

This item is the archived peer-reviewed author-version of:

Optimal transport and integer partitions

Reference:

Hohloch Sonja.- Optimal transport and integer partitions

Discrete applied mathematics - ISSN 0166-218X - 190(2015), p. 75-85

Full text (Publishers DOI): <http://dx.doi.org/doi:10.1016/j.dam.2015.04.002>

To cite this reference: <http://hdl.handle.net/10067/1251480151162165141>

Optimal transport and integer partitions

Sonja Hohloch

Universiteit Antwerpen

email: sonja.hohloch@uantwerpen.be

phone: +32-3-265-3231

fax: +32-3-265-3777

April 26, 2015

Abstract

We link the theory of optimal transportation to the theory of integer partitions. Let $\mathcal{P}(n)$ denote the set of integer partitions of $n \in \mathbb{N}$ and write partitions $\pi \in \mathcal{P}(n)$ as $(n_1, \dots, n_{k(\pi)})$. Using terminology from optimal transport, we characterize certain classes of partitions like symmetric partitions and those in Euler's identity

$$|\{\pi \in \mathcal{P}(n) \mid \text{all } n_i \text{ distinct}\}| = |\{\pi \in \mathcal{P}(n) \mid \text{all } n_i \text{ odd}\}|.$$

Then we sketch how optimal transport might help to understand higher dimensional partitions.

1 Introduction

In this paper, we apply the theory of optimal transport to the classical topic of integer partitions. Both fields have been studied independently since the 18th century, but have – up to our knowledge – not yet been linked.

1.1 Optimal transport and integer partitions

The theory of optimal transport goes back to a problem posed by Monge [Mo] in 1781 and reformulated by Kantorovich [K42, K48] in 1942. Monge asked the following question: Let μ^- and μ^+ denote two heaps of sand with $\text{vol}(\mu^-) = \text{vol}(\mu^+)$ as

in Figure 1. Imagine both heaps as consisting of grains. Is there a map $\varphi : \mu^- \rightarrow \mu^+$ with $\varphi(\mu^-) = \mu^+$ minimizing $\sum_{x \in \mu^-} \text{dist}_{\text{eucl}}(x, \varphi(x))$, i.e. can we ‘transport’ μ^- into μ^+ such that the sum of the transported distances is minimal? And what if we assume the heaps to consist of ‘continuous’ matter instead of grains? Or what if we replace the Euclidean distance by a more general ‘cost function’ $c : \mu^- \times \mu^+ \rightarrow \mathbb{R}$?

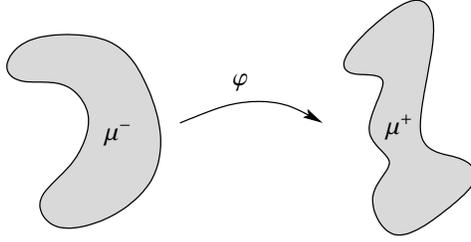


Figure 1: Transporting μ^- into μ^+ .

Reformulated in modern language, we consider two finite measures μ^- and μ^+ in \mathbb{R}^n with $\mu^-(\mathbb{R}^n) = \mu^+(\mathbb{R}^n)$ and we are looking for a map $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with image measure $\varphi(\mu^-) = \mu^+$ minimizing $\int_{\mathbb{R}^n} c(x, \varphi(x)) d\mu^-(x)$ for a given cost function $c : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$. The infimum (hopefully minimum) is denoted by $C(\mu^-, \mu^+)$. For more details, in particular Kantorovich’s formulation, and references we refer the reader to Section 2.

Integer partitions seem to have fascinated humans already since the stone age (cf. Andrews & Eriksson [AE]). Partitions describe the way to decompose an integer into (a sum of) integers: for instance, if we take the number 4, there are the five partitions $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$. One considers partitions up to reordering, i.e. $3 + 1$ and $1 + 3$ are the same partition. Denote by $\mathcal{P}(n)$ the set of partitions of the (positive) integer $n \in \mathbb{N}$ and by $p(n)$ its cardinality. Partitions $\pi \in \mathcal{P}(n)$ are usually written as ordered tuples $\pi = (n_1, \dots, n_{k(\pi)})$ where we often abbreviate $k(\pi) = k$, i.e. $\pi = (n_1, \dots, n_k)$, if no confusion is possible. For instance, we have $\mathcal{P}(4) = \{(4), (3, 1), (2, 2), (2, 1, 1), (1, 1, 1, 1)\}$. There are several ways to display partitions visually, like Ferrer graphs and Young tableaux, see Figure 2.

Using the so-called ‘generating function method’, Euler proved that the elements of the sequence $(p(n))_{n \in \mathbb{N}}$ are the coefficients of the expansion

$$\prod_{i \geq 1} \frac{1}{1 - x^i} = \sum_{n \geq 0} p(n) x^n$$

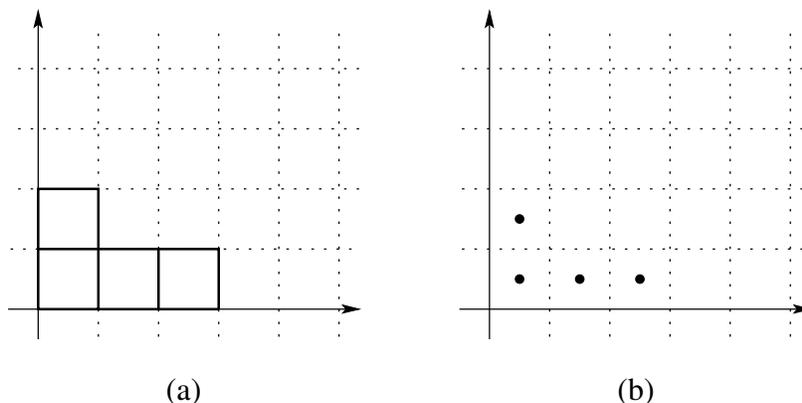


Figure 2: (a) Young tableau and (b) Ferrer graph of $(2, 1, 1) \in \mathcal{P}(4)$.

which is called the *generating function* of $p(n)$. The generating function method is an important tool, see Andrews [An] and Andrews & Eriksson [AE] for more details. Denote by $\mathcal{P}(n | A) \subseteq \mathcal{P}(n)$ the subset of partitions with property A and by $p(n | A)$ its cardinality. Among many other ‘partition identities’, Euler proved by means of generating functions

$$p(n | \text{all } n_i \text{ odd}) = p(n | \text{all } n_i \text{ mutually distinct}). \quad (1)$$

Integer partitions as described above are called ‘one dimensional’ integer partitions: Intuitively one can see a Ferrer graph or a Young tableau as ‘graph’ over the one dimensional real axis. Consequently, ‘two dimensional’ partitions of an integer n can be displayed as Ferrer graphs or Young tableaux over the two dimensional plane \mathbb{R}^2 . We denote by $\mathcal{P}_2(n)$ the set of two dimensional partitions of $n \in \mathbb{N}$ and by $p_2(n)$ its cardinality. Figure 3 (a) displays the Young tableau of $[\begin{smallmatrix} 1 \\ 2 \\ 1 \end{smallmatrix}] \in \mathcal{P}_2(4)$ and Figure 3 (b) shows the Young tableau of $[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}] \in \mathcal{P}_2(2)$. Two dimensional partitions have been studied by MacMahon who found

$$\prod_{k \geq 1} \frac{1}{(1 - x^k)^k} = \sum_{n \geq 0} p_2(n) x^n$$

as generating function for the two dimensional partitions (cf. Andrews [An] and Andrews & Eriksson [AE]).

Analogously one can define the set of m -dimensional partitions $\mathcal{P}_m(n)$ with $n \in \mathbb{N}$. A look at the literature of the last 250 years shows that there has been lots of

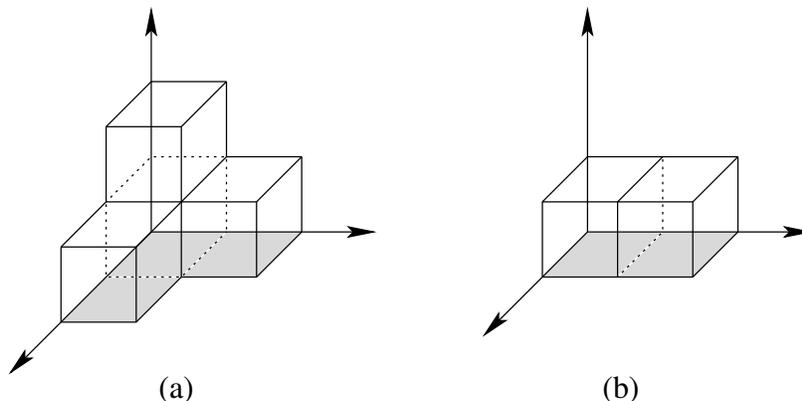


Figure 3: Two dimensional Young tableaux.

research on one dimensional partitions, some research on two dimensional partitions (also called ‘plane partitions’), but not much on m -dimensional partitions for $m \geq 3$. This might be due to the fact that higher dimensional partitions are difficult to display such that it is hard to obtain a good intuition – many one dimensional partition identities have been found by manipulating the associated Ferrer graphs in a tricky way.

1.2 Main results

We hope to have found a way to overcome the ‘dimension problem’ or at least make higher dimensional partitions easier to approach. Let us first line out our ideas for one dimensional partitions.

We want to link optimal mass transportation to integer partitions. How do we do this? Well, let us have a look at the Ferrer board or Young tableau of a partition. Instead of thinking of a Ferrer graph as marked points in the plane, interpret it as the sum of point measures. If $\delta_{(x,y)}$ denotes the point measure with mass one at $(x, y) \in \mathbb{R}^2$ we can see a partition $\pi = (n_1, \dots, n_k) \in \mathcal{P}(n)$ as $\delta_\pi := \sum_{i=1}^k \sum_{\alpha=1}^{n_i} \delta_{(i,\alpha)}$. Analogously one can proceed with the Lebesgue measure restricted to squares and Young tableaux. We have $\delta_\pi(\mathbb{R}^2) = n$ for all $\pi \in \mathcal{P}(n)$. Note that the support of each δ_π for $\pi \in \mathcal{P}(n)$ consists of exactly n points. Thus, given two partitions $\pi^-, \pi^+ \in \mathcal{P}(n)$, we can set

$$\mu^- := \delta_{\pi^-} \quad \text{and} \quad \mu^+ := \delta_{\pi^+}$$

and look for an ‘optimal’ map φ ‘transforming’ π^- into π^+ in an ‘optimal way’ w.r.t. a given cost function c , i.e. we look for φ with $\varphi(\delta_{\pi^-}) = \delta_{\pi^+}$ and

$$C(\delta_{\pi^-}, \delta_{\pi^+}) = \int_{\text{spt}(\delta_{\pi^-})} c((x, y), \varphi(x, y)) d\delta_{\pi^-}(x, y)$$

minimal. Since the support of the involved measures is finite there is always a map realizing the minimum.

There are several subsets of partitions which can nicely be characterized by optimal transport. For example, symmetric partitions and self-symmetric partitions (see Definition 17 and Figure 5) are easy to describe.

Theorem. *Let π be a partition and $\text{sym}(\pi)$ its symmetric partition and the Euclidean distance the cost function. Then the function which is the identity on $\text{spt}(\delta_{\pi}) \cap \text{spt}(\delta_{\text{sym}(\pi)})$ and the reflection on the $x = y$ axis is optimal for δ_{π} and $\delta_{\text{sym}(\pi)}$. Moreover, a partition is selfsymmetric if and only if $\text{spt}(\delta_{\pi}) = \text{spt}(\delta_{\text{sym}(\pi)})$, i.e. the identity function is optimal.*

This result is proven in Theorem 19 and Corollary 20. Now let us have a look at Euler’s identity (1). The left hand side as well as the right hand side can be nicely described by optimal transport. To $\pi \in \mathcal{P}_1(n)$, we associate another two types of measures. There is $\hat{\delta}_{\pi}$ and, for a given permutation σ , the measure δ_{π}^{σ} (for details, see Subsection 4.3).

Theorem. *Let the cost function be a metric and S the reflection on the x -axis. Then*

$$1) \pi \in \mathcal{P}_1(n \mid \text{all } n_i \text{ odd}) \Leftrightarrow S(\hat{\delta}_{\pi}) = \hat{\delta}_{\pi} \Leftrightarrow C(\hat{\delta}_{\pi}, S(\hat{\delta}_{\pi})) = 0.$$

$$2) \pi \in \mathcal{P}_1(n \mid \text{all } n_i \text{ distinct}) \Leftrightarrow \forall \sigma \neq \text{Id} : \delta_{\pi} \neq \delta_{\pi}^{\sigma} \Leftrightarrow \forall \sigma \neq \text{Id} : C(\delta_{\pi}, \delta_{\pi}^{\sigma}) \neq 0.$$

This is proven in Proposition 23 and Proposition 24. Andrews & Eriksson [AE] give an explicit algorithm which turns a partition with distinct n_i into a partition with only odd n_i , thus proving Euler’s identity (1). The essential part of the algorithm is expressed by a certain map φ from $\mathcal{P}_1(n \mid \text{all } n_i \text{ distinct}) =: \mathcal{D}$ to the slightly generalized space $\mathcal{P}_1^{\text{perm}}(n \mid \text{all } n_i \text{ odd}) =: \mathcal{O}$ whose exact definition is stated after Proposition 24. In Theorem 25, we show that

Theorem. *There is a cost function $\mathcal{C} : \mathcal{D} \times \mathcal{O} \rightarrow \mathbb{R}_+$ for which φ is optimal.*

We believe that an analogous statement can always be proven if there is an explicit bijection or algorithm for a partition identity.

When identifying partitions with measures, the dimension of the partition is irrelevant: A partition $\pi \in \mathcal{P}_\ell(n)$ can be seen as the measure δ_π on $\mathbb{R}^{\ell+1}$ given by

$$\delta_\pi = \sum_{i_1=1}^{k_1} \cdots \sum_{i_\ell=1}^{k_\ell} \sum_{\alpha=1}^{n_{i_1 \dots i_\ell}} \delta_{(i_1, \dots, i_\ell, \alpha)}$$

where the $n_{i_1 \dots i_\ell}$ are monotone decreasing in each coordinate (see Definition 9). Similarly, we can use the Young tableaux and the restriction of the Lebesgue measure to cubes in $\mathbb{R}^{\ell+1}$. As in the one dimensional case, we can formulate the optimal transport setting for two partitions $\pi^-, \pi^+ \in \mathcal{P}_\ell(n)$ via

$$\mu^- := \delta_{\pi^-} \quad \text{and} \quad \mu^+ := \delta_{\pi^+}$$

and look for a map φ with $\varphi(\delta_{\pi^-}) = \delta_{\pi^+}$ sending π^- to π^+ in an optimal way, i.e. minimizing $\int_{\text{spt}(\delta_{\pi^-})} c(z, \varphi(z)) d\delta_{\pi^-}(z)$.

When investigating higher dimensional partitions in future works, we would like to use results from the theory of optimal transportation. Since most results are stated for Lebesgue continuous measures μ^- , it might be advisable to work with the Lebesgue measure induced from the cubes of the Young tableau rather than with the discrete measure coming from the Ferrer graph. Moreover, the possibility to dualize the problem as done in (5) might be of help.

It is interesting to note, as pointed out to us by Leonid Polterovich, that discrete measures and the Monge-Kantorovich distance find applications in statistics and computer science (in particular in artificial intelligence), cf. Cuturi [Cu].

Organisation of the paper

Section 2 recalls important facts from the theory of optimal transportation. Section 3 defines integer partitions in any dimension by means of measure theory and formulates the setting for optimal transport of partitions. Section 4 applies the theory of optimal transportation to partitions.

2 Optimal transport

The literature on mass transportation problems is vast, see for instance Villani [Vi] or Rachev & Rüschendorf [RR] for an overview. In modern language, Monge's

problem [Mo] can be formulated as follows. Let M be a finite dimensional manifold and denote by $\mathfrak{S}(M)$ its Borel σ -algebra and by $\mathfrak{M}(M)$ the space of finite, positive Borel measures on M . Given a measurable map $\psi : M \rightarrow M$ and $\mu \in \mathfrak{M}(M)$, the *image* or *push forward measure* $\psi(\mu)$ is defined via $\psi(\mu(B)) := \mu(\psi^{-1}(B))$ for all measurable $B \subset M$.

Problem 2 (Monge).

Given: $\mu^-, \mu^+ \in \mathfrak{M}(M)$ with $\mu^-(M) = \mu^+(M)$ and a measurable ‘*cost function*’ $c : M \times M \rightarrow \mathbb{R}^{\geq 0}$.

Wanted: A measurable ‘*optimal map*’ $\varphi : M \rightarrow M$ which realizes the minimum of

$$C(\mu^-, \mu^+) := \inf \left\{ \int_M c(x, \varphi(x)) d\mu^-(x) \mid \varphi \text{ Borel, } \varphi(\mu^-) = \mu^+ \right\}.$$

In particular, in case $M = \mathbb{R}^m$ with point measures $\mu^- = \sum_{i=1}^{\ell} \delta_{x_i}$ and $\mu^+ = \sum_{i=1}^{\ell} \delta_{y_i}$ and $c(x, y) = d(x, y)$ is a distance function, the map φ is transporting the mass of μ^- to the mass of μ^+ while minimizing the ‘sum’ $\sum_{i=1}^{\ell} d(x_i, \varphi(x_i))$ of the total transported distances. But when working with point measures, we have to be cautious:

Remark 3. *If μ^- contains point measures, there does not necessarily exist a transport map: For $\mu^- = \delta_x$ and $\mu^+ = \frac{1}{2}\delta_{y_1} + \frac{1}{2}\delta_{y_2}$ there is no map φ with $\varphi(\mu^-) = \mu^+$.*

In 1979, Sudakov [Su] proposed a proof of Monge’s problem in case of \mathbb{R}^m with the Euclidean distance as a cost function. Unfortunately the proof turned out to have a gap (cf. Ambrosio [Am1, p. 137], [Am2, chapter 6]) which can only be mended under stronger assumptions.

Kantorovich [K42, K48] came up with another approach to Monge’s problem which is much easier to handle. Denote by $p^- : M \times M \rightarrow M$, $p^-(x, y) = x$ the projection on the first factor and by p^+ the projection on the second one. Let $\mu^-, \mu^+ \in \mathfrak{M}(M)$ with $\mu^-(M) = \mu^+(M)$ and define

$$\mathfrak{M}(\mu^-, \mu^+) := \{ \mu \in \mathfrak{M}(M \times M) \mid p^-(\mu) = \mu^- \text{ and } p^+(\mu) = \mu^+ \}.$$

Problem 4 (Kantorovich).

Given: $\mu^-, \mu^+ \in \mathfrak{M}(M)$ and a ‘cost function’ $c : M \times M \rightarrow \mathbb{R}^{\geq 0}$.

Wanted: An ‘optimal measure’ $\mu \in \mathfrak{M}(\mu^-, \mu^+)$ which realizes the minimum of

$$K(\mu^-, \mu^+) := \inf \left\{ \int_{M \times M} c(x, y) d\mu(x, y) \mid \mu \in \mathfrak{M}(\mu^-, \mu^+) \right\}.$$

In contrast to Monge’s problem, Kantorovich’s setting is linear in μ and $\mathfrak{M}(\mu^-, \mu^+)$ is convex. Moreover, under reasonable assumptions on μ^-, μ^+ and c , there exists always an optimal measure on M (via a standard compactness argument using the calculus of variations).

Monge’s and Kantorovich’s problem are linked as follows. If there is a measurable map $\varphi : M \rightarrow M$ with $\varphi(\mu^-) = \mu^+$, then define $\text{Id} \times \varphi : M \rightarrow M \times M$, $x \mapsto (x, \varphi(x))$. The image measure satisfies $(\text{Id} \times \varphi)(\mu^-) \in \mathfrak{M}(\mu^-, \mu^+)$ and its support lies in the graph of φ . We compute

$$\begin{aligned} \inf_{\varphi \text{ with } \varphi(\mu^-) = \mu^+} \int_M c(x, \varphi(x)) d\mu^-(x) &= \inf_{\varphi \text{ with } \varphi(\mu^-) = \mu^+} \int_{M \times M} c(x, y) d(\text{Id} \times \varphi)(\mu^-)(x, y) \\ &\geq \min_{\mu \in \mathfrak{M}(\mu^-, \mu^+)} \int_{M \times M} c(x, y) d\mu(x, y) \end{aligned}$$

such that Kantorovich’s problem yields a lower bound for Monge’s problem. The question when an optimal μ is of the form $\text{Id} \times \varphi$ was studied by Gangbo & McCann [GM] on \mathbb{R}^m with Lebesgue continuous μ^- for strictly convex and strictly concave cost functions.

If the cost function is of the form $c(x, y) = h(x - y)$ where h is strictly convex and satisfies certain growth conditions then Gangbo & McCann [GM] find a unique optimal $\mu \in \mathfrak{M}(\mu^-, \mu^+)$ for Kantorovich’s problem which turns out to be of the form $(\text{Id} \times \varphi)(\mu^-)$. Thus they obtain also an optimal map for Monge’s problem. This φ is explicitly given by $\varphi(x) = x - \nabla h^{-1}(\nabla \psi(x))$ where $\psi : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{-\infty\}$ is a so-called c -concave function, i.e. ψ is not identical $-\infty$ and $\psi(x) := \inf_{(y,r) \in A} \{c(x, y) + r\}$ for a subset $A \subseteq \mathbb{R}^m \times \mathbb{R}$.

If $c(x, y) = f(|x - y|)$ (where the function f is non negative) is strictly concave then the cost function induces a metric such that a minimal measure does not ‘move’ the intersection set of the support of μ^- and μ^+ . Thus one only obtains a map if

the two measures have disjoint support. Otherwise one only gets a map for the positive parts of the Jordan decomposition $[\mu^- - \mu^+]_+$ and $[\mu^+ - \mu^-]_+$.

An essential tool in Gangbo & McCann's work is the so-called c -cyclic monotonicity (cf. Definition 15) which was introduced by Smith & Knott [SK] and Rüschemdorf [Rü], [RR].

There is another way to approach Kantorovich's problem: Since it is a 'convex problem' Kantorovich [K42] (and later others) dualized it, i.e. they consider

$$\min_{\mu \in \mathfrak{M}(m^-, \mu^+)} \int_{M \times M} c(x, y) d\mu(x, y) = \sup \left\{ \int_M h_-(x) d\mu^-(x) + \int_M h_+(y) d\mu^+(y) \right\} \quad (5)$$

where the supremum is taken over all $(h_-, h_+) \in L^1(\mu^-) \times L^1(\mu^+)$ with $h_-(x) + h_+(y) \leq c(x, y)$.

3 Integer Partitions

Integer partitions naturally arise in many places throughout mathematics, physics and computer sciences. They have been studied already by Euler in the 18th century and later by Legendre, Ramanujan, and Hardy to name just a few. We refer the reader to the two monographs by Andrews [An] and Andrews & Eriksson [AE] for further information on the history of integer partitions.

3.1 Partitions as measures

Since, on the one hand, 1-dimensional partitions are much more intuitive and, on the other hand, most of Section 4 deals with 1-dimensional partitions, we first present the new concept for 1-dimensional partitions before we generalize it to the higher dimensional case.

Definition 6. *Let $n \in \mathbb{N}$. Then a **one dimensional partition** of n is a tuple of ordered integers (n_1, \dots, n_k) with $n \geq n_1 \geq \dots \geq n_k \geq 1$ for some integer $1 \leq k \leq n$ such that $\sum_{i=1}^k n_i = n$. The **set of one dimensional partitions** of n is denoted by $\mathcal{P}(n) := \mathcal{P}_1(n)$ and $p(n) := p_1(n) := |\mathcal{P}_1(n)|$ is its cardinality. We often abbreviate 'one dimensional partitions' by 'partitions'.*

Let us introduce some notations.

Definition 7. Let $n \in \mathbb{N}$ and $\pi \in \mathcal{P}_1(n)$. We call $|\pi| := n$ the **cardinality** of π . Instead of $\pi = (n_1, \dots, n_{k(\pi)})$ with $\sum_{i=1}^{k(\pi)} n_i = n$ we often write by abuse of notation $\pi = (n_1, \dots, n_k)$ abbreviating $k = k(\pi)$. The number $k = k(\pi)$ is the **length** of π . If E denotes a certain property, we define

$$\mathcal{P}(n | E) := \mathcal{P}_1(n | E) := \{\pi \in \mathcal{P}_1(n) \mid \pi \text{ has property } E\}$$

and denote by $p(n | E) := p_1(n | E)$ its cardinality.

For a subset $A \subset \mathbb{R}^\ell$, define its *characteristic function* $\chi_A : \mathbb{R}^\ell \rightarrow \{0, 1\}$ via $\chi(x) = 1$ if $x \in A$ and $\chi(x) = 0$ if $x \notin A$. Finite sums of characteristic functions are called *stair functions*. \mathcal{L}_ℓ is the ℓ -dimensional Lebesgue measure and δ_z denotes the point measure at a point $z \in \mathbb{R}^\ell$. If μ is a measure then $\text{spt}(\mu)$ denotes its support. We abbreviate $\mathbb{R}_+^\ell := ([0, \infty])^\ell$.

Let $p_x : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x$ be the projection on the first and $p_y : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto y$ be the projection on the second coordinate. The idea of this paper is to identify partitions with measures such that we can apply optimal transport theory.

Proposition 8. *The following statements are equivalent.*

- 1) $(n_1, \dots, n_k) \in \mathcal{P}_1(n)$.
- 2) $\pi := (n_1, \dots, n_k) \in \mathbb{N}^k$ and $\rho_\pi : \mathbb{R}_+ \rightarrow \{n_1, \dots, n_k\}$, $\rho_\pi := \sum_{i=1}^k n_i \chi_{[i-1, i]}$ is a monotone decreasing stair function with $\int_{\mathbb{R}_+} \rho_\pi(x) d\mathcal{L}_1(x) = n$. We set $\mathcal{L}_\pi := \rho_\pi \mathcal{L}_1$.
- 3) Let $\pi := (n_1, \dots, n_k) \in \mathbb{N}^k$ and denote by $\xi_{i,\alpha}$ the center of the square $[i-1, i] \times [\alpha-1, \alpha]$. The measure $\delta_\pi := \sum_{i=1}^k \sum_{\alpha=1}^{n_i} \delta_{\xi_{i,\alpha}}$ satisfies $\delta_\pi(\mathbb{R}^2) = n$ and the image measure $p_x(\delta_\pi)$ has a monotone decreasing density function on $\frac{1}{2} + \mathbb{N}_0$.
- 4) Let $\pi := (n_1, \dots, n_k) \in \mathbb{N}^k$. Then $\tilde{\delta}_\pi := \sum_{i=1}^k \sum_{\alpha=1}^{n_i} \delta_{(i,\alpha)}$ satisfies $\tilde{\delta}_\pi(\mathbb{R}^2) = n$ and the image measure $p_x(\tilde{\delta}_\pi)$ has a monotone decreasing density function on \mathbb{N} .

The proof is obvious. Instead of using discrete measures in item 3) and 4), we can as well consider the restriction of the Lebesgue measure to the square $[i-1, i] \times [\alpha-1, \alpha]$ and obtain a continuous measure characterizing the partition.

Note that, for $\pi = (n_1, \dots, n_k) \in \mathcal{P}(n)$, the push forward measure $p_x(\mathcal{L}_1|_{Y(\pi)})$ has ρ_π as density function. Conversely, $\mathcal{L}_1|_{Y(\pi)}$ can be recovered from ρ_π .

Now we consider m -dimensional partitions. We abbreviate multi-indices like

$$\binom{1 \leq i_1 \leq k_1}{\vdots} \\ 1 \leq i_m \leq k_m$$

by $1 \leq i_1, \dots, i_m \leq k_1, \dots, k_m$.

Definition 9. Let $n \in \mathbb{N}$. An m -dimensional partition of n is an array consisting of $n_{i_1 \dots i_m} \in \mathbb{N}$ where $1 \leq i_1, \dots, i_m \leq k_1, \dots, k_m$ for some natural numbers $1 \leq k_1, \dots, k_m \leq n$ such that for each index $i_j = 1, \dots, k_j$ with $1 \leq j \leq m$ the natural numbers $n_{i_1 \dots i_m}$ are a monotone decreasing sequence with $n \geq \max_{i_j \in \{1, \dots, k_j\}} n_{i_1, \dots, i_m}$ and $\min_{i_j \in \{1, \dots, k_j\}} n_{i_1, \dots, i_m} \geq 1$ and $\sum_{i_1=1}^{k_1} \dots \sum_{i_m=1}^{k_m} n_{i_1 \dots i_m} = n$. $\mathcal{P}_m(n)$ denotes the set of m -dimensional partitions and $p_m(n) := |\mathcal{P}_m(n)|$ denotes its cardinality.

As an example, consider Figure 3 where (a) displays the Young tableau of $[\begin{smallmatrix} 1 & \\ & 2 \end{smallmatrix}] \in \mathcal{P}_2(4)$ and (b) shows the Young tableau of $[\begin{smallmatrix} 1 & \\ & 1 \end{smallmatrix}] \in \mathcal{P}_2(2)$.

A function $g : A \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ is monotone decreasing if the functions $x_i \mapsto g(x_1, \dots, x_m)$ are monotone decreasing for all $1 \leq i \leq m$. Moreover, given two monotone decreasing functions $g_1, g_2 : A \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$, then their sum $g_1 + g_2$ is monotone decreasing. Denote by $p_{x_1 \dots x_m} : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$, $(x_1, \dots, x_{m+1}) \mapsto (x_1, \dots, x_m)$ the projection which forgets the last coordinate.

Proposition 10. The following statements are equivalent.

- 1) $\pi = (n_{i_1 \dots i_m})_{1 \leq i_1, \dots, i_m \leq k_1, \dots, k_m} \in \mathcal{P}_m(n)$.
- 2) Let $\pi = (n_{i_1 \dots i_m})_{1 \leq i_1, \dots, i_m \leq k_1, \dots, k_m} \in \mathbb{N}^{k_1 + \dots + k_m}$ and set $\rho_\pi : \mathbb{R}_+^m \rightarrow \{n_{i_1 \dots i_m} \in \mathbb{N} \mid 1 \leq i_1, \dots, i_m \leq k_1, \dots, k_m\}$ with $\rho_\pi = \sum_{i_1=1}^{k_1} \dots \sum_{i_m=1}^{k_m} n_{i_1 \dots i_m} \chi_{C_{i_1 \dots i_m}}$ where $C_{i_1 \dots i_m} := [i_1 - 1, i_1] \times \dots \times [i_m - 1, i_m]$. Then ρ_π is a monotone decreasing stair function with $\int_{\mathbb{R}_+^m} \rho_\pi(x_1, \dots, x_m) d\mathcal{L}_m(x_1, \dots, x_m) = n$. We set $\mathcal{L}_\pi := \rho_\pi \mathcal{L}_m$.
- 3) Let $\pi = (n_{i_1 \dots i_m})_{1 \leq i_1, \dots, i_m \leq k_1, \dots, k_m} \in \mathbb{N}^{k_1 + \dots + k_m}$ and denote by $\xi_{i_1 \dots i_m \alpha}$ the center of the cube $[i_1 - 1, i_1] \times \dots \times [i_m - 1, i_m] \times [\alpha - 1, \alpha]$. The measure $\delta_\pi := \sum_{i_1=1}^{k_1} \dots \sum_{i_m=1}^{k_m} \sum_{\alpha=1}^{n_{i_1 \dots i_m}} \delta_{\xi_{i_1 \dots i_m \alpha}}$ satisfies $\delta_\pi(\mathbb{R}^{m+1}) = n$ and the image measure $p_{x_1 \dots x_m}(\delta_\pi)$ has a monotone decreasing density function on $(\frac{1}{2} + \mathbb{N}_0)^m$.
- 4) Let $\pi = (n_{i_1 \dots i_m})_{1 \leq i_1, \dots, i_m \leq k_1, \dots, k_m} \in \mathbb{N}^{k_1 + \dots + k_m}$. The measure $\tilde{\delta}_\pi := \sum_{i_1=1}^{k_1} \dots \sum_{i_m=1}^{k_m} \sum_{\alpha=1}^{n_{i_1 \dots i_m}} \delta_{(i_1, \dots, i_m, \alpha)}$ satisfies $\tilde{\delta}_\pi(\mathbb{R}^m) = n$ and the image measure $p_{x_1 \dots x_m}(\tilde{\delta}_\pi)$ has a monotone decreasing density function on \mathbb{N}^m .

Using Proposition 8 for intuition, the proof is obvious. Let $p_{x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_m} : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1}$ be the projection which forgets the j th coordinate.

Remark 11. 1) Given a partition $\pi = (n_{i_1 \dots i_m})_{1 \leq i_1, \dots, i_m \leq k_1, \dots, k_m} \in \mathcal{P}_m(n)$, the support of δ_π is called the **Ferrer graph** and the union of cubes $Y(\pi) := \bigcup_{i_1=1}^{k_1} \dots \bigcup_{i_m=1}^{k_m} \bigcup_{\alpha=1}^{n_{i_1 \dots i_m}} [i_1 - 1, i_1] \times \dots \times [i_m - 1, i_m] \times [\alpha - 1, \alpha]$ is called the **Young tableau or Ferrer board** of π .

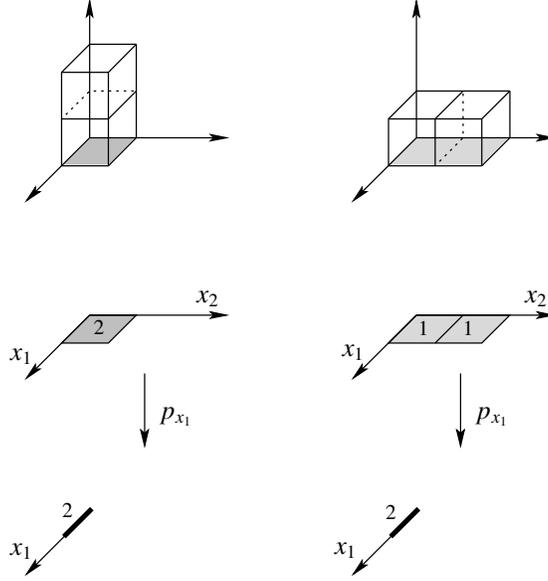


Figure 4: Different 2-dimensional partitions can have the same p_{x_1} -projection.

- 2) \mathcal{L}_π determines the m push forward measures $p_{x_1, \dots, x_{m-1}}(\mathcal{L}_\pi), \dots, p_{x_2, \dots, x_m}(\mathcal{L}_\pi)$ uniquely and, conversely, these m push forward measures determine \mathcal{L}_π uniquely.
- 3) Let $\pi \in \mathcal{P}_m(n)$. Then the measures $p_{x_1, \dots, x_{m-1}}(\mathcal{L}_\pi), \dots, p_{x_2, \dots, x_m}(\mathcal{L}_\pi)$ represent the partitions $p_{x_1, \dots, x_{m-1}}(\pi), \dots, p_{x_2, \dots, x_m}(\pi)$. This is of particular interest to the Kantorovich problem.

Note that we really need all m push forward measures in item 2) in order to determine \mathcal{L}_π uniquely. Otherwise it is not true, cf. Figure 4.

3.2 Optimal transport of integer partitions

Let $m, n \in \mathbb{N}$ and $\pi \in \mathcal{P}_m(n)$. The partition π can be displayed by the continuous measure \mathcal{L}_π as well as by the discrete measures δ_π and $\tilde{\delta}_\pi$. Let us first consider the latter ones. Using the above notation, we have

$$\begin{aligned} \text{spt}(\delta_\pi) &= \{\xi_{i_1 \dots i_m \alpha} \mid 1 \leq i_1, \dots, i_m, \alpha \leq k_1, \dots, k_m, n_{i_1 \dots i_m}\}, \\ \text{spt}(\tilde{\delta}_\pi) &= \{(i_1, \dots, i_m, \alpha) \mid 1 \leq i_1, \dots, i_m, \alpha \leq k_1, \dots, k_m, n_{i_1 \dots i_m}\} \end{aligned}$$

which are sets of isolated points. For all $\pi \in \mathcal{P}_m(n)$, the cardinality of the supports is $|\text{spt}(\delta_\pi)| = n = |\text{spt}(\tilde{\delta}_\pi)|$. Thus, given two partitions $\pi^-, \pi^+ \in \mathcal{P}_m(n)$, there *always exists* a bijective map $f : \text{spt}(\pi^-) \rightarrow \text{spt}(\pi^+)$, i.e. there are no obstructions as in Remark 3. We denote the space of such maps by $\mathcal{F}(\pi^-, \pi^+)$ and its cardinality is *finite*. In this context, Monge's problem translates into looking for

$$C(\pi^-, \pi^+) := \min_{f \in \mathcal{F}(\pi^-, \pi^+)} \int_{\text{spt}(\delta_{\pi^-})} c(z, f(z)) d\delta_\pi(z) = \min_{f \in \mathcal{F}(\pi^-, \pi^+)} \sum_{z \in \text{spt}(\delta_{\pi^-})} c(z, f(z))$$

We will be working a lot with the following type of cost functions.

Definition 12. A measurable cost function $c : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ is *metric-like* if c induces a metric on \mathbb{R}^m , i.e.

- a) $c \geq 0$ with $c(x, y) = 0$ if and only if $x = y$.
- b) $c(x, y) = c(y, x)$ for all $x, y \in \mathbb{R}^m$.
- c) $c(x, z) \leq c(x, y) + c(y, z)$ for all $x, y, z \in \mathbb{R}^m$.

Gangbo & McCann [GM], using Lebesgue continuous measures and concave cost functions, showed that the optimal map can be chosen to be the identity on the intersection set $\text{spt}(\mu^-) \cap \text{spt}(\mu^+)$. This holds true for our metric-like cost functions:

Proposition 13. Let $\pi^-, \pi^+ \in \mathcal{P}_m(n)$ and let c be a metric-like cost function. Then an optimal $f \in \mathcal{F}(\pi^-, \pi^+)$ can be chosen to be the identity on $\text{spt}(\delta_{\pi^-}) \cap \text{spt}(\delta_{\pi^+})$.

Proof. We study where the points in $\text{spt}(\delta_{\pi^-})$ are mapped to and if one can minimize the transport costs.

Case 1: Let $z \in \text{spt}(\delta_{\pi^-})$ be mapped to $f(z) \in \text{spt}(\delta_{\pi^-}) \cap \text{spt}(\delta_{\pi^+})$. Since $f(z) \in \text{spt}(\delta_{\pi^-})$, it is mapped to $f(f(z))$. If $f(f(z)) \neq f(z)$, then we have $c(z, f(z)) + c(f(z), f(f(z))) \geq c(z, f(f(z))) + c(f(z), f(z)) = c(z, f(f(z)))$. Now define a new map $\tilde{f} \in \mathcal{F}(\pi^-, \pi^+)$ via $\tilde{f}(z) := f(f(z))$ and $\tilde{f}(f(z)) = f(z)$. We obtain $\int_{\text{spt}(\delta_{\pi^-})} c(z, f(z)) d\delta_{\pi^-}(z) \geq \int_{\text{spt}(\delta_{\pi^-})} c(z, \tilde{f}(z)) d\delta_{\pi^-}(z)$. Iterating this procedure, we obtain a function which leaves all points in $\text{spt}(\delta_{\pi^-}) \cap \text{spt}(\delta_{\pi^+})$ fixed and has equal or lower transport costs than any function which does not fix the points in $\text{spt}(\delta_{\pi^-}) \cap \text{spt}(\delta_{\pi^+})$.

Case 2: Let $z \in \text{spt}(\delta_{\pi^-})$ be mapped to $f(z) \in \text{spt}(\delta_{\pi^+}) \setminus \text{spt}(\delta_{\pi^-})$. The point $f(z)$ is not in the support of δ_{π^-} , so $f(z)$ is mapped nowhere by f . There is no need to modify f . \square

Moreover, we can define a distance for partitions.

Remark 14. Let $\pi^-, \pi^+ \in \mathcal{P}_m(n)$ and let c be a metric-like cost function. Then $\text{dist}(\pi^-, \pi^+) := \min_{f \in \mathcal{F}(\pi^-, \pi^+)} \int_{\text{spt}(\delta_{\pi^-})} c(z, f(z)) d\delta_{\pi^-}(z)$ is a metric on $\mathcal{P}_m(n)$.

The following definition is important. It goes back to Smith & Knott [SK] and Rüschemdorf [Rü].

Definition 15. Let c be a metric-like cost function. A subset $B \subset \mathbb{R}^m \times \mathbb{R}^m$ is *c-cyclic monotone* if $\sum_{i=1}^k c(x_i, y_i) \leq \sum_{i=1}^k c(x_{\sigma(i)}, y_i)$ for all $k \in \mathbb{N}$ and $(x_i, y_i) \in B$ for $1 \leq i \leq k$ and $\sigma \in \text{Perm}(k)$.

Gangbo & McCann [GM] proved that the support of an optimal measure for the Kantorovich problem is c -cyclic monotone and that the graph of an optimal map of the Monge problem always lies in a c -cyclic monotone set. The following is the according statement for our situation.

Theorem 16. Let $\pi^-, \pi^+ \in \mathcal{P}_m(n)$ and let c be a metric-like cost function. Then $\varphi \in \mathcal{F}(\pi^-, \pi^+)$ is optimal for $C(\pi^-, \pi^+)$ if and only if $\text{graph}(\varphi)$ is c -cyclic monotone.

4 Optimal transport applied to integer partitions

4.1 Symmetric partitions in $\mathcal{P}_1(n)$

Let us consider a special type of partitions.

Definition 17. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (y, x)$ be the reflection on the $x = y$ line and let $\pi \in \mathcal{P}_1(n)$ with Young tableau $Y(\pi)$. The *symmetric or conjugate partition* $\text{sym}(\pi) \in \mathcal{P}_1(n)$ of π is the partition with the Young tableau $Y(\text{sym}(\pi)) = T(Y(\pi))$, i.e. we obtain $\text{sym}(\pi)$ by reflecting π on the $x = y$ line. Partitions $\pi \in \mathcal{P}_1(n)$ with $\text{sym}(\pi) = \pi$ are called *self-symmetric or self-conjugate*.

An example for a partition and its symmetric partition is sketched in Figure 5.

Let $p_x : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x$ be the projection on the x -axis and $p_y : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto y$ the projection on the y -axis.

Remark 18. (1) Consider $\pi = (n_1, \dots, n_k) \in \mathcal{P}_1(n)$ with associated measure $\delta_\pi = \sum_{i=1}^k \sum_{\alpha=1}^{n_i} \delta_{\xi_{i\alpha}}$. Then $\text{sym}(\pi)$ has the associated measure $\delta_{\text{sym}(\pi)} = \sum_{i=1}^k \sum_{\alpha=1}^{n_i} \delta_{T(\xi_{i\alpha})} = \sum_{i=1}^k \sum_{\alpha=1}^{n_i} \delta_{\xi_{\alpha i}}$.

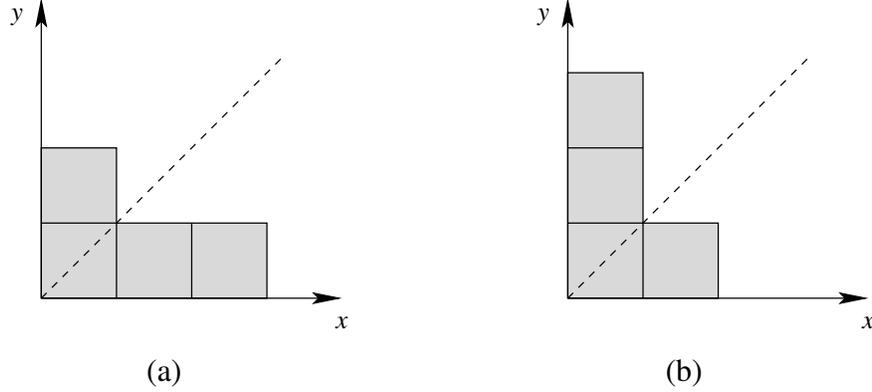


Figure 5: A partition and its symmetric partition.

(2) *The push forward measures $p_y(\mathcal{L}_\pi)$ resp. $p_y(\delta_\pi)$ have the same density functions as $p_x(\mathcal{L}_{sym(\pi)})$ resp. $p_x(\delta_{sym(\pi)})$.*

Now we investigate the relation between a partition and its symmetric partition by means of optimal transport. We are looking for a function which realizes minimal transport costs in Monge's problem for the partitions π and $sym(\pi)$, more precisely for the measures $\mu^- = \delta_\pi$ and $\mu^+ = \delta_{sym(\pi)}$.

Theorem 19. *Let $\pi \in \mathcal{P}_1(n)$ and let c be the Euclidean distance. Then $f \in \mathcal{F}(\delta_\pi, \delta_{sym(\pi)})$ given by $f = \text{Id}$ on $\text{spt}(\delta_\pi) \cap \text{spt}(\delta_{sym(\pi)})$ and $f = T$ elsewhere minimizes $C(\delta_\pi, \delta_{sym(\pi)})$.*

Proof. Due to Proposition 13, we can set $f = \text{Id}$ on $\text{spt}(\delta_\pi) \cap \text{spt}(\delta_{sym(\pi)})$. Now consider $z \in \text{spt}(\pi) \setminus \text{spt}(sym(\pi))$. If we consider the graph of T , we observe that it is c -cyclic monotone since none of the segments between z and $T(z)$ cross each other, cf. Figure 6. Thus Theorem 16 implies the claim. \square

Corollary 20. *$\pi \in \mathcal{P}_1(n)$ is self-symmetric if and only if $C(\delta_\pi, \delta_{sym(\pi)}) = 0$ with the Euclidean distance as cost function.*

4.2 Generalized symmetric partitions in $\mathcal{P}_k(n)$

Symmetric partitions have a natural generalization to higher dimensions. For $m \in \mathbb{N}$, denote by $Perm(m)$ the permutation group of the set $\{1, \dots, m\}$. Given $\sigma \in Perm(m)$, we write $\sigma(1, \dots, m) = (\sigma(1), \dots, \sigma(m))$.

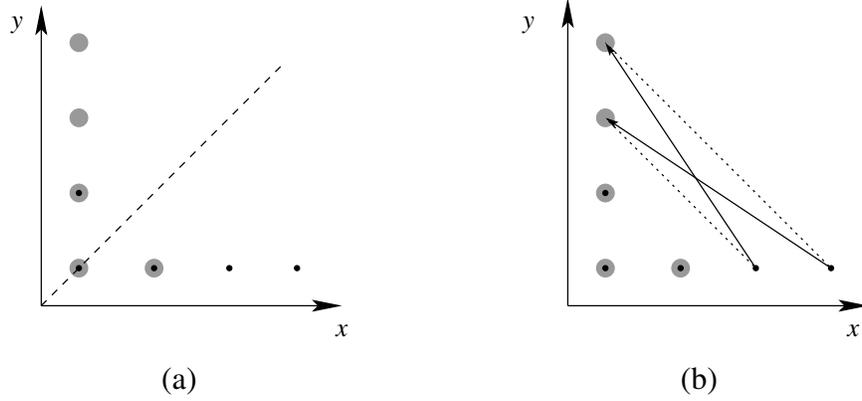


Figure 6: (a) $\text{spt}(\delta_\pi)$ is displayed by small black points and $\text{spt}(\delta_{\text{sym}(\pi)})$ by larger greyish points. (b) $f = T$ on $\text{spt}(\delta_\pi) \cap \text{spt}(\delta_{\text{sym}(\pi)})$ due to c -cyclicity.

Definition 21. Let $\sigma \in \text{Perm}(k+1)$ and $T_\sigma : \mathbb{R}^{k+1} \rightarrow \mathbb{R}^{k+1}$ be linear with $T_\sigma = (e_{\sigma(1)}, \dots, e_{\sigma(k+1)})$ as matrix w.r.t. the standard basis e_1, \dots, e_{k+1} of \mathbb{R}^{k+1} . The σ -**symmetric partition** of $\pi \in \mathcal{P}_k(n)$ is the partition $\text{sym}_\sigma(\pi)$ with the Young tableau $Y(\text{sym}_\sigma(\pi)) = T_\sigma(Y(\pi))$. Partitions $\pi \in \mathcal{P}_k(n)$ with $\pi = \text{sym}_\sigma(\pi)$ are called σ -**selfsymmetric**.

We believe that σ -symmetric partitions behave similar as symmetric partitions:

Conjecture 22. 1) The map T_σ induces an optimal transport map for π and $\text{sym}_\sigma(\pi)$.

2) $\pi \in \mathcal{P}_k(n)$ is σ -selfsymmetric if and only if $C(\delta_\pi, \delta_{\text{sym}_\sigma(\pi)}) = 0$ where c is a metric-like cost functions.

4.3 The Euler identity

When studying subsets of $\mathcal{P}_1(n)$, Euler proved the identity

$$p_1(n \mid \text{all } n_i \text{ odd}) = p_1(n \mid \text{all } n_i \text{ mutually distinct}).$$

We investigate if these subsets can be characterized by means of optimal transport. For $r \in \mathbb{R}$, set $\lfloor r \rfloor := \max\{l \in \mathbb{Z} \mid l \leq r\}$. Let $\pi = (n_1, \dots, n_k) \in \mathcal{P}_1(n)$ and introduce

$$\hat{\delta}_\pi := \sum_{i=1}^k \sum_{\alpha=1}^{n_i} \delta_{(i, \lfloor -\frac{n_i}{2} \rfloor + \alpha)}$$

which displays π centered on the x -axis, see Figure 7. We obtain

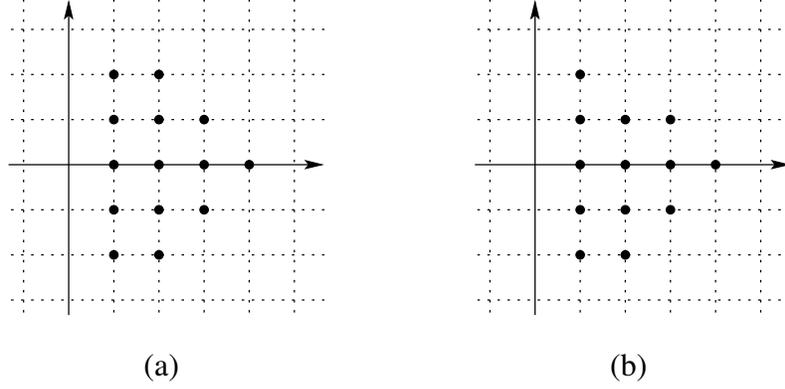


Figure 7: (a) The support of $\hat{\delta}_\pi$ for $\pi = (5, 5, 3, 1)$. (b) The support of $\hat{\delta}_\pi$ for $\pi = (5, 4, 3, 1)$.

Proposition 23. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(x, y) \mapsto (x, -y)$ be the reflection on the x -axis and let c be a metric-like cost function. Then

- (1) $\pi \in \mathcal{P}_1(n \mid \text{all } n_i \text{ odd})$ if and only if $T(\hat{\delta}_\pi) = \hat{\delta}_\pi$.
- (2) $\pi \in \mathcal{P}_1(n \mid \text{all } n_i \text{ odd})$ if and only if $C(\hat{\delta}_\pi, T(\hat{\delta}_\pi)) = 0$.

Proof. (1) follows from the definition of $\hat{\delta}_\pi$. (2) follows from Proposition 13 since $\text{spt}(\hat{\delta}_\pi) = \text{spt}(T(\hat{\delta}_\pi))$ if and only if $\pi \in \mathcal{P}_1(n \mid \text{all } n_i \text{ odd})$. \square

Let $\pi = (n_1, \dots, n_k) \in \mathcal{P}_1(n)$ and $\sigma \in \text{Perm}(k)$. We define

$$\delta_\pi^\sigma := \sum_{i=1}^k \sum_{\alpha=1}^{n_i} \delta_{(\sigma(i), \alpha)}.$$

For $\sigma = \text{Id}$, we recover $\delta_\pi^{\text{Id}} = \delta_\pi$.

Proposition 24. Let c be a metric-like cost function. Then

- (1) $\pi \in \mathcal{P}_1(n \mid \text{not all } n_i \text{ mutually distinct})$.
 - \Leftrightarrow There is $\sigma \in \text{Perm}(k(\pi)) \setminus \{\text{Id}\}$ with $\delta_\pi = \delta_\pi^\sigma$.
 - \Leftrightarrow There is $\sigma \in \text{Perm}(k(\pi)) \setminus \{\text{Id}\}$ with $C(\delta_\pi, \delta_\pi^\sigma) = 0$.

(2) $\pi \in \mathcal{P}_1(n \mid \text{all } n_i \text{ mutually distinct})$.

\Leftrightarrow For all $\sigma \in \text{Perm}(k(\pi)) \setminus \{\text{Id}\}$ holds $\delta_\pi \neq \delta_\pi^\sigma$.

\Leftrightarrow For all $\sigma \in \text{Perm}(k(\pi)) \setminus \{\text{Id}\}$ holds $C(\delta_\pi, \delta_\pi^\sigma) \neq 0$.

Proof. (1) $\pi \in \mathcal{P}_1(n \mid \text{not all } n_i \text{ mutually distinct})$ if and only if there is $1 \leq i_1 \leq i_2 \leq k(\pi)$ with $n_{i_1} = n_{i_2}$ if and only if there is $\sigma \in \text{Perm}(k(\pi))$, namely the permutation exchanging i_1 and i_2 and leaving the other indices fixed, with $\delta_\pi = \delta_\pi^\sigma$ if and only if $C(\delta_\pi, \delta_\pi^\sigma) = 0$ due to Proposition 13.

(2) is the negation of the first item. \square

Now we want to see how we can characterize Euler's identity (1) by means of optimal transport. Andrews & Eriksson [AE] give an algorithm in order to show how a partition in $\mathcal{P}_1(n \mid \text{all } n_i \text{ mutually distinct})$ is transformed into a partition in $\mathcal{P}_1(n \mid \text{all } n_i \text{ odd})$ and back. We will write this algorithm as an explicit bijection, but first we need some notation. Recall that each natural number $m \in \mathbb{N}$ has a unique decomposition into prime factors

$$m = \prod_{p \text{ prime}} p^{\lambda(m,p)} = 2^{\lambda(m,2)} \prod_{2 < p \text{ prime}} p^{\lambda(m,p)}.$$

This decomposition into even and odd part suggests the definition of the functions

$$\begin{aligned} g : \mathbb{N} &\rightarrow \mathbb{N}, & g(m) &:= 2^{\lambda(m,2)}, \\ u : \mathbb{N} &\rightarrow \mathbb{N}, & u(m) &:= \prod_{2 < p \text{ prime}} p^{\lambda(m,p)}. \end{aligned}$$

Note that, for $\pi = (n_1, \dots, n_{k(\pi)}) \in \mathcal{P}(n)$, the number $k(\pi)$ is the number of 'columns' n_i of $(n_1, \dots, n_{k(\pi)})$. We introduce

$$\mathcal{P}_1^{\text{perm}} := \{\sigma(\pi) := (n_{\sigma(1)}, \dots, n_{\sigma(k)}) \mid \pi = (n_1, \dots, n_k) \in \mathcal{P}_1(n), \sigma \in \text{Perm}(k)\}$$

which are 'generalized partitions' since the n_i are *not necessarily* monotone decreasing. We set

$$\varphi : \mathcal{P}_1(n \mid \text{all } n_i \text{ mutually distinct}) \longrightarrow \mathcal{P}_1^{\text{perm}}(n \mid \text{all } n_i \text{ odd})$$

given by

$$\varphi(\pi) = \varphi(n_1, \dots, n_k) := \left(\underbrace{u(n_1), \dots, u(n_1)}_{g(n_1)}, \dots, \underbrace{u(n_k), \dots, u(n_k)}_{g(n_k)} \right).$$

$\varphi(\pi)$ is only a ‘generalized partition’ since its entries $u(n_1), \dots, u(n_k)$ are not necessarily monotonically decreasing. Ordering them monotonically decreasing leads to a partition which we call $\bar{\varphi}(\pi)$. It lies by construction in $\mathcal{P}_1(n \mid \text{all } n_i \text{ odd})$. Andrews & Eriksson [AE] show that this construction induces a bijection

$$\bar{\varphi} : \mathcal{P}_1(n \mid \text{all } n_i \text{ mutually distinct}) \longrightarrow \mathcal{P}_1(n \mid \text{all } n_i \text{ odd}).$$

They also give the inverse construction, but we do not need it here.

Theorem 25. *Abbreviate $\mathcal{P}_1(n \mid \text{all } n_i \text{ mutually distinct}) =: \mathcal{D}$ and $\mathcal{P}_1^{\text{perm}}(n \mid \text{all } n_i \text{ odd}) =: \mathcal{O}$ and denote by $\mathcal{F}(\mathcal{D}, \mathcal{O})$ the space of maps from \mathcal{D} to \mathcal{O} . Let c be a metric-like cost function. Then there is a cost function $\mathcal{C} : \mathcal{D} \times \mathcal{O} \rightarrow \mathbb{R}_+$ such that $\mathcal{C}(\pi, \varphi(\pi)) = \min\{C(\pi, \psi(\pi)) \mid \psi \in \mathcal{F}(\mathcal{D}, \mathcal{O})\}$.*

Proof. The idea is to construct \mathcal{C} in such a way that $\mathcal{C}(\pi, \varphi(\pi)) = 0$ and $\mathcal{C}(\pi, \psi(\pi)) > 0$ for $\psi \neq \varphi$. We note $k(\varphi(\pi)) \geq k(\pi)$. Given $\pi = (n_1, \dots, n_{k(\pi)}) \in \mathcal{D}$ we write (abusing the notation)

$$\varphi(n_i) := \underbrace{(u(n_i), \dots, u(n_i))}_{g(n_i)}$$

and we have $k(\varphi(n_i)) = 1$ if and only if $g(n_i) = 2^0 = 1$. Moreover, for $i \in \mathbb{N}$, define the entry functions $i : \mathcal{P}_1(n) \rightarrow \mathbb{N}$, $i(\pi) = i(n_1, \dots, n_{k(\pi)}) := n_i$ if $1 \leq i \leq k(\pi)$ and 0 otherwise. For $\pi \in \mathcal{P}_1(n)$ define the ‘column measure’

$$\delta_{i(\pi)} := \sum_{\alpha=1}^{i(\pi)} \delta_{(i, \alpha)}$$

and the ‘odd-even prime decomposition measure’ of an entry $i(\pi)$

$$\delta_{gu(i(\pi))} := \sum_{j_1=1}^{g(i(\pi))} \sum_{j_2=1}^{u(i(\pi))} \delta_{(j_1, j_2)}.$$

Now pick $\psi \in \mathcal{F}(\mathcal{D}, \mathcal{O})$ and apply it to $\pi \in \mathcal{D}$. $\psi(\pi) \in \mathcal{O}$ is a generalized partition, but we can assign a measure $\delta_{\psi(\pi)}$ analogously to partitions. Then $\psi(\delta_\pi) = \delta_{\psi(\pi)}$. But $\psi(\delta_{i(\pi)})$ is not even a ‘generalized partition’ since its support lies just somewhere in the support of $\delta_{\psi(\pi)}$. Worse, we do not even know which point goes where since the push forward of a sum of point measures does not care about it as long as the supports are mapped bijectively to each other. To circumvent this problem,

we will minimize over all possibilities in the following way. Let c be a metric-like cost function. Set

$$d(\varphi(n_i), \psi(n_i)) := \min C(\delta_{\varphi(n_i)}, \delta_{\psi(n_i)})$$

where the minimum is taken over all (finite) possibilities to place $\psi(n_i)$ in $\psi(\pi)$. Then $d(\varphi(n_i), \psi(n_i)) = 0$ if and only if the support of the measures $\delta_{\varphi(n_i)}$ and $\delta_{\psi(n_i)}$ coincides. Now define $d(\varphi(\pi), \psi(\pi)) := \sum_{1 \leq i \leq \ell(\pi)} d(\varphi(n_i), \psi(n_i))$. Then $d(\varphi(\pi), \varphi(\pi)) = 0$ and $d(\varphi(\pi), \psi(\pi)) \geq 0$ for $\psi \neq \varphi$. Now set $\mathcal{C}(\pi, \psi(\pi)) := d(\varphi(\pi), \psi(\pi))$ and the claim follows. \square

References

- [Am1] AMBROSIO, L.: *Optimal transport maps in Monge-Kantorovich problem*. ICM 2002, Vol. III, 131 – 140.
- [Am2] AMBROSIO, L.: *Lecture notes on optimal transport problems*. In: *Mathematical Aspects of Evolving Interfaces, Lecture Notes in Mathematics Volume 1812*, Springer 2003, pp. 1 – 52.
- [An] ANDREWS, G.: *The theory of partitions*, Cambridge University Press 1998.
- [AE] ANDREWS, G.; ERIKSSON, K.: *Integer partitions*, Cambridge University Press 2004.
- [Cu] CUTURI, M.: *Permanents, Transportation Polytopes and Positive Definite Kernels on Histograms*, International Joint Conference on Artificial Intelligence (IJCAI) 2007.
- [FM] FELDMAN, M.; McCANN, R.: *Monge’s transport problem on a Riemannian manifold*, Trans. Amer. Math. Soc. 354 (2002), 1667 – 1697.
- [GM] GANGBO, W.; McCANN, R.: *The geometry of optimal transportation*, Acta Math. 177 (1996), no. 2, 113 – 161.
- [K42] KANTOROVICH, L.: *On the translocation of masses*. C.R. (Doklady) Acad. Sci. URSS (N.S.) **37**, 199 – 201 (1942).
- [K48] KANTOROVICH, L.: *On a problem of Monge (In Russian)*. Uspekhi Mat. Nauk. **3**, 225 – 226 (1948), no. 2 (24) March – April.

- [Mo] MONGE, G.: *Mémoire sur la théorie des déblais et des remblais*. Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année. Pages 666 – 704 (1781).
- [RR] RACHEV, S.; RÜSCHENDORF, L.: *Mass Transportation problems*, Volume I + II, Probability and its Applications, Springer 1998.
- [Rü] RÜSCHENDORF, L.: *Fréchet-bounds and their applications*. In: G. Dall'Aglio et al., editor, *Advances in Probability Distributions with given Marginals*, **67** (1991) *Math. Appl.*, 151 – 187. Kluwer Academic Publishers, Dordrecht 1991.
- [SK] SMITH, C.; KNOTT, M.: *On Hoeffding-Fréchet bounds and cyclic monotone relations*. *J. Multivariate Anal.* **40**, 328 – 334 (1992).
- [Su] SUDAKOV, V.: *Geometric problems in the theory of infinite-dimensional probability distributions*. *Proc. Steklov Inst. Math.* **141**, 1 – 178 (1979).
- [Vi] VILLANI, C.: *Optimal transport: old and new*, Springer 2009.