=\mathrm{supp}(\eta_{\beta'}^\omega)$ and

$$\tilde q_{\beta|\beta'}^\omega(dx,dy) \;:=\; 1_{\mathbb{R}^d\times A_{\beta'}^\omega}\tilde q_\beta^\omega(dx,dy).$$

This is equivariant as $A_{\beta'}^{\theta_g\omega}=\mathrm{supp}(\eta_{\beta'}^{\theta_g\omega})$. Then, also the measures

$$\lambda_{\beta'}^\omega(dx) \;:=\; \tilde q_\beta^\omega(dx,\complement A_{\beta'}^\omega)$$

and

$$\xi_{\beta'}^\omega(dy) \;:=\; \eta_{\beta'}^\omega(dy)-\tilde q_{\beta|\beta'}^\omega(\mathbb{R}^d,dy)$$

are equivariant. $\lambda_{\beta'}^\omega$ is the part of the Lebesgue measure that becomes free if we restrict to $\mathbb{R}^d\times A_{\beta'}^\omega$. $\xi_{\beta'}^\omega$ is all the Poisson mass from $\eta_{\beta'}^\omega$ which is not completely satisfied under $\tilde q_\beta^\omega$, i.e. which is still looking for some Lebesgue points.

The strategy is to find a good semicoupling $\kappa_{\beta'}^\bullet$ between $\lambda_{\beta'}^\bullet$ and $\xi_{\beta'}^\bullet$ whose cost can be controlled nicely. Then, $\hat q_{\beta'}^\bullet:=\tilde q_{\beta|\beta'}^\bullet+\kappa_{\beta'}^\bullet$ is a semicoupling between $\mathcal{L}$ and $\eta_{\beta'}^\bullet$ with $\mathfrak{C}(\hat q_{\beta'}^\bullet)\ge\mathfrak{C}(\tilde q_{\beta'}^\bullet)$. Hence, showing that $\mathfrak{C}(\kappa_{\beta'}^\bullet)$ converges to zero as $\beta'\to\beta$ proves the desired left continuity because $\mathfrak{C}(\tilde q_{\beta|\beta'}^\bullet)\to\mathfrak{C}(\tilde q_\beta^\bullet)$ as $\beta'\to\beta$ by monotone convergence.

Put $\Theta^a(\beta')=\Omega^a(\beta)\cap\Omega^a(\beta')$. Then, it holds that $\mathbb{P}[\Theta^a(\beta')]\ge 1-2\cdot C\exp\left(-\beta' a^d\cdot\frac{\rho^2}{4}\right)$. For $\omega\in\Omega^a(\beta)$ we have for any $B_z=z+[-5^k a/2,5^k a/2)^d$ with z as above that

$$\eta_\beta^\omega(B_z)>(1-\rho)\cdot\beta\cdot\mathcal{L}(B_z).$$

β and β' are assumed to be large such that $(1-\rho)\cdot\beta'$ is still very large, say bigger than 42. We claim that $\xi_{\beta'}^\omega(B_z)$ is still much larger than $\mathcal{L}(B_z)$. Indeed, if this was not the case on a set of positive measure $D\subset\Omega^a(\beta')$, $\tilde q_{\beta|\beta'}^\omega$ would need to transport Lebesgue measure from far away into B_z. However, as there is much more Poisson mass than Lebesgue mass in each of the boxes $B_{z'}$ it is easy to construct locally a semicoupling with less cost on the set D. Due to invariance we can do this on every translate of B_z and therefore produce globally a semicoupling with less cost contradicting optimality of $\tilde{q}^\bullet_\beta$. Hence, by possibly choosing β, β' even bigger we can assume that for $\omega \in \Theta^a(\beta')$

$$\xi^\omega_{\beta'}(B_z) > \frac{1}{2} \cdot (1-\rho) \cdot \beta \cdot \mathcal{L}(B_z).$$

Therefore, even if $\lambda^\omega_{\beta'}$ was equal to the Lebesgue measure $\mathcal{L}$ in any cube $B_z = z + [-5^k a/2, 5^k a/2]$ there is still much more $\xi^\omega_{\beta'}$ measure than Lebesgue measure. Thus, we can find for any $x \in B_z \cap \mathrm{supp}(\lambda^\bullet_{\beta'})$ a Poisson partner at most distance $\sqrt{d} \cdot 5^{k+1} \cdot a$ apart, that is a partner in the same box or one of the surrounding boxes. In particular, we can find a partner for $x \in [-a/2, a/2)^d$ at distance at most $\sqrt{d} \cdot 5 \cdot a$. Put $\tilde{a} = \frac{a}{5\sqrt{d}}$. Then, we can find (e.g. by a variant of the stable marriage algorithm [HHP06]) an equivariant semicoupling $\kappa^\bullet_{\beta'} = (id, S)_* \lambda^\bullet_{\beta'}$ with $\mathbb{P}[|x - S(x)| > a] \leq \mathbb{P}[\complement\Theta^{\tilde{a}}(\beta')] \leq C \exp(-ca^d)$ for $x \in [-\tilde{a}/2, \tilde{a}/2)^d$. In particular, $\mathfrak{C}(\kappa^\bullet_{\beta'}) = \frac{1}{\tilde{a}^d}\mathbb{E}\left[\int_{[-\tilde{a}/2,\tilde{a}/2)^d \times \mathbb{R}^d} |x - S(x)|^p \ \lambda^\bullet_{\beta'}(dx)\right] < \infty$. To estimate the mean transportation cost of $\kappa^\bullet_{\beta'}$ we split the integral into two parts. First,

$$\mathbb{E}\left[1_{\Theta^{\tilde{a}}(\beta')} \frac{1}{\tilde{a}^d} \int_{[-\tilde{a}/2,\tilde{a}/2)^d \times \mathbb{R}^d} |x - S(x)|^p \ \lambda^\bullet_{\beta'}(dx)\right] \leq a^p(\beta - \beta') \cdot (5 \cdot \sqrt{d})^d$$

The term a^p comes from the upper bound on the transportation distance as argued earlier. $1_{[-\tilde{a}/2,\tilde{a}/2]}\lambda^\omega_{\beta'}$ – the mass that becomes free if we restrict q^ω_β to $\eta^\omega_{\beta'}$ – can be bounded from above by $1 \cdot (\eta^\omega_\beta([-a/2,a/2)^d) - \eta^\omega_{\beta'}([-a/2,a/2)^d))$ on $\Theta^{\tilde{a}}(\beta')$. The reason is same as for the lower bound of $\xi^\omega_{\beta'}$. There are so many Poisson points close by that the Lebesgue mass will not be transported far. Using the bound $|x-y|^p \leq a^p$ and taking expectations yields the claim (Note that $\eta^\omega_\beta - \eta^\omega_{\beta'} = \eta^\omega_{\beta-\beta'} \geq 0$. Hence, enlarging the domain of integration can just increase the integral.).

The second part we have to estimate is $\mathbb{E}\left[1_{\complement\Theta^{\tilde{a}}(\beta')} \frac{1}{\tilde{a}^d} \int_{[-\tilde{a}/2,\tilde{a}/2)^d \times \mathbb{R}^d} |x - S(x)|^p \ \lambda^\bullet_{\beta'}(dx)\right]$. Fixing $\delta > 0$, we can make a so large that this expectation can be bounded by $\delta/2$ because $\mathfrak{C}(\kappa^\bullet_{\beta'}) < \infty$ and $\mathbb{P}[\complement\Theta^{\tilde{a}}(\beta')] \leq 2 \cdot C \exp\left(-ca^d\right)$. Hence, we have

$$\mathfrak{C}(\kappa^\bullet_{\beta'}) \leq a^p \cdot (\beta - \beta') \cdot (5 \cdot \sqrt{d})^d + \delta/2.$$

For any $\beta' < \gamma < \beta$ we put $A^\omega_\gamma = \mathrm{supp}(\eta^\omega_\gamma)$ with $\eta^\bullet_\beta = \eta^\bullet_{\beta'} + \eta^\bullet_{\gamma-\beta'} + \eta^\bullet_{\beta-\gamma}$ and $\eta^\bullet_\gamma = \eta^\bullet_{\beta'} + \eta^\bullet_{\gamma-\beta'}$ for independent Poisson processes as above. Set

$$\lambda^\omega_\gamma(dx) \ := \ \tilde{q}^\omega_\beta(dx, \complement A^\omega_\gamma)$$

and

$$\kappa^\omega_\gamma \ := \ (id, S)_* \lambda^\omega_\gamma.$$

That is we take the *same* semicoupling as before but transport less mass because some of the $\lambda^\bullet_{\beta'}$ mass is again used by $\tilde{q}^\bullet_{\beta|\gamma}$. Note that the Poisson points which get mass from $\kappa^\bullet_{\beta'}$ do not get mass from any of the $\tilde{q}^\bullet_\gamma$ transports by construction. Hence, we can estimate the mean transportation cost as above using the *same* set $\Theta^{\tilde{a}}(\beta')$ yielding

$$\mathfrak{C}(\kappa^\bullet_\gamma) \ \leq a^d \cdot (\beta - \gamma) + \delta/2.$$

Taking $\beta - \delta/2a^p(5 \cdot \sqrt{d})^d < \gamma < \beta$ yields

$$\mathfrak{C}(\kappa^\bullet_\gamma) < \delta.$$

As δ was chosen arbitrary, this proves the claim. Therefore, the map $C : \beta \mapsto \mathfrak{C}(q^\bullet_\beta)$ is continuous for large $\beta < \infty$.

To show that $q_\beta^\bullet \to q_\infty^\bullet$ vaguely for $\beta \to \infty$ we need to twist the last argument slightly. We will not argue using the mean transportation cost but we will show that for any compact set K we have

$$\lim_{\beta\to\infty} \mathcal{L}\otimes\mathbb{P}[(x,\omega)\in K\times\Omega : T_\beta(x,\omega)\neq T_\infty(x,\omega)] = 0.$$

Applying Lemma 5.9 proves the claim. To show this claim let $\Omega(a,\rho)$ with $a>1, 0<\rho<2$ be the event such that

$$\left|\mu^\omega\left(z+[-a/2,a/2)^d\right) - a^d\right| \;\leq\; \rho\cdot a^d$$

$\forall\ z\in a\cdot\left[\mathbb{Z}^d\cap[-2,2]^d\right]$ and

$$\left|\mu^\omega\left(z+[-5^k\cdot a/2, 5^k\cdot a/2)^d\right) - 5^{kd}a^d\right| \;\leq\; \rho\cdot\cdot 5^{kd}a^d$$

$\forall\ z\in 5^k a\cdot\left[\mathbb{Z}^d\cap[-2,2]^d\setminus\{0\}\right],\ \forall k\in\mathbb{N}$. Just as above we can estimate

$$\mathbb{P}[\complement\Omega(a,\rho)] \leq C\cdot\exp(-a^d\rho^2/4).$$

Let any compact set K be given. By equivariance, we can assume w.l.o.g. that $K\subset[-a/2,a/2)^d$. For $\beta > a^d(5^d\cdot 3^d)/(1-\rho)$ and $\omega\in\Omega(a,\rho)$ we claim that $T_\beta(x,\omega) = T_\infty(x,\omega)$ for all $x\in[-a/2,a/2)^d$. First of all note that for any $B_z = z+[-R^k a/2, R^k a/2)^d$ with z as above we have for $\omega\in\Omega(a,\rho)$

$$\mu_\beta^\omega(B_z) > \beta\cdot(1-\rho)\mathcal{L}(B_z) > (5^d\cdot 3^d)\mathcal{L}(B_z).$$

The last quantity is a bound on the volume of the union of the boxes surrounding B_z. This implies that any Lebesgue point will be transported to a Poisson point which is at most one box away, i.e. if $x\in B_z$ than $T_\beta(x,\omega)$ lies in B_z or a box surrounding B_z. In particular, any Poisson point in $B_0=[-a/2,a/2)^d$ wants to have much more mass than there is available in B_0 and its surrounding boxes. Hence, any Lebesgue point chooses the Poisson point which is closest. This yields (locally) the Voronoi diagram. Therefore,

$$\mathcal{L}\otimes\mathbb{P}[(x,\omega)\in K\times\Omega : T_\beta(x,\omega)\neq T_\infty(x,\omega)] \;\leq\; \mathcal{L}(K)\cdot\mathbb{P}[\complement\Omega(a,\rho)] \;\leq\; a^d\cdot C\exp(-a^d\rho^2/4),$$

which can be made arbitrarily small by making a – therefore also β – big. Hence, the claim follows.

Bibliography

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

[AHA92] F. Aurenhammer, F. Hoffman, and B. Aronov. Minkowski-type theorems and least-squares partitioning. In *SCG '92: Proceedings of the eighth annual symposium on Computational geometry*, pages 350–357, New York, NY, USA, 1992. ACM.

[AKT84] M. Ajtai, J. Komlós, and G. Tusnády. On optimal matchings. *Combinatorica*, 4(4):259–264, 1984.

[Amb03] L. Ambrosio. Lecture notes on optimal transport problems. In *Mathematical aspects of evolving interfaces (Funchal, 2000)*, volume 1812 of *Lecture Notes in Math.*, pages 1–52. Springer, Berlin, 2003.

[APRS05] A. Abdulkadiroğlu, P.A. Pathak, A.E. Roth, and T. Sönmez. The boston public school match. *American Economic Review*, pages 368–371, 2005.

[ASZ09] L. Ambrosio, G. Savaré, and L. Zambotti. Existence and stability for Fokker-Planck equations with log-concave reference measure. *Probab. Theory Related Fields*, 145(3-4):517–564, 2009.

[Aur91] F. Aurenhammer. Voronoi diagramsa survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, 23(3):345–405, 1991.

[Bau01] H. Bauer. *Measure and integration theory*, volume 26. Walter de Gruyter, 2001.

[BB11] F. Barthe and C. Bordenave. Combinatorial optimization over two random point sets. *Arxiv preprint arXiv:1103.2734*, 2011.

[BLPS99] I. Benjamini, R. Lyons, Y. Peres, and O. Schramm. Group-invariant percolation on graphs. *Geometric And Functional Analysis*, 9(1):29–66, 1999.

[Bow06] B.H. Bowditch. A course on geometric group theory. 2006.

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

[CPPRa] S. Chatterjee, R. Peled, Y. Peres, and D. Romik. Gravitational allocation to Poisson points. *Annals of Mathematics*, 172.

[CPPRb] S. Chatterjee, R. Peled, Y. Peres, and D. Romik. Phase Transitions in Gravitational Allocation. *Geometric And Functional Analysis*, pages 1–48.

[Dec08] L. Decreusefond. Wasserstein distance on configuration space. *Potential Analysis*, 28(3):283–300, 2008.

[Del75] C. Dellacherie. Ensembles analytiques: Théorèmes de séparation et applications. *Séminaire de Probabilités IX Université de Strasbourg*, pages 336–372, 1975.

[DHT09] A. Daly, S. Hess, and K. Train. Assuring finite moments for willingness to pay in random coefficient models. *Transportation*, pages 1–13, 2009.

[DM78] C. Dellacherie and P.A. Meyer. *Probabilities and potential*, volume 29. Elsevier, 1978.

[Doo53] J.L. Doob. Stochastic processes. 1953.

[DSS11] S. Dereich, M. Scheutzow, and R. Schottstedt. Constructive quantization: Approximation by empirical measures. *Arxiv preprint arXiv:1108.5346*, 2011.

[Dud02] R.M. Dudley. *Real analysis and probability*. Cambridge Univ Pr, 2002.

[DVJ07] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes: General theory and structure*, volume 2. Springer Verlag, 2007.

[Fig10] A. Figalli. The optimal partial transport problem. *Archive for rational mechanics and analysis*, 195(2):533–560, 2010.

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

[GP01] N.L. Garcia and J.L. Palacios. On inverse moments of nonnegative random variables. *Statistics & probability letters*, 53(3):235–239, 2001.

[GS62] D. Gale and L.S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[HHP06] C. Hoffman, A.E. Holroyd, and Y. Peres. A stable marriage of Poisson and Lebesgue. *The Annals of Probability*, 34(4):1241–1272, 2006.

[HL01] A.E. Holroyd and T.M. Liggett. How to find an extra head: optimal random shifts of Bernoulli and Poisson random fields. *Annals of Probability*, pages 1405–1425, 2001.

[Hol09] A.E. Holroyd. Geometric properties of poisson matchings. *Probability Theory and Related Fields*, pages 1–17, 2009.

[HP05] A. E. Holroyd and Y. Peres. Extra heads and invariant allocations. *Ann. Probab.*, 33(1):31–52, 2005.

[HPPS09] A.E. Holroyd, R. Pemantle, Y. Peres, and O. Schramm. Poisson matching. *Ann. Inst. Henri Poincaré Probab. Stat*, 45(1):266–287, 2009.

[HS10] M. Huesmann and K.T. Sturm. Optimal transport from Lebesgue to Poisson. *Arxiv preprint arXiv:1012.3845*, 2010.

[Kal97] O. Kallenberg. *Foundations of modern probability*. Springer Verlag, 1997.

[Kan06] LV Kantorovich. On the translocation of masses. *Journal of Mathematical Sciences*, 133(4):1381–1382, 2006.

[Kat92] S. Katok. *Fuchsian groups*. University of Chicago Press, 1992.

[KM07] Y.-K. Kim and R. J. McCann. Continuity, curvature, and the general covariance of optimal transportation, 2007.

[Kri07] M. Krikun. Connected allocation to Poisson points in R^2. *Electron. Comm. Probab*, 12:140–145, 2007.

[Loe09] G. Loeper. On the regularity of solutions of optimal transportation problems. *Acta Mathematica*, 202(2):241–283, 2009.

[LT08] G. Last and H. Thorisson. Invariant transports of stationary random measures and mass-stationarity. *Annals of probability*, 2008.

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

[LZ08] C. Lautensack and S. Zuyev. Random Laguerre tessellations. *Adv. in Appl. Probab.*, 40(3):630–650, 2008.

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

[Mon81] G. Monge. *Mémoire sur la théorie des déblais et des remblais.* De l'Imprimerie Royale, 1781.

[MT11] R. Markó and Á. Timár. A poisson allocation of optimal tail. *Arxiv preprint arXiv:1103.5259*, 2011.

[MTW05] X.-N. Ma, N. S. Trudinger, and X.-J. Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.

[NSV07] F. Nazarov, M. Sodin, and A. Volberg. Transportation to random zeroes by the gradient flow. *Geometric And Functional Analysis*, 17(3):887–935, 2007.

[Oht09] S. Ohta. Uniform convexity and smoothness, and their applications in Finsler geometry. *Math. Ann.*, 343(3):669–699, 2009.

[OS09] S. Ohta and K.-T. Sturm. Heat flow on Finsler manifolds. *Comm. Pure Appl. Math.*, 62(10):1386–1433, 2009.

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

[OV00] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[Pat00] A.L.T. Paterson. *Amenability*, volume 29. Amer Mathematical Society, 2000.

[Pru66] W. Pruitt. Summability of independent random variables. *Indiana Univ. Math. J.*, 15:769–776, 1966.

[Rot] A.E. Roth. Bibliography of matching and market design. http://kuznets.harvard.edu/ aroth/matchbib.html.

[RR90] L. Rüschendorf and ST Rachev. A characterization of random variables with minimum l2-distance. *Journal of multivariate analysis*, 32(1):48–54, 1990.

[RR98] S. T. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, 1998. Theory.

[RSUU05] A.E. Roth, T. Sonmez, and M. Utku Unver. Pairwise kidney exchange. *Journal of Economic Theory*, 125(2):151–188, 2005.

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

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

[Stu09] K.-T. Sturm. Entropic Measure on Multidimensional Spaces. *eprint arXiv: 0901.1815*, 2009.

[Stu11] K.T. Sturm. Generalized orlicz spaces and wasserstein distances for convex-concave scale functions. *Arxiv preprint arXiv:1104.4223*, 2011.

[Tal94] M. Talagrand. The transportation cost from the uniform measure to the empirical measure in dimension ≥ 3. *Ann. Probab.*, 22(2):919–959, 1994.

[Tou56] J. Touchard. Nombres exponentiels et nombres de Bernoulli. *Canad. J. Math.*, 8:305–320, 1956.

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

[Vil09] C. Villani. *Optimal transport: old and new*. Springer Verlag, 2009.

[vN09] G.T. von Nessi. On the regularity of optimal transportation potentials on round spheres. *Arxiv preprint arXiv:0901.2700*, 2009.