Title

Decomposition of Transport Paths: Map-Compatibility and Capacity Constraint

Author

Sun, Haotian

Publication Date

2023

Decomposition of Transport Paths: Map-Compatibility and Capacity Constraint

By

Haotian Sun
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Qinglan Xia, Chair

Bruno Nachtergaele

Adam J. Jacob

Committee in Charge

2023

Contents

Abstract

Optimal transportation studies the transportation of a given mass distribution to a designated mass distribution so that a given transport cost function reaches minimum. Under different formulations of transport problems, including the Monge transport problem, the Kantorovich transport problem, and the ramified transport problem, transportation has various characterizations. In this dissertation, we use transport paths from ramified transportation to characterize transportation, and use measures and currents to characterize transport paths. We show that a good decomposition of a transport path can be refined into a better decomposition that is more cycle "sensitive". Using better decomposition, we show that cycle-free transport paths can be decomposed into map-compatible transport paths components, and we also prove similar results when transport paths are under capacity constraint. These decomposition results describe properties of optimal transport paths, and using these properties we can narrow down the scope of finding an optimal transport path. The notion of capacity constraint on transport paths is a generalization of the usual transport paths, and it makes transport paths more relevant and applicable to real life transportation.

In Chapter 1, we first review concepts related to measures, then introducing the Monge and the Kantorovich transport problems. In the Monge and the Kantorovich transport problem, transportation is characterized by functions defined on sources and targets, rather than the actual transport path connecting them. In the next chapter, we will see another characterization of transportation using transport paths.

In Chapter 2, we introduce ramified transportation, which uses the actual transport paths from sources to targets to characterize the transportation between two mass distributions. Transport paths in ramified transportation can be described using rectifiable 1-currents, and this is where we start in this chapter.

In Chapter 3, we first revise good decompositions of a transport path into better decompositions which are used later to decompose cycle-free transport paths based on targets. Then we show the components of previously decomposed transport paths are compatible with certain transport maps and plans. Finally we consider a special type of transport paths, the stair-shaped transport paths, which can be decomposed as the difference of two map-compatible transport paths.

In Chapter 4, we study transport paths under capacity constraint. In this case, each transport path is defined through multiple components such that the total mass transported on each transport

path component is no more than the predetermined capacity. Then we start to analyze the amount of components needed in a transportation, the existence of admissible optimal transport paths, regularities among different transport path components, and existence of map-compatibility.

Acknowledgments

I'd like to thank my dearest advisor, Qinglan Xia, for his unremitting support in my academic research and his guidance on learning new subjects. His remarkable insight, consistent patience, and easygoing personality are the key factors of what I achieved today.

I'd like to thank Bruno Nachtergaele, Adam Jacob, Naoki Saito, and Michael Zhang for being on my qualifying exam committee. I'd like to thank Bruno's lectures in analysis, Adam's lectures in measure theory, that help me build up foundations for future research. Moreover, I'd like to thank all the faculty members that I took classes with, since they constituted each step of progress I made until today.

I'd like to thank Tina Denena, Diana Bond, Victoria Whistler, and all the staff members in the math department for their vital efforts in doing the paper works, logistics, coordination, not just for me but for the whole math department.

I'd like to thank the Galois Group (an organization for graduate students) for their informative website and all the supporting resources on it.

I'd like to thank my family for their consistent support and encouragement throughout all these years.

CHAPTER 1

Introduction to Monge-Kantorovich optimal transportation

This chapter is based on content from [**1**] and [**4**]. In this chapter, we will first review some basic definitions and notations for functions and measures in section 1.1. In section 1.2, we will see the definition and basic results of the Monge optimal transportation problem. In section 1.3, we will move to the Kantorovich optimal transport problem, and its related results, i.e. existence, examples, compactness. In general, optimal transport can be applied in various areas including economics, image processing, PDEs, probability and statistics. [**7**]

1.1. Basic notations and concepts

In optimal transport problems, functions and measures are the basic elements that describe these problems. Functions are used to describe the way that a mass is being transported and the total cost of a transportation. Measures tell us how these "ready to ship" masses are distributed before the transportation, and how do we want these masses to be distributed after the transportation. This brings the need of knowing properties of functions and measures. In this section we will review some basic concepts and properties of functions and measures from [**1**] and [**4**].

Denote $\overline{\mathbb{R}}$ as the extended real line. The characteristic function $\chi_E : X \to \{0,1\}$ is defined by

$$\chi_E(x) := \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \in X \setminus E. \end{cases}$$

The Lebesgue measure in $\mathbb{R}^n$ will be denoted by $\mathcal{L}^n$.

Let (X,d) be a metric space, then denote $C(X)$ as the space of continuous functions $f : X \to \mathbb{R}$, and denote $C_b(X)$ as the subspace of bounded continuous functions. Let $\mathrm{Lip}(X)$ and $\mathrm{Lip}_b(X)$ denote the spaces of Lipschitz and bounded Lipschitz functions respectively, with

$$\mathrm{Lip}(f) := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$$

denotes the Lipschitz constant.

Given an open set $U \subseteq \mathbb{R}^n$ and $k \in \mathbb{N}$, denote $C^k(U)$ as the set of functions with k continuous derivatives in U, and denote elements of $C^k(\overline{U})$ as the restriction to $\overline{U}$ of functions in $C^k(\mathbb{R}^n)$. Let

$$C^\infty(U) := \bigcap_{k=1}^{\infty} C^k(U), \text{ and } C^\infty(\overline{U}) := \bigcap_{k=1}^{\infty} C^k(\overline{U}).$$

Denote elements in $C_c(U), C_c^k(U), C_c^\infty(U)$, as functions with compact support.

In a metric space (X, d), let $\mathcal{B}(X)$ be its Borel σ-algebra and let $\mathcal{M}(X)$ be the set of the σ-additive functions $\mu : \mathcal{B}(X) \to \mathbb{R}$. Furthermore, denote by

$$\mathcal{M}_+(X) := \{\mu \in \mathcal{M}(X) : \mu \geq 0\}, \ \mathcal{P}(X) = \{\mu \in \mathcal{M}_+(X) : \mu(X) = 1\}.$$

the subsets of nonnegative and probability measures, respectively.

Definition 1.1.1. Given $\mu \in \mathcal{M}(X)$, the total variation measure $|\mu|$ is the set function defined on $\mathcal{B}(X)$ by

$$|\mu|(B) := \sup \left\{ \sum_{i \in \mathbb{N}} |\mu(B_i)| : \{B_i\}_{i \in \mathbb{N}} \text{ is a Borel partition of } B \right\},$$

and for $E \in \mathcal{B}(X)$, the restriction $\mu\lfloor_E$ of μ on E is defined by

$$\mu\lfloor_E(B) := \mu(E \cap B), \qquad B \in \mathcal{B}(X).$$

Sometimes we write $\chi_E \mu$ to denote $\mu\lfloor_E$.

Definition 1.1.2. Given $\mu \in \mathcal{M}(X)$, its support is the closed set defined by

$$\text{supp } \mu := \{x \in X : |\mu|(U) > 0 \text{ for all } U \text{ open, such that } x \in U\}.$$

We say that μ is concentrated on $A \in \mathcal{B}(X)$ if $|\mu|(X \setminus A) = 0$.

Given two measures, we define the push forward operator in the following definition, and we will see this operator is used to defined the Monge transport problem in the next section.

Definition 1.1.3. Given a Borel function $f : X \to Y$, we define the push forward operator $f_\# : \mathcal{M}(X) \to \mathcal{M}(Y)$ by

$$f_\#\mu(B) := \mu(f^{-1}(B)), \quad \forall B \in \mathcal{B}(Y),$$

and call $f_\#\mu$ the push forward measure of μ by f.

Integration that involve the push forward operator of measures is demonstrated in the following two propositions.

PROPOSITION 1.1.4. *For any Borel function $f : X \to Y$ and any Borel function $\phi : Y \to [0, \infty]$ one has*

$$\int_Y \phi \ \mathrm{d}f_{\#}\mu = \int_X (\phi \circ f) \ \mathrm{d}\mu.$$

It follows that $\psi : X \to \overline{\mathbb{R}}$ is $f_{\#}\mu$-integrable if and only if $\psi \circ f$ is μ-integrable.

PROPOSITION 1.1.5. *For $T : X \to Y$ Borel, one has $T_{\#}\mu = \nu$ if and only if*

$$\int_Y \phi \ \mathrm{d}\nu = \int_X (\phi \circ T) \ \mathrm{d}\mu, \quad \phi \in C_b(Y).$$

Given any two measures μ, ν, we define the upper density $\overline{D}_\mu \nu(x)$ and lower density $\underline{D}_\mu \nu(x)$ of ν with respect to μ as follows.

DEFINITION 1.1.6. Let μ and ν be Radon measures on $\mathbb{R}^n$. For each point $x \in \mathbb{R}^n$, define

$$\overline{D}_\mu \nu(x) := \begin{cases} \limsup\limits_{r \to 0} \frac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0 \end{cases}$$

and

$$\underline{D}_\mu \nu(x) := \begin{cases} \liminf\limits_{r \to 0} \frac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0 \end{cases}$$

DEFINITION 1.1.7. If $\overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x) < +\infty$, we say ν is differentiable with respect to μ at x and write

$$D_\mu \nu := \overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x).$$

$D_\mu \nu$ is the derivative of ν with respect to μ. We also call $D_\mu \nu$ the density of ν with respect to μ.

DEFINITION 1.1.8. Let μ and ν be Borel measures on $\mathbb{R}^n$.

(i) The measure ν is absolutely continuous with respect to μ, written

$$\nu \ll \mu,$$

provided $\mu(A) = 0$ implies $\nu(A) = 0$ for all $A \subseteq \mathbb{R}^n$.

(ii) The measure ν and μ are mutually singular, written

$$\nu \perp \mu,$$

if there exists a Borel subset $B \subseteq \mathbb{R}^n$ such that

$$\mu(\mathbb{R}^n \setminus B) = \nu(B) = 0.$$

Using concepts and notations of density of measures, we have the following measure decomposition theorem.

THEOREM 1.1.9 (Lebesgue Decomposition Theorem). *Let ν and μ be Radon measures on $\mathbb{R}^n$.*

(i) *Then*

$$\nu = \nu_{ac} + \nu_s,$$

where ν_{ac}, ν_s are Radon measures on $\mathbb{R}^n$ with

$$\nu_{ac} \ll \mu, \;\; \nu_s \perp \mu.$$

(ii) *Furthermore,*

$$D_\mu \nu = D_\mu \nu_{ac}, \;\; D_\mu \nu_s = 0 \quad \mu - a.e.,$$

and consequently

$$\nu(A) = \int_A D_\mu \nu d\mu \; + \nu_s(A)$$

for each Borel set $A \subseteq \mathbb{R}^n$.

DEFINITION 1.1.10. We call ν_{ac} the absolutely continuous part and ν_s the singular part of ν with respect to μ.

Now, we may introduce the notion of (weak) convergence of measures.

THEOREM 1.1.11 (Weak convergence of measures). *Let μ, μ_k, $(k = 1, 2, \ldots)$ be Radon measures on $\mathbb{R}^n$. The following three statements are equivalent:*

(i) $\lim_{k\to\infty} \int_{\mathbb{R}^n} f d\mu_k = \int_{\mathbb{R}^n} f d\mu$ *for all $f \in C_c(\mathbb{R}^n)$.*

(ii) $\limsup_{k\to\infty} \mu_k(K) \le \mu(K)$ *for each compact set $K \subseteq \mathbb{R}^n$ and $\mu(U) \le \liminf_{k\to\infty} \mu_k(U)$ for each open set $U \subseteq \mathbb{R}^n$.*

(iii) $\lim_{k\to\infty} \mu_k(B) = \mu(B)$ *for each bounded Borel set* $B \subseteq \mathbb{R}^n$ *with* $\mu(\partial B) = 0$.

DEFINITION 1.1.12. If (i), (ii), (iii) hold, we say the measures $\{\mu_k\}_{k=1}^{\infty}$ converge weakly to the measure μ, written as

$$\mu_k \rightharpoonup \mu.$$

1.2. The Monge optimal transport problem

The Monge optimal transport problem (see [**1**]) is to find a way to transport mass from a given distribution to another distribution, such that a given cost function is minimized. In the current description of Monge optimal transport problem, suppose $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$, a Borel map $T : X \to Y$ such that $T_{\#}\mu = \nu$ is called a *transport map* from μ to ν, and we denote

$$Map(\mu, \nu) := \{T : X \to Y \text{ Borel}, \ T_{\#}\mu = \nu\}.$$

The Monge optimal transport problem is

$$\inf\left\{\int_X c(x, T(x))\ \mathrm{d}\mu(x) : \ T \in Map(\mu, \nu)\right\}, \tag{1.2.1}$$

where $c(x, y) : X \times Y \to [0, \infty]$ is a Borel function, which gives the cost of transporting a unit of mass from x to y.

In Monge's original formulation, $X = Y$ were Euclidean spaces and $c(x, y) = |x - y|$. Let $\mathcal{C}_\mu(T)$ be the transport cost $\int_X c(x, T(x))\ \mathrm{d}\mu(x)$, we will omit μ when it is clear from the context. In the following example, we will first see the existence of optimal transport maps for some mass distributions.

EXAMPLE 1. *Given two measures*

$$\mu = \sum_{i=1}^{N} \frac{1}{N}\delta_{x_i}, \ \nu = \sum_{j=1}^{N} \frac{1}{N}\delta_{y_j},$$

with discrete spaces $X = \{x_1, \ldots, x_N\}$ *and* $Y = \{y_1, \ldots, y_N\}$*, such that both have cardinality* N*. A function* $T : X \to Y$ *is a bijection if and only if* $T_{\#}\mu = \nu$*, since*

$$T_{\#}\mu = \sum_{j=1}^{N} \frac{1}{N}\delta_{T(x_i)}.$$

For any choice of cost function $c(x, y)$, since the class of admissible transport maps T is a finite set, we have the existence of an optimal transport map.

Monge transport problem may fail to have a solution. For instance, when $\mu = \delta_0$, and $\nu = \frac{1}{2}(\delta_{-1} + \delta_1)$ there is no transport map, since we cannot map one point $\{0\}$, to two points $\{-1, 1\}$. Nevertheless, the following result gives the existence of an optimal transport map when measures are supported on $\mathbb{R}$, and the source measure is atom free.

THEOREM 1.2.1. *If $\mu, \nu \in \mathcal{P}(\mathbb{R})$ and μ has no atom, then there exists $T : \mathbb{R} \to [-\infty, \infty]$ non-decreasing pushing μ into ν and any other map S with these properties coincides with T on $\operatorname{supp} \mu$, with at most countably many exceptions.*

If $c(x, y) = \phi(|y - x|)$ with $\phi : [0, \infty) \to [0, \infty)$ convex and non-decreasing, and if $\mathcal{C}_\mu(T) < \infty$, then T is an optimal map. If ϕ is strictly convex, T is the unique optimal map.

PROOF. Denote $F_\mu(x) := \mu((-\infty, x])$ as the cumulative distribution function of a probability measure μ in $\mathbb{R}$. Then one can check that the desired transport map is given by

$$T(x) := \inf\{y \in \mathit{supp}\, \nu : F_\nu(y) \geq F_\mu(x)\}.$$

$\square$

When ϕ is not strictly convex, we will see from the following example that optimal transport map is not unique. i.e. when $p = 1$ both T_1 and T_2 are optimal in the following example.

EXAMPLE 2. *Given an integer $M \geq 2$, consider*

$$\mu = \frac{1}{M} \mathcal{L}^1 \lfloor_{[0,M]}, \;\; \nu = \frac{1}{M} \mathcal{L}^1 \lfloor_{[1,M+1]},$$

and the cost function $c(x, y) = |x - y|^p$ with $0 < p < \infty$. Let T_1, T_2 be two admissible transport maps where

$$T_1(x) := x + 1, \quad T_2(x) := \begin{cases} x + M & \text{if } 0 \leq x \leq 1, \\ x & \text{otherwise.} \end{cases}$$

By Theorem 1.2.1, we have T_1 is an optimal map for $p \geq 1$. Moreover, when $p = 1$, both T_1 and T_2 are optimal. When $p = 1$,

$$\mathcal{C}(T_1) = \frac{1}{M}\int_0^M |(x+1) - x|\, dx = 1,\ \mathcal{C}(T_2) = \frac{1}{M}\int_0^1 |(x+M) - x|\, dx = 1.$$

For every admissible T,

$$\begin{aligned} \mathcal{C}(T) &= \int_{\mathbb{R}} |T(x) - x|\, d\mu(x) \geq \int_{\mathbb{R}} T(x)\, d\mu(x) - \int_{\mathbb{R}} x\, d\mu(x) = \int_{\mathbb{R}} y\, d\nu(y) - \int_{\mathbb{R}} x\, d\mu(x) \\ &= \frac{1}{M}\int_1^{M+1} y\, dy - \frac{1}{M}\int_0^M x\, dx = 1. \end{aligned}$$

From the following example, we will see the infimum in the Monge formulation of transport map is not necessarily the minimum. i.e. in this example, there is no transport map such that the corresponding transport cost equals its infimum.

Example 3. *Consider*

$$\mu = \mathcal{H}^1 \lfloor_{\{0\}\times[0,1]} \in \mathcal{P}(\mathbb{R}^2),\ \nu = \frac{1}{2}\mathcal{H}^1 \lfloor_{\{-1\}\times[0,1]} + \frac{1}{2}\mathcal{H}^1 \lfloor_{\{1\}\times[0,1]},$$

and the cost function is $c(x,y) := |y-x|$. Here, $\mathcal{H}^1$ denotes the one-dimensional Hausdorff measure in $\mathbb{R}^2$.

Let T be any admissible transport map, then for any $x \in supp\, \mu$, and $y \in supp\, \nu$,

$$\mathcal{C}(T) = \int_{\mathbb{R}^2} |y - x|\, d\mu \geq \int_{\mathbb{R}^2} 1\, d\mu = 1.$$

Next, divide the segment $\{0\} \times [0,1]$ into $2N$ equal pieces, and divide the segments $\{\pm 1\} \times [0,1]$ into N equal pieces. Let T_N be the map that maps linearly the $(2i+1)$-th piece of $\{0\} \times [0,1]$ to the $(i+1)$-th piece of $\{-1\} \times [0,1]$, and maps linearly the $(2i+2)$-th piece of $\{0\} \times [0,1]$ to the $(i+1)$-th piece $\{1\} \times [0,1]$, for $i = 0, 1, \ldots, N-1$. Then,

$$\mathcal{C}(T_N) = 2N\int_0^{\frac{1}{2N}} \sqrt{1+y^2}\, dy \leq 2N\int_0^{\frac{1}{2N}} 1 + y\, dy = 1 + \frac{1}{4N},$$

Hence, as $N \to \infty$, $\mathcal{C}(T_N) \to 1$.

However, no optimal transport map T exists. Indeed, assume there exists T such that

$$\int_{\mathbb{R}^2} (|T(x) - x| - 1)\, d\mu(x) = 0.$$

Since $|T(x)-x|\geq 1$ by definition of μ,ν, we have $|T(x)-x|=1$ μ-a.e. Hence, for $\mathcal{L}^1$-a.e. and $t\in[0,1]$ either $T((0,t))=(1,t)$ or $T((0,t))=(-1,t)$. Denote

$$A_+=\{t\in[0,1]:T((0,t))=(1,t)\},\ A_-=\{t\in[0,1]:T((0,t))=(-1,t)\},$$

then

$$T_\#\mu=\mathcal{H}^1\lfloor_{\{-1\}\times A_-}+\mathcal{H}^1\lfloor_{\{-1\}\times A_-},$$

which implies $T_\#\mu\neq\nu$, a contradiction.

1.3. The Kantorovich optimal transport problem

In the Kantorovich formulation of optimal transport problem, it uses transport plans to characterize transportation, this will resolve the "non-splitting" issue of the source measure in the Monge transport problem. Again, the definitions and some major results of Kantorovich optimal transport problem are from [**1**].

Definition 1.3.1 (Transport Plans). Given $\mu\in\mathcal{P}(X)$ and $\nu\in\mathcal{P}(Y)$, define

$$Plan(\mu,\nu):=\{\pi\in\mathcal{P}(X\times Y):\pi(A\times Y)=\mu(A),\ \pi(X\times B)=\nu(B),\ \text{for } A,B \text{ Borel}\}.$$

Denoting by $p_X:X\times Y\to X$, $p_Y:X\times Y\to Y$ the coordinate projections, definition of transport plan is equivalent to

$$(p_X)_\#\pi=\mu,\ (p_Y)_\#\pi=\nu.$$

Transport plans represent a way to transport mass, $\pi(A\times B)$, from A to B. Kantorovich formulation of the optimal transport problem is to find

$$\inf\left\{\int_{X\times Y}c(x,y)\,d\pi(x,y):\pi\in Plan(\mu,\nu)\right\}. \tag{1.3.1}$$

For each $\pi\in Plan(\mu,\nu)$, its transportation cost is denoted by

$$\mathcal{C}(\pi):=\int_{X\times Y}c(x,y)\,d\pi(x,y).$$

Unlike in the Monge problem, it is much easier to get the existence of an optimal transport plan for the Kantorovich problem. When c is lower semi-continuous, by [**1**, Theorem 2.6] we have

$\pi \to \mathcal{C}(\pi)$ is also lower semi-continuous. Since [**1**, Corollary 2.9] gives the set $Plan(\mu,\nu)$ is compact with respect to the weak topology, we have the following existence result.

THEOREM 1.3.2. *[**1**, Theorem 2.10] Let X, Y be Polish spaces and let $c : X \times Y \to [0,\infty]$ be lower semicontinuous. Then the minimum (infimum) in (1.3.1) is attained.*

Given a transport map T (as illustrated in 1.2.1), we can define the corresponding transport plan $\pi_T := (id \times T)_{\#}\mu$, where $id \times T : X \to X \times Y$ is the map $x \mapsto (x, T(x))$. By Proposition 1.1.4, we have

$$\mathcal{C}(\pi_T) = \int_X c(x, T(x))\, d\mu = \mathcal{C}(T).$$

Thus, by definition from (1.2.1) and (1.3.1),

$$\inf\left\{\mathcal{C}(T) : T : X \to Y \text{ Borel }, T_{\#}\mu = \nu\right\} \geq \inf\left\{\mathcal{C}(\pi) : \pi \in Plan(\mu,\nu)\right\},$$

which gives the infimum in the Monge problem is larger or equal to the infimum in the Kantorovich problem. By the following theorem, we can still reach equality under suitable conditions.

THEOREM 1.3.3. *[**1**, Theorem 2.2 (Pratelli)] If μ has no atom and $c : X \times Y \to [0,\infty)$ is continuous, then*

$$\inf\left\{\int_X c(x, T(x))\ \mathrm{d}\mu(x) :\ T \in Map(\mu,\nu)\right\} = \min\left\{\int_{X\times Y} c(x,y)\, d\pi(x,y) : \pi \in Plan(\mu,\nu)\right\}.$$

When $c(x,y)$ is strictly convex (e.g. $\frac{1}{2}|x-y|^2$) and μ is absolutely continuous with respect to Lebesgue measure, we can find optimal transport map by using Kantorovich duality. More precisely, we have the following results.

DEFINITION 1.3.4. Given a metric space (X,d) and $p \in [1,\infty)$, we define

$$\mathcal{P}_p(X) := \{\mu \in \mathcal{P}(x) : \int_X d(x,x_0)^p\, d\mu(x) < \infty \text{ for some } x_0 \in X\}.$$

THEOREM 1.3.5. *([**1**, Theorem 5.2]) Assume that $X = Y = \mathbb{R}^n$, $c(x,y) = \frac{1}{2}|x-y|^2$, $\mu,\nu \in \mathcal{P}_2(\mathbb{R}^n)$, and $\mu \ll \mathcal{L}^n$. Then*

(i) *the Kantorovich optimal transport problem in (1.3.1) has a unique solution π. In addition, π is induced by a transport map T, which is a unique solution to the Monge optimal* *transport problem in (1.2.1), and* $T = \nabla\psi$*, where* $\psi : \mathbb{R}^n \to (-\infty, \infty]$ *is a lower semi-continuous convex function differentiable* μ*-a.e.;*

(ii) *conversely, if* ψ *convex, lower semi-continuous, differentiable* μ*-a.e. with* $|\nabla\psi| \in L^2(\mu)$*, then* $T := \nabla\psi$ *is optimal from* μ *to* $\nu := T_{\#}\mu \in \mathcal{P}_2(\mathbb{R}^n)$*;*

(iii) *if* $\nu \ll \mathcal{L}^n$*, denoting by* $T^{\mu\to\nu}$ *(respectively,* $T^{\nu\to\mu}$*) the unique optimal transport map from* μ *to* ν *(respectively, from* ν *to* μ*), we get that*

$$T^{\nu\to\mu} \circ T^{\mu\to\nu} = id \quad \mu - a.e.\ in\ \mathbb{R}^n, \quad T^{\mu\to\nu} \circ T^{\nu\to\mu} = id \quad \nu - a.e.\ in\ \mathbb{R}^n.$$

The map T in (i) is usually called the Brenier map, and has many applications in the Monge-Ampère equation and proof of geometric and Gaussian inequalities. [**10**]

CHAPTER 2

Introduction to ramified optimal transportation

In this Chapter we will review some basic definitions of geometric measures theory ([**5**], [**8**]), and ramified optimal transportation ([**11**], [**12**], [**13**], [**14**], [**17**], [**15**]).

2.1. Differential forms & Rectifiable currents

We start with the concept of covectors. By convention 0-covectors is defined as scalars, i.e.

$$\bigwedge^0(\mathbb{R}^p) := \mathbb{R}.$$

Denote

$$\bigwedge^1(\mathbb{R}^p) = \{\omega : \omega \text{ is a linear functional from } \mathbb{R}^p \text{ to } \mathbb{R}\}$$

as the dual space of $\mathbb{R}^p$. Let $dx^1, \ldots, dx^p \in \bigwedge^1(\mathbb{R}^p)$ be the basis dual to the standard basis $e_1, \ldots, e_p$ of $\mathbb{R}^p$. i.e. If $v = (a_1, \ldots, a_p) \in \mathbb{R}^p$, then $dx^j(v) = a_j$, for $j = 1, 2, \ldots, p$.

When $k \geq 2$, $\bigwedge^k(\mathbb{R}^p)$ denotes the space of k-covectors, which are alternating k-linear functions on

$$\underbrace{\mathbb{R}^p \times \mathbb{R}^p \times \cdots \times \mathbb{R}^p}_{k \text{ factors}}.$$

Here, elements $\omega \in \bigwedge^k(\mathbb{R}^p)$ means $\omega(v_1, \ldots, v_k)$ is linear in each $v_j \in \mathbb{R}^p$, and

$$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k), \quad \text{for each } i \neq j.$$

Let $\omega_1, \ldots, \omega_n \in \bigwedge^1(\mathbb{R}^p)$, then the wedge product $\omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_n \in \bigwedge^n(\mathbb{R}^p)$ is defined as

$$\begin{aligned} \omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_n(v_1, v_2, \cdots, v_n) &= \sum_\sigma sgn(\sigma) \cdot \omega_{\sigma(1)}(v_1)\omega_{\sigma(2)}(v_2) \cdots \omega_{\sigma(n)}(v_n) \\ &= det\left(\omega_i(v_j)\right), \end{aligned} \tag{2.1.1}$$

where the sum in equation (2.1.1) is over all permutations σ of $\{1, 2, \ldots, n\}$, and $sgn(\sigma)$ is the sign of the permutation σ.

For each $k > 0$, the space $\bigwedge^k(\mathbb{R}^p)$ is a vector space of dimension $\binom{p}{k}$ with basis

$$\{dx^\alpha = dx^{\alpha(1)} \wedge \ldots \wedge dx^{\alpha(k)}, \alpha \in I_{k,p}\},$$

where

$$I_{k,p} = \{\alpha = (i_1, i_2, \ldots, i_k) \in \mathbb{Z}_+^k : 1 \leq i_1 < \cdots < i_k \leq p\}.$$

Using this basis, each $\omega \in \bigwedge^k(\mathbb{R}^p)$ can be represented as

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq p} \omega_{i_1 \ldots i_k}\ dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \sum_{\alpha \in I_{k,p}} \omega_\alpha\ dx^\alpha,$$

where $\omega_{i_1 \ldots i_k} = \omega(e_{i_1}, \ldots, e_{i_k})$. For any $\omega = \sum_{\alpha \in I_{n,p}} \omega_\alpha\ dx^\alpha \in \bigwedge^n(\mathbb{R}^p)$ and $\eta = \sum_{\beta \in I_{m,p}} \eta_\beta\ dx^\beta \in \bigwedge^m(\mathbb{R}^p)$, then

$$\omega \wedge \eta = \sum_{\alpha \in I_{n,p},\, \beta \in I_{m,p}} \omega_\alpha\, \eta_\beta\, dx^\alpha \wedge dx^\beta \in \bigwedge^{n+m}(\mathbb{R}^p).$$

A k-covector is *simple* if it is the wedge product of k numbers of 1-covectors. Note that, for $k = 0, 1, p$, all k-covectors are simple. Indeed, when $k = 0$, a 0-covector is just a scalar in $\mathbb{R}$, which is the wedge product of 0 1-covector. When $k = 1$, a 1-covector is of the form

$$\sum_{i=1}^p \omega_i dx^i,$$

which is the wedge product of the 1-covector itself. When $k = p$, a p-covector is of the form

$$\sum_{\alpha \in I_{p,p}} \omega_\alpha dx^\alpha = \sum_{\alpha \in I_{p,p}} \omega_\alpha dx^1 \wedge dx^2 \wedge \cdots \wedge dx^p.$$

Hence, it is the wedge product of p 1-covectors, dx^i's.

Definition 2.1.1. The dual space $\bigwedge_k(\mathbb{R}^p)$ of $\bigwedge^k(\mathbb{R}^p)$ is called the space of k-vectors, it has the dual basis

$$\{e_\alpha = e_{\alpha(1)} \wedge \ldots \wedge e_{\alpha(k)} : \alpha \in I_{k,p}\}.$$

The spaces $\bigwedge_k(\mathbb{R}^p)$ and $\bigwedge^k(\mathbb{R}^p)$ have inner products induced from $\mathbb{R}^p$ as follows:

$$\left\langle \sum_{\alpha \in I_{k,p}} \omega_\alpha^1 dx^\alpha, \sum_{\alpha \in I_{k,p}} \omega_\alpha^2 dx^\alpha \right\rangle = \sum_{\alpha \in I_{k,p}} \omega_\alpha^1 \cdot \omega_\alpha^2, \quad \left\langle \sum_{\alpha \in I_{k,p}} v_1^\alpha e_\alpha, \sum_{\alpha \in I_{k,p}} v_2^\alpha e_\alpha \right\rangle = \sum_{\alpha \in I_{k,p}} v_1^\alpha \cdot v_2^\alpha.$$

The length of ω is given by $|\omega| = \langle \omega, \omega \rangle^{1/2}$, and similarly, the length of v is given by $|v| = \langle v, v \rangle^{1/2}$.

Moreover, for any k-covector $\omega = \sum_{\alpha \in I_{k,p}} \omega_\alpha dx^\alpha$ and k-vector $v = \sum_{\alpha \in I_{k,p}} v^\alpha e_\alpha$, the inner product of ω and v is defined as:

$$\langle \omega, v \rangle = \sum_{\alpha \in I_{k,p}} \omega_\alpha \cdot v^\alpha. \tag{2.1.2}$$

The comass norm of ω is defined as

$$\|\omega\| := \sup_v \{\langle \omega, v \rangle, v \text{ is simple}, |v| \leq 1, v \in \bigwedge\nolimits_k(\mathbb{R}^p)\}. \tag{2.1.3}$$

Given $\ell : \mathbb{R}^p \to \mathbb{R}^q$ linear, the "pull-back" map $\ell^\# : \bigwedge^k(\mathbb{R}^q) \to \bigwedge^k(\mathbb{R}^p)$ is defined as

$$(\ell^\# \omega)(v_1, v_2, \ldots, v_k) := \omega(\ell(v_1), \ell(v_2), \ldots, \ell(v_k)), \quad v_1, v_2, \ldots, v_k \in \mathbb{R}^p. \tag{2.1.4}$$

The "push-forward" map $\ell_\# : \bigwedge_k(\mathbb{R}^p) \to \bigwedge_k(\mathbb{R}^q)$ is defined as

$$\langle \ell^\# \omega, w \rangle = \langle \omega, \ell_\# w \rangle, \quad \omega \in \bigwedge^k(\mathbb{R}^q), w \in \bigwedge\nolimits_k(\mathbb{R}^p).$$

Here, the inner product is defined according to equation (2.1.2). In this case, we have

$$\ell^\# : \bigwedge^k(\mathbb{R}^q) \to \bigwedge^k(\mathbb{R}^p) \text{ and } \ell_\# : \bigwedge\nolimits_k(\mathbb{R}^p) \to \bigwedge\nolimits_k(\mathbb{R}^q),$$

such that

$$\ell^\#(\omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_k) = (\omega_1 \circ \ell) \wedge (\omega_2 \circ \ell) \wedge \cdots \wedge (\omega_k \circ \ell), \text{ for } \omega_1, \ldots, \omega_k \in \bigwedge^1(\mathbb{R}^q) \tag{2.1.5}$$

and

$$\ell_\#(v_1 \wedge v_2 \wedge \cdots \wedge v_k) = \ell(v_1) \wedge \ell(v_2) \wedge \cdots \wedge \ell(v_n), \text{ for } v_1, \ldots, v_k \in \mathbb{R}^p.$$

Definition 2.1.2. A k-form (respectively k-vector field) on an open subset U of $\mathbb{R}^p$ is a function $\omega : U \to \bigwedge^k(\mathbb{R}^p)$ (respectively, $\xi : U \to \bigwedge_k(\mathbb{R}^p)$). The set of all C^∞ k-forms on U is denoted by $\mathcal{E}^k(U)$. The set of all C^∞ k-forms with compact supports in U is denoted by $\mathcal{D}^k(U)$, where

$$spt\left(\sum_\alpha \omega_\alpha dx^\alpha\right) = \bigcup_\alpha spt\,(\omega_\alpha)\,.$$

Note that $\omega \in \mathcal{E}^k(U)$ means that

$$\omega = \sum_{\alpha \in I_{k,p}} \omega_\alpha dx^\alpha, \text{ where } \omega_\alpha \in C^\infty(U).$$

The value of

$$\omega(x) = \sum_{\alpha \in I_{k,p}} \omega_\alpha(x) dx^\alpha, \ x \in U,$$

can also be denoted as $\omega|_x$. We also denote the comass norm of $\omega \in \mathcal{E}^k(U)$ by

$$||\omega|| := \sup_{x \in U} ||\omega(x)||, \tag{2.1.6}$$

where $||\omega(x)||$ denotes the comass norm of the covector $\omega(x)$ as defined in (2.1.3).

Also, the space $\mathcal{D}^k(U)$ has the topology where the sequence $\{\omega_i\} \subseteq \mathcal{D}^k(U)$ converges to $\omega \in \mathcal{D}^k(U)$ as $i \to \infty$ if and only if

$$U \cap \text{closure}\left(\bigcup_{i=1}^{\infty} spt(\omega_i)\right) \text{ is compact,}$$

and

$$|D^\beta \omega_i - D^\beta \omega| \to 0 \text{ uniformly on } U,$$

for every multi-index β as $i \to \infty$.

Let $U \subseteq \mathbb{R}^p$ be an open set, and

$$\omega = \sum_{\alpha \in I_{k,p}} \omega_\alpha dx^\alpha \in \mathcal{E}^k(U),$$

then the exterior derivative $d : \mathcal{E}^k(U) \to \mathcal{E}^{k+1}(U)$ is defined as:

$$d\omega := \sum_{j=1}^{p} \sum_{\alpha \in I_{k,p}} \frac{\partial \omega_\alpha}{\partial x^j} \, dx^j \wedge dx^\alpha.$$

Since

$$\frac{\partial^2 \omega_\alpha}{\partial x^i \partial x^j} = \frac{\partial^2 \omega_\alpha}{\partial x^j \partial x^i}, \text{ and } dx^i \wedge dx^j = -dx^j \wedge dx^i,$$

then

$$d^2\omega = d\left(\sum_{j=1}^{p} \sum_{\alpha \in I_{k,p}} \frac{\partial \omega_\alpha}{\partial x^j} \, dx^j \wedge dx^\alpha\right) = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{\alpha \in I_{k,p}} \frac{\partial^2 \omega_\alpha}{\partial x^i \partial x^j} \, dx^i \wedge dx^j \wedge dx^\alpha = 0, \tag{2.1.7}$$

for all $\omega \in \mathcal{E}^k(U)$.

Also, given $V \subseteq \mathbb{R}^q$ open set, with

$$\omega = \sum_{\alpha \in I_{k,q}} \omega_\alpha(y) dy^\alpha \in \mathcal{E}^k(V),$$

and let $f = (f^1, f^2, \ldots, f^q) : U \to V$ be a smooth map. The "pull back" form $f^{\#}\omega \in \mathcal{E}^k(U)$ is defined as

$$f^{\#}\omega = \sum_{\alpha=(i_1,i_2,\ldots,i_k)\in I_{k,q}} (\omega_\alpha \circ f) df^{i_1} \wedge df^{i_2} \wedge \cdots \wedge df^{i_n},$$

such that for each $j = 1, 2, \ldots, q$,

$$df^j := \sum_{i=1}^{p} \frac{\partial f^j}{\partial x^i} dx^i.$$

In other words,

$$\left(f^{\#}\omega\right)|_x = (df_x)^{\#}\left(\omega|_{f(x)}\right).$$

Indeed, note that for each $j = 1, 2, \ldots, q$,

$$df_x^j = \sum_{i=1}^{p} \frac{\partial f^j}{\partial x^i} dx^i,$$

and this gives

$$df_x = \left(\sum_{i=1}^{p} \frac{\partial f^1}{\partial x^i} dx^i, \sum_{i=1}^{p} \frac{\partial f^2}{\partial x^i} dx^i, \ldots, \sum_{i=1}^{p} \frac{\partial f^q}{\partial x^i} dx^i\right).$$

Using results from equation (2.1.4) and (2.1.5), we have

$$(df_x)^{\#}\left(\omega|_{f(x)}\right) = \sum_{\alpha=(i_1,i_2,\ldots,i_k)\in I_{k,q}} (\omega_\alpha \circ f(x)) dy^{i_1}(df_x) \wedge dy^{i_2}(df_x) \wedge \cdots \wedge dy^{i_k}(df_x).$$

Since $dy^j(df_x) = df_x^j$, we get $\left(f^{\#}\omega\right)|_x = (df_x)^{\#}\left(\omega|_{f(x)}\right)$.

PROPOSITION 2.1.3. *The exterior derivative commutes with the pull back operator,*

$$df^{\#} = f^{\#}d.$$

PROOF. We may verify this equality by assuming

$$\omega = \sum_{\alpha \in I_{k,q}} \omega_\alpha(y) dy^\alpha \in \mathcal{E}^k(V),$$

and $f : U \to V$ smooth, with $U \subseteq \mathbb{R}^p$, $V \subseteq \mathbb{R}^q$.

Then,

$$\begin{aligned}
df^{\#}(\omega) &= d\left(\sum_{\alpha=(i_1,i_2,\ldots,i_k)\in I_{k,q}} (\omega_\alpha \circ f) df^{i_1} \wedge df^{i_2} \wedge \cdots \wedge df^{i_k}\right) \\
&= d\left(\sum_{\alpha=(i_1,i_2,\ldots,i_k)\in I_{k,q}} \sum_{n_1,n_2,\ldots,n_k=1}^{p} (\omega_\alpha \circ f)\frac{\partial f^{i_1}}{\partial x^{n_1}}\frac{\partial f^{i_2}}{\partial x^{n_2}} \cdots \frac{\partial f^{i_n}}{\partial x^{n_k}} dx^{n_1} \wedge dx^{n_2} \wedge \cdots \wedge dx^{n_k}\right) \\
&= \sum_{\alpha=(n_1,n_2,\ldots,n_k)\in I_{k,q}} \sum_{n_1,n_2,\ldots,n_k=1}^{p} \sum_{n_0=1}^{p} A \cdot dx^{n_0} \wedge dx^{n_1} \wedge dx^{n_2} \wedge \cdots \wedge dx^{n_k},
\end{aligned}$$

where

$$\begin{aligned}
A &= \frac{\partial}{\partial x^{n_0}}\left((\omega_\alpha \circ f)\frac{\partial f^{i_1}}{\partial x^{n_1}}\frac{\partial f^{i_2}}{\partial x^{n_2}} \cdots \frac{\partial f^{i_n}}{\partial x^{n_k}}\right) \\
&= \sum_{j=1}^{q}\left(\frac{\partial \omega_\alpha}{\partial y^j} \circ f(x) \cdot \frac{\partial f^j}{\partial x^{n_0}}\frac{\partial f^{i_1}}{\partial x^{n_1}} \cdots \frac{\partial f^{i_n}}{\partial x^{n_k}}\right) + \left(\omega_\alpha \circ f(x) \cdot \frac{\partial^2 f^{i_1}}{\partial x^{n_0}\partial x^{n_1}}\frac{\partial f^{i_2}}{\partial x^{n_2}} \cdots \frac{\partial f^{i_n}}{\partial x^{n_k}}\right) \\
&\qquad + \cdots + \left(\omega_\alpha \circ f(x) \cdot \frac{\partial f^{i_1}}{\partial x^{n_1}}\frac{\partial f^{i_2}}{\partial x^{n_2}} \cdots \frac{\partial^2 f^{i_n}}{\partial x^{n_0}\partial x^{n_k}}\right).
\end{aligned}$$

Notice that each $n_0, n_1, n_2, \ldots, n_k$ are from 1 to p, with

$$dx^{n_0} \wedge dx^{n_1} = -dx^{n_1} \wedge dx^{n_0}, \cdots, dx^{n_0} \wedge dx^{n_k} = -dx^{n_k} \wedge dx^{n_0},$$

and

$$\frac{\partial^2 f^{i_1}}{\partial x^{n_0}\partial x^{n_1}} = \frac{\partial^2 f^{i_1}}{\partial x^{n_1}\partial x^{n_0}}, \ \frac{\partial^2 f^{i_2}}{\partial x^{n_0}\partial x^{n_2}} = \frac{\partial^2 f^{i_2}}{\partial x^{n_2}\partial x^{n_0}}, \ \ldots, \frac{\partial^2 f^{i_n}}{\partial x^{n_0}\partial x^{n_k}} = \frac{\partial^2 f^{i_n}}{\partial x^{n_k}\partial x^{n_0}}.$$

These imply

$$df^{\#}(\omega) = \sum_{\alpha=(n_1,n_2,\ldots,n_k)\in I_{k,q}} \sum_{n_1,n_2,\ldots,n_k=1}^{p} \sum_{n_0=1}^{p} B \cdot dx^{n_0} \wedge dx^{n_1} \wedge dx^{n_2} \wedge \cdots \wedge dx^{n_k},$$

with

$$B = \sum_{j=1}^{q}\left(\frac{\partial \omega_\alpha}{\partial y^j} \circ f(x) \cdot \frac{\partial f^j}{\partial x^{n_0}}\frac{\partial f^{i_1}}{\partial x^{n_1}} \cdots \frac{\partial f^{i_n}}{\partial x^{n_k}}\right).$$

On the other side of the equation, we have

$$\begin{aligned} f^{\#}d(\omega) &= f^{\#}\left(\sum_{\alpha=(i_1,i_2,\ldots,i_k)\in I_{k,q}}\sum_{j=1}^{q}\frac{\partial\omega_\alpha}{\partial y^j}dy^j\wedge dy^{i_1}\wedge dy^{i_2}\wedge\cdots\wedge dy^{i_k}\right)\\ &= \sum_{\alpha=(i_1,i_2,\ldots,i_k)\in I_{k,q}}\sum_{j=1}^{q}\frac{\partial\omega_\alpha}{\partial y^j}\circ f(x)df^j\wedge df^{i_1}\wedge df^{i_2}\wedge\cdots\wedge df^{i_k}. \end{aligned}$$

Since

$$df^j=\sum_{n_0=1}^{p}\frac{\partial f^j}{\partial x^{n_0}}dx^{n_0}$$

and

$$df^{i_1}\wedge df^{i_2}\wedge\cdots\wedge df^{i_k}=\sum_{n_1,n_2,\ldots,n_k=1}^{p}\frac{\partial f^{i_1}}{\partial x^{n_1}}\frac{\partial f^{i_2}}{\partial x^{n_2}}\cdots\frac{\partial f^{i_n}}{\partial x^{n_k}}dx^{n_1}\wedge dx^{n_2}\wedge\cdots\wedge dx^{n_k},$$

then

$$\begin{aligned} f^{\#}d(\omega) &= \sum_{\alpha=(n_1,n_2,\ldots,n_k)\in I_{k,q}}\sum_{n_1,n_2,\ldots,n_k=1}^{p}\sum_{n_0=1}^{p}B\cdot dx^{n_0}\wedge dx^{n_1}\wedge dx^{n_2}\wedge\cdots\wedge dx^{n_k}\\ &= df^{\#}(\omega). \end{aligned}$$

$\square$

In the rest of this section, we will turn our page to concepts that are related to currents. We will see some of the operations on currents is related to the corresponding operations on differential forms.

Definition 2.1.4. The space of k-currents $\mathcal{D}_k(U)$ is the dual space of $\mathcal{D}^k(U)$.

Here, a k-current is a continuous linear functional from $\mathcal{D}^k(U)$ or $\mathcal{D}^k(\mathbb{R}^p)$ to $\mathbb{R}$, and when $k=0$, $\mathcal{D}^k(U)=C_c^\infty(U)$, $\mathcal{D}^k(\mathbb{R}^p)=C_c^\infty(\mathbb{R}^p)$. Hence, 0-currents are continuous linear functionals on $C_c^\infty(U)$ or $C_c^\infty(\mathbb{R}^p)$, and in other words, 0-currents are (Schwartz) distributions on these sets.

In general, a k-current ($k\geq 1$) can be identified as a generalized k-dimensional oriented submanifold, such that its $\mathcal{H}^n$ measure is locally finite.

Example 4. *Let $M\subseteq U\subseteq\mathbb{R}^p$, and M is an oriented k-dimensional C^1-submanifold of $\mathbb{R}^p$, with orientation $\xi(x)=\pm\tau_1\wedge\tau_2\wedge\cdots\wedge\tau_k$. Here, $\tau_1,\tau_2,\ldots,\tau_k$ is an orthonormal basis. Using M,* *we can define a k-current,* $[\![M]\!] \in \mathcal{D}_k(U)$*, as*

$$[\![M]\!](\omega) := \int_M \langle \omega(x), \xi(x) \rangle d\mathcal{H}^k, \qquad \forall \omega \in \mathcal{D}^k(U).$$

Definition 2.1.5. For $T \in \mathcal{D}_k(U)$ $(k \geq 1), U \subseteq \mathbb{R}^p$, the boundary of T is $\partial T \in \mathcal{D}_{k-1}(U)$, where

$$\partial T(\omega) := T(d\omega), \quad \omega \in \mathcal{D}^{k-1}(U).$$

Note that $\partial^2 T := \partial(\partial T) = 0$, since $d^2 = 0$ by equation (2.1.7).

Definition 2.1.6. The support of a k-current $T \in \mathcal{D}_k(U), U \subseteq \mathbb{R}^p$ is defined by

$$spt(T) = U \setminus \cup\{V \subseteq U, V \text{ open} \mid spt(\omega) \subseteq V \implies T(\omega) = 0\}.$$

Definition 2.1.7. Given $T \in \mathcal{D}_k(U)$ and an open subset $W \subseteq U$, the mass of T in W is defined by

$$\mathbf{M}_W(T) = \sup\{T(\omega) : \|\omega\| \leq 1, \omega \in \mathcal{D}^k(U), spt(\omega) \subseteq W\},$$

where $\|\omega\|$ denotes the comass norm (2.1.6) of ω.

When $W = U$, we may omit the subscript, W, so that mass of T is written as $\mathbf{M}(T)$. Suppose $T \in \mathcal{D}_k(U)$, and $\mathbf{M}_W(T) < \infty$ for all $W \subset\subset U$. Riesz representation theorem gives that there exists a positive Radon measure μ_T on U, and a μ_T-measurable map $\xi : U \to \bigwedge_k(\mathbb{R}^p)$ with $\|\xi\| = 1$ μ_T-a.e. such that

$$T(\omega) = \int_U \langle \omega(x), \xi(x) \rangle \, d\mu_T(x), \quad \omega \in \mathcal{D}^k(U).$$

Here, μ_T is defined (according to Riesz representation theorem) as

$$\mu_T(W) = \mathbf{M}_W(T) = \sup\{T(\omega) : \|\omega\| \leq 1, \omega \in \mathcal{D}^k(U), spt(\omega) \subseteq W\},$$

which also implies $\mu_T(U) = \mathbf{M}(T)$. For any μ_T-measurable subset $E \subseteq U$, the restriction $T\lfloor_E \in \mathcal{D}_k(U)$ on the set E is defined as

$$(T\lfloor_E)(\omega) = \int_E \langle \omega(x), \xi(x) \rangle d\mu_T(x), \quad \omega \in \mathcal{D}^k(U).$$

In general, for any locally μ_T-integrable function φ on U, the restriction $T\lfloor_\varphi \in \mathcal{D}_k(U)$ is defined as

$$(T\lfloor_\varphi)(\omega) = \int_U \langle \omega, \xi \rangle \varphi d\mu_T.$$

In the following two propositions, we will see the connection between mass of currents and convergence of a sequence of currents.

Definition 2.1.8. Let $T_j, T \in \mathcal{D}_k(U), U \subseteq \mathbb{R}^p$, for $j = 1, 2, \cdots$. We say T_j converges weakly to T if and only if

$$\lim_{j\to\infty} T_j(\omega) = T(\omega), \text{ for all } \omega \in \mathcal{D}^k(U),$$

and denote it as $T_j \rightharpoonup T$.

The following proposition says that $\mathbf{M}_W$ is lower semi-continuous.

Proposition 2.1.9. *Let $T_j, T \in \mathcal{D}_k(U)$, for $j = 1, 2, \cdots$. Suppose T_j converges weakly to T, then for any open subset $W \subseteq U$,*

$$\mathbf{M}_W(T) \leq \liminf_{j\to\infty} \mathbf{M}_W(T_j).$$

Proof. For any $\omega \in \mathcal{D}^k(U)$ with $\|\omega\| \leq 1$ and $spt(\omega) \subseteq W$, by definition of weak convergence,

$$T(\omega) = \lim_{j\to\infty} T_j(\omega).$$

Thus,

$$\mathbf{M}_W(T) = \sup_\omega T(\omega) = \sup_\omega \lim_{j\to\infty} T_j(\omega) \leq \liminf_{j\to\infty} \left(\sup_\omega T_j(\omega) \right) = \liminf_{j\to\infty} \mathbf{M}_W(T_j)$$

$\square$

Proposition 2.1.10. *If $\{T_j\} \subseteq \mathcal{D}_k(U), U \subseteq \mathbb{R}^p$, and $\mathbf{M}_W(T_j) < \infty$ for each $W \subset\subset U$, then there exists a subsequence $\{T_{j_k}\}$, and $T \in \mathcal{D}_k(U)$ such that $T_{j_k} \rightharpoonup T$ as $j_k \to \infty$ in U.*

Proof. Direct application of Banach-Alaoglu Theorem in the space

$$\mathcal{M}_k(W) = \{T \in \mathcal{D}_k(W) : \mathbf{M}_W(T) < \infty\}.$$

$\square$

The following result for currents can be used to show mass minimality for some known currents.

Proposition 2.1.11. *Suppose $T \in \mathcal{D}_k(U), U \subseteq \mathbb{R}^p$, $\mathbf{M}(T) < \infty$, and there exists a k-form $\omega^* = d\varphi$ with $|\omega^*| \leq 1$ and $\mathbf{M}(T) = T(\omega^*)$. Then $\mathbf{M}(T) \leq \mathbf{M}(S)$ for any $S \in \mathcal{D}_k(U)$ with $\partial S = \partial T$.*

Proof. Direct calculation and using definition of boundary of currents, we get that

$$\mathbf{M}(T) = T(\omega^*) = T(d\varphi) = \partial T(\varphi) = \partial S(\varphi) = S(d\varphi) = S(\omega^*) \leq \mathbf{M}(S).$$

□

Since we have defined the push forward and pull back operator in differential forms, we may also define the push forward operator for currents in the following definition.

Definition 2.1.12. Given open sets $U \subseteq \mathbb{R}^p$, $V \subseteq \mathbb{R}^q$, and $f : U \to V$ is a smooth map. The push forward $f_\#$ induced by f is defined as $f_\# : \mathcal{D}_k(U) \to \mathcal{D}_k(V)$, such that

$$(f_\# T)(\omega) = T(f^\#(\omega)), \text{ for any } T \in \mathcal{D}_k(U), \omega \in \mathcal{D}^k(V),$$

whenever $spt(T)$ is compact.

Proposition 2.1.13. *Using notations and conditions as in Definition 2.1.12, then*

$$\partial f_\# T = f_\# \partial T.$$

More precisely,

$$\partial(f_\# T) = f_\#(\partial T).$$

Proof. By Proposition 2.1.3, $df^\# = f^\# d$, then

$$\partial f_\# T(\omega) = f_\# T(d\omega) = T(f^\# d\omega) = T(df^\# \omega) = \partial T(f^\# \omega) = f_\# \partial T(\omega).$$

□

Definition 2.1.14. Let $T \in \mathcal{D}_k(U), U \subseteq \mathbb{R}^p$, T is normal if $spt(T)$ is compact and

$$\mathbf{M}(T) + \mathbf{M}(\partial T) < \infty, \text{ when } k > 0, \text{ or } \mathbf{M}(T) < \infty, \text{ when } k = 0.$$

Before giving the definition for rectifiable currents, we need to first introduce the definition for rectifiable sets and approximate tangent space.

Definition 2.1.15. A set $M \subseteq \mathbb{R}^{n+k}$ is said to be countably n-rectifiable if

$$M \subseteq M_0 \cup \left(\bigcup_{j=1}^{\infty} F_j(\mathbb{R}^n) \right),$$

where $\mathcal{H}^n(M_0) = 0$, and $F_j : \mathbb{R}^n \to \mathbb{R}^{n+k}$ are Lipschitz functions for $j = 1, 2, \ldots$.

Definition 2.1.16. Let M be an $\mathcal{H}^n$-measurable subset of $\mathbb{R}^{n+k}$ with $\mathcal{H}^n(M \cap K) < \infty$ for any compact K. An n-dimensional subspace, $P \subseteq \mathbb{R}^{n+k}$, is called the approximate tangent space for M at $x \in \mathbb{R}^{n+k}$ if

$$\lim_{\lambda \to 0} \int_{\eta_{x,\lambda}(M)} f(y) d\mathcal{H}^n(y) = \int_P f(y) d\mathcal{H}^n(y)$$

for any $f \in C_c^0(\mathbb{R}^{n+k})$, where $\eta_{x,\lambda} : \mathbb{R}^{n+k} \to \mathbb{R}^{n+k}$ is

$$\eta_{x,\lambda}(y) = \frac{1}{\lambda}(y - x),$$

for $x, y \in \mathbb{R}^{n+k}$, $\lambda > 0$.

Theorem 2.1.17. *Suppose M is $\mathcal{H}^n$-measurable with $\mathcal{H}^n(H \cap K) < \infty$ for any compact $K \subseteq \mathbb{R}^{n+k}$. Then M is countably n-rectifiable if and only if the approximate tangent space $T_x M$ exists for $\mathcal{H}^n$-a.e. $x \in M$.*

Definition 2.1.18. Let U be an open set in $\mathbb{R}^{n+k}$. An n-current $T \in \mathcal{D}_n(U)$ is called rectifiable if for each $\omega \in \mathcal{D}^n(U)$,

$$T(\omega) = \int_M \langle \omega(x), \xi(x) \rangle \theta(x) d\mathcal{H}^n(x),$$

where

(1) M is $\mathcal{H}^n$-measurable, countable n-rectifiable subset of U,
(2) θ is a locally $\mathcal{H}^n$-integrable positive function,
(3) $\xi : M \to \bigwedge_n(\mathbb{R}^{n+k})$ is an $\mathcal{H}^n$-measurable function such that for $\mathcal{H}^n$-a.e. $x \in M$,

$$\xi(x) = \tau_1 \wedge \cdots \wedge \tau_n,$$

where $\{\tau_1, \ldots, \tau_n\}$ is an orthonormal basis for the approximate tangent space $T_x M$.

The rectifiable current T will be denoted by $\underline{\underline{\tau}}(M,\theta,\xi)$. Moreover, when θ is integer valued, T is called an integer multiplicity (rectifiable) n-current, and θ is called the multiplicity.

When $T=\underline{\underline{\tau}}(M,\theta,\xi)$ is a rectifiable k-current, its mass equals to

$$\mathbf{M}(T)=\int_M \theta(x)\,d\mathcal{H}^k(x).$$

Finally, we have the important compactness theorem for rectifiable currents.

THEOREM 2.1.19. *Suppose T_h is a sequence of rectifiable n-currents in $\mathbb{R}^{n+k}$ with corresponding density functions θ_h. If for some $R>0$, $\bigcup_h spt(T_h)\subseteq B_R$,*

$$\sup_h\{\mathbf{M}(T_h)+\mathbf{M}(\partial T_h)\}\le R,$$

and

$$\theta_h\ge\frac{1}{R},\qquad \mu_{T_h}-a.e.\ in\ \mathbb{R}^{n+k},$$

then there exists a subsequence T_{h_j} and a rectifiable n-current T, such that $T_{h_j}\rightharpoonup T$. If each T_h is integer multiplicity, then T is also integer multiplicity.

2.2. Ramified and Branched transport

This section is based on [**11**], [**14**], and [**15**]. We will first see the ramified optimal transport problem in the discrete case, where the starting and ending measures are atomic measures. Next, we will characterize ramified optimal transport in the continuous case, and see how these "transport paths" related to the rectifiable currents introduced in the previous section. In the end, we discuss some theoretical results of ramified optimal transportation. We start by introducing the definition of ramified optimal transportation in the discrete settings.

Let X be a convex compact subset in a Euclidean space $\mathbb{R}^d$. For any $x\in X$, let δ_x be the Dirac measure centered at x. An atomic measure in X is in the form of

$$\sum_{i=1}^k m_i\delta_{x_i}$$

with distinct points $x_i\in X$, and $m_i>0$ for each $i=1,\dots,k$. For any $\Lambda>0$, let $\mathcal{A}_\Lambda(X)$ be the space of all atomic measures on X with total mass Λ.

DEFINITION 2.2.1. For any $\Lambda > 0$, and any two atomic measures

$$\mathbf{a} = \sum_{i=1}^{k} m_i \delta_{x_i} \ , \ \mathbf{b} = \sum_{j=1}^{\ell} n_j \delta_{y_j} \in \mathcal{A}_\Lambda(X) \tag{2.2.1}$$

a transport path from $\mathbf{a}$ to $\mathbf{b}$ is a weighted directed graph G consisting of a vertex set $V(G)$, a directed edge set $E(G)$ and a weight function

$$w : E(G) \to (0, +\infty)$$

such that $\{x_1, x_2, \ldots, x_k\} \cup \{y_1, y_2, \ldots, y_\ell\} \subseteq V(G)$ and for any vertex $v \in V(G)$, there is a balance equation:

$$\sum_{\substack{e \in E(G) \\ e^- = v}} w(e) \ = \sum_{\substack{e \in E(G) \\ e^+ = v}} w(e) \ + \begin{cases} m_i & \text{if } v = x_i \text{ for some } i = 1, \ldots, k \\ -n_j & \text{if } v = y_j \text{ for some } j = 1, \ldots, \ell \\ 0 & \text{otherwise} \end{cases}$$

where e^- and e^+ denote the starting and ending point of the edge $e \in E(G)$.

For any real number $\alpha \in [0, 1]$, the $\mathbf{M}_\alpha$ cost of

$$G = \{V(G), E(G), w : E(G) \to (0, \infty)\}$$

is defined by

$$\mathbf{M}_\alpha(G) := \sum_{e \in E(G)} w(e)^\alpha length(e), \tag{2.2.2}$$

where $length(e)$ denotes the Euclidean distance between endpoints e^- and e^+ of e.

For any two atomic measures $\mathbf{a}, \mathbf{b}$ on X of equal mass, let $Path(\mathbf{a}, \mathbf{b})$ be the space of all transport paths from $\mathbf{a}$ to $\mathbf{b}$. The ramified optimal transport problem is: Minimize $\mathbf{M}_\alpha(G)$ among all $G \in Path(\mathbf{a}, \mathbf{b})$.

An $\mathbf{M}_\alpha$ minimizer in $Path(\mathbf{a}, \mathbf{b})$ is called an *α-optimal transport path* from $\mathbf{a}$ to $\mathbf{b}$.

A weighted directed graph $G = \{V(G), E(G), w : E(G) \to (0, \infty)\}$ contains a *cycle* if for $k \geq 3$, there exist vertices $\{v_1, v_2, \ldots, v_k\}$ in $V(G)$ such that for each $i = 1, 2, \ldots, k$, either the segment $[v_i, v_{i+1}]$ or $[v_{i+1}, v_i]$ is a directed edge in $E(G)$, and $v_{k+1} = v_1$.

PROPOSITION 2.2.2. *Given* $\mathbf{a}, \mathbf{b}$ *as in (2.2.1), and let* $G \in Path(\boldsymbol{a}, \boldsymbol{b})$. *Then, there exists* $\tilde{G} \in Path(\boldsymbol{a}, \boldsymbol{b})$ *such that* $V(\tilde{G}) \subseteq V(G)$, $\mathbf{M}_\alpha(\tilde{G}) \leq \mathbf{M}_\alpha(G)$, *and* $\tilde{G}$ *contains no cycle.*

The above result implies that when trying to find optimal transport paths, we may restrict to the set of acyclic transport paths, denoted by

$$Path_0(\mathbf{a}, \mathbf{b}) = \{G \in Path(\mathbf{a}, \mathbf{b}) : G \text{ contains no cycles}\}.$$

The following result gives an upper bound of the number of branching points for acyclic transport paths.

PROPOSITION 2.2.3. *Suppose* $G \in Path_0(\mathbf{a}, \mathbf{b})$, *where* $\mathbf{a}, \mathbf{b}$ *as in (2.2.1), then*

$$|\{v : deg(v) \geq 3\}| \leq k + \ell - 2,$$

where k *and* ℓ *are the cardinality of* $\mathbf{a}, \mathbf{b}$ *respectively, and* $deg(v)$ *denotes the number of edges having an endpoint* v.

LEMMA 2.2.4. *Suppose* $G \in Path_0(\mathbf{a}, \mathbf{b})$, *where* $\mathbf{a}, \mathbf{b}$ *as in (2.2.1), then for any edge* $e \in E(G)$, $0 < w(e) \leq \Lambda$, *and*

$$\frac{\mathbf{M}_\alpha(G)}{\Lambda^\alpha} \geq \frac{\mathbf{M}_1(G)}{\Lambda}.$$

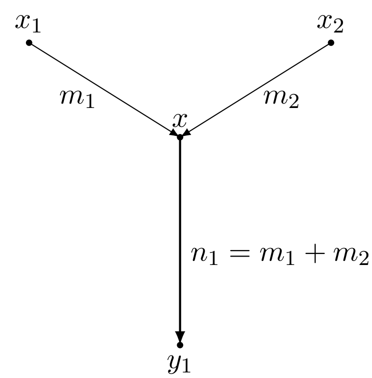
For any atomic measures $\mathbf{a}, \mathbf{b}$ on X of equal mass, define the minimum $\mathbf{M}_\alpha$ cost as

$$d_\alpha(\mathbf{a}, \mathbf{b}) := \min\{\mathbf{M}_\alpha(G) : G \in Path(\mathbf{a}, \mathbf{b})\}. \tag{2.2.3}$$

Based on results from [**11**], d_α is a metric on the space of atomic measures of equal mass, and for each $\lambda > 0$, $d_\alpha(\lambda\mathbf{a}, \lambda\mathbf{b}) = \lambda^\alpha \cdot d_\alpha(\mathbf{a}, \mathbf{b})$.

In the following example, we will illustrate transport paths and transport cost in ramified transportation.

EXAMPLE 5. *Let* $\mathbf{a} = m_1\delta_{x_1} + m_2\delta_{x_2}$, $\mathbf{b} = n_1\delta_{y_1}$, *such that* $m_1 + m_2 = n_1$. *In the non-trivial case, the optimal transport path from* $\mathbf{a}$ *to* $\mathbf{b}$ *with the cost function* $\mathbf{M}_\alpha$ *is a "Y" shaped graph.*



By definition, the total $\mathbf{M}_\alpha$ *cost for the above transport path is*

$$m_1^\alpha |x_1 - x| + m_2^\alpha |x_2 - x| + n_1^\alpha |y_1 - x|.$$

By taking partial derivatives with respect to coordinates and set the partial derivatives equals to 0*, we get*

$$m_1^\alpha \frac{x_1 - x}{|x_1 - x|} + m_2^\alpha \frac{x_2 - x}{|x_2 - x|} + n_1^\alpha \frac{y_1 - x}{|y_1 - x|} = 0.$$

Denote θ_1 *as the angle between*

$$\frac{x_1 - x}{|x_1 - x|} \text{ and } -\frac{y_1 - x}{|y_1 - x|},$$

denote θ_2 *as the angle between*

$$\frac{x_2 - x}{|x_2 - x|} \text{ and } -\frac{y_1 - x}{|y_1 - x|},$$

and let

$$k_1 = \frac{m_1}{m_1 + m_2}, \ k_2 = \frac{m_2}{m_1 + m_2} = 1 - k_1.$$

By using cosine formula, we have

$$\cos\theta_1 = \frac{k_1^{2\alpha} + 1 - k_2^{2\alpha}}{2k_1^\alpha}, \ \cos\theta_2 = \frac{k_2^{2\alpha} + 1 - k_1^{2\alpha}}{2k_2^\alpha}, \ \text{and } \cos(\theta_1 + \theta_2) = \frac{1 - k_1^{2\alpha} - k_2^{2\alpha}}{2k_1^\alpha k_2^\alpha}.$$

When $m_1 = m_2$*, then* $\theta_1 + \theta_2 = \arccos(2^{2\alpha - 1} - 1)$*, and when* $\alpha = 0$*, we have* $\theta_1 = \theta_2 = \pi/3$*.*

Let $\mathcal{M}_\Lambda(X)$ be the space of Radon measures μ on X with total mass $\mu(X) = \Lambda$, and let $\mu^-, \mu^+ \in \mathcal{M}_\Lambda(X)$. Suppose there are two sequences of atomic measures, $\{\mathbf{a}_i\}, \{\mathbf{b}_i\} \in \mathcal{M}_\Lambda(X)$, and $G_i \in Path(\mathbf{a}_i, \mathbf{b}_i)$ such that

$$\mathbf{a}_i \rightharpoonup \mu^-, \mathbf{b}_i \rightharpoonup \mu^+, G_i \rightharpoonup T,$$

then we call T a *transport path* from μ^- to μ^+. The sequence of triples $\{\{\mathbf{a}_i\}, \{\mathbf{b}_i\}, G_i\}$ is called an *approximating graph sequence* for T.

DEFINITION 2.2.5. Let $Path(\mu^-, \mu^+)$ be the space of all transport paths from μ^- to μ^+. For $\alpha \in [0, 1], T \in Path(\mu^-, \mu^+)$, define its $\mathbf{M}_\alpha$ cost as:

$$\mathbf{M}_\alpha(T) := \inf_{\{\mathbf{a}_i, \mathbf{b}_i, G_i\}} \liminf_{i \to \infty} \mathbf{M}_\alpha(G_i),$$

where the infimum is taken over the set of all possible approximating graph sequence $\{\{\mathbf{a}_i\}, \{\mathbf{b}_i\}, G_i\}$ of T.

Here, transport paths in ramified transportation can also be expressed in terms of rectifiable 1-currents. Using Definition 2.1.18, and suppose

$$T = \underline{\underline{\tau}}(M, \theta, \xi).$$

In this case, T can be regarded as a transport path in ramified transportation as follows. The rectifiable set M is equivalent to the set of curves or edges that are in a ramified transport path. The locally integrable function $\theta(x)$ represents the weights that are transported through the set M at the point x. In the discrete case, $\theta(x) = w(e)$ for all $x \in e$, where e is an edge in a ramified transport path. The k-vector valued function $\xi(x)$ is the direction of transportation in a transport path. Hence, we may use rectifiable currents to define transport paths in ramified transportation problems.

Suppose $T \in Path(\mathbf{a}, \mathbf{b})$, $\mathbf{a}, \mathbf{b}$ are atomic measures, and T is a rectifiable 1-currents. Then using boundary of currents (Definition 2.1.5), we may express

$$\partial T = \mathbf{b} - \mathbf{a}. \tag{2.2.4}$$

Since T is a 1-current, ∂T is a 0-current, which is equivalent to a linear functional on $C_c^\infty(\mathbb{R}^n)$. A measure, μ, can also be defined as a linear functional on $C_c^\infty(\mathbb{R}^n)$ as

$$\mu(f) := \int_{\mathbb{R}^n} f d\mu.$$

This implies notations on both side of equation (2.2.4) are well-defined.

In general, given Radon measures μ^-, μ^+ of equal mass, and $T \in Path(\mu^-, \mu^+)$. Let $T = \underline{\underline{\tau}}(M, \theta, \xi)$ be a rectifiable 1-current, with $\partial T = \mu^+ - \mu^-$. For $\alpha \in [0, 1]$, the $\mathbf{M}_\alpha$ cost is

$$\mathbf{M}_\alpha(T) = \int_M \theta(x)^\alpha d\mathcal{H}^1(x).$$

The first important result of this section is the existence result of optimal transport paths, as stated in the following theorem.

THEOREM 2.2.6. *Given Radon measures $\mu^-, \mu^+ \in \mathcal{M}_\Lambda(X)$ on $X \subseteq \mathbb{R}^m$ and $\alpha \in (1 - 1/m, 1]$, there exists an optimal transport path S with least $\mathbf{M}_\alpha$ cost among all transport paths in the family $Path(\mu^-, \mu^+)$. Moreover,*

$$\mathbf{M}_\alpha(S) \leq \frac{\Lambda^\alpha}{2^{1-m(1-\alpha)}} \frac{\sqrt{m}d}{2},$$

where d is the diameter of the convex hull of the supports of μ^- and μ^+.

Using optimal transport paths, we can define a metric on the space of probability measures. For $\alpha \in (1 - 1/m, 1]$, and two Radon measures $\mu^-, \mu^+ \in \mathcal{M}_\Lambda(X)$, we may define

$$d_\alpha(\mu^-, \mu^+) := \min\{\mathbf{M}_\alpha(T) : T \in Path(\mu^-, \mu^+)\}.$$

Note that for any $\Lambda > 0$, and $\mu^-, \mu^+ \in \mathcal{M}_\Lambda(X)$,

$$d_\alpha(\mu^-, \mu^+) = \Lambda^\alpha d_\alpha\left(\frac{\mu^-}{\Lambda}, \frac{\mu^+}{\Lambda}\right).$$

THEOREM 2.2.7. *d_α is a metric on $\mathcal{M}_1(X)$ and metrizes the weak$-*$ topology of $\mathcal{M}_1(X)$. Moreover, the space $(\mathcal{M}_1, d_\alpha)$ is a length space in the sense that for any $\mu^-, \mu^+ \in \mathcal{M}_1(X)$, each $\alpha-$optimal transport path T corresponds to a continuous map*

$$\psi : [0, d_\alpha(\mu^-, \mu^+)] \to \mathcal{M}_1(X)$$

such that $\psi(0) = \mu^-$, $\psi(d_\alpha(\mu^-, \mu^+)) = \mu^+$ and for any $0 \leq s_1 \leq s_2 \leq d_\alpha(\mu^-, \mu^+)$,

$$d_\alpha(\psi(s_1), \psi(s_2)) = s_2 - s_1.$$

We recall the definition of transport plans between two measures from (1.3.1). When both measures are atomic, we have the following characterization.

Suppose $\mathbf{a}$ and $\mathbf{b}$ are two atomic measures on X as in (2.2.1). A *transport plan* from $\mathbf{a}$ to $\mathbf{b}$ is an atomic measure

$$q = \sum_{i=1}^{k} \sum_{j=1}^{\ell} q_{ij} \delta_{(x_i, y_j)}$$

on the product space $X \times X$ such that for each i and j, $q_{ij} \geq 0$,

$$\sum_{i=1}^{k} q_{ij} = n_j, \text{ and } \sum_{j=1}^{\ell} q_{ij} = m_i.$$

We denote $Plan(\mathbf{a}, \mathbf{b})$ as the space of all transport plans from $\mathbf{a}$ to $\mathbf{b}$. We now consider the compatibility between transport plans and transport paths.

DEFINITION 2.2.8. Let $G \in Path(\mathbf{a}, \mathbf{b})$ be a transport path such that for each x_i and y_j there exists at most one directed polyhedral curve g_{ij} from x_i to y_j, and $q \in Plan(\mathbf{a}, \mathbf{b})$ be a transport plan. The pair (G, q) is compatible if $q_{ij} = 0$ whenever $g_{ij} = 0$ and

$$G = \sum_{i=1}^{k} \sum_{j=1}^{\ell} q_{ij} \cdot g_{ij}.$$

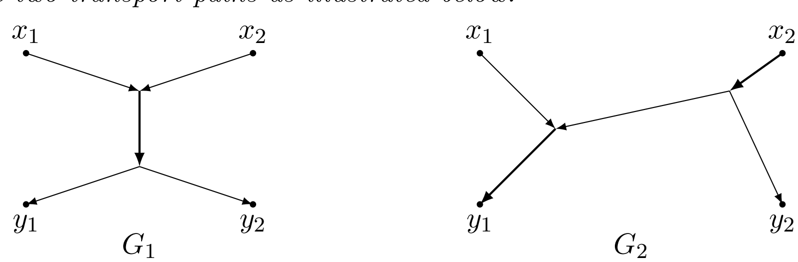
Here, $g_{ij} = 0$ represents no directed polyhedral curve exists, and $q_{ij} \cdot g_{ij}$ represents a mass of q_{ij} is transported along the polyhedral curve g_{ij} from x_i to y_j. Using notation of edges, $e \in E(G)$, we have

$$\sum_{e \subseteq g_{ij}} q_{ij} = w(e).$$

EXAMPLE 6. *Let* $\mathbf{a} = \frac{1}{4}\delta_{x_1} + \frac{3}{4}\delta_{x_2}$, $\mathbf{b} = \frac{5}{8}\delta_{y_1} + \frac{3}{8}\delta_{y_2}$, *and suppose there exists a transport plan as follows:*

$$q = \frac{1}{8}\delta_{(x_1,y_1)} + \frac{1}{8}\delta_{(x_1,y_2)} + \frac{1}{2}\delta_{(x_2,y_1)} + \frac{1}{4}\delta_{(x_2,y_2)} \in Plan(\mathbf{a}, \mathbf{b}).$$

Let G_1 and G_2 be two transport paths as illustrated below:



From the above two transport paths, q is compatible with G_1 but not compatible with G_2, since there is no directed curve from x_1 to y_2 in G_2.

2.3. Applications in ramified transport

Ramified transportation can be applied in various situations, and we will demonstrate some applications in this section.

2.3.1. Application in the formation of a tree leaf.

This subsection is based on content from [**13**] and [**15**]. In this subsection, we will see how ramified transportation is used to simulate the growth of a tree leaf. We can visualize this from the following pictures from [**13**].

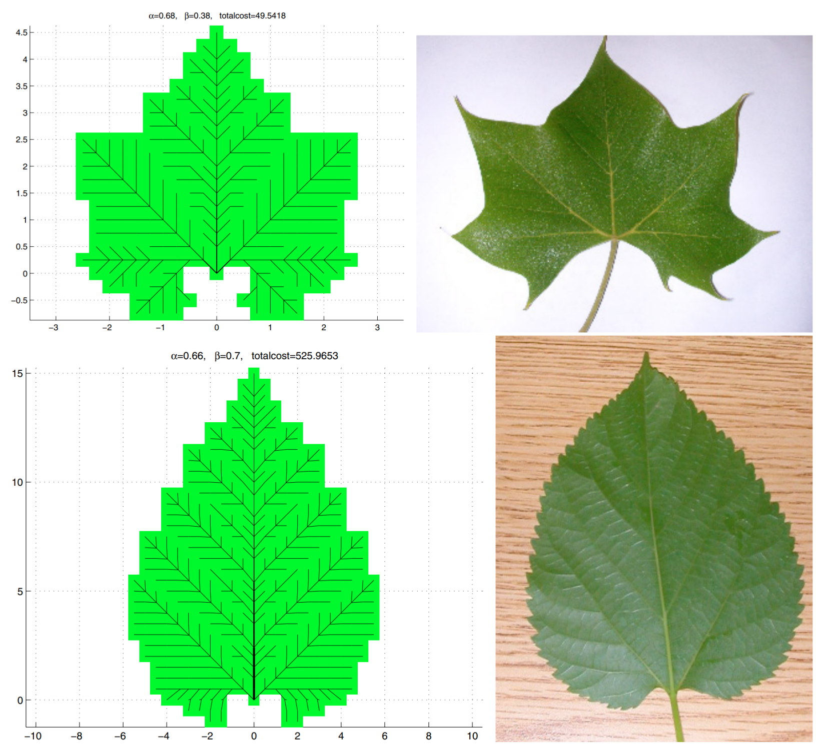


FIGURE 2.1. Formation of a tree leaf.

A leaf is defined as a finite union of squares centered on a given grid. Let $h > 0$, $m, n \in \mathbb{Z}$, and define

$$\Gamma_h = \{(mh, nh) : m, n \in \mathbb{Z}\}$$

as the grid of size h. Let

$$C_{m,n} = \left[mh - \frac{h}{2}, mh + \frac{h}{2}\right) \times \left[nh - \frac{h}{2}, nh + \frac{h}{2}\right)$$

be the cell of size h, and centered at (mh, nh). Let the origin $O = (0, 0) \in \Gamma_h$ be the root of a leaf, and $\vec{e}_O = (0, 1)$ be the initial transport direction of water that coming out of the root, O.

Let $\Omega = \{x_1, x_2, \ldots, x_k\} \subseteq \Gamma_h$ be the positions on the grid, which represents a potential tree leaf. One assumption in the formation of a tree leaf is that each cell need water to survive, and the amount of water needed is proportional to its area, and we may assume it is h^2. This implies a transport system for a leaf can be modeled as a transport path G from the root O to $\sum_{i=1}^{k} h^2 \cdot \delta_{x_i}$.

Definition 2.3.1. A transport system of Ω is a weighted directed graph $G = \{V(G), E(G), \omega\}$ consists of a finite set of vertices, $V(G) \subseteq \Gamma_h$, a directed edge set $E(G)$, and a weight function $\omega : E(G) \to (0, +\infty)$ such that

(1) $\Omega \cup \{O\} \subseteq V(G)$

(2) G is connected and contains no cycle.

(3) The weight function ω satisfies the balance equation

$$\sum_{\substack{e \in E(G) \\ e^+ = v}} \omega(e) = \sum_{\substack{e \in E(G) \\ e^- = v}} \omega(e) + \begin{cases} h^2 & \text{if } v \in \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

at each vertex $v \in V(G) \setminus \{O\}$.

Let $v \in V(G) \setminus \{O\}$, since G is connected and contains no cycle, there exists a unique path from O to v, and we denote it as $P_v := \{v_1, v_2, , \ldots, v_i, v_{i+1}, \ldots, v_k\}$ with $v_1 = O$ and $v_k = v$. We define $p(v_{i+1}) = v_i$, so that $p(v)$ is the "parent" or "previous" vertex of v. The directed edge $(p(v), v) \in E(G)$ is denoted by e_v.

When calculating the total cost of a transport path, shipping in bulk has lower cost than shipping individually, this can be noticed in the original cost function (2.2.2). Also, branching structures in ramified transportation tends to transport items or mass in a given transport direction. This brings the need to consider the cost of rotating a transport direction. For any $\beta > 0$, let $H_\beta : \mathbb{S}^1 \times \mathbb{S}^1 \to (0, \infty]$, such that

$$H_\beta(u, v) = \begin{cases} |u \cdot v|^{-\beta}, & \text{if } u \cdot v > 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Definition 2.3.2. Let $\alpha \in [0,1)$, $\beta > 0$, and $G = \{V(G), E(G), \omega\}$ be a transport system of Ω. Let $m_\beta(O) = 1$, $\vec{e}_O = (0,1)$, and for each $v \in V(G)$, define

$$m_\beta(v) = m_\beta(p(v)) H_\beta(\vec{e}_v, \vec{e}_{p(v)}).$$

The cost of G is defined as :

$$\mathbf{F}(G) := \sum_{e \in E(G)} m_\beta(e^+)\omega(e)^\alpha \text{length}(e) = \sum_{v \in V(G)\setminus\{O\}} m_\beta(v)\omega(e_v)^\alpha \text{length}(e_v).$$

After describing transport paths and transport cost for a prospective leaf, ramified transportation can also describe the dynamic formation of a tree leaf.

Let $P_G : [0,\infty) \times V(G) \to (0,\infty]$, such that

$$P_G(x,v) = \sum_{u \in P_v \setminus \{O\}} m_\beta(u)[(\omega(e_u) + x)^\alpha - \omega(e_u)^\alpha]\text{length}(e_u),$$

for $x \in [0,\infty)$ and $v \in V(G) \setminus \{O\}$, where P_v is the unique path in G from O to v. This function gives the increment of cost when adding weight x to the point v. Let

$$\mathcal{A}_h := \{(\Omega, G) : \Omega \subseteq \Gamma_h, G \text{ is an optimal transport system of } \Omega \text{ under the } \mathbf{F} \text{ cost}\},$$

and for any $(\Omega, G) \in \mathcal{A}_h$, a point $q \in \Omega$ is called a *boundary point* of Ω if at least one of its eight neighboring cells in Γ_h is not in Ω.

For any $x \in \Gamma_h \setminus \Omega \in \mathcal{A}_h$, $b \in B$, the transport cost of x through b is defined as:

$$C_\Omega(x,b) := h^{2\alpha}|x-b| m_\beta(b) H_\beta\left(\frac{x-b}{|x-b|}, \vec{e}_b\right) + P_G(h^2, b),$$

and the cost of adding 1 cell located at x of mass h^2 to the transport system G is

$$C_\Omega(x) := \min_{b \in B} C_\Omega(x,b) = C_\Omega(x, b(x)), \text{ for some } b(x) \in B.$$

Here, we made the assumption that a new cell is generated only if the expense $C_\Omega(x)$ is less than the revenue ϵh^2 it produces. Given $\epsilon > 0$, let

$$\tilde{\Omega} = \{x \in \Gamma_h \setminus \Omega : C_\Omega(x) \le \epsilon h^2\} \cup \Omega, \ \tilde{V} = V(G) \cup \tilde{\Omega}, \ \bar{E} = E(G) \cup \{[x, b(x)] : x \in \tilde{\Omega} \setminus \Omega\}.$$

Let $\tilde{G}$ be the new optimal transport path for the transport system on $\tilde{\Omega}$.

Now, define $L_{\epsilon,h} : \mathcal{A}_h \to \mathcal{A}_h$ as $L_{\epsilon,h}(\Omega, G) = (\tilde{\Omega}, \tilde{G})$, and define

$$\mathcal{A}_{\epsilon,h} := \{(\Omega, G) \in \mathcal{A}_h : \Omega \subseteq B_{R(\epsilon,\alpha)}(O)\},$$

where $R(\epsilon, \alpha)$ is a constant depending on ϵ and α. Then we have the following proposition:

PROPOSITION 2.3.3. *Let* $\alpha \in (1/2, 1)$*, and* $(\Omega, G) \in \mathcal{A}_{\epsilon,h}$*. Suppose* $L_{\epsilon,h}(\Omega, G) = (\tilde{\Omega}, \tilde{G})$*, then* $(\tilde{\Omega}, \tilde{G}) \in \mathcal{A}_{\epsilon,h}$*, and* $\mathbf{F}(\tilde{G}) \leq \mathbf{F}(G) + \epsilon h^2 \|\tilde{\Omega} \setminus \Omega\|$.

Hence, for $\alpha \in (1/2, 1)$, and $(\Omega, G) = (\Omega_0, G_0)$, we can inductively define $(\Omega_n, G_n) = L_{\epsilon,h} : \mathcal{A}_h(\Omega_{n-1}, G_{n-1})$ for $n \geq 1$. Since

$$\Omega_0 \subseteq \Omega_1 \subseteq \Omega_2 \subseteq \cdots \subseteq B_{R(\epsilon,\alpha)}(O) \cap \Gamma_h,$$

and $B_{R(\epsilon,\alpha)}(O) \cap \Gamma_h$ is a finite set, we have Ω_i's converges to some set, i.e. for some N, $i > N$ implies $\Omega_i = \Omega_N$. This shows a tree leaf will not grow forever under the ramified transportation model defined above.

2.3.2. Application in ramified optimal allocation problem.

This subsection in based on [**17**] and [**15**]. In this subsection we will see how ramified transportation can be applied in economics and give an optimal resource allocation plan. In the Monge-Kantorovich and ramified transport problem, a starting measure (source) and an ending measure (target) is given in the first place. However, these two predetermined parameters are not always given when considering allocation problems.

In a product allocation problem, suppose there are k factories and ℓ households located in different regions. Given the demand of ℓ households, the supply of k factories is determined by the demand of households and their relative locations to factories. In this case, we need to generate a production plan among the k factories in this allocation production problem. Therefore, the ultimate goal is finding an optimal production plan and its corresponding optimal transport path, so that the total transport cost from factories to households is minimized.

In this allocation problem, there is only 1 product, and let each household $j = 1, 2, \ldots, \ell$ has demand n_j for this particular product, which can be represented as

$$\mathbf{b} = \sum_{j=1}^{\ell} n_j \delta_{y_j}. \tag{2.3.1}$$

Definition 2.3.4. Let $\mathbf{x} = \{x_1, x_2, \ldots, x_k\}$ be a finite subset of X, which represents locations of factories, and $\mathbf{b}$ be an atomic probability measure that represents demands of households. An allocation plan from $\mathbf{x}$ to $\mathbf{b}$ is a probability measure

$$q = \sum_{i=1}^{k} \sum_{j=1}^{\ell} q_{ij} \delta_{(x_i, y_j)}$$

on $X \times X$ such that $q_{ij} \geq 0$ for each i, j and

$$\sum_{i=1}^{k} q_{ij} = n_j \text{ for each } j = 1, 2, \ldots, \ell.$$

Denote $Plan[\mathbf{x}, \mathbf{b}]$ as the set of all allocation plans from $\mathbf{x}$ to $\mathbf{b}$.

Let $q \in Plan[\mathbf{x}, \mathbf{b}]$, then there exists $\mathbf{a}(q)$ and $\mathbf{b}$ such that $q \in Plan(\mathbf{a}(q), \mathbf{b})$. Here, $\mathbf{a}(q)$ is the probability measure supported on $\mathbf{x}$, which represents the supply of k factories, and

$$\mathbf{a}(q) := \sum_{i=1}^{k} m_i(q) \delta_{x_i}, \text{ and } m_i(q) = \sum_{j=1}^{\ell} q_{ij}, \text{ for } i = 1, 2, \ldots, k.$$

After defining the allocation plan, we may proceed to define its associated cost.

Definition 2.3.5. For any allocation plan $q \in Plan[\mathbf{x}, \mathbf{b}]$ and $\alpha \in [0, 1)$, the ramified transportation cost of q is

$$\mathbf{T}_\alpha(q) := \min\{\mathbf{M}_\alpha(G) : G \in Path(\mathbf{a}(q), \mathbf{b}), (G, q) \text{ is compatible}\}.$$

The $\mathbf{M}_\alpha$ cost is defined as in (2.2.2), and compatibility is defined as in Definition 2.2.8. An allocation plan $q^* \in Plan[\mathbf{x}, \mathbf{b}]$ is optimal if

$$\mathbf{T}_\alpha(q^*) \leq \mathbf{T}_\alpha(q), \text{ for any } q \in Plan[\mathbf{x}, \mathbf{b}].$$

Hence, the ramified optimal allocation problem is to find an optimal allocation plan.

Given any allocation plan q, there exists a transport path $G_q \in Path(\mathbf{a}(q), \mathbf{b})$ such that G_q is compatible with q, and $\mathbf{T}_\alpha(q) = \mathbf{M}_\alpha(G_q)$. This means finding an optimal allocation plan can be transferred into finding an optimal transport path.

Definition 2.3.6. An allocation path from $\mathbf{x}$ to $\mathbf{b}$ is a transport path $G \in Path(\mathbf{a}, \mathbf{b})$ for some atomic probability measure $\mathbf{a}$ supported on $\mathbf{x}$. Denote $Path[\mathbf{x}, \mathbf{b}]$ as the set of all allocation paths from $\mathbf{x}$ to $\mathbf{b}$. An allocation path $G^* \in Path[\mathbf{x},\mathbf{b}]$ is optimal if

$$\mathbf{M}_\alpha(G^*) \leq \mathbf{M}_\alpha(G), \text{ for any } G \in Path[\mathbf{x},\mathbf{b}].$$

Using allocation path we have the following important result:

THEOREM 2.3.7. *Given $G \in Path[\boldsymbol{x},\boldsymbol{b}]$, there exists an allocation path $\tilde{G} \in Path[\boldsymbol{x},\boldsymbol{b}]$, such that*

$$\boldsymbol{M}(\tilde{G}) \leq \boldsymbol{M}(G),$$

and for any $r \neq s \in \{1,2,\ldots,k\}$, x_r and x_s do not belong to the same connected component of $\tilde{G}$.

This theorem gives the information that any two factories belong to 2 disconnected transport paths, and each household will only receive product from 1 factory. Therefore, each $x_i \in \mathbf{x}$ belongs to a connected component $\tilde{G}_i$ of $\tilde{G}$, and

$$\tilde{G} = \sum_{i=1}^{k} \tilde{G}_i, \text{ with } \mathbf{M}_\alpha(\tilde{G}) = \sum_{i=1}^{k} \mathbf{M}_\alpha(\tilde{G}_i).$$

Since each y_j is connected to a unique x_i, this gives a map $S : \{1,2,\ldots,\ell\} \to \{1,2,\ldots,k\}$ such that $S(j) = i$. Figure 2.2 from [**17**] gives an illustration of allocation paths.

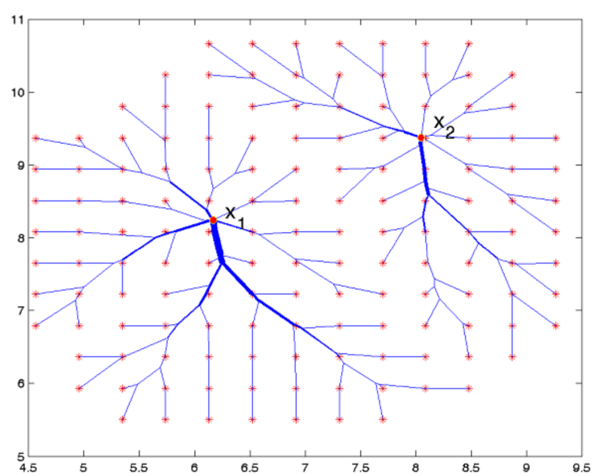


FIGURE 2.2. Allocation path and its connected components.

DEFINITION 2.3.8. An assignment map is a function $S : \{1,2,\ldots,\ell\} \to \{1,2,\ldots,k\}$, and let $\mathrm{Map}[\ell,k]$ be the set of all assignment maps. For any $S \in \mathrm{Map}[\ell,k]$, $\alpha \in [0,1)$, and any given $\mathbf{x}$ and

$\mathbf{b}$, define

$$\mathbf{E}_\alpha(S;\mathbf{x},\mathbf{b}) := \sum_{i=1}^{k} d_\alpha(\mathbf{a}_i,\mathbf{b}_i), \text{ with } \mathbf{a}_i = \left(\sum_{j\in S^{-1}(i)} n_j\right)\delta_{x_i}, \ \mathbf{b}_i = \sum_{j\in S^{-1}(i)} n_j\delta_{y_j},$$

where d_α is defined in equation (2.2.3). An assignment map $S^* \in \mathrm{Map}[\ell,k]$ is optimal if

$$\mathbf{E}_\alpha(S^*;\mathbf{x},\mathbf{b}) \le \mathbf{E}_\alpha(S;\mathbf{x},\mathbf{b}), \text{ for any } S \in \mathrm{Map}[\ell,k].$$

The main result for allocation problem is as follows:

THEOREM 2.3.9. *Given* $\boldsymbol{x} = \{x_1, x_2, \ldots, x_k\}$ *in* X*, an atomic probability measure* $\boldsymbol{b}$ *as in* (2.3.1)*, and* $\alpha \in [0,1)$*.*

(1) *An allocation plan* $q \in Plan[\boldsymbol{x},\boldsymbol{b}]$ *is optimal if and only if there exists an optimal assignment map* $S \in \mathrm{Map}[\ell,k]$ *such that*

$$q = q_S = \sum_{j=1}^{\ell} n_j \delta_{(x_{S(j)},y_j)}$$

(2) *An allocation path* $G \in Path[\boldsymbol{x},\boldsymbol{b}]$ *is optimal if and only if there exists an optimal assignment map* $S \in \mathrm{Map}[\ell,k]$ *such that* $G = G_S$*, where* $G_S = \sum_{i=1}^{k} G_i \in Path[\boldsymbol{x},\boldsymbol{b}]$ *with each* $G_i \in Path(\boldsymbol{a}_i,\boldsymbol{b}_i)$ *being an optimal transport path.*

(3) *Moreover,*

$$\min_{q\in Plan[\boldsymbol{x},\boldsymbol{b}]} \boldsymbol{T}_\alpha(q) = \min_{S\in \mathrm{Map}[\ell,k]} \boldsymbol{E}_\alpha(S;\boldsymbol{x},\boldsymbol{b}) = \min_{G\in Path[\boldsymbol{x},\boldsymbol{b}]} \boldsymbol{M}_\alpha(G).$$

CHAPTER 3

Map-compatible decomposition of transport paths in discrete case

3.1. Introduction

This chapter is based on the paper [**16**]. In the well-known Monge-Kantorovich transport problem (see Chapter 1), the transport cost is expressed in terms of transport maps or transport plans. The existence of optimal transport maps, especially the Brenier map in the case of quadratic cost, leads to numerous applications of optimal transportation theory in PDEs, Probability theory, Machine learning, etc. A variant of the Monge-Kantorovich transport problem is ramified optimal transportation (see Chapter 2). Through the lens of economy of scales, ramified optimal transportation aims at studying the branching structures that appeared in many living or non-living transport systems. In contrast to the classical Monge-Kantorovich transport problems, where the transport cost relies on transport maps and plans, the transport cost in the ramified transport problem is assessed across the entire branching transport system, referred to as transport paths.

Since transport maps/plans only utilize information from the initial/target measures, knowing only transport maps/plans is insufficient for describing the transport cost that appears in ramified optimal transportation problem. In general, two transport paths (e.g. a "Y-shaped" and a "V-shaped" path) may have different transportation costs while sharing the same transport map/plan. Nevertheless, motivated by the significance of transport maps in the context of the Monge-Kantorovich problem, when a transport path is given, one may wonder if there exists a hidden transport map or plan that is compatible with this specific transport path. This compatible transport map/plan tells one how the initial measure is distributed to the target measure via the given transport path. For simplicity, we will only considers the case of atomic measures, deferring the exploration of other scenarios for future endeavors. We want to provide a decomposition of transport paths such that each component in the decomposition is compatible with some transport map or transport plan.

Roughly speaking, main results of this chapter are :

- **Theorem 3.4.8:** Every cycle-free [1] transport path T can be decomposed as a sum of subcurrents $T = T_0 + T_1 + \cdots + T_N$ such that each $T_1, T_2, \cdots, T_N$ has a single target and T_0 has at most $\binom{N}{2}$ sources[2].
- **Theorem 3.5.6:** Every cycle-free transport path T can be decomposed as a sum of subcurrents $T = T_\varphi + T_\pi$ such that T_φ is compatible with some transport map φ and T_π is compatible with some transport plan π.
- **Theorem 3.6.8:** Every stair-shaped transport path T can be decomposed as a sum of subcurrents $T = T_1 + T_2$ such that both T_1 and $-T_2$ are compatible with some transport maps.

In Section 3.2, we recall some related concepts in geometric measure theory, the classical Monge-Kantorovich transport problem, and the ramified optimal transport problem. In particular, the *good decomposition* (i.e., Smirnov decomposition) of acyclic normal 1-currents.

In general, the family of atoms (i.e., supporting curves) of a good decomposition is not necessarily linearly independent. This fact brings a non-unique representation of vanishing currents and causes a technical obstacle for the proof of Theorem 3.4.8. To overcome this, we generalize the notion of "good decomposition" to "better decomposition" (Definition 3.3.1) of transport paths in Section 3.3. A better decomposition η of a transport path T prohibits combinations of any four supporting curves of η to form a non-trivial cycle on the support of T. We showed in Theorem 3.3.3 that any good decomposition of a transport path has a better decomposition that is absolutely continuous with respect to the original good decomposition.

In Section 3.4, we introduce the concept of cycle-free transport paths, which are transport paths with no non-trivial cycles on[3] them. Then, we use the "better decomposition" achieved in Theorem 3.3.3 to give a decomposition of cycle-free transport paths, described in Theorem 3.4.8.

In Section 3.5, we consider the concept of "compatibility" between transport paths and transport plans/maps. This concept was first introduced in [**11**, Definition 7.1] for cycle-free transport paths to describe whether a given transport plan is practically possible for transportation along the

[1]A transport path T is called cycle-free if there are no nonzero cycles on T. See Definition 3.4.2.

[2]Here, N is the number of targets in the target measure μ^+.

[3]The concept cycle-free is different to the concept "acyclic" defined using subcurrents. As in Definition 3.4.1, a current S is "on" another current T does not mean that S is a subcurrent of T. When S is on T, unlike being a subcurrent, it is possible that S has a reverse orientation with T on their intersections.

given transport path. We first generalize this concept, in a more general setting, to the compatibility between transport paths and transport plans/maps. Then, using Theorem 3.4.8, we decompose a cycle-free transport path into the sum of a map-compatible path and a plan-compatible path, which gives Theorem 3.5.6.

In Section 3.6, we proceed to study stair-shaped transport paths. We first show in Theorem 3.6.4 that each matrix[4] with non-negative entries can be transformed into a stair-shaped matrix, and in Algorithm 3.6.5, we provide an algorithm for calculating the stair-shaped matrix. A transport path is called stair-shaped if it has a good decomposition that is represented by a stair-shaped matrix. A stair-shaped transport path is not necessarily cycle-free, but it still has a better decomposition. Our main result for the section is Theorem 3.6.8, which says that any stair-shaped transport path can be decomposed into the difference of two map-compatible transport paths. Note that some cycle-free transport paths are also stair-shaped. They can be decomposed not only as the sum of a map-compatible path and a plan-compatible path by Theorem 3.5.6, but also as the sum of two map-compatible transport paths by Theorem 3.6.8. We further investigate some sufficient conditions under which cycle-free transport paths are stair-shaped. An illustrating example is provided at the end.

3.2. Preliminaries

3.2.1. Basic concepts in geometric measure theory.

Using notations and definitions from Section 2.1, we recall some other concepts in literature that are particularly related to this Chapter.

We first recall the concept of subcurrents, which was introduced by Paolini and Stepanov in [**6**]. For any $T, S \in \mathcal{D}_k(U)$, S is called a *subcurrent* of T if

$$\mathbf{M}(T - S) + \mathbf{M}(S) = \mathbf{M}(T).$$

A normal current $T \in \mathcal{D}_k(\mathbb{R}^m)$ is *acyclic* if there is no non-trivial subcurrent S of T such that $\partial S = 0$.

Also, in [**9**], Smirnov showed that every acyclic normal 1-current can be written as the weighted average of simple Lipschitz curves in the following sense. Let Γ be the space of 1-Lipschitz curves

[4] The size of this matrix may be countably infinite.

$\gamma : [0,\infty) \to \mathbb{R}^m$, which are eventually constant. For $\gamma \in \Gamma$, we denote

$$t_0(\gamma) := \sup\{t : \gamma \text{ is constant on } [0,t]\},\ t_\infty(\gamma) := \inf\{t : \gamma \text{ is constant on } [t,\infty)\},$$

and $p_0(\gamma) := \gamma(0)$, $p_\infty(\gamma) := \gamma(\infty) = \lim_{t\to\infty}\gamma(t)$. A curve $\gamma \in \Gamma$ is simple if $\gamma(s) \neq \gamma(t)$ for every $t_0(\gamma) \le s < t \le t_\infty(\gamma)$. For each simple curve $\gamma \in \Gamma$, we may associate it with the following rectifiable 1-current,

$$I_\gamma := \underline{\underline{\tau}}\left(\mathrm{Im}(\gamma), \frac{\gamma'}{|\gamma'|}, 1\right), \tag{3.2.1}$$

where $Im(\gamma)$ denotes the image of γ in $\mathbb{R}^m$.

Definition 3.2.1. Let T be a normal 1-current in $\mathbb{R}^m$ and let η be a finite positive measure on Γ such that

$$T = \int_\Gamma I_\gamma \, d\eta(\gamma) \tag{3.2.2}$$

in the sense that for every smooth compactly supported 1-form $\omega \in \mathcal{D}^1(\mathbb{R}^m)$, it holds that

$$T(\omega) = \int_\Gamma I_\gamma(\omega)\, d\eta(\gamma). \tag{3.2.3}$$

We say that η is a *good decomposition* of T (see [**2**], [**3**], [**9**]) if η is supported on non-constant, simple curves and satisfies the following equalities:

(a) $\mathbf{M}(T) = \int_\Gamma \mathbf{M}(I_\gamma) d\eta(\gamma) = \int_\Gamma \mathcal{H}^1(Im(\gamma)) d\eta(\gamma)$;

(b) $\mathbf{M}(\partial T) = \int_\Gamma \mathbf{M}(\partial I_\gamma) d\eta(\gamma) = 2\eta(\Gamma)$.

Moreover, if η is a good decomposition of T, the following statements hold [**2**, Proposition 3.6]:

- $$\mu^- = \int_\Gamma \delta_{\gamma(0)}\, d\eta(\gamma),\ \mu^+ = \int_\Gamma \delta_{\gamma(\infty)}\, d\eta(\gamma). \tag{3.2.4}$$

- If $T = \underline{\underline{\tau}}(M,\theta,\xi)$ is rectifiable, then

$$\theta(x) = \eta(\{\gamma \in \Gamma : x \in \mathrm{Im}(\gamma)\}) \tag{3.2.5}$$

for $\mathcal{H}^1$-a.e. $x \in M$.

- For every $\tilde{\eta} \leq \eta$, the representation

$$\tilde{T} = \int_\Gamma I_\gamma d\tilde{\eta}(\gamma)$$

is a good decomposition of $\tilde{T}$. Moreover, if $T = \underline{\underline{\tau}}(M, \theta, \xi)$ is rectifiable, then $\tilde{T}$ can be written as $\tilde{T} = \underline{\underline{\tau}}(M, \tilde{\theta}, \xi)$ with

$$\tilde{\theta}(x) \leq \min\{\theta(x), \tilde{\eta}(\Gamma)\} \tag{3.2.6}$$

for $\mathcal{H}^1$-a.e. $x \in M$.

In the following contexts, we adopt the notations: for any points $x, y \in \mathbb{R}^m$ and subset $A \subseteq \mathbb{R}^m$, denote

$$\Gamma_x = \{\gamma \in \Gamma : x \in \mathrm{Im}(\gamma)\}, \tag{3.2.7}$$

$$\Gamma_{x,y} = \{\gamma \in \Gamma : p_0(\gamma) = x,\ p_\infty(\gamma) = y\}, \tag{3.2.8}$$

$$\Gamma_{A,y} = \{\gamma \in \Gamma : p_0(\gamma) \in A,\ p_\infty(\gamma) = y\}. \tag{3.2.9}$$

3.2.2. Basic concepts in optimal transportation theory.

In the following results, we will focus on transportation between atomic measures. Let

$$\mu^- = \sum_{i=1}^M m_i' \delta_{x_i} \text{ and } \mu^+ = \sum_{j=1}^N m_j \delta_{y_j} \text{ with } \sum_{i=1}^M m_i' = \sum_{j=1}^N m_j < \infty \tag{3.2.10}$$

be two finite atomic measures on X of equal mass with $M, N \in \mathbb{N} \cup \{\infty\}$. In this case, the concepts of Monge-Kantorovich transport problem in Chapter 1 and the concepts of Ramified transport problem in Chapter 2 have simplified forms:

- A transport map $\varphi \in Map(\mu^-, \mu^+)$ corresponds to a map

$$\varphi : \{1, 2, \cdots, M\} \to \{1, 2, \cdots, N\}$$

such that for each $j = 1, 2, \cdots, N$,

$$m_j = \sum_{i \in \varphi^{-1}(\{j\})} m_i'.$$

The corresponding transport cost is

$$I_C(\varphi) = \sum_{i=1}^{M} C(x_i, y_{\varphi(i)}) m_i'.$$

- A transport plan $\pi \in Map(\mu^-, \mu^+)$ corresponds to an $M \times N$ matrix $\pi = [\pi_{ij}]$ such that for each i, j, it holds that

$$\sum_i \pi_{ij} = m_j \text{ and } \sum_j \pi_{ij} = m_i'.$$

  The corresponding transport cost is

$$J_C(\pi) = \sum_{i=1}^{M} \sum_{j=1}^{N} c_{ij} \pi_{ij}$$

  where $c_{ij} = C(x_i, y_j)$.

- A transport path $T \in Path(\mu^-, \mu^+)$ corresponds to a weighted directed graph T consisting of a vertex set V, a directed edge set E and a weight function $w : E \to (0, +\infty)$ such that $\{x_1, x_2, \ldots, x_M\} \cup \{y_1, y_2, \ldots, y_N\} \subseteq V$ and for any vertex $v \in V$, there is a balance equation:

$$\sum_{e \in E, e^- = v} w(e) \;=\; \sum_{e \in E, e^+ = v} w(e) \;+\; \begin{cases} m_i & \text{if } v = x_i \text{ for some } i = 1, \ldots, M \\ -n_j & \text{if } v = y_j \text{ for some } j = 1, \ldots, N \\ 0 & \text{otherwise,} \end{cases}$$

  where e^- and e^+ denote the starting and ending point of the edge $e \in E$. The corresponding transport $\mathbf{M}_\alpha$-cost of T is

$$\mathbf{M}_\alpha(T) = \sum_{e \in E} w(e)^\alpha length(e)$$

  where the length $length(e)$ of the edge e equals to $\mathcal{H}^1(e)$.

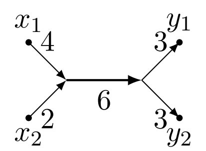
3.3. Better decomposition of acyclic transport paths

Let μ^- and μ^+ be two atomic measures as given in (3.2.10), T be an acyclic transport path from μ^- to μ^+, and let η be a good decomposition (i.e., Smirnov decomposition) of T. Observe that as shown in the following example, with respect to the good decomposition η, it is possible that the family

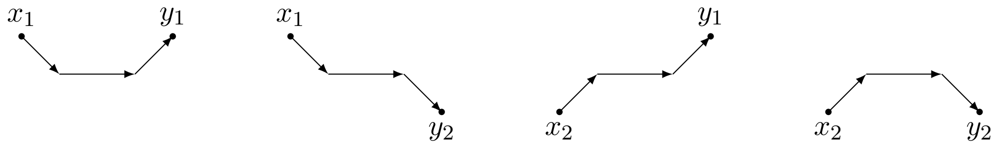
$$\{I_\gamma : \eta(\{\gamma\}) > 0\}$$

is linearly dependent.

EXAMPLE 7. *Let T be a transport path from $\mu^- = 4\delta_{x_1} + 2\delta_{x_2}$ to $\mu^+ = 3\delta_{y_1} + 3\delta_{y_2}$, as shown in the following figure*

$T =$ 

For each (i, j), let γ_{x_i,y_j} be the corresponding curve from x_i to y_j on T:



γ_{x_1,y_1} γ_{x_1,y_2} γ_{x_2,y_1} γ_{x_2,y_2}

Then

$$\eta = 2\delta_{\gamma_{x_1,y_1}} + 2\delta_{\gamma_{x_1,y_2}} + \delta_{\gamma_{x_2,y_1}} + \delta_{\gamma_{x_2,y_2}}$$

is a good decomposition of T. But

$$I_{\gamma_{x_1,y_1}} - I_{\gamma_{x_1,y_2}} - I_{\gamma_{x_2,y_1}} + I_{\gamma_{x_2,y_2}}$$

is the zero 1-current.

The linear dependence of the family $\{I_\gamma : \eta(\{\gamma\}) > 0\}$ brings a non-unique representation of vanishing currents and causes an obstacle later for the proof of Theorem 3.4.8. To overcome this, we introduce the concept of "better decomposition" of T as follows.

For each $i = 1, 2, \cdots, M$, $j = 1, 2, \cdots, N$, as given in (3.2.8), let Γ_{x_i,y_j} denote all 1-Lipschitz curves in Γ from x_i to y_j. Also, for any finite positive measure η on Γ, denote

$$S_{i,j}(\eta) := \begin{cases} \frac{1}{\eta(\Gamma_{x_i,y_j})} \int_{\Gamma_{x_i,y_j}} I_\gamma d\eta, & \text{if } \eta(\Gamma_{x_i,y_j}) > 0 \\ 0, & \text{if } \eta(\Gamma_{x_i,y_j}) = 0. \end{cases} \tag{3.3.1}$$

DEFINITION 3.3.1. Let T be a transport path from μ^- to μ^+ where μ^- and μ^+ are given in (3.2.10). Suppose η is a good decomposition of T. We say that η is a *better decomposition* of T if for any pairs $1 \leq i_1 < i_2 \leq M$ and $1 \leq j_1 < j_2 \leq N$,

$$S_{i_1,j_1}(\eta) - S_{i_1,j_2}(\eta) - S_{i_2,j_1}(\eta) + S_{i_2,j_2}(\eta) = 0$$

implies that

$$\eta(\Gamma_{x_{i_1},y_{j_1}}) = \eta(\Gamma_{x_{i_1},y_{j_2}}) = \eta(\Gamma_{x_{i_2},y_{j_1}}) = \eta(\Gamma_{x_{i_2},y_{j_2}}) = 0.$$

EXAMPLE 8. *In Example 7,*

$$\eta = 2\delta_{\gamma_{x_1,y_1}} + 2\delta_{\gamma_{x_1,y_2}} + \delta_{\gamma_{x_2,y_1}} + \delta_{\gamma_{x_2,y_2}}$$

is a good but not better decomposition of T. Indeed,

$$S_{1,1}(\eta) - S_{1,2}(\eta) - S_{2,1}(\eta) + S_{2,2}(\eta) = I_{\gamma_{x_1,y_1}} - I_{\gamma_{x_1,y_2}} - I_{\gamma_{x_2,y_1}} + I_{\gamma_{x_2,y_2}} = 0,$$

but

$$\eta(\Gamma_{x_1,y_1}) = 2, \eta(\Gamma_{x_1,y_2}) = 2, \eta(\Gamma_{x_2,y_1}) = 1, \text{ and } \eta(\Gamma_{x_2,y_2}) = 1.$$

To realize T using η, all four transportation need to be used.

On the other hand,

$$\tilde{\eta} = 3\delta_{\gamma_{x_1,y_1}} + \delta_{\gamma_{x_1,y_2}} + 2\delta_{\gamma_{x_2,y_2}}$$

is a better decomposition of T. In this case,

$$S_{1,1}(\tilde{\eta}) - S_{1,2}(\tilde{\eta}) - S_{2,1}(\tilde{\eta}) + S_{2,2}(\tilde{\eta}) = I_{\gamma_{x_1,y_1}} - I_{\gamma_{x_1,y_2}} + I_{\gamma_{x_2,y_2}} \neq 0$$

despite that

$$\tilde{\eta}(\Gamma_{x_1,y_1}) = 3, \tilde{\eta}(\Gamma_{x_1,y_2}) = 1, \tilde{\eta}(\Gamma_{x_2,y_1}) = 0, \tilde{\eta}(\Gamma_{x_2,y_2}) = 2.$$

Using this new decomposition, to realize the same T, one only needs to arrange three transportation.

DEFINITION 3.3.2. For any two finite measures η and $\tilde{\eta}$ on Γ, we say $\tilde{\eta} \prec\!\!\prec \eta$ if for each pair (i,j),

$$\int_{\Gamma_{x_i,y_j}} I_\gamma d\tilde{\eta} = a_{i,j} \int_{\Gamma_{x_i,y_j}} I_\gamma d\eta \tag{3.3.2}$$

for some $a_{i,j} \geq 0$.

Our main result for this section is the following theorem:

Theorem 3.3.3. *Let T be a transport path from μ^- to μ^+ where μ^- and μ^+ are given in (3.2.10). For any good decomposition η of T, there exists a better decomposition η_∞ of T such that $\eta_\infty \prec\!\!\prec \eta$.*

We first give an equivalent definition of $\tilde{\eta} \prec\!\!\prec \eta$ as follows.

Lemma 3.3.4. *For any two finite measures η and $\tilde{\eta}$ on Γ, $\tilde{\eta} \prec\!\!\prec \eta$ if and only if they satisfy the condition*

$$\text{if } \tilde{\eta}(\Gamma_{x_i,y_j}) > 0 \text{ for some } (i,j), \text{ then } \eta(\Gamma_{x_i,y_j}) > 0 \text{ and } S_{i,j}(\tilde{\eta}) = S_{i,j}(\eta). \tag{3.3.3}$$

Remark 3.3.5. By Lemma 3.3.4, it follows that $\tilde{\eta}(\Gamma_{x_i,y_j}) = 0$ whenever $\eta(\Gamma_{x_i,y_j}) = 0$. We use the notation $\tilde{\eta} \prec\!\!\prec \eta$ to mimic the absolute continuity notation $\ll$ of measures.

Proof. Suppose $\tilde{\eta} \prec\!\!\prec \eta$. By taking the boundary operator on both sides of (3.3.2), it follows that

$$\int_{\Gamma_{x_i,y_j}} (\delta_{y_j} - \delta_{x_i}) d\tilde{\eta} = a_{i,j} \int_{\Gamma_{x_i,y_j}} (\delta_{y_j} - \delta_{x_i}) d\eta.$$

That is,

$$\tilde{\eta}(\Gamma_{x_i,y_j})(\delta_{y_j} - \delta_{x_i}) = a_{i,j}\eta(\Gamma_{x_i,y_j})(\delta_{y_j} - \delta_{x_i}),$$

which implies that $\tilde{\eta}(\Gamma_{x_i,y_j}) = a_{i,j}\eta(\Gamma_{x_i,y_j})$. Thus, $\tilde{\eta}(\Gamma_{x_i,y_j}) > 0$ implies $a_{ij} > 0$ and $\eta(\Gamma_{x_i,y_j}) > 0$. Moreover,

$$S_{i,j}(\tilde{\eta}) = \frac{1}{\tilde{\eta}(\Gamma_{x_i,y_j})} \int_{\Gamma_{x_i,y_j}} I_\gamma d\tilde{\eta} = \frac{1}{a_{i,j}\eta(\Gamma_{x_i,y_j})} \cdot a_{i,j} \int_{\Gamma_{x_i,y_j}} I_\gamma d\eta = S_{i,j}(\eta).$$

On the other hand, suppose (3.3.3) holds. If $\tilde{\eta}(\Gamma_{x_i,y_j}) = 0$, then $a_{i,j} = 0$ will give (3.3.2). If $\tilde{\eta}(\Gamma_{x_i,y_j}) > 0$, then (3.3.3) implies $\eta(\Gamma_{x_i,y_j}) > 0$ and $S_{i,j}(\tilde{\eta}) = S_{i,j}(\eta)$. By setting

$$a_{i,j} = \frac{\tilde{\eta}(\Gamma_{x_i,y_j})}{\eta(\Gamma_{x_i,y_j})},$$

equation (3.3.1) gives that

$$\int_{\Gamma_{x_i,y_j}} I_\gamma d\tilde{\eta} = \tilde{\eta}(\Gamma_{x_i,y_j}) S_{i,j}(\tilde{\eta}) = (a_{i,j}\eta(\Gamma_{x_i,y_j})) S_{i,j}(\eta) = a_{i,j} \int_{\Gamma_{x_i,y_j}} I_\gamma d\eta.$$

$\square$

Note that, by using the sign function

$$(3.3.4) \qquad sgn(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x = 0 \\ -1, & \text{if } x < 0, \end{cases}$$

equation (3.3.1) gives

$$(3.3.5) \qquad \partial S_{i,j}(\eta) = \begin{cases} \delta_{y_j} - \delta_{x_i}, & \text{if } \eta(\Gamma_{x_i,y_j}) > 0, \\ 0, & \text{if } \eta(\Gamma_{x_i,y_j}) = 0 \end{cases} = sgn(\eta(\Gamma_{x_i,y_j}))(\delta_{y_j} - \delta_{x_i}).$$

For any pairs $1 \leq i_1 < i_2 \leq M$ and $1 \leq j_1 < j_2 \leq N$, define

$$(3.3.6) \qquad C[(i_1, j_1), (i_2, j_2), \eta] := S_{i_1,j_1}(\eta) - S_{i_1,j_2}(\eta) - S_{i_2,j_1}(\eta) + S_{i_2,j_2}(\eta).$$

Direct calculation gives

$$\begin{aligned} \partial C[(i_1, j_1), (i_2, j_2), \eta] \quad = \quad & \left(sgn(\eta(\Gamma_{x_{i_1},y_{j_2}}) - sgn(\eta(\Gamma_{x_{i_1},y_{j_1}})\right) \delta_{x_{i_1}} \\ & + \left(sgn(\eta(\Gamma_{x_{i_2},y_{j_1}}) - sgn(\eta(\Gamma_{x_{i_2},y_{j_2}})\right) \delta_{x_{i_2}} \\ & + \left(sgn(\eta(\Gamma_{x_{i_1},y_{j_1}}) - sgn(\eta(\Gamma_{x_{i_2},y_{j_1}})\right) \delta_{y_{j_1}} \\ & + \left(sgn(\eta(\Gamma_{x_{i_2},y_{j_2}}) - sgn(\eta(\Gamma_{x_{i_1},y_{j_2}})\right) \delta_{y_{j_2}}. \end{aligned}$$

Hence, it follows that $\partial C[(i_1, j_1), (i_2, j_2), \eta] = 0$ if and only if

$$(3.3.7) \qquad sgn(\eta(\Gamma_{x_{i_1},y_{j_1}})) = sgn(\eta(\Gamma_{x_{i_1},y_{j_2}})) = sgn(\eta(\Gamma_{x_{i_2},y_{j_1}})) = sgn(\eta(\Gamma_{x_{i_2},y_{j_2}})) = c,$$

where $c = 0$ or 1. We denote this common value, c, by $s[(i_1, j_1), (i_2, j_2), \eta]$.

Definition 3.3.6. For any finite positive measure η on Γ, define

$$\mathcal{A}_\eta(i^*, j^*) = \{(i, j) : i^* < i \leq M, j^* < j \leq N,\ C[(i^*, j^*), (i, j), \eta] = 0 \text{ and } s[(i^*, j^*), (i, j), \eta] = 1\}.$$

Using this definition, saying a good decomposition η of T is a better decomposition of T is equivalent to $\mathcal{A}_\eta(i, j) = \emptyset$ for all pairs (i, j).

We now consider the graded lexicographical order on $\mathbb{N}^2$, namely

$$(a, b) < (c, d) \text{ if } a + b < c + d \text{ or } a = c \text{ but } b < d.$$

Under this order, $\mathbb{N}^2$ is listed in the order of

$$\{(i_n, j_n)\}_{n=1}^{\infty} = \{(1, 1), (1, 2), (2, 1), (1, 3), \ldots, (i_n, j_n), (i_{n+1}, j_{n+1}), \ldots\}. \tag{3.3.8}$$

Lemma 3.3.7. *For any good decomposition η of T, there exists a good decomposition $\tilde{\eta}$ of T such that $\tilde{\eta} \prec\!\!\prec \eta$ and $\mathcal{A}_{\tilde{\eta}}(1, 1) = \emptyset$.*

Proof. When $\eta(\Gamma_{x_1,y_1}) = 0$, by (3.3.7), the condition $C[(1, 1), (i, j), \eta] = 0$ implies

$$s[(1, 1), (i, j), \eta] = 0,$$

and hence $\mathcal{A}_\eta(1, 1) = \emptyset$. Setting $\tilde{\eta} := \eta$ gives us the desired results.

When $\eta(\Gamma_{x_1,y_1}) \neq 0$, we inductively define a sequence of good decomposition $\{\eta_n\}$ of T with $\eta_n(\Gamma_{x_1,y_1}) > 0$, and whose limit is our desired measure $\tilde{\eta}$. Set $\eta_1 = \eta$.

If $\mathcal{A}_{\eta_n}(1, 1) = \emptyset$ for some $n \geq 1$, set $\eta_m = \eta_n$ for all $m \geq n$ and set $\tilde{\eta} = \eta_n$ as well.

If $\mathcal{A}_{\eta_n}(1, 1)$ is non-empty for all $n \geq 1$, we construct $\tilde{\eta}$ from $\{\eta_n\}$ via the following steps.

Step 1: Construct a sequence of good decomposition $\{\eta_n\}$ of T.

For each $n \geq 1$, assume that η_n is a good decomposition of T with $\eta_n(\Gamma_{x_1,y_1}) > 0$. Let (i_n, j_n) be the minimum element in $\mathcal{A}_{\eta_n}(1, 1)$ which is a subset of $\mathbb{N}^2$ with the graded lexicographical order. Define

$$\eta_{n+1} := \eta_n + \min\{\eta_n(\Gamma_{x_1,y_{j_n}}), \eta_n(\Gamma_{x_{i_n},y_1})\} \left(\frac{\eta_n \lfloor_{\Gamma_{x_1,y_1}}}{\eta_n(\Gamma_{x_1,y_1})} - \frac{\eta_n \lfloor_{\Gamma_{x_1,y_{j_n}}}}{\eta_n(\Gamma_{x_1,y_{j_n}})} - \frac{\eta_n \lfloor_{\Gamma_{x_{i_n},y_1}}}{\eta_n(\Gamma_{x_{i_n},y_1})} + \frac{\eta_n \lfloor_{\Gamma_{x_{i_n},y_{j_n}}}}{\eta_n(\Gamma_{x_{i_n},y_{j_n}})} \right)$$

Here, the denominators in the above equation are positive because $s[(1, 1), (i_n, j_n), \eta_n] = 1$. Without loss of generality, we may assume that

$$0 < \eta_n(\Gamma_{x_1,y_{j_n}}) \leq \eta_n(\Gamma_{x_{i_n},y_1}).$$

Under this construction, we have for each i, j,

$$\eta_{n+1} \lfloor_{\Gamma_{x_i,y_j}} = (1 + \lambda_{n,i,j}) \eta_n \lfloor_{\Gamma_{x_i,y_j}} \tag{3.3.9}$$

for some real number $\lambda_{n,i,j} \geq -1$. In particular, it follows that

$$\eta_{n+1}(\Gamma_{x_1,y_1}) > \eta_n(\Gamma_{x_1,y_1}) > 0,\ \eta_n(\Gamma_{x_1,y_{j_n}}) > \eta_{n+1}(\Gamma_{x_1,y_{j_n}}) = 0, \tag{3.3.10}$$

$$\eta_n(\Gamma_{x_{i_n},y_1}) > \eta_{n+1}(\Gamma_{x_{i_n},y_1}) \geq 0,\ \eta_{n+1}(\Gamma_{x_{i_n},y_{j_n}}) > \eta_n(\Gamma_{x_{i_n},y_{j_n}}) > 0, \tag{3.3.11}$$

and

$$\eta_{n+1}(\Gamma_{x_i,y_j}) = \eta_n(\Gamma_{x_i,y_j}) \text{ for all other } i, j. \tag{3.3.12}$$

Since η_n is a good decomposition of T, we have

$$T = \int_\Gamma I_\gamma d\eta_n,\ \mathbf{M}(T) = \int_\Gamma \mathbf{M}(I_\gamma) d\eta_n(\gamma) \text{ and } \mathbf{M}(\partial T) = \int_\Gamma \mathbf{M}(\partial I_\gamma) d\eta_n(\gamma).$$

In particular, $\mathbf{M}(T) = \int_\Gamma \mathbf{M}(I_\gamma) d\eta_n(\gamma)$ implies that

$$\mathbf{M}(S_{1,1}(\eta_n) + S_{i_n,j_n}(\eta_n)) = \mathbf{M}(S_{1,1}(\eta_n)) + \mathbf{M}(S_{i_n,j_n}(\eta_n)),$$

and

$$\mathbf{M}(S_{1,j_n}(\eta_n) + S_{i_n,1}(\eta_n)) = \mathbf{M}(S_{1,j_n}(\eta_n)) + \mathbf{M}(S_{i_n,1}(\eta_n)).$$

By assumption,

$$C[(1,1),(i_n,j_n),\eta_n] = S_{1,1}(\eta_n) - S_{1,j_n}(\eta_n) - S_{i_n,1}(\eta_n) + S_{i_n,j_n}(\eta_n) = 0,$$

i.e., $S_{1,1}(\eta_n) + S_{i_n,j_n}(\eta_n) = S_{1,j_n}(\eta_n) + S_{i_n,1}(\eta_n)$. Thus,

$$\begin{aligned} & \mathbf{M}(S_{1,1}(\eta_n)) + \mathbf{M}(S_{i_n,j_n}(\eta_n)) = \mathbf{M}(S_{1,1}(\eta_n) + S_{i_n,j_n}(\eta_n)) \\ = \quad & \mathbf{M}(S_{1,j_n}(\eta_n) + S_{i_n,1}(\eta_n)) = \mathbf{M}(S_{1,j_n}(\eta_n)) + \mathbf{M}(S_{i_n,1}(\eta_n)). \end{aligned}$$

Now, by the construction of η_{n+1},

$$\int_\Gamma I_\gamma d\eta_{n+1} - \int_\Gamma I_\gamma d\eta_n = \min\{\eta_n(\Gamma_{x_1,y_{j_n}}), \eta_n(\Gamma_{x_{i_n},y_1})\} \cdot C[(1,1),(i_n,j_n),\eta_n] = 0,$$

and

$$
\begin{aligned}
&\int_\Gamma \mathbf{M}(I_\gamma) d\eta_{n+1}(\gamma) - \int_\Gamma \mathbf{M}(I_\gamma) d\eta_n(\gamma) \\
= \quad & \min\{\eta_n(\Gamma_{x_1, y_{j_n}}), \eta_n(\Gamma_{x_{i_n}, y_1})\} \left(\mathbf{M}(S_{1,1}) - \mathbf{M}(S_{1,j_n}) - \mathbf{M}(S_{i_n,1}) + \mathbf{M}(S_{i_n,j_n})\right) = 0.
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
&\int_\Gamma \mathbf{M}(\partial I_\gamma) d\eta_{n+1}(\gamma) - \int_\Gamma \mathbf{M}(\partial I_\gamma) d\eta_n(\gamma) \\
= \quad & \min\{\eta_n(\Gamma_{x_1, y_{j_n}}), \eta_n(\Gamma_{x_{i_n}, y_1})\} \left(\mathbf{M}(\partial S_{1,1}) - \mathbf{M}(\partial S_{1,j_n}) - \mathbf{M}(\partial S_{i_n,1}) + \mathbf{M}(\partial S_{i_n,j_n})\right) \\
= \quad & \min\{\eta_n(\Gamma_{x_1, y_{j_n}}), \eta_n(\Gamma_{x_{i_n}, y_1})\} (2 - 2 - 2 + 2) = 0.
\end{aligned}
$$

As a result, since η_n is a good decomposition of T, η_{n+1} is a good decomposition of T as well.

Step 2: Show that the sequence $\{\eta_n\}$ converges to a good decomposition $\tilde{\eta}$ of T.

Note that for each $1 \leq i \leq M$ and $1 \leq j \leq N$, the sequence $\{\eta_n \lfloor_{\Gamma_{x_i, y_j}}\}_{n=1}^\infty$ is a monotonic sequence of measures with bounded mass. Indeed, by the construction above and by equations (3.3.10), (3.3.11) and (3.3.12),

- if $i = 1, j = 1$, then $\{\eta_n \lfloor_{\Gamma_{x_i, y_j}}\}_{n=1}^\infty$ is monotone increasing;
- if $i = 1, j > 1$, then $\{\eta_n \lfloor_{\Gamma_{x_i, y_j}}\}_{n=1}^\infty$ is monotone decreasing;
- if $i > 1, j = 1$, then $\{\eta_n \lfloor_{\Gamma_{x_i, y_j}}\}_{n=1}^\infty$ is monotone decreasing;
- if $i > 1, j > 1$, then $\{\eta_n \lfloor_{\Gamma_{x_i, y_j}}\}_{n=1}^\infty$ is monotone increasing, and eventually constant.

As a result, the sequence, $\{\eta_n \lfloor_{\Gamma_{x_i, y_j}}\}_{n=1}^\infty$, converges to some measure η_{ij} for each (i, j). Define

$$
\tilde{\eta} := \sum_{i=1}^M \sum_{j=1}^N \eta_{ij}.
$$

Hence, as $n \to \infty$,

$$
\eta_n = \sum_{i=1}^M \sum_{j=1}^N \eta_n \lfloor_{\Gamma_{x_i, y_j}} \longrightarrow \tilde{\eta} = \sum_{i=1}^M \sum_{j=1}^N \eta_{ij}.
$$

Since each η_n is a good decomposition of T, it follows that

$$\int_\Gamma I_\gamma d\tilde{\eta} = \lim_{n\to\infty}\int_\Gamma I_\gamma d\eta_n = T,$$
$$\int_\Gamma \mathbf{M}(I_\gamma) d\tilde{\eta} = \lim_{n\to\infty}\int_\Gamma \mathbf{M}(I_\gamma) d\eta_n = \mathbf{M}(T),$$
$$\int_\Gamma \mathbf{M}(\partial I_\gamma) d\tilde{\eta} = \lim_{n\to\infty}\int_\Gamma \mathbf{M}(\partial I_\gamma) d\eta_n = \mathbf{M}(\partial T).$$

As a result, $\tilde{\eta}$ is also a good decomposition of T.

Step 3: Show that $\tilde{\eta} \prec\!\!\prec \eta$.

Suppose $\tilde{\eta}(\Gamma_{x_i,y_j}) > 0$ for some pair (i,j). Then, $\eta_n(\Gamma_{x_i,y_j}) > 0$ when n is large enough. By (3.3.9),

$$\eta_n\lfloor_{\Gamma_{x_i,y_j}} = \prod_{k=1}^{n-1}(1+\lambda_{k,i,j})\eta\lfloor_{\Gamma_{x_i,y_j}},\ \text{for some } \lambda_{k,i,j} \geq -1 \text{ for each } k.$$

That is,

$$\eta_n = \left(\prod_{k=1}^{n-1}(1+\lambda_{k,i,j})\right)\eta \text{ on } \Gamma_{x_i,y_j}.$$

As a result, $\eta_n(\Gamma_{x_i,y_j}) > 0$ implies $\eta(\Gamma_{x_i,y_j}) > 0$ and $S_{i,j}(\eta_n) = S_{i,j}(\eta)$. Since $\tilde{\eta}$ is the limit of η_n,

$$S_{i,j}(\tilde{\eta}) = \lim_{n\to\infty} S_{i,j}(\eta_n) = S_{i,j}(\eta).$$

This proves $\tilde{\eta} \prec\!\!\prec \eta$.

Step 4: Show that $\mathcal{A}_{\eta_{n+1}}(1,1) \subsetneqq \mathcal{A}_{\eta_n}(1,1)$ for each n.

Note that $(i_n, j_n) \in \mathcal{A}_{\eta_n}(1,1) \setminus \mathcal{A}_{\eta_{n+1}}(1,1)$. Indeed, if $(i_n,j_n) \in \mathcal{A}_{\eta_{n+1}}(1,1)$, then

$$C[(1,1),(i_n,j_n),\eta_{n+1}] = 0 \text{ and } s[(1,1),(i_n,j_n),\eta_{n+1}] = 1.$$

This implies $sgn(\eta_{n+1}(\Gamma_{x_1,y_{j_n}})) = 1$, which contradicts with $\eta_{n+1}(\Gamma_{x_1,y_{j_n}}) = 0$ as given in (3.3.10).

We now show that $\mathcal{A}_{\eta_{n+1}}(1,1) \subseteq \mathcal{A}_{\eta_n}(1,1)$. For any $(i_0,j_0) \in \mathcal{A}_{\eta_{n+1}}(1,1)$, by definition,

$$C[(1,1),(i_0,j_0),\eta_{n+1}] = 0 \text{ and } s[(1,1),(i_0,j_0),\eta_{n+1}] = 1.$$

The condition $s[(1,1),(i_0,j_0),\eta_{n+1}] = 1$ indicates that

$$\eta_{n+1}(\Gamma_{x_1,y_1}) > 0, \eta_{n+1}(\Gamma_{x_1,y_{j_0}}) > 0, \eta_{n+1}(\Gamma_{x_{i_0},y_1}) > 0, \eta_{n+1}(\Gamma_{x_{i_0},y_{j_0}}) > 0.$$

By equations (3.3.10)–(3.3.12), and $(i_0, j_0) \neq (i_n, j_n)$,

$$\eta_n(\Gamma_{x_1,y_1}) > 0,\ \eta_n(\Gamma_{x_1,y_{j_0}}) \geq \eta_{n+1}(\Gamma_{x_1,y_{j_0}}) > 0,$$

$$\eta_n(\Gamma_{x_{i_0},y_1}) \geq \eta_{n+1}(\Gamma_{x_{i_0},y_1}) > 0,\ \eta_n(\Gamma_{x_{i_0},y_{j_0}}) = \eta_{n+1}(\Gamma_{x_{i_0},y_{j_0}}) > 0.$$

By (3.3.9), for each i, j, when both $\eta_n(\Gamma_{x_i,y_j}) > 0$ and $\eta_{n+1}(\Gamma_{x_i,y_j}) > 0$, then

$$S_{i,j}(\eta_n) = S_{i,j}(\eta_{n+1}).$$

As a result,

$$C[(1,1),(i_0,j_0),\eta_n] = C[(1,1),(i_0,j_0),\eta_{n+1}] = 0.$$

Therefore, $(i_0, j_0) \in \mathcal{A}_{\eta_n}(1,1)$ and hence $\mathcal{A}_{\eta_{n+1}}(1,1) \subseteq \mathcal{A}_{\eta_n}(1,1)$.

Step 5: Show that $\mathcal{A}_{\tilde{\eta}}(1,1) = \emptyset$.

Assume that there exists $(i', j') \in \mathcal{A}_{\tilde{\eta}}(1,1)$, i.e. $C[(1,1),(i',j'),\tilde{\eta}] = 0$ and $s[(1,1),(i',j'),\tilde{\eta}] = 1$. For any $(i,j) \in \{(1,1),(1,j'),(i',1),(i',j')\}$, since $s[(1,1),(i',j'),\tilde{\eta}] = 1$, it follows that

$$\lim_{n\to\infty} \eta_n(\Gamma_{x_i,y_j}) = \tilde{\eta}(\Gamma_{x_i,y_j}) > 0.$$

Thus, there exists an $N_0 \in \mathbb{N}$ such that $\eta_n(\Gamma_{x_i,y_j}) > 0$ for all $n \geq N_0$. By (3.3.9), this implies that the normalized current $S_{i,j}(\eta_n)$ is independent of n, and hence $S_{i,j}(\eta_n) = S_{i,j}(\tilde{\eta})$ for all $n \geq N_0$. As a result, for each $n \geq N_0$,

$$C[(1,1),(i',j'),\eta_n] = C[(1,1),(i',j'),\tilde{\eta}] = 0 \text{ and } s[(1,1),(i',j'),\eta_n] = s[(1,1),(i',j'),\tilde{\eta}] = 1.$$

This shows that $(i', j') \in \mathcal{A}_{\eta_n}(1,1)$. On the other hand, since $\{\mathcal{A}_{\eta_n}(1,1)\}$ is a sequence of nested subsets in $\mathbb{N}^2$ with $\mathcal{A}_{\eta_{n+1}}(1,1) \subsetneqq \mathcal{A}_{\eta_n}(1,1)$ for each n. When n is larger than the order of the fixed element (i', j'), it is not possible for $(i', j') \in \mathcal{A}_{\eta_n}(1,1)$. A contradiction.

$\square$

We now extend Lemma 3.3.7 to a more general case:

Lemma 3.3.8. *For any good decomposition η of T, there exists a sequence of good decomposition $\{\eta_n\}_{n=0}^{\infty}$ of T with $\eta_0 = \eta$ such that for each $n \geq 1$, $\eta_n \prec\prec \eta_{n-1}$ and $\mathcal{A}_{\eta_n}(i_k, j_k) = \emptyset$ for all $1 \leq k \leq n$, where $\{(i_k, j_k)\}$ is given in (3.3.8).*

Proof. We will prove these results by induction. Lemma 3.3.7 provides the base case when $n = 1$. For each $n \geq 2$, assume that there exists a good decomposition η_{n-1} of T such that $\eta_{n-1} \prec\!\!\prec \eta_{n-2}$ and $\mathcal{A}_{\eta_{n-1}}(i_k, j_k) = \emptyset$ for all $1 \leq k \leq n-1$. Using η_{n-1}, we construct η_n as follows.

Denote

$$\tilde{\Gamma}_n = \bigcup_{i_n \leq i, j_n \leq j} \Gamma_{x_i, y_j}.$$

Let $\tilde{\eta}_n$ be the measure $\tilde{\eta}$ achieved in Lemma 3.3.7 with η being replaced by $\eta_{n-1}\lfloor_{\tilde{\Gamma}_n}$ and T being replaced by $\tilde{T} := \int_{\tilde{\Gamma}_n} I_\gamma d\eta_{n-1}$. Define

$$\eta_n := \eta_{n-1}\lfloor_{\Gamma \setminus \tilde{\Gamma}_n} + \tilde{\eta}_n.$$

We first claim that η_n is a good decomposition of T. Indeed, since both $\tilde{\eta}_n$ and $\eta_{n-1}\lfloor_{\tilde{\Gamma}_n}$ are good decompositions of $\tilde{T}$,

$$\int_\Gamma I_\gamma d\eta_n - \int_\Gamma I_\gamma d\eta_{n-1} = \int_\Gamma I_\gamma d\tilde{\eta}_n - \int_{\tilde{\Gamma}_n} I_\gamma d\eta_{n-1} = 0,$$

$$\int_\Gamma \mathbf{M}(I_\gamma) d\eta_n(\gamma) - \int_\Gamma \mathbf{M}(I_\gamma) d\eta_{n-1}(\gamma) = \int_\Gamma \mathbf{M}(I_\gamma) d\tilde{\eta}_n - \int_{\tilde{\Gamma}_n} \mathbf{M}(I_\gamma) d\eta_{n-1} = 0,$$

and

$$\int_\Gamma \mathbf{M}(\partial I_\gamma) d\eta_n(\gamma) - \int_\Gamma \mathbf{M}(\partial I_\gamma) d\eta_{n-1}(\gamma) = \int_\Gamma \mathbf{M}(\partial I_\gamma) d\tilde{\eta}_n - \int_{\tilde{\Gamma}_n} \mathbf{M}(\partial I_\gamma) d\eta_{n-1} = 0.$$

As a result, since η_{n-1} is a good decomposition of T, η_n is also a good decomposition of T.

We now show that $\eta_n \prec\!\!\prec \eta_{n-1}$. Suppose $\eta_n(\Gamma_{x_i, y_j}) > 0$ for some $1 \leq i \leq M, 1 \leq j \leq N$.

- When $i < i_n$ or $j < j_n$, definition of η_n gives $\eta_n\lfloor_{\Gamma_{x_i,y_j}} = \eta_{n-1}\lfloor_{\Gamma_{x_i,y_j}}$. Therefore,

$$\eta_{n-1}(\Gamma_{x_i,y_j}) = \eta_n(\Gamma_{x_i,y_j}) > 0 \text{ and } S_{i,j}(\eta_{n-1}) = S_{i,j}(\eta_n).$$

- When $i \geq i_n$ and $j \geq j_n$, definition of η_n gives $\eta_n\lfloor_{\Gamma_{x_i,y_j}} = \tilde{\eta}_n\lfloor_{\Gamma_{x_i,y_j}}$, so that

$$\tilde{\eta}_n(\Gamma_{x_i,y_j}) = \eta_n(\Gamma_{x_i,y_j}) > 0.$$

  Since $\tilde{\eta}_n \prec\!\!\prec \eta_{n-1}\lfloor_{\tilde{\Gamma}_n}$ by Lemma 3.3.7, it follows that

$$\eta_{n-1}(\Gamma_{x_i,y_j}) > 0 \text{ and } S_{i,j}(\eta_{n-1}) = S_{i,j}(\tilde{\eta}_n) = S_{i,j}(\eta_n).$$

In both cases, $\eta_{n-1}(\Gamma_{x_i,y_j}) > 0$ and $S_{i,j}(\eta_{n-1}) = S_{i,j}(\eta_n)$. That is, $\eta_n \prec\!\!\prec \eta_{n-1}$.

We now show that $\mathcal{A}_{\eta_n}(i_k, j_k) = \emptyset$ for all $1 \leq k \leq n$. When $k = n$, $\mathcal{A}_{\eta_n}(i_n, j_n) = \emptyset$ by Lemma 3.3.7. Suppose $k < n$, and for contradiction, we assume $\mathcal{A}_{\eta_n}(i_k, j_k) \neq \emptyset$. Thus, there exists $(i^*, j^*) \in \mathcal{A}_{\eta_n}(i_k, j_k)$, i.e.,

$$C[(i_k, j_k), (i^*, j^*), \eta_n] = 0 \text{ and } s[(i_k, j_k), (i^*, j^*), \eta_n] = 1.$$

Now, for any $(i, j) \in \{(i_k, j_k), (i_k, j^*), (i^*, j_k), (i^*, j^*)\}$, since $s[(i_k, j_k), (i^*, j^*), \eta_n] = 1$, it follows that $\eta_n(\Gamma_{x_i,y_j}) > 0$. By the definition of η_n, when $i < i_n$ or $j < j_n$, $\eta_n = \eta_{n-1}$ on Γ_{x_i,y_j}. Thus,

$$\eta_{n-1}(\Gamma_{x_i,y_j}) = \eta_n(\Gamma_{x_i,y_j}) > 0 \text{ and } S_{i,j}(\eta_n) = S_{i,j}(\eta_{n-1}). \tag{3.3.13}$$

When $i \geq i_n$ and $j \geq j_n$,

$$\tilde{\eta}_n(\Gamma_{x_i,y_j}) = \eta_n(\Gamma_{x_i,y_j}) > 0.$$

Since $\tilde{\eta}_n \prec\!\!\prec \eta_{n-1}\lfloor_{\tilde{\Gamma}_n}$, then equations in (3.3.13) still hold. As a result,

$$C[(i_k, j_k), (i^*, j^*), \eta_{n-1}] = C[(i_k, j_k), (i^*, j^*), \eta_n] = 0 \text{ and } s[(i_k, j_k), (i^*, j^*), \eta_{n-1}] = 1.$$

Therefore, $(i^*, j^*) \in \mathcal{A}_{\eta_{n-1}}(i_k, j_k)$, which contradicts with $\mathcal{A}_{\eta_{n-1}}(i_k, j_k) = \emptyset$ whenever $k \leq n-1$. $\square$

We now give the proof of Theorem 3.3.3 by showing that for any good decomposition η of T, there exists a good decomposition η_∞ of T such that $\eta_\infty \prec\!\!\prec \eta$ and $\mathcal{A}_{\eta_\infty}(i, j) = \emptyset$ for all $1 \leq i \leq M, 1 \leq j \leq N$.

Proof of Theorem 3.3.3. Let $\{\eta_n\}$ be the sequence of good decomposition of T constructed in the proof of Lemma 3.3.8. Observe that by the construction of the sequence $\{\eta_n\}$, it follows that for any $k \in \mathbb{N}$,

$$\eta_n\lfloor_{\Gamma_{x_{i_k},y_{j_k}}} = \eta_k\lfloor_{\Gamma_{x_{i_k},y_{j_k}}} \tag{3.3.14}$$

for all $n \geq k$. Define $\eta_\infty : \Gamma \to \mathbb{R}$ by setting

$$\eta_\infty := \eta_k \quad \text{on } \Gamma_{x_{i_k},y_{j_k}}, \forall k \in \mathbb{N}. \tag{3.3.15}$$

We first show that $\{\eta_n\}$ converges to η_∞ with respect to the total variation distance $\|\cdot\|$. Indeed, by (3.3.14),

$$\begin{aligned}
\|\eta_n - \eta_\infty\| &= \|\sum_{k\geq 1}(\eta_n - \eta_k)\lfloor_{\Gamma_{x_{i_k},y_{j_k}}}\| = \|\sum_{k\geq n+1}(\eta_n - \eta_k)\lfloor_{\Gamma_{x_{i_k},y_{j_k}}}\| \\
&\leq \sum_{k\geq n+1}\eta_n(\Gamma_{x_{i_k},y_{j_k}}) + \sum_{k\geq n+1}\eta_k(\Gamma_{x_{i_k},y_{j_k}}) \\
&\leq \sum_{i_k+j_k\geq i_n+j_n}\eta_n(\Gamma_{x_{i_k},y_{j_k}}) + \sum_{k\geq n+1}\eta_k(\Gamma_{x_{i_k},y_{j_k}}) \\
&\leq \sum_{i_k\geq\sqrt{i_n j_n}}\sum_{j_k=1}^{N}\eta_n(\Gamma_{x_{i_k},y_{j_k}}) + \sum_{j_k\geq\sqrt{i_n j_n}}\sum_{i_k=1}^{M}\eta_n(\Gamma_{x_{i_k},y_{j_k}}) + \sum_{k\geq n+1}\eta_k(\Gamma_{x_{i_k},y_{j_k}}) \\
&= \sum_{i_k\geq\sqrt{i_n j_n}} m'_{i_k} + \sum_{j_k\geq\sqrt{i_n j_n}} m_{j_k} + \sum_{k\geq n+1}\eta_k(\Gamma_{x_{i_k},y_{j_k}}),
\end{aligned}$$

and

$$\eta_\infty(\Gamma) = \sum_{k=1}^{\infty}\eta_k(\Gamma_{x_{i_k},y_{j_k}}) = \lim_{n\to\infty}\sum_{k=1}^{n}\eta_k(\Gamma_{x_{i_k},y_{j_k}}) = \lim_{n\to\infty}\sum_{k=1}^{n}\eta_n(\Gamma_{x_{i_k},y_{j_k}}) \leq \lim_{n\to\infty}\eta_n(\Gamma) = \eta(\Gamma) < \infty.$$

Thus, since $\lim_{n\to\infty} i_n j_n = \infty$ and $\sum_{i=1}^{M} m'_i = \sum_{j=1}^{N} m_j < \infty$, it follows that $\lim_{n\to\infty}\|\eta_n - \eta_\infty\| = 0$. Since η_n is a good decomposition for each n, it follows that its limit η_∞ is also a good decomposition of T.

Moreover, if $\eta_\infty(\Gamma_{x_{i_k},y_{j_k}}) > 0$ for some k, then $\eta_k(\Gamma_{x_{i_k},y_{j_k}}) > 0$ by (3.3.15). Thus, by Lemma 3.3.8 and transitivity of "$\prec\!\!\prec$", we have $\eta_k \prec\!\!\prec \eta$, which implies

$$\eta(\Gamma_{x_{i_k},y_{j_k}}) > 0 \text{ and } S_{i_k,j_k}(\eta_\infty) = S_{i_k,j_k}(\eta_k) = S_{i_k,j_k}(\eta).$$

Therefore, $\eta_\infty \prec\!\!\prec \eta$.

We now show that $\mathcal{A}_{\eta_\infty}(i_k, j_k) = \emptyset$ for each k. Assume that for some k, $\mathcal{A}_{\eta_\infty}(i_k, j_k)$ contains an element (i_n, j_n). Then the definition of $\mathcal{A}_{\eta_\infty}(i_k, j_k)$ implies $n > k$ and

$$C[(i_k, j_k), (i_n, j_n), \eta_\infty] = 0 \text{ and } s[(i_k, j_k), (i_n, j_n), \eta_\infty] = 1.$$

By (3.3.14) and (3.3.15), since (i_n, j_n) has the largest order among the elements

$$\{(i_k, j_k), (i_k, j_n), (i_n, j_k), (i_n, j_n)\},$$

it follows that $\eta_\infty = \eta_n$ on Γ_{x_i,y_j} for each (i,j) of these four elements. Thus,

$$C[(i_k, j_k), (i_n, j_n), \eta_n] = 0 \text{ and } s[(i_k, j_k), (i_n, j_n), \eta_n] = 1.$$

This shows $(i_n, j_n) \in \mathcal{A}_{\eta_n}(i_k, j_k)$, a contradiction with $\mathcal{A}_{\eta_n}(i_k, j_k) = \emptyset$ due to Lemma 3.3.8. $\square$

3.4. Decomposition of cycle-free transport paths

In this section, we will prove the decomposition theorem in Theorem 3.4.8 using the better decomposition η_∞ achieved from Theorem 3.3.3.

We first recall a concept that was introduced in [**18**, Definition 4.6].

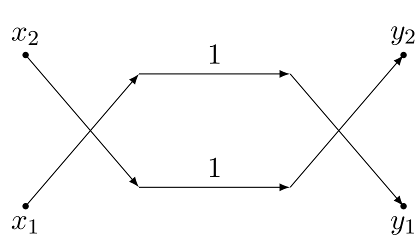
Definition 3.4.1. Let $T = \underline{\underline{\tau}}(M, \theta, \xi)$ and $S = \underline{\underline{\tau}}(N, \phi, \zeta)$ be two real rectifiable k-currents. We say S is on T if $\mathcal{H}^k(N \setminus M) = 0$, and $\phi(x) \leq \theta(x)$ for $\mathcal{H}^k$ almost all $x \in N$.

Note that when $S = \underline{\underline{\tau}}(N, \phi, \zeta)$ is on $T = \underline{\underline{\tau}}(M, \theta, \xi)$, then $\xi(x) = \pm\zeta(x)$ for $\mathcal{H}^k$ almost all $x \in N$, since two rectifiable sets have the same tangent almost everywhere on their intersection. Using it, we now introduce the concept of "cycle-free" currents as follows:

Definition 3.4.2. Let T and S be two real rectifiable k-currents. S is called a cycle on T if S is on T and $\partial S = 0$. Also, T is called *cycle-free* if except for the zero current, there is no other cycle on T.

The zero current is called the trivial cycle on T.

Remark 3.4.3. The concept of "cycle-free" is different from "acyclic". A cycle-free current is automatically acyclic, but not vice versa. For instance, let T be a transport path (which is a 1-current) from $\mu^- = \delta_{x_1} + \delta_{x_2}$ to $\mu^+ = \delta_{y_1} + \delta_{y_2}$ as shown below.



T

Then T is acyclic but not cycle-free.

As an example, we first show that each optimal transport path is cycle-free. To do so, we start with an analogous result to [**18**, Theorem 4.7] as follows.

PROPOSITION 3.4.4. *Let $T \in Path(\mu^-, \mu^+)$ with $\mathbf{M}_\alpha(T) < \infty$ for some $0 < \alpha < 1$. Suppose there exists a rectifiable 1-current S such that S is on T and $\partial S = 0$, then for any $\epsilon \in [-1, 1]$, $T + \epsilon S \in Path(\mu^-, \mu^+)$ and*

$$\min\{\mathbf{M}_\alpha(T + S), \mathbf{M}_\alpha(T - S)\} \leq \mathbf{M}_\alpha(T)$$

with the equality holds only when $S = 0$.

PROOF. The statements clearly hold if $S = 0$. Thus, in the following, we may assume that S is non-zero. Since $T \in Path(\mu^-, \mu^+)$ and $\partial S = 0$, it holds that $\partial(T+\epsilon S) = \partial T + \epsilon \partial S = \partial T = \mu^+ - \mu^-$. That is, $T + \epsilon S \in Path(\mu^-, \mu^+)$.

Let $T = \underline{\underline{\tau}}(M, \theta, \xi)$ and $S = \underline{\underline{\tau}}(N, \phi, \zeta)$. Since S is on T, we have $\mathcal{H}^1(N \setminus M) = 0$, and $\phi(x) \leq \theta(x)$ for $\mathcal{H}^1$ almost all $x \in N$. One may assume that $N = M$ by extending $\phi(x) = 0$ and $\zeta(x) = \xi(x)$ for $x \in M \setminus N$.

For $\epsilon \in [-1, 1]$, we now consider the function

$$g(\epsilon) = \mathbf{M}_\alpha(T + \epsilon S) = \int_M (\theta(x) + \epsilon\phi(x)\langle\xi(x), \zeta(x)\rangle)^\alpha \, d\mathcal{H}^1(x).$$

Here, the value of the inner product is $\langle\xi(x), \zeta(x)\rangle = \pm 1$ for $\mathcal{H}^1 - a.e.\ x \in M$. Since $\mathbf{M}_\alpha(T) = \int_M \theta^\alpha d\mathcal{H}^1 < \infty$ and $\phi(x) \leq \theta(x)$ for $\mathcal{H}^1$ almost all $x \in M$, we have for any $\epsilon \in (-1, 1)$,

$$g'(\epsilon) = \alpha \int_M (\theta(x) + \epsilon\phi(x)\langle\xi(x), \zeta(x)\rangle)^{\alpha-1} \phi(x)\langle\xi(x), \zeta(x)\rangle d\mathcal{H}^1(x)$$

and

$$g''(\epsilon) = \alpha(\alpha - 1) \int_M (\theta(x) + \epsilon\phi(x)\langle\xi(x), \zeta(x)\rangle)^{\alpha-2} \phi(x)^2 d\mathcal{H}^1(x) < 0,$$

because $0 < \alpha < 1$ and S is non-zero. This shows that $g(\epsilon)$ is a strictly concave function on $(-1, 1)$. By the lower semi-continuity of $\mathbf{M}_\alpha$, $g(\epsilon)$ is lower semi-continuous at $\epsilon = \pm 1$. Thus, $\min\{g(-1), g(1)\} < g(0)$. That is, $\min\{\mathbf{M}_\alpha(T + S), \mathbf{M}_\alpha(T - S)\} < \mathbf{M}_\alpha(T)$ whenever S is on T, nonzero and $\partial S = 0$. $\square$

COROLLARY 3.4.5. *Suppose T is an α-optimal transport path from μ^- to μ^+ for $0 < \alpha < 1$. Then T is cycle-free.*

PROOF. Since T is α-optimal, it is acyclic and hence it has a good decomposition. Suppose S is on T and $\partial S = 0$. Assume S is non-zero, then $\min\{\mathbf{M}_\alpha(T+S), \mathbf{M}_\alpha(T-S)\} < \mathbf{M}_\alpha(T)$, which contradicts with the $\mathbf{M}_\alpha$ optimality of T. Therefore, S must be zero. Hence, T is cycle-free. $\square$

To characterize cycle-free transport paths, we consider their better decomposition.

PROPOSITION 3.4.6. *Each cycle-free transport path $T \in Path(\mu^-, \mu^+)$ has at least a better decomposition.*

PROOF. By definition, each cycle-free transport path is acyclic and hence has a good decomposition. By Theorem 3.3.3, it has a better decomposition. $\square$

PROPOSITION 3.4.7. *Let $T \in Path(\mu^-, \mu^+)$ be a cycle-free transport path, and let η be a better decomposition of T. For each $y_j \in \{y_1, y_2, \dots, y_N\}$, denote*

$$X_j(\eta) := \{x_i \in X : \eta(\Gamma_{x_i, y_j}) > 0\}. \tag{3.4.1}$$

Then for each pair $1 \le j_1 < j_2 \le N$,

$$|X_{j_1}(\eta) \cap X_{j_2}(\eta)| \le 1, \tag{3.4.2}$$

i.e., the intersection $X_{j_1}(\eta) \cap X_{j_2}(\eta)$ is either empty or a single point.

PROOF. Assume $|X_{j_1}(\eta) \cap X_{j_2}(\eta)| > 1$. Then there exist two distinct points $x_{i_1}, x_{i_2} \in X_{j_1}(\eta) \cap X_{j_2}(\eta)$ with $i_1 < i_2$. Thus,

$$\eta(\Gamma_{x_{i_1}, y_{j_1}}) > 0,\ \eta(\Gamma_{x_{i_1}, y_{j_2}}) > 0,\ \eta(\Gamma_{x_{i_2}, y_{j_1}}) > 0, \text{ and } \eta(\Gamma_{x_{i_2}, y_{j_2}}) > 0. \tag{3.4.3}$$

By (3.3.7), this implies that $C[(i_1, j_1), (i_2, j_2), \eta]$ defined in (3.3.6) is a cycle. Since η is a better decomposition of T, by (3.4.3), it follows that $C[(i_1, j_1), (i_2, j_2), \eta]$ is non-vanishing. Pick

$$0 < \epsilon_0 \le \frac{1}{4} \min\{\eta(\Gamma_{x_{i_1}, y_{j_1}}), \eta(\Gamma_{x_{i_1}, y_{j_2}}), \eta(\Gamma_{x_{i_2}, y_{j_1}}), \eta(\Gamma_{x_2, y_{j_2}})\},$$

and observe that

$$S = \epsilon_0 \cdot C[(i_1, j_1), (i_2, j_2), \eta]$$

is a non-vanishing cycle on T. Indeed, assume $T = \underline{\underline{\tau}}(M, \theta, \xi)$ and $S = \underline{\underline{\tau}}(N, \phi, \zeta)$, then $N \subseteq M$ and for $\mathcal{H}^1$-a.e. x,

$$\begin{aligned}
\phi(x) &\leq \epsilon_0 \left(\frac{\eta\lfloor_{\Gamma_{x_{i_1}, y_{j_1}}}}{\eta(\Gamma_{x_{i_1}, y_{j_1}})} + \frac{\eta\lfloor_{\Gamma_{x_{i_1}, y_{j_2}}}}{\eta(\Gamma_{x_{i_1}, y_{j_2}})} + \frac{\eta\lfloor_{\Gamma_{x_{i_2}, y_{j_1}}}}{\eta(\Gamma_{x_{i_2}, y_{j_1}})} + \frac{\eta\lfloor_{\Gamma_{x_{i_2}, y_{j_2}}}}{\eta(\Gamma_{x_{i_2}, y_{j_2}})} \right) (\{\gamma \in \Gamma : x \in \mathrm{Im}(\gamma)\}) \\
&\leq \epsilon_0 \left(\frac{1}{\eta(\Gamma_{x_{i_1}, y_{j_1}})} + \frac{1}{\eta(\Gamma_{x_{i_1}, y_{j_2}})} + \frac{1}{\eta(\Gamma_{x_{i_2}, y_{j_1}})} + \frac{1}{\eta(\Gamma_{x_{i_2}, y_{j_2}})} \right) \eta\left(\{\gamma \in \Gamma : x \in \mathrm{Im}(\gamma)\}\right) \\
&\leq \eta\left(\{\gamma \in \Gamma : x \in \mathrm{Im}(\gamma)\}\right) = \theta(x),
\end{aligned}$$

by equation (3.2.5). This shows that S is a non-vanishing cycle on T. A contradiction with T is cycle-free. □

THEOREM 3.4.8. *Let T be a cycle-free transport path from μ^- to μ^+, where μ^- and μ^+ are given in (3.2.10). Then there exists a decomposition*

$$T = \sum_{j=0}^{N} T_j \tag{3.4.4}$$

such that

(a) *The set $\{x_1, x_2, \cdots, x_M\}$ can be expressed as the disjoint union of its subsets $\{B_j\}_{j=0}^N$ with the cardinality $|B_0| \leq \binom{N}{2}$;*

(b) *For each $j = 1, 2, \cdots, N$, T_j is a single-target transport path from*

$$\mu_j^- := \mu^- \lfloor_{B_j} \text{ to } \mu_j^+ = \tilde{m}_j \delta_{y_j}$$

for some $0 \leq \tilde{m}_j := \mu^-(B_j) \leq m_j$. Each T_j is a subcurrent of T.

(c) *T_0 is a transport path from*

$$\mu_0^- := \mu^- \lfloor_{B_0} \text{ to } \mu_0^+ = \sum_{j=1}^{N} (m_j - \tilde{m}_j) \delta_{y_j}.$$

T_0 is also a subcurrent of T.

Note that, by Theorem 3.4.8 , it follows that

$$\mu^- = \sum_{j=0}^{N} \mu_j^- \text{ and } \mu^+ = \sum_{j=0}^{N} \mu_j^+. \tag{3.4.5}$$

PROOF. Let η be a better decomposition of T, and $X_j(\eta)$ be the set as defined in (3.4.1). Denote

$$B_0 := \bigcup_{1 \le j_1 < j_2 \le N} (X_{j_1}(\eta) \cap X_{j_2}(\eta)) \tag{3.4.6}$$

and for each $1 \le j \le N$, denote

$$B_j := X_j(\eta) \setminus B_0.$$

Then $\{B_j\}_{j=0}^N$ are pairwise disjoint. Moreover, by (3.4.2), $|B_0| \le \binom{N}{2}$.

Define

$$T_0 := \sum_{j=1}^N \sum_{x_i \in B_0} \int_{\Gamma_{x_i,y_j}} I_\gamma \, d\eta,$$

and for each $1 \le j \le N$, denote

$$T_j := \sum_{x_i \in B_j} \int_{\Gamma_{x_i,y_j}} I_\gamma \, d\eta.$$

Then each T_j is a subcurrent of T for $0 \le j \le N$ and

$$\begin{aligned} T &= \sum_{j=1}^N \sum_{i=1}^M \int_{\Gamma_{x_i,y_j}} I_\gamma \, d\eta = \sum_{j=1}^N \left(\sum_{x_i \in B_j} \int_{\Gamma_{x_i,y_j}} I_\gamma \, d\eta + \sum_{x_i \in B_0} \int_{\Gamma_{x_i,y_j}} I_\gamma \, d\eta \right) \\ &= \sum_{j=1}^N \sum_{x_i \in B_j} \int_{\Gamma_{x_i,y_j}} I_\gamma \, d\eta + \sum_{j=1}^N \sum_{x_i \in B_0} \int_{\Gamma_{x_i,y_j}} I_\gamma \, d\eta \\ &= \sum_{j=1}^N T_j \; + T_0 = \sum_{j=0}^N T_j. \end{aligned}$$

For each $1 \le j \le N$, T_j is a single-target transport path with

$$\partial T_j = \sum_{x_i \in B_j} \int_{\Gamma_{x_i,y_j}} (\delta_{y_j} - \delta_{x_i}) \, d\eta = \left(\sum_{x_i \in B_j} \eta(\Gamma_{x_i,y_j}) \right) \delta_{y_j} - \sum_{x_i \in B_j} \eta(\Gamma_{x_i,y_j}) \delta_{x_i}.$$

Note that when $x_i \in B_j$, since $\{B_k\}$'s are pairwise disjoint, it follows that $\eta(\Gamma_{x_i,y_k}) = 0$ for all $k \ne j$. So,

$$\sum_{x_i \in B_j} \eta(\Gamma_{x_i,y_j}) \delta_{x_i} = \sum_{x_i \in B_j} \left(\sum_{k=1}^N \eta(\Gamma_{x_i,y_k}) \right) \delta_{x_i} = \sum_{x_i \in B_j} \mu^-(\{x_i\}) \delta_{x_i} = \mu^- \lfloor_{B_j} = \mu_j^-,$$

and

$$\left(\sum_{x_i\in B_j}\eta(\Gamma_{x_i,y_j})\right)\delta_{y_j}=\mu^-(B_j)\delta_{y_j}=\mu_j^+.$$

As a result, $\partial T_j=\mu_j^+-\mu_j^-$.

Moreover, we have the result,

$$\begin{aligned}
\partial T_0 &= \sum_{j=1}^{N}\sum_{x_i\in B_0}\int_{\Gamma_{x_i,y_j}}(\delta_{y_j}-\delta_{x_i})\,d\eta \\
&= \sum_{j=1}^{N}\left(\sum_{x_i\in B_0}\eta(\Gamma_{x_i,y_j})\right)\delta_{y_j}-\sum_{x_i\in B_0}\left(\sum_{j=1}^{N}\eta(\Gamma_{x_i,y_j})\right)\delta_{x_i} \\
&= \sum_{j=1}^{N}\left(\sum_{x_i\in B_0\cap X_j(\eta)}\eta(\Gamma_{x_i,y_j})\right)\delta_{y_j}-\sum_{x_i\in B_0}\mu^-(\{x_i\})\delta_{x_i} \\
&= \sum_{j=1}^{N}\left(\sum_{x_i\in X_j(\eta)}\eta(\Gamma_{x_i,y_j})-\sum_{x_i\in B_j}\eta(\Gamma_{x_i,y_j})\right)\delta_{y_j}-\mu^-\lfloor_{B_0} \\
&= \sum_{j=1}^{N}\left(\sum_{i=1}^{M}\eta(\Gamma_{x_i,y_j})-\sum_{x_i\in B_j}\eta(\Gamma_{x_i,y_j})\right)\delta_{y_j}-\mu^-\lfloor_{B_0} \\
&= \sum_{j=1}^{N}\big(m_j-\mu^-(B_j)\big)\delta_{y_j}-\mu^-\lfloor_{B_0} \\
&= \mu_0^+-\mu_0^-.
\end{aligned} \tag{3.4.7}$$

□

3.5. Transport paths induced transport maps and transport plans

In this section, we will decompose a cycle-free transport path into the sum of two transport paths, the first one is induced by a compatible transport map, while the second one is induced by a compatible transport plan. We first recall the concept of compatibility introduced in [**11**, Definition 7.1], and rewrite it in terms of our current contexts.

Suppose μ^- and μ^+ are two atomic measures of equal finite mass as given in (3.2.10). Let $Path_0(\mu^-,\mu^+)$ denote the family of all cycle-free transport paths from μ^- to μ^+.

Remark 3.5.1. In [**11**, Definition 7.1], we used $Path_0(\mu^-,\mu^+)$ to denote the family of all "acyclic" transport paths from μ^- to μ^+. In [**11**], a transport path G is called "acyclic" if it satisfies the following condition: *for any polyhedral 1-chain* $\tilde{G}$ *with the support of* $\tilde{G}$ *contained in the support of* G*, if* $\partial\tilde{G} = 0$ *then* $\tilde{G} = 0$. In the current context, G is an "acyclic" transport path simply means that it is cycle-free. To avoid confusion between the term "acyclic" used in [**11**] and the acyclic concept defined using subcurrents in [**6**], we opt for the term "cycle-free" to name the term "acyclic" used in [**11**].

Observe that for any $G \in Path_0(\mu^-, \mu^+)$ and for each x_i and y_j, there exists at most one directed polyhedral curve g_{ij} from x_i to y_j, supported on the support of G. Thus, we associate each $G \in Path_0(\mu^-, \mu^+)$ with a $M \times N$ polyhedral 1-chain valued matrix $g = \left[I_{g_{ij}}\right]$, such that $I_{g_{ij}} = 0$ when g_{ij} does not exist.

Definition 3.5.2. ([**11**, Definition 7.1]) Let $G \in Path_0(\mu^-, \mu^+)$ and $q \in Plan(\mu^-, \mu^+)$ with associated matrices $\left[I_{g_{ij}}\right]$ and $\left[q_{ij}\right]$ respectively. The pair (G, q) is called *compatible* if $q_{ij} = 0$ whenever $I_{g_{ij}} = 0$ and

$$G = \sum_{i=1}^{M}\sum_{j=1}^{N} q_{ij} I_{g_{ij}} \text{ and } q = \sum_{i=1}^{M}\sum_{j=1}^{N} q_{ij}\delta_{(x_i,y_j)} \tag{3.5.1}$$

as polyhedral 1-chains.

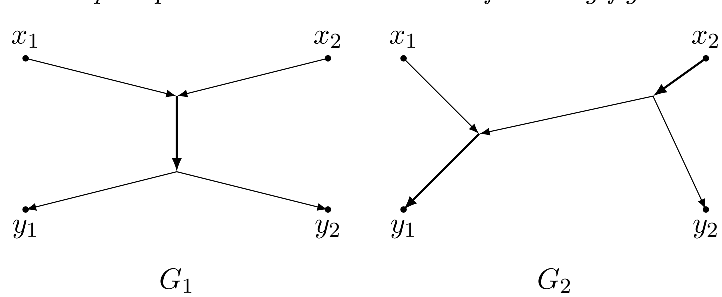
Example 9. *For instance, let*

$$\mu^- = \frac{1}{4}\delta_{x_1} + \frac{3}{4}\delta_{x_2},\ \ \mu^+ = \frac{5}{8}\delta_{y_1} + \frac{3}{8}\delta_{y_2},$$

and consider the following transport plan,

$$q = \frac{1}{8}\delta_{(x_1,y_1)} + \frac{1}{8}\delta_{(x_1,y_2)} + \frac{1}{2}\delta_{(x_2,y_1)} + \frac{1}{4}\delta_{(x_2,y_2)} \in Plan(\mu^-, \mu^+).$$

Let G_1 *and* G_2 *be two transport paths as illustrated in the following figure.*



Then (G_1, q) is compatible but (G_2, q) is not, since $q_{12} = \frac{1}{8} \neq 0$ and there is no directed curve g_{12} from x_1 to y_2 on the support of G_2.

Now, we generalize the compatibility of atomic measures μ^-, μ^+ stated above to those of general measures.

Definition 3.5.3. Let μ and ν be two Radon measures on X of equal total mass. Given $T \in Path(\mu, \nu)$, and $\pi \in Plan(\mu, \nu)$, we say the pair (T, π) is compatible if there exists a finite Borel measure η on Γ such that

$$T = \int_\Gamma I_\gamma d\eta, \text{ and } \pi = \int_\Gamma \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta.$$

Moreover, given $T \in Path(\mu, \nu)$ and $\varphi \in Map(\mu, \nu)$, we say the pair (T, φ) is compatible if (T, π_φ) is compatible, where $\pi_\varphi = (id \times \varphi)_\# \mu$.

The following Proposition says that Definition 3.5.3 is a generalization of Definition 3.5.2.

Proposition 3.5.4. *Let μ^- and μ^+ be two atomic measures of equal mass as given in (3.2.10). Let $G \in Path_0(\mu^-, \mu^+)$ and $q \in Plan(\mu^-, \mu^+)$. Then (G, q) is compatible in the sense of Definition 3.5.2 if and only if (G, q) is compatible in the sense of Definition 3.5.3.*

Proof. Suppose (G, q) is compatible in the sense of Definition 3.5.2. By setting

$$\eta = \sum_{i=1}^{M} \sum_{j=1}^{N} q_{ij} \delta_{g_{ij}}$$

over all $\{1 \leq i \leq M, 1 \leq j \leq N\}$ with g_{ij} exists, equation (3.5.1) gives that

$$G = \int_\Gamma I_\gamma d\eta \text{ and } q = \int_\Gamma \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta.$$

Therefore, (G, q) is also compatible in the sense of Definition 3.5.3.

On the other hand, suppose (G, q) is compatible in the sense of Definition 3.5.3, then there exists a Borel measure η on Γ such that

$$G = \int_\Gamma I_\gamma d\eta \text{ and } q = \int_\Gamma \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta.$$

Since $q \in Plan(\mathbf{a}, \mathbf{b})$, we may write

$$q = \sum_{i=1}^{M} \sum_{j=1}^{N} q_{ij} \delta_{(x_i, y_j)}$$

for some $q_{ij} \geq 0$. Denote

$$J_q := \{(i,j) : 1 \leq i \leq M, 1 \leq j \leq N, \text{ with } q_{ij} > 0\}.$$

and

$$\tilde{\Gamma} := \bigcup_{(i,j) \in J_q} \Gamma_{x_i, y_j}.$$

Since

$$\begin{aligned} \int_{\Gamma \setminus \tilde{\Gamma}} \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta + \int_{\tilde{\Gamma}} \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta &= \int_{\Gamma} \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta \\ &= q = \sum_{i=1}^{M} \sum_{j=1}^{N} q_{ij} \delta_{(x_i, y_j)} = \sum_{(i,j) \in J_q} q_{ij} \delta_{(x_i, y_j)}, \end{aligned}$$

it follows that

$$\int_{\Gamma \setminus \tilde{\Gamma}} \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta = 0 \text{ and } \int_{\tilde{\Gamma}} \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta = \sum_{(i,j) \in J_q} q_{ij} \delta_{(x_i, y_j)}.$$

Thus, $\eta(\Gamma \setminus \tilde{\Gamma}) = 0$ and

$$q = \sum_{(i,j) \in J_q} \int_{\Gamma_{x_i, y_j}} \delta_{(p_0(\gamma), p_\infty(\gamma))} d\eta = \sum_{(i,j) \in J_q} q_{ij} \delta_{(x_i, y_j)}.$$

Hence for each $1 \leq i \leq M, 1 \leq j \leq N$,

$$\eta(\Gamma_{x_i, y_j}) = q_{ij} \text{ if } (i,j) \in J_q \text{ and } \eta(\Gamma_{x_i, y_j}) = 0 \text{ if not.}$$

Now, for each $(i,j) \in J_q$, since $\eta(\Gamma_{x_i, y_j}) = q_{ij} > 0$ and

$$G = \int_{\Gamma} I_\gamma d\eta = \sum_{(i,j) \in J_q} \int_{\Gamma_{x_i, y_j}} I_\gamma d\eta,$$

it follows that there exists a polyhedral 1-curve g_{ij} supported on the support of G. Let

$$\tilde{G} = \sum_{(i,j)\in J_q} q_{ij} I_{g_{ij}},$$

then

$$\partial(G - \tilde{G}) = \partial\left(\sum_{(i,j)\in J_q} \int_{\Gamma_{x_i,y_j}} I_\gamma d\eta - \sum_{(i,j)\in J_q} q_{ij} I_{g_{ij}}\right) = \sum_{(i,j)\in J_q} \left(\eta\left(\Gamma_{x_i,y_j}\right) - q_{ij}\right)\left(\delta_{y_j} - \delta_{x_i}\right) = 0,$$

so that $G - \tilde{G}$ is a cycle supported on the support of G. Since $G \in Path_0(\mu^-, \mu^+)$, we have $G - \tilde{G} = 0$. Therefore,

$$G = \tilde{G} = \sum_{(i,j)\in J_q} q_{ij} I_{g_{ij}}.$$

Note also that whenever $I_{g_{ij}} = 0$, it follows that $(i,j) \notin J_q$, and thus $q_{ij} = 0$. As a result, (G, q) is compatible in the sense of Definition 3.5.2. $\square$

PROPOSITION 3.5.5. *Let μ^- and μ^+ be two atomic measures of equal mass as given in (3.2.10), $T \in Path(\mu^-, \mu^+)$ is a optimal transport paths, and let η be a good decomposition of T. Suppose for any $1 \le j_1 < j_2 \le N$, $|X_{j_1}(\eta) \cap X_{j_2}(\eta)| = 0$, then there exists a transport map φ, such that (T, φ) is compatible.*

PROOF. We first recall the definition of $X_j(\eta)$,

$$X_j(\eta) := \{x_i \in X : \eta(\Gamma_{x_i,y_j}) > 0\}.$$

In this case, we may define

$$\varphi : \bigcup_{j=1}^{N} X_j(\eta) \to \{y_1, y_2, \ldots, y_N\}, \text{ such that for } x \in X_j(\eta),\ \varphi(x) := y_j.$$

Since for any $1 \le j_1 < j_2 \le N$, $|X_{j_1}(\eta) \cap X_{j_2}(\eta)| = 0$, the function φ defined above is well-defined.

Direct calculation gives that for each $j = 1, 2, \ldots, N$,

$$\varphi_{\#}\mu^-(\{y_j\}) = \mu^-(X_j(\eta)) = \sum_{i=1}^{N} \eta(\Gamma_{x_i,y_j}) = \mu^+(\{y_j\}),$$

which implies $\varphi_{\#}\mu^- = \mu^+$. Therefore, we have

$$\begin{aligned}(id\times\varphi)_{\#}\mu^- &= \sum_{i=1}^{M}\sum_{j=1}^{N}\eta(\Gamma_{x_i,y_j})(id\times\varphi)_{\#}\delta_{x_i} = \sum_{j=1}^{N}\sum_{x_i\in X_j(\eta)}\eta(\Gamma_{x_i,y_j})\delta_{(x_i,y_j)}\\ &= \sum_{j=1}^{N}\sum_{x_i\in X_j(\eta)}\int_{\Gamma_{x_i,y_j}}\delta_{(p_0(\gamma),p_\infty(\gamma))}\,d\eta = \sum_{j=1}^{N}\sum_{i=1}^{M}\int_{\Gamma_{x_i,y_j}}\delta_{(p_0(\gamma),p_\infty(\gamma))}\,d\eta,\end{aligned}$$

which implies (T,φ) is compatible. □

By Theorem 3.4.8, we now have the following theorem:

THEOREM 3.5.6. *Let $T \in Path(\mu^-,\mu^+)$ be a cycle-free transport path, where μ^- and μ^+ are given in (3.2.10). Then there exist*

(a) *decomposition*

$$\mu^- = \mu^-_\pi + \mu^-_\varphi,\ \mu^+ = \mu^+_\pi + \mu^+_\varphi,\ \textit{with}\ \mu^-_\pi(X) = \mu^+_\pi(X),\ \mu^-_\varphi(X) = \mu^+_\varphi(X)$$

where μ^-_π and μ^-_φ have disjoint supports and $|spt(\mu^-_\pi)| \le \binom{N}{2}$ with $|A|$ denoting the cardinality of the set A;

(b) *$T = T_\pi + T_\varphi$ for some $T_\pi \in Path\,(\mu^-_\pi,\mu^+_\pi)$ and $T_\varphi \in Path\,(\mu^-_\varphi,\mu^+_\varphi)$. Both T_π and T_φ are subcurrents of T;*

(c) *a transport map $\varphi \in Map\,(\mu^-_\varphi,\mu^+_\varphi)$ such that (T_φ,φ) is compatible;*

(d) *a transport plan $\pi \in Plan\,(\mu^-_\pi,\mu^+_\pi)$ such that (T_π,π) is compatible;*

(e) *For each x_i with $\mu^-_\pi(\{x_i\}) > 0$, there are at least two y_{j_1}, y_{j_2}, such that*

$$\pi(\{x_i\}\times\{y_{j_1}\}) > 0, \pi(\{x_i\}\times\{y_{j_2}\}) > 0.$$

PROOF. We continue with the same notations used in Theorem 3.4.8. Part (a),(b) follows from (3.4.4) and (3.4.5) by setting

$$\mu^-_\pi := \mu^-_0,\ \mu^-_\varphi := \sum_{j=1}^{N}\mu^-_j,\ \mu^+_\pi := \mu^+_0,\ \mu^+_\varphi := \sum_{j=1}^{N}\mu^+_j,\ T_\pi := T_0,\ T_\varphi := \sum_{j=1}^{N}T_j.$$

For part (c), we define

$$\varphi := \sum_{j=1}^{N} y_j\chi_{B_j},$$

where B_j's are subsets of $\{x_1, x_2, \cdots, x_M\}$ given in Theorem 3.4.8. Since $\mu_j^- = \mu^- \lfloor_{B_j}$, $\mu_j^+ = \tilde{m}_j \delta_{y_j}$, and B_j's are pairwise disjoint for $j = 1, 2, \ldots, N$, we get

$$\varphi_{\#}(\mu_\varphi^-) = \varphi_{\#}\left(\sum_{j=1}^N \mu_j^-\right) = \varphi_{\#}\left(\sum_{j=1}^N \mu^- \lfloor_{B_j}\right) = \sum_{j=1}^N \mu^-(B_j)\delta_{y_j} = \sum_{j=1}^N \tilde{m}_j \delta_{y_j} = \mu_\varphi^+.$$

Therefore, φ is a transport map from μ_φ^- to μ_φ^+.

We now show that (T_φ, φ) is compatible. Since

$$T_\varphi := \sum_{j=1}^N T_j = \sum_{j=1}^N \sum_{x_i \in B_j} \int_{\Gamma_{x_i, y_j}} I_\gamma \, d\eta,$$

it is sufficient to show that

$$\pi_\varphi := (id \times \varphi)_{\#}\, \mu^- = \sum_{j=1}^N \sum_{x_i \in B_j} \int_{\Gamma_{x_i, y_j}} \delta_{(p_0(\gamma), p_\infty(\gamma))} \, d\eta. \tag{3.5.2}$$

Indeed, for any measurable rectangle $Q \times R$ in $X \times X$,

$$\begin{aligned}
\pi_\varphi(Q \times R) &= (id \times \varphi)_{\#}\mu^-(Q \times R) = \mu^-(\{x : x \in Q, \varphi(x) \in R\}) \\
&= \sum_{j=1}^N \mu^-(\{x : x \in Q, \varphi(x) = y_j, y_j \in R\}) = \sum_{j=1}^N \chi_R(y_j)\mu^-(\{x : x \in Q, \varphi(x) = y_j\}) \\
&= \sum_{j=1}^N \chi_R(y_j)\mu^-(\{x : x \in Q, x \in B_j\}) = \sum_{j=1}^N \chi_R(y_j)\mu^-(Q \cap B_j) \\
&= \sum_{j=1}^N \chi_R(y_j)\left((p_0)_{\#}\eta\right)(Q \cap B_j) = \sum_{j=1}^N \chi_R(y_j)\eta(p_0^{-1}(Q \cap B_j)) \\
&= \sum_{j=1}^N \chi_R(y_j)\eta(\{\gamma \in \Gamma, p_0(\gamma) \in Q \cap B_j\}) = \sum_{j=1}^N \chi_R(y_j) \sum_{x_i \in B_j} \int_{\Gamma_{x_i, y_j}} \chi_Q(p_0(\gamma)) d\eta \\
&= \sum_{j=1}^N \sum_{x_i \in B_j} \int_{\Gamma_{x_i, y_j}} \chi_Q(p_0(\gamma)) \cdot \chi_R(y_j) d\eta = \sum_{j=1}^N \sum_{x_i \in B_j} \int_{\Gamma_{x_i, y_j}} \chi_Q(p_0(\gamma)) \cdot \chi_R(p_\infty(\gamma)) d\eta \\
&= \sum_{j=1}^N \sum_{x_i \in B_j} \int_{\Gamma_{x_i, y_j}, x_i \in Q, y_j \in R} \delta_{p_0(\gamma)} \cdot \delta_{p_\infty(\gamma)} d\eta \\
&= \sum_{j=1}^N \sum_{x_i \in B_j} \int_{\Gamma_{x_i, y_j}} \delta_{(p_0(\gamma), p_\infty(\gamma))} \, d\eta(Q \times R).
\end{aligned}$$

Therefore, (3.5.2) holds and hence (T_φ, φ) is compatible.

For part (d), we define

$$\pi := \sum_{x_i \in B_0} \sum_{j=1}^{N} \eta\left(\Gamma_{x_i,y_j}\right) \delta_{(x_i,y_j)}.$$

As shown in (3.4.7),

$$\mu_\pi^+ - \mu_\pi^- = \mu_0^+ - \mu_0^- = \sum_{j=1}^{N} \left(\sum_{x_i \in B_0} \eta(\Gamma_{x_i,y_j}) \right) \delta_{y_j} - \sum_{x_i \in B_0} \left(\sum_{j=1}^{N} \eta(\Gamma_{x_i,y_j}) \right) \delta_{x_i}.$$

This shows that π is a transport plan from μ_π^- to μ_π^+. Note that since

$$T_0 = \sum_{j=1}^{N} \sum_{x_i \in B_0} \int_{\Gamma_{x_i,y_j}} I_\gamma \, d\eta$$

and

$$\pi = \sum_{x_i \in B_0} \sum_{j=1}^{N} \eta\left(\Gamma_{x_i,y_j}\right) \delta_{(x_i,y_j)} = \sum_{j=1}^{N} \sum_{x_i \in B_0} \int_{\Gamma_{x_i,y_j}} \delta_{(p_0(\gamma),p_\infty(\gamma))} \, d\eta,$$

we have (T_π, π) is compatible.

For part (e), by definition of μ_π^-, $x_i \in B_0$ which is defined in Theorem 3.4.8. The result in (e) then follows from the definition of B_0 given in (3.4.6). $\square$

3.6. Stair-shaped matrices and decomposition of stair-shaped transport paths

In Theorem 3.5.6, we decomposed a cycle-free transport path as the sum of a map-compatible path and a plan-compatible path. In this section, we aim to decompose some transport paths as the difference of two map-compatible paths. The family of transport paths that we are interested in are stair-shaped transport paths. To do this, we start with the study of stair-shaped matrices.

3.6.1. Stair-shaped matrices.

Given $M, N \in \mathbb{N} \cup \{\infty\}$, let $\mathcal{A}_{M,N}$ denote the collection of all $M \times N$ matrices with non-negative entries.

Definition 3.6.1. A matrix $A \in \mathcal{A}_{M,N}$ is called stair-shaped if there exists two non-decreasing sequences of natural numbers $\{r_1, r_2, \cdots, r_{M+N-1}\}$ and $\{c_1, c_2, \cdots, c_{M+N-1}\}$ with $r_k + c_k = k+1$ for each $k = 1, 2, \cdots, M+N-2$, and entries of A that are not located in the positions $\{(r_k, c_k)\}_{k=1}^{M+N-1}$ equal to zero.

Note that when $A \in \mathcal{A}_{M,N}$ is stair-shaped, then $(r_1, c_1) = (1,1)$ and $(r_{M+N-1}, c_{M+N-1}) = (M, N)$.

Definition 3.6.2. For each $k = 1, 2, \cdots, M+N-1$, a matrix $A \in \mathcal{A}_{M,N}$ is called k-stairable if it is in the form of

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1,c-1} & a_{1,c} & 0 & \cdots & 0 & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \\ a_{r-1,1} & \cdots & a_{r-1,c-1} & a_{r-1,c} & 0 & \cdots & 0 & \cdots \\ a_{r,1} & \cdots & a_{r,c-1} & a_{r,c} & a_{r,c+1} & \cdots & a_{r,j} & \cdots \\ 0 & \cdots & 0 & a_{r+1,c} & a_{r+1,c+1} & \cdots & a_{r+1,j} & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \\ 0 & \cdots & 0 & a_{i,c} & a_{i,c+1} & \cdots & a_{i,j} & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \end{bmatrix},$$

where the leading (i.e., upper left corner) sub-matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1,c-1} & a_{1,c} \\ \vdots & & \vdots & \vdots \\ a_{r-1,1} & \cdots & a_{r_k-1,c-1} & a_{r-1,c} \\ a_{r,1} & \cdots & a_{r,c-1} & a_{r,c} \end{bmatrix}$$

is stair-shaped and $k = r + c - 1$.

In particular, each matrix $A \in \mathcal{A}_{M,N}$ is at least 1-stairable, and each stair-shaped matrix $A \in \mathcal{A}_{M,N}$ is $(M+N-1)$-stairable.

For each $1 \le i_1 < i_2 \le M$ and $1 \le j_1 < j_2 \le N$, denote $E[(i_1, j_1), (i_2, j_2)]$ as the $M \times N$ matrix with 1 at (i_1, j_1) and (i_2, j_2) entries, with -1 at (i_1, j_2) and (i_2, j_1) entries, and 0 at all other entries. Each $E[(i_1, j_1), (i_2, j_2)]$ is called an elementary matrix.

Definition 3.6.3. For any two matrices $A, B \in \mathcal{A}_{M,N}$, we say $A \cong B$ if there exists a list of real numbers $\{t_k\}_{k=1}^K$ and a list of elementary matrices $\{E_k\}_{k=1}^K$ such that $B = A + \sum_{k=1}^K t_k E_k$ for some $K \in \mathbb{N} \cup \{\infty\}$.

THEOREM 3.6.4. *For any matrix $A \in \mathcal{A}_{M,N}$, there exists a stair-shaped matrix $B \in \mathcal{A}_{M,N}$ such that $A \cong B$.*

PROOF. **Step 1:** Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots \\ \vdots & \vdots & & \vdots & \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix},$$

and

$$u_1 = \sum_{i=2}^{M} a_{i1} \text{ and } v_1 = \sum_{j=2}^{N} a_{1j}.$$

If $u_1 = 0$, and since all entries in A are non-negative, then we get

$$A_1 = A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots \\ 0 & a_{22} & \cdots & a_{2j} & \cdots \\ \vdots & \vdots & & \vdots & \\ 0 & a_{i2} & \cdots & a_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix}.$$

If $u_1 \neq 0$, and $u_1 \geq v_1$ then we do the following transformation and denote

$$A_1 = A + \sum_{i=2}^{\infty}\sum_{j=2}^{\infty} \frac{a_{i1}a_{1j}}{u_1} E[(1,1),(i,j)].$$

This implies

$$A_1 = \begin{bmatrix} a_{11} + \sum_{i=2}^{\infty}\sum_{j=2}^{\infty} \frac{a_{i1}a_{1j}}{u_1} & a_{12} - \sum_{i=2}^{\infty} \frac{a_{i1}a_{12}}{u_1} & \cdots & a_{1j} - \sum_{i=2}^{\infty} \frac{a_{i1}a_{1j}}{u_1} & \cdots \\ a_{21} - \sum_{j=2}^{\infty} \frac{a_{21}a_{1j}}{u_1} & a_{22} + \frac{a_{21}a_{12}}{u_1} & \cdots & a_{2j} + \frac{a_{21}a_{1j}}{u_1} & \cdots \\ \vdots & \vdots & & \vdots & \\ a_{i1} - \sum_{j=2}^{\infty} \frac{a_{i1}a_{1j}}{u_1} & a_{i2} + \frac{a_{i1}a_{12}}{u_1} & \cdots & a_{ij} + \frac{a_{i1}a_{1j}}{u_1} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+v_1 & 0 & \cdots & 0 & \cdots \\ \left(1-\frac{v_1}{u_1}\right)a_{21} & a_{22}+\frac{a_{21}a_{12}}{u_1} & \cdots & a_{2j}+\frac{a_{21}a_{1j}}{u_1} & \cdots \\ \vdots & \vdots & & \vdots & \\ \left(1-\frac{v_1}{u_1}\right)a_{i1} & a_{i2}+\frac{a_{i1}a_{12}}{u_1} & \cdots & a_{ij}+\frac{a_{i1}a_{1j}}{u_1} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix}.$$

If $u_1 \neq 0$, and $u_1 \leq v_1$, we consider the following transformation:

$$A_1 = A + \sum_{i=2}^{\infty}\sum_{j=2}^{\infty} \frac{a_{i1}a_{1j}}{v_1} E[(1,1),(i,j)],$$

and

$$A_1 = \begin{bmatrix} a_{11}+\sum_{i=2}^{\infty}\sum_{j=2}^{\infty}\frac{a_{i1}a_{1j}}{v_1} & a_{12}-\sum_{i=2}^{\infty}\frac{a_{i1}a_{12}}{v_1} & \cdots & a_{1j}-\sum_{i=2}^{\infty}\frac{a_{i1}a_{1j}}{v_1} & \cdots \\ a_{21}-\sum_{j=2}^{\infty}\frac{a_{21}a_{1j}}{v_1} & a_{22}+\frac{a_{21}a_{12}}{v_1} & \cdots & a_{2j}+\frac{a_{21}a_{1j}}{v_1} & \cdots \\ \vdots & \vdots & & \vdots & \\ a_{i1}-\sum_{j=2}^{\infty}\frac{a_{i1}a_{1j}}{v_1} & a_{i2}+\frac{a_{i1}a_{12}}{v_1} & \cdots & a_{ij}+\frac{a_{i1}a_{1j}}{v_1} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+u_1 & \left(1-\frac{u_1}{v_1}\right)a_{12} & \cdots & \left(1-\frac{u_1}{v_1}\right)a_{1j} & \cdots \\ 0 & a_{22}+\frac{a_{21}a_{12}}{v_1} & \cdots & a_{2j}+\frac{a_{21}a_{1j}}{v_1} & \cdots \\ \vdots & \vdots & & \vdots & \\ 0 & a_{i2}+\frac{a_{i1}a_{12}}{v_1} & \cdots & a_{ij}+\frac{a_{i1}a_{1j}}{v_1} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix}.$$

Hence, $A \cong A_1$ where A_1 is of the form:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots \\ 0 & a_{22} & \cdots & a_{2j} & \cdots \\ \vdots & \vdots & & \vdots & \\ 0 & a_{i2} & \cdots & a_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix} \text{ or } \begin{bmatrix} a_{11} & 0 & \cdots & 0 & \cdots \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots \\ \vdots & \vdots & & \vdots & \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix}$$

and $(r_1, c_1) = (1, 1)$. Here and in the following steps, for simplicity of notations, we continue using the same notation, a_{ij}'s, to denote non-negative entries.

Step 2: Set $A_1 = f(A)$, note that $A_1 \cong A$ is 1-stairable. For each $k \in \mathbb{N}$, if $A_k \cong A$ is k-stairable, we construct a $(k+1)$-stairable matrix $A_{k+1} \cong A$ as follows. Given

$$A_k = \begin{bmatrix} a_{11} & \cdots & a_{1,c_k-1} & a_{1,c_k} & 0 & \cdots & 0 & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \\ a_{r_k-1,1} & \cdots & a_{r_k-1,c_k-1} & a_{r_k-1,c_k} & 0 & \cdots & 0 & \cdots \\ a_{r_k 1} & \cdots & a_{r_k c_k-1} & a_{r_k c_k} & a_{r_k,c_k+1} & \cdots & a_{r_k,j} & \cdots \\ 0 & \cdots & 0 & a_{r_k+1,c_k} & a_{r_k+1,c_k+1} & \cdots & a_{r_k+1,j} & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \\ 0 & \cdots & 0 & a_{i,c_k} & a_{i,c_k+1} & \cdots & a_{ij} & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \end{bmatrix},$$

where the upper left corner sub-matrix

$$S = \begin{bmatrix} a_{11} & \cdots & a_{1,c_k-1} & a_{1,c_k} \\ \vdots & & \vdots & \vdots \\ a_{r_k-1,1} & \cdots & a_{r_k-1,c_k-1} & a_{r_k-1,c_k} \\ a_{r_k 1} & \cdots & a_{r_k c_k} & a_{r_k c_k} \end{bmatrix}$$

is stair-shaped (which implies that $r_k + c_k - 1 = k$), $S \in \mathcal{A}_{r_k,c_k}$, and let

$$B = f\left(\begin{bmatrix} a_{r_k c_k} & a_{r_k,c_k+1} & \cdots & a_{r_k,j} & \cdots \\ a_{r_k+1,c_k} & a_{r_k+1,c_k+1} & \cdots & a_{r_k+1,j} & \cdots \\ \vdots & \vdots & & \vdots & \\ a_{i,c_k} & a_{i,c_k+1} & \cdots & a_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix} \right) = \begin{bmatrix} b_{r_k c_k} & b_{r_k,c_k+1} & \cdots & b_{r_k,j} & \cdots \\ b_{r_k+1,c_k} & b_{r_k+1,c_k+1} & \cdots & b_{r_k+1,j} & \cdots \\ \vdots & \vdots & & \vdots & \\ b_{i,c_k} & b_{i,c_k+1} & \cdots & b_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \end{bmatrix}.$$

Then we define

$$A_{k+1} = \begin{bmatrix} a_{11} & \cdots & a_{1,c_k-1} & a_{1,c_k} & 0 & \cdots & 0 & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \\ a_{r_k-1,1} & \cdots & a_{r_k-1,c_k-1} & a_{r_k-1,c_k} & 0 & \cdots & 0 & \cdots \\ a_{r_k 1} & \cdots & a_{r_k c_k-1} & b_{r_k c_k} & b_{r_k,c_k+1} & \cdots & b_{r_k,j} & \cdots \\ 0 & \cdots & 0 & b_{r_k+1,c_k} & b_{r_k+1,c_k+1} & \cdots & b_{r_k+1,j} & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \\ 0 & \cdots & 0 & b_{i,c_k} & b_{i,c_k+1} & \cdots & b_{ij} & \cdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \end{bmatrix}.$$

By definition of f, two sequences $(r_k)_{k=1}^{\infty}$ and $(c_k)_{k=1}^{\infty}$ can be constructed as follows:

(1) If

$$\begin{bmatrix} b_{r_k,c_k+1} & \dots & b_{r_k,j} & \dots \end{bmatrix} \neq \begin{bmatrix} 0 & \dots & 0 & \dots \end{bmatrix},$$

then $(r_{k+1}, c_{k+1}) = (r_k, c_k + 1)$;

(2) If

$$\begin{bmatrix} b_{r_k,c_k+1} & \dots & b_{r_k,j} & \dots \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & \dots \end{bmatrix}$$

and

$$\begin{bmatrix} b_{r_k+1,c_k} & \dots & b_{i,c_k} & \dots \end{bmatrix}^T \neq \begin{bmatrix} 0 & \dots & 0 & \dots \end{bmatrix}^T,$$

then $(r_{k+1}, c_{k+1}) = (r_k + 1, c_k)$;

(3) If

$$\begin{bmatrix} b_{r_k,c_k+1} & \dots & b_{r_k,j} & \dots \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & \dots \end{bmatrix}$$

and

$$\begin{bmatrix} b_{r_k+1,c_k} & \dots & b_{i,c_k} & \dots \end{bmatrix}^T = \begin{bmatrix} 0 & \dots & 0 & \dots \end{bmatrix}^T,$$

then $(r_{k+1}, c_{k+1}) = (r_k, c_k + 1)$.

This gives $(r_k)_{k=1}^{\infty}$, $(c_k)_{k=1}^{\infty}$ are non-decreasing sequences with $r_{k+1} + c_{k+1} = r_k + c_k + 1 = k + 2$. By doing so, we get a $(k + 1)$-stairable matrix A_{k+1} with $A \cong A_k \cong A_{k+1}$. Note that in this construction we have

(3.6.1) $$A_{k+1}(i, j) = A_k(i, j), \quad \text{for } i < r_k \text{ or } j < c_k.$$

Moreover,

$$A_{k+1} = A_k + \sum_{l=1}^{\infty} t_{k,l} E_{k,l}$$

for some $t_{k,l} \in \mathbb{R}$, and $E_{k,l}$'s are elementary matrices. Set

$$A_\infty := A_1 + \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} t_{k,l} E_{k,l},$$

then

$$A_\infty \cong A_1 \cong A.$$

Note that for each $i = 1, \ldots, M$ and $j = 1, \ldots, N$, by (3.6.1) and $r_k + c_k - 1 = k$, the sequence $A_k(i,j)$ is eventually constant when k is large enough. Thus, $A_\infty(i,j) = \lim_{k\to\infty} A_k(i,j)$ is well defined, stair-shaped, with non-negative entries. $\square$

After knowing the existence of the stair-shaped matrix B using Theorem 3.6.4, one may use the following algorithm to recursively find its entries.

Algorithm 3.6.5.

Input: A matrix $A = [a_{ij}] \in \mathcal{A}_{M,N}$.

Output: A stair-shaped matrix $B = [b_{ij}] \in \mathcal{A}_{M,N}$ with $B \cong A$.

Algorithm: One may recursively calculate the entries of B as follows:

- Step 1: Start with $i_0 = 1, j_0 = 1$, set

  $$R = \sum_{j=1}^{N} a_{1j} \text{ and } C = \sum_{i=1}^{M} a_{i1}.$$

  If $R \leq C$, then $b_{11} = R,\ b_{1j} = 0$ for all $j > 1$. Otherwise, $b_{11} = C$ and $b_{i1} = 0$ for all $i > 1$.

- Step 2: For each (i_0, j_0) with b_{i_0,j_0} unknown and b_{ij} is known for all $i < i_0$ and $j < j_0$, let

  $$R = \sum_{j=1}^{N} a_{i_0,j} - \sum_{j<j_0} b_{i_0,j},\ C = \sum_{i=1}^{M} a_{i,j_0} - \sum_{i<i_0} b_{i,j_0}.$$

  If $R \leq C$, set

  $$b_{i_0,j_0} = R,\ b_{i_0,j} = 0 \text{ for all } j > j_0.$$

Otherwise, when $R > C$, set

$$b_{i_0,j_0} = C,\ b_{i,j_0} = 0 \text{ for all } i > i_0.$$

Using Step 2 recursively, one can calculate all entries of the stair-shaped matrix B.

3.6.2. Stair-shaped good decomposition.

DEFINITION 3.6.6. Let η be a finite measure on Γ with $(p_0)_\#\eta = \mu^-$ and $(p_\infty)_\#\eta = \mu^+$. The representing matrix of η is the matrix $A = [a_{ij}] \in \mathcal{A}_{M,N}$ such that $a_{ij} = \eta(\Gamma_{x_i,y_j})$ for each i, j. We say that η is stair-shaped if its representing matrix A is stair-shaped. A transport path $T \in Path(\mu^-, \mu^+)$ is called stair-shaped if there exists a good decomposition η of T such that η is stair-shaped.

PROPOSITION 3.6.7. *Any stair-shaped good decomposition η of T is a better decomposition of T.*

PROOF. By Definition 3.3.1, suppose there exist $1 \leq i_1 < i_2 \leq M$ and $1 \leq j_1 < j_2 \leq N$, with

$$S_{i_1,j_1}(\eta) - S_{i_1,j_2}(\eta) - S_{i_2,j_1}(\eta) + S_{i_2,j_2}(\eta) = 0,$$

then direct calculation from (3.3.7) gives either

$$\eta(\Gamma_{i_1,j_1}) = \eta(\Gamma_{i_1,j_2}) = \eta(\Gamma_{i_2,j_1}) = \eta(\Gamma_{i_2,j_2}) = 0,$$

or

$$\eta(\Gamma_{i_1,j_1}) > 0, \eta(\Gamma_{i_1,j_2}) > 0, \eta(\Gamma_{i_2,j_1}) > 0, \eta(\Gamma_{i_2,j_2}) > 0.$$

The latter case cannot appear since η is stair-shaped and there is no way to align the indexes

$$(i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2),$$

such that both two coordinates are non-decreasing sequences. As a result, η is a better decomposition.

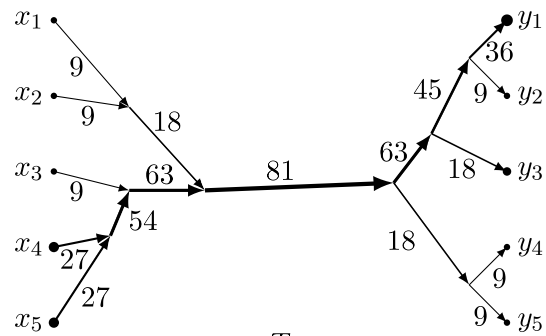
□

A stair-shaped path is not necessarily cycle-free. For instance, the transport path T given in Remark 3.4.3 is stair-shaped because $\eta = \delta_{\gamma_{x_1,y_1}} + \delta_{\gamma_{x_2,y_2}}$ is a stair-shaped good decomposition of T. However, T is not cycle-free.

EXAMPLE 10. *Let T be a transport path from*

$$\mu^- = 9\delta_{x_1} + 9\delta_{x_2} + 9\delta_{x_3} + 27\delta_{x_4} + 27\delta_{x_5} \text{ to } \mu^+ = 36\delta_{y_1} + 9\delta_{y_2} + 18\delta_{y_3} + 9\delta_{y_4} + 9\delta_{y_5}$$

given as shown in the following figure.



T

For each (i,j), let $\gamma_{x_i,y_j} \in \Gamma$ be the unique polyhedral curve from x_i to y_j on T, and $a_{i,j}$ be the (i,j)-entry of the matrix

$$A = \begin{bmatrix} 4 & 1 & 2 & 1 & 1 \\ 4 & 1 & 2 & 1 & 1 \\ 4 & 1 & 2 & 1 & 1 \\ 12 & 3 & 6 & 3 & 3 \\ 12 & 3 & 6 & 3 & 3 \end{bmatrix}.$$

Then

$$\eta_A := \sum_{i,j=1}^{5} a_{ij}\delta_{\gamma_{x_i,y_j}}$$

is a good but not a better decomposition of T. Using Algorithm 3.6.5, the corresponding stair-shaped matrix of A is given by

$$B = \begin{bmatrix} 9 & 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 \\ 9 & 9 & 9 & 0 & 0 \\ 0 & 0 & 9 & 9 & 9 \end{bmatrix}.$$

The corresponding measure

$$\eta_B := \sum_{i,j=1}^{5} b_{ij}\delta_{\gamma_{x_i,y_j}}$$

on Γ is a stair-shaped good decomposition of T, which is automatically a better decomposition of T.

The following theorem says that any stair-shaped transport path can be decomposed as the sum of two subcurrents generated by two transport maps.

THEOREM 3.6.8. *Let $T \in Path(\mu^-, \mu^+)$ be a stair-shaped transport path, where μ^- and μ^+ are given in (3.2.10). Then there exist decomposition*

$$\mu^- = \mu_1^- + \mu_2^-, \mu^+ = \mu_1^+ + \mu_2^+, \text{ and } T = T_1 + T_2$$

such that

(a) *for each $i = 1, 2$, T_i is a subcurrent of T and $T_i \in Path(\mu_i^-, \mu_i^+)$,*

(b) *there exists transport maps $\varphi \in Map(\mu_1^-, \mu_1^+)$ and $\psi \in Map(\mu_2^+, \mu_2^-)$ such that both (T_1, φ) and $(-T_2, \psi)$ are compatible.*

PROOF. Since T is stair-shaped, there exists a good decomposition η whose representing matrix $A = [a_{ij}]$ is a stair-shaped matrix. We now write A as the sum of $B = [b_{ij}]$ and $C = [c_{ij}]$ as follows. For each i and j, if $a_{ij} = 0$, set $b_{ij} = 0$ and $c_{ij} = 0$. When $a_{ij} > 0$,

- if a_{ij} is the last non-zero entry in the i-th row of A, (i.e., $a_{ij'} = 0$ for all $j' \geq j + 1$,) we set $b_{ij} = a_{ij}$ and $c_{ij} = 0$;
- if a_{ij} is not the last non-zero entry in the i-th row of A, since A is stair-shaped, a_{ij} is the last non-zero entry in the j-th column of A. In this case, we set $b_{ij} = 0$ and $c_{ij} = a_{ij}$.

By doing so, we write $A = B + C$ such that each row of $B = [b_{ij}]$ and each column of $C = [c_{ij}]$ contain at most one non-zero entry. Note that for each (i, j), $a_{ij} = b_{ij} + c_{ij}$ and $a_{ij} > 0$ means either $b_{ij} > 0$ or $c_{ij} > 0$ but not both. Define

$$\mu_1^- = \sum_i \left(\sum_j b_{ij}\right)\delta_{x_i},\ \mu_1^+ = \sum_j \left(\sum_i b_{ij}\right)\delta_{y_j},\ \mu_2^- = \sum_i \left(\sum_j c_{ij}\right)\delta_{x_i},\ \mu_2^+ = \sum_j \left(\sum_i c_{ij}\right)\delta_{y_j}.$$

Then $\mu^- = \mu_1^- + \mu_2^-$ and $\mu^+ = \mu_1^+ + \mu_2^+$. Let

$$T_1 := \int_{\{\gamma \in \Gamma_{x_i,y_j} :\, b_{ij}>0\}} I_\gamma \, d\eta, \text{ and } T_2 := \int_{\{\gamma \in \Gamma_{x_i,y_j} :\, c_{ij}>0\}} I_\gamma \, d\eta.$$

Both T_1 and T_2 are subcurrents of T, and

$$\partial T_1 = \int_{\{\gamma \in \Gamma_{x_i,y_j} :\, b_{ij}>0\}} (\delta_{y_j} - \delta_{x_i}) \, d\eta = \sum_{i,j} b_{ij}(\delta_{y_j} - \delta_{x_i}) = \mu_1^+ - \mu_1^-,$$

$$\partial T_2 = \int_{\{\gamma \in \Gamma_{x_i,y_j} :\, c_{ij}>0\}} (\delta_{y_j} - \delta_{x_i}) \, d\eta = \sum_{i,j} c_{ij}(\delta_{y_j} - \delta_{x_i}) = \mu_2^+ - \mu_2^-,$$

which gives $T_i \in Path(\mu_i^-, \mu_i^+)$ for $i = 1, 2$. Then,

$$T = \int_\Gamma I_\gamma \, d\eta = \int_{\{\gamma \in \Gamma_{x_i,y_j} :\, a_{ij}>0\}} I_\gamma \, d\eta = \int_{\{\gamma \in \Gamma_{x_i,y_j} :\, b_{ij}>0\}} I_\gamma \, d\eta + \int_{\{\gamma \in \Gamma_{x_i,y_j} :\, c_{ij}>0\}} I_\gamma \, d\eta = T_1 + T_2.$$

Denote

$$X_1 = \{x_i \in X : \mu_1^-(\{x_i\}) > 0\}, \; Y_1 = \{y_j \in X : \mu_1^+(\{y_j\}) > 0\},$$

$$X_2 = \{x_i \in X : \mu_2^-(\{x_i\}) > 0\}, \; Y_2 = \{y_j \in Y : \mu_2^+(\{y_j\}) > 0\}.$$

Observe that since A is stair-shaped, by the construction of b_{ij}, for each i, there exists at most one j (i.e. the largest j with $a_{ij} > 0$) such that $b_{ij} > 0$. This leads to a map: $\varphi : X_1 \to Y_1$ given by

$$\varphi(x_i) = y_j \text{ if } b_{ij} > 0.$$

Similarly, for each j, there exists at most one i (i.e. the largest i with $a_{ij} > 0$) such that $c_{ij} > 0$. This leads to a map: $\psi : Y_2 \to X_2$ given by

$$\psi(y_j) = x_i \text{ if } c_{ij} > 0.$$

By definition of φ, for each $y_j \in Y_1$,

$$\varphi_\# \mu_1^-(\{y_j\}) = \mu_1^-(\varphi^{-1}(y_j)) = \mu_1^-(\{x_i : b_{ij} > 0\}) = \sum_{b_{ij}>0} \mu_1^-(\{x_i\}) = \sum_i b_{ij} = \mu_1^+(\{y_j\}).$$

Therefore, $\varphi_\# \mu_1^- = \mu_1^+$, and similarly, $\mu_2^- = \psi_\# \mu_1^+$. Also, direct calculation gives

$$\pi_\varphi := (id \times \varphi)_\# \mu_1^- = \int_{\{\gamma \in \Gamma_{x_i,y_j} :\, b_{ij}>0\}} \delta_{(x_i,y_j)} \, d\eta,$$

and

$$\pi_\psi := (id \times \psi)_{\#}\mu_2^+ = \int_{\{\gamma \in \Gamma_{x_i,y_j}:\, c_{ij}>0\}} \delta_{(y_j,x_i)}\, d\eta.$$

Hence, (T_1, φ) and $(-T_2, \psi)$ are compatible. $\square$

We now provide an example to illustrate Theorem 3.6.8.

Example 11. *Let T, μ^-, μ^+, A, B, η_A, η_B be the same values as defined in Example 10. By Theorem 3.6.8, we have*

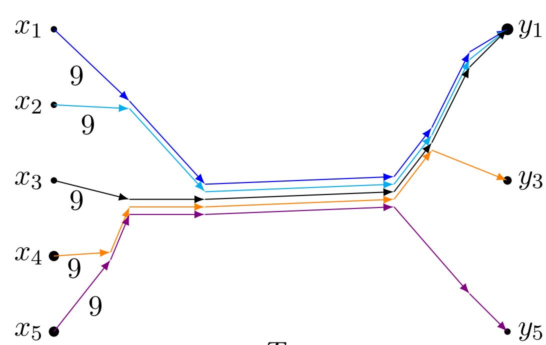
$$B_1 = \begin{bmatrix} 9 & 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 9 & 9 & 0 & 0 & 0 \\ 0 & 0 & 9 & 9 & 0 \end{bmatrix},$$

so that $B = B_1 + B_2$. By matrix B_1, we get a transport path T_1, with

$$\mu_1^- = 9\delta_{x_1} + 9\delta_{x_2} + 9\delta_{x_3} + 9\delta_{x_4} + 9\delta_{x_5},\ \mu_1^+ = 27\delta_{y_1} + 9\delta_{y_3} + 9\delta_{y_5},$$

and $\varphi : \{x_1, x_2, x_3, x_4, x_5\} \to \{y_1, y_3, y_5\}$, such that

$$\varphi(x_1) = \varphi(x_2) = \varphi(x_3) = y_1,\ \varphi(x_4) = y_3,\ \varphi(x_5) = y_5.$$



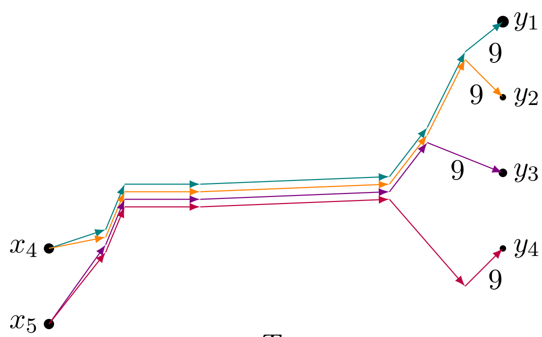
T_1

By matrix B_2, we get a transport path T_2, with

$$\mu_2^- = 18\delta_{x_4} + 18\delta_{x_5},\ \mu_2^+ = 9\delta_{y_1} + 9\delta_{y_2} + 9\delta_{y_3} + 9\delta_{y_4},$$

and $\psi : \{y_1, y_2, y_3, y_4\} \to \{x_4, x_5\}$*, such that*

$$\psi(y_1) = \psi(y_2) = x_4,\ \ \psi(y_3) = \psi(y_4) = x_5$$



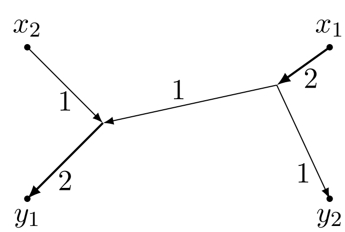
T_2

Then, T *is decomposed as the sum of* T_1 *and* T_2*.*

3.6.3. Cycle-free stair-shaped transport paths.

To use Theorem 3.6.8, for a given transport path, one may want to find a stair-shaped good decomposition of it. However, the stair-shaped matrix generated by Algorithm 3.6.5 does not necessarily correspond to a good decomposition, even if we start with a good decomposition, as demonstrated by the following example.

Example 12. *Let* T *be the graph given in the following figure, and* $\gamma_{i,j}$ *be the curve on* T *from* x_i *to* y_j *for each* i, j*.*



T

Then,

$$\eta = \delta_{\gamma_{1,1}} + \delta_{\gamma_{1,2}} + \delta_{\gamma_{2,1}}$$

is a good decomposition of T *with the representing matrix*

$$A = [a_{ij}] = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Algorithm 3.6.5 gives the stair-shaped matrix

$$B = [b_{ij}] = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

However, the corresponding measure,

$$\eta_B := 2\delta_{\gamma_{1,1}} + \delta_{\gamma_{2,2}}$$

is not a good decomposition of T anymore.

To overcome this issue, we introduce the following concepts:

Definition 3.6.9. Given $A \in \mathcal{A}_{M,N}$, an elementary matrix $E[(i_1, j_1), (i_2, j_2)]$ is called admissible to A if $a_{ij} > 0$ for all $(i,j) \in \{(i_1, j_1), (i_2, j_2), (i_1, j_2), (i_2, j_1)\}$. For any two matrices $A, B \in \mathcal{A}_{M,N}$, we say $A \triangleq B$ if there exists a list of real numbers $\{t_k\}_{k=1}^K$ and a list of elementary matrices $\{E_k\}_{k=1}^K$ admissible to A such that $B = A + \sum_{k=1}^K t_k E_k$ for some $K \in \mathbb{N} \cup \{\infty\}$.

Lemma 3.6.10. *Suppose A is the representing matrix of a finite measure η_A on Γ satisfying $(p_0)_\# \eta_A = \mu^-$ and $(p_\infty)_\# \eta_A = \mu^+$. For any matrix $B = [b_{ij}]$ with $A \triangleq B$, define*

$$\eta_B := \sum_{\substack{i,j \\ \text{with } a_{ij} > 0}} \frac{b_{ij}}{a_{ij}} \eta_A \lfloor_{\Gamma_{x_i, y_j}}. \tag{3.6.2}$$

Then η_B is a finite measure on Γ with $(p_0)_\# \eta_B = \mu^-$ and $(p_\infty)_\# \eta_B = \mu^+$. Moreover, B is the representing matrix of η_B and $\eta_B \prec\!\!\prec \eta_A$.

Proof. The condition $A \triangleq B$ gives

$$B = A + \sum_k t_k E_k,$$

for some real numbers t_k and elementary matrices $E_k = E[(i_k, j_k), (i'_k, j'_k)]$ that are *admissible* to A.

Note that

$$\begin{aligned}\eta_B(\Gamma) &= \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} \frac{b_{ij}}{a_{ij}} \eta_A \lfloor_{\Gamma_{x_i,y_j}}(\Gamma) = \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} \frac{b_{ij}}{a_{ij}} \eta_A(\Gamma_{x_i,y_j}) = \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} b_{ij} \\ &= \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} (a_{ij} + t_k(E_k)_{ij}) = \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} a_{ij} = \eta_A(\Gamma) < \infty.\end{aligned}$$

Moreover,

$$\begin{aligned}(p_0)_\# \eta_B &= \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} \frac{b_{ij}}{a_{ij}} \eta_A(\Gamma_{x_i,y_j}) \delta_{x_i} = \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} b_{ij} \delta_{x_i} \\ &= \sum_i \left(\sum_{\substack{j \\ \text{with } a_{ij}>0}} \left(a_{ij} + \sum_k t_k (E_k)_{ij} \right) \right) \delta_{x_i} \\ &= \sum_i \left(\sum_{\substack{j \\ \text{with } a_{ij}>0}} a_{ij} \right) \delta_{x_i} = \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} a_{ij} \delta_{x_i} = (p_0)_\# \eta_A = \mu^-.\end{aligned}$$

Similarly, $(p_\infty)_\# \eta_B = \mu^+$.

We now show that B is the representing matrix of η_B, i.e., $\eta_B(\Gamma_{x_{i'},y_{j'}}) = b_{i'j'}$ for each pair (i', j'). If $a_{i'j'} = 0$, then $\eta_B(\Gamma_{x_{i'},y_{j'}}) = 0$ since the sum is over all $a_{ij} > 0$. Also, since E_k's are admissible to A, this gives $(E_k)_{i'j'} = 0$ for all k, so that $b_{i'j'} = 0 = \eta_B(\Gamma_{x_{i'},y_{j'}})$. If $a_{i'j'} > 0$, then since $\eta_A(\Gamma_{x_{i'},y_{j'}}) = a_{i'j'}$,

$$\eta_B(\Gamma_{x_{i'},y_{j'}}) = \sum_{\substack{i,j \\ \text{with } a_{ij}>0}} \frac{b_{ij}}{a_{ij}} \eta_A \lfloor_{\Gamma_{x_i,y_j}}(\Gamma_{x_{i'},y_{j'}}) = b_{i'j'}.$$

Therefore, B is the representing matrix of η_B.

In the end, we show $\eta_B \prec\!\!\prec \eta_A$ by using Lemma 3.3.4. Suppose $\eta_B(\Gamma_{x_{i'},y_{j'}}) = b_{i'j'} > 0$, then previous argument gives $a_{i'j'} > 0$. Also, by definition of η_B,

$$\int_{\Gamma_{x_{i'},y_{j'}}} I_\gamma d\eta_B = \frac{b_{i'j'}}{a_{i'j'}} \int_{\Gamma_{x_{i'},y_{j'}}} I_\gamma d\eta_A, \text{ and hence } \frac{1}{b_{i'j'}} \int_{\Gamma_{x_{i'},y_{j'}}} I_\gamma d\eta_B = \frac{1}{a_{i'j'}} \int_{\Gamma_{x_{i'},y_{j'}}} I_\gamma d\eta_A.$$

As a result, $S_{i'j'}(\eta_B) = S_{i'j'}(\eta_A)$ as desired. $\square$

PROPOSITION 3.6.11. *Let T be a cycle-free transport path from μ^- to μ^+. Suppose η_A is a good decomposition of T, then for any matrix $B = [b_{ij}]$ with $A \triangleq B$, η_B given in (3.6.2) is also a good decomposition of T.*

PROOF. Let $A = [a_{ij}] \in \mathcal{A}_{M,N}$, $B = [b_{ij}] \in \mathcal{A}_{M,N}$, then $A \triangleq B$ gives

$$B = A + \sum_k t_k E_k,$$

for some real numbers t_k and elementary matrices $E_k = E[(i_k, j_k), (i'_k, j'_k)]$ that are *admissible* to A. Using $S_{i,j}(\eta)$ defined in (3.3.1), we have

$$\begin{aligned}
\int_\Gamma I_\gamma d(\eta_B - \eta_A) &= \int_\Gamma I_\gamma \, d \left(\sum_{i,j} \frac{b_{ij}}{a_{ij}} \eta_A \lfloor_{\Gamma_{x_i,y_j}} - \sum_{i,j} \eta_A \lfloor_{\Gamma_{x_i,y_j}} \right) \\
&= \sum_{i,j} \frac{b_{ij} - a_{ij}}{a_{ij}} \int_{\Gamma_{x_i,y_j}} I_\gamma d\eta_A \\
&= \sum_{i,j} (b_{ij} - a_{ij}) S_{i,j}(\eta_A) = \sum_k t_k \sum_{i,j} (E_k)_{ij} S_{i,j}(\eta_A) \\
&= \sum_k t_k \cdot \left(S_{i_k,j_k}(\eta_A) - S_{i_k,j'_k}(\eta_A) - S_{i'_k,j_k}(\eta_A) + S_{i'_k,j'_k}(\eta_A) \right).
\end{aligned}$$

Since E_k's are admissible to A, then $a_{ij} > 0$ for $(i,j) \in \{(i_k, j_k)), (i_k, j'_k)), (i'_k, j_k)), (i'_k, j'_k))\}$. Since

$$S_{i_k,j_k}(\eta_A) - S_{i_k,j'_k}(\eta_A) - S_{i'_k,j_k}(\eta_A) + S_{i'_k,j'_k}(\eta_A)$$

is on T and $a_{ij} > 0$, direct calculation gives

$$\partial \left(S_{i_k,j_k}(\eta_A) - S_{i_k,j'_k}(\eta_A) - S_{i'_k,j_k}(\eta_A) + S_{i'_k,j'_k}(\eta_A) \right) = 0.$$

By Definition 3.4.2, T is a cycle-free transport path implies

$$S_{i_k,j_k}(\eta_A) - S_{i_k,j'_k}(\eta_A) - S_{i'_k,j_k}(\eta_A) + S_{i'_k,j'_k}(\eta_A) = 0.$$

Hence,

$$\int_\Gamma I_\gamma d\eta_B = \int_\Gamma I_\gamma d\eta_A.$$

By using an analogous argument as in the proof of Step 1 in Lemma 3.3.7, it follows that η_B is also a good decomposition of T. $\square$

Given a matrix A with non-negative entries, Theorem 3.6.4 gives a stair-shaped matrix B, such that $A \cong B$, which by definition says $B = A + \sum_k t_k E_k$ for some elementary matrices E_k. In general, $A \cong B$ does not imply $A \triangleq B$, since it is possible that some E_k's are not admissible to A. However, when each entries of A is positive (as illustrated in Example 10), $A \cong B$ implies $A \triangleq B$. In general, when A satisfies certain conditions as stated in the following corollary, we have both $A \cong B$ and $A \triangleq B$, so that the η_B in (3.6.2) is a stair-shaped good decomposition.

Suppose $A = [a_{ij}]$, let $A[(i_0, j_0), (i'_0, j'_0)]$ be the "sub-matrix" of A with entries a_{ij}'s such that $i_0 \le i \le i'_0$, $j_0 \le j \le j'_0$.

Corollary 3.6.12. *Let T be a cycle-free transport path from μ^- to μ^+. Let $A = [a_{ij}]$ be the representing matrix of a good decomposition η_A of T. If there exist a list of sub-matrices $A_k = A[(i_k, j_k), (i'_k, j'_k)]$ of A such that*

(a) $(i_1, j_1) = (1, 1)$ *and* $i'_k \le i_{k+1} \le i'_k + 1$, $j'_k \le j_{k+1} \le j'_k + 1$ *for each* k,

(b) *all elements of the sub-matrix A_k are positive for each k,*

(c) *all elements of A not in any of the sub-matrices are* 0,

then there exists a stair-shaped good decomposition η_B of T with $\eta_B \prec\prec \eta_A$. Hence, T is stair-shaped.

Proof. We construct the desired stair-shaped matrix by using induction. We first apply Theorem 3.6.4 to the sub-matrix

$$A_1 = A[(i_1, j_1), (i'_1, j'_1)]$$

and get a stair-shaped A'_1. Then replace entries in A with entries in A'_1 in their corresponding original positions in A, and denote this new matrix as B_1. Inductively, for each $k \ge 1$, apply Theorem 3.6.4 to the sub-matrix

$$B_k[(i_{k+1}, j_{k+1}), (i'_{k+1}, j'_{k+1})]$$

of B_k and get a stair-shaped A'_{k+1}. Then replace entries in B_k with entries in A'_{k+1} in their corresponding original positions in B_k, and denote this matrix as B_{k+1}. Note that for each k, by condition (a), the sub-matrix $B_k[(i_1, j_1), (i'_k, j'_k)]$ is stair-shaped and

$$B_K[(i_1, j_1), (i'_k, j'_k)] = B_k[(i_1, j_1), (i'_k, j'_k)], \text{ for each } K \ge k+2. \tag{3.6.3}$$

As a result, for each (i,j), the limit $\lim_{k\to\infty} B_k(i,j)$ exists and equals the value of $B_k(i,j)$ when k is large enough.

Let B be the limit matrix of $\{B_k\}$ whose (i,j)-entry $B(i,j) = \lim_{k\to\infty} B_k(i,j)$ for each (i,j). By (3.6.3), $B[(i_1,j_1),(i'_k,j'_k)] = B_k[(i_1,j_1),(i'_k,j'_k)]$ for each k. Since $B_k[(i_1,j_1),(i'_k,j'_k)]$ is stair-shaped, B is also stair-shaped. Since B is a stair-shaped matrix, its corresponding measure η_B as defined in (3.6.2) is stair-shaped. By (b) and definition of *admissible matrices*, we have $A \triangleq B$. Therefore, Proposition 3.6.11 gives η_B is a good decomposition with $\eta_B \prec\prec \eta_A$. $\square$

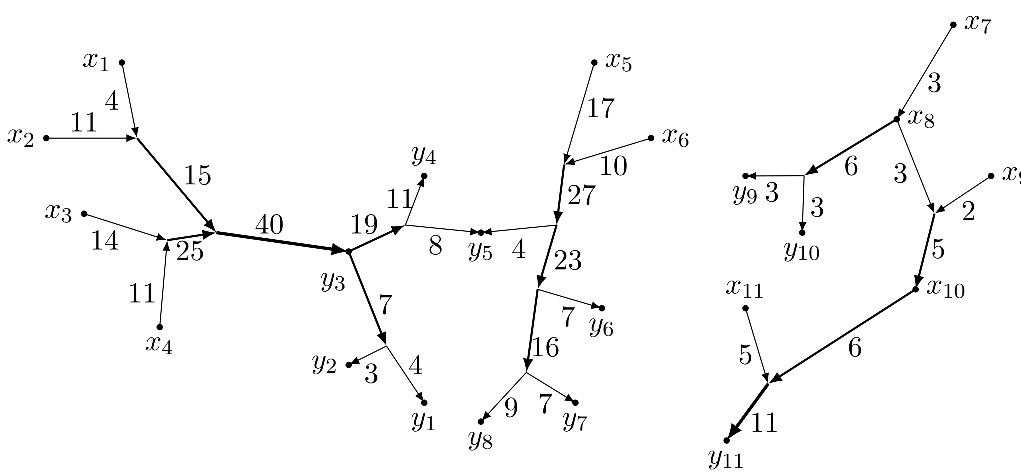
In the end, we provide a typical matrix of finite size satisfying conditions $(a), (b), (c)$ in Corollary 3.6.12, and see how to decompose the corresponding cycle-free stair-shaped transport path into the difference of two map-compatible paths.

Example 13. *Let*

$$\mu^- = 4\delta_{x_1} + 11\delta_{x_2} + 14\delta_{x_3} + 11\delta_{x_4} + 17\delta_{x_5} + 10\delta_{x_6} + 3\delta_{x_7} + 6\delta_{x_8} + 2\delta_{x_9} + \delta_{x_{10}} + 5\delta_{x_{11}},$$

$$\mu^+ = 4\delta_{y_1} + 3\delta_{y_2} + 14\delta_{y_3} + 11\delta_{y_4} + 12\delta_{y_5} + 7\delta_{y_6} + 7\delta_{y_7} + 9\delta_{y_8} + 3\delta_{y_9} + 3\delta_{y_{10}} + 11\delta_{y_{11}},$$

and T be a cycle-free transport path from μ^- to μ^+ illustrated by the following diagram:



Transport Path T

Let

$$A = \begin{bmatrix} 1 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 2 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 2 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 & 6 & 7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 4 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}.$$

Then, $A = [a_{ij}]$ *is the corresponding matrix of a good decomposition* η_A *of* T*, namely*

$$\eta_A := \sum_{i,j} a_{ij} \delta_{\gamma_{x_i, y_j}}.$$

Here, A *satisfies conditions* $(a), (b), (c)$ *in Corollary 3.6.12 with*

$$A_1 = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix},\ A_2 = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 7 & 1 \\ 5 & 2 & 4 \end{bmatrix},\ A_3 = \begin{bmatrix} 1 & 3 & 6 & 7 \\ 3 & 4 & 1 & 2 \end{bmatrix},\ A_4 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix},\ \text{and } A_5 = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 5 \end{bmatrix}.$$

Using algorithm 3.6.5, we have

$$A_1' = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 3 \end{bmatrix},\ A_2' = \begin{bmatrix} 8 & 0 & 0 \\ 6 & 8 & 0 \\ 0 & 3 & 8 \end{bmatrix},\ A_3' = \begin{bmatrix} 4 & 7 & 6 & 0 \\ 0 & 0 & 1 & 9 \end{bmatrix},\ A_4' = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix},\ \text{and } A_5' = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 5 \end{bmatrix}.$$

By Corollary 3.6.12,

$$\eta_B := \sum_{i,j} b_{ij} \delta_{\gamma_{x_i, y_j}}$$

is a stair-shaped good decomposition of T *with* $\eta_B \prec\!\!\prec \eta_A$*, where the matrix*

$$B = [b_{ij}] = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 7 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

is stair-shaped.

Now, by the proof of Theorem 3.6.8, one may decompose the stair-shaped matrix B *into* $B = B_1 + B_2$ *where*

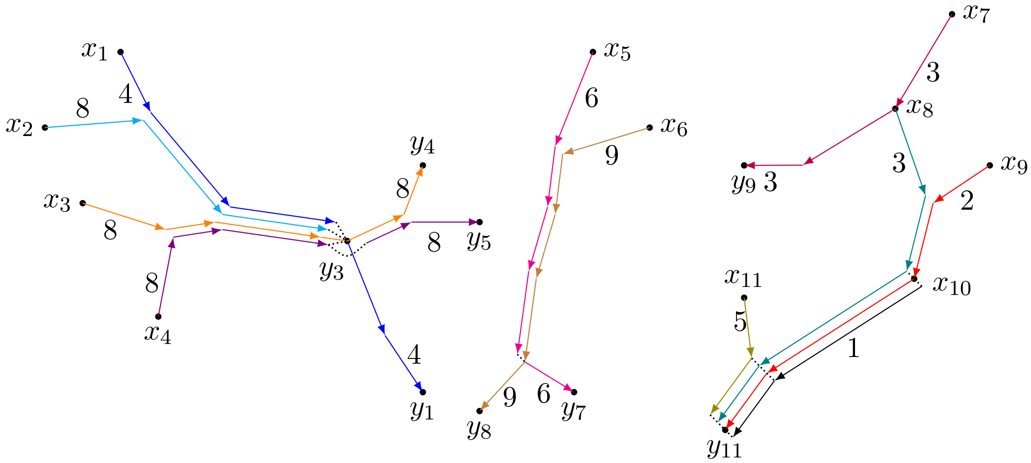
$$B_1 = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix} \text{ and } B_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

From matrix B_1 and the transport path T, we may construct the corresponding transport path $T_1 \in Path(\mu_1^-, \mu_1^+)$ illustrated below, where

$$\mu_1^- = 4\delta_{x_1} + 8\delta_{x_2} + 8\delta_{x_3} + 8\delta_{x_4} + 6\delta_{x_5} + 9\delta_{x_6} + 3\delta_{x_7} + 3\delta_{x_8} + 2\delta_{x_9} + \delta_{x_{10}} + 5\delta_{x_{11}},$$

and

$$\mu_1^+ = 4\delta_{y_1} + 8\delta_{y_3} + 8\delta_{y_4} + 8\delta_{y_5} + 6\delta_{y_7} + 9\delta_{y_8} + 3\delta_{y_9} + 11\delta_{y_{11}}.$$



Transport Path T_1

Note that from the non-zero entries of B_1, there exists a transport map

$$\varphi_1 : \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}\} \longrightarrow \{y_1, y_3, y_4, y_5, y_7, y_8, y_9, y_{11}\},$$

where

$$\varphi_1(x_1) = y_1,\ \varphi_1(x_2) = y_3,\ \varphi_1(x_3) = y_4,\ \varphi_1(x_4) = y_5,\ \varphi_1(x_5) = y_7,\ \varphi_1(x_6) = y_8,$$

$$\varphi_1(x_7) = y_9,\ \varphi_1(x_8) = y_{11},\ \varphi_1(x_9) = y_{11},\ \varphi_1(x_{10}) = y_{11},\ \varphi_1(x_{11}) = y_{11}.$$

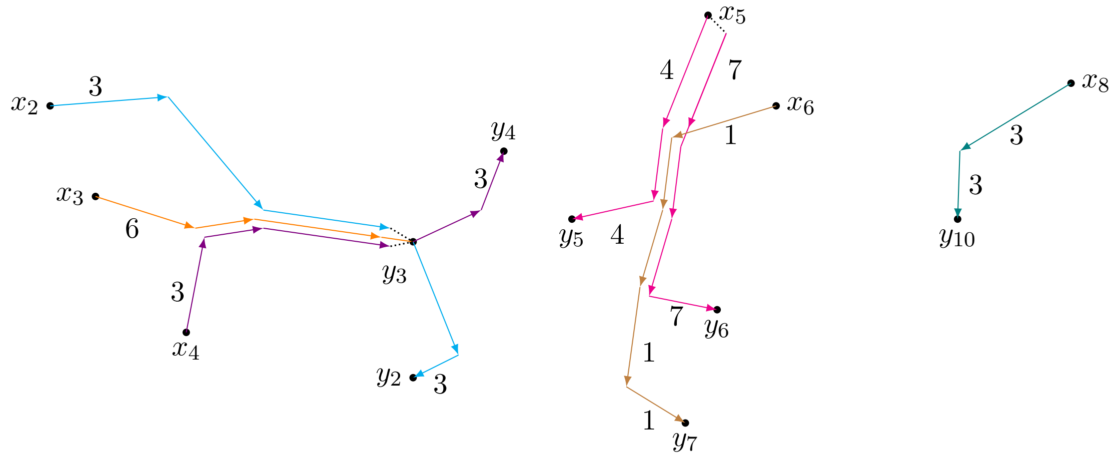
Here, $\varphi_{1\#}\mu_1^- = \mu_1^+$, and (T_1, φ_1) is compatible.

Similarly, using matrix B_2 and transport path T, we may construct the corresponding transport path $T_2 \in Path(\mu_2^-, \mu_2^+)$ as illustrated below, where

$$\mu_2^- = 3\delta_{x_2} + 6\delta_{x_3} + 3\delta_{x_4} + 11\delta_{x_5} + \delta_{x_6} + 3\delta_{x_8},$$

and

$$\mu_2^+ = 3\delta_{y_2} + 6\delta_{y_3} + 3\delta_{y_4} + 4\delta_{y_5} + 7\delta_{y_6} + \delta_{y_7} + 3\delta_{y_{10}}.$$



Transport Path T_2

Again, using the non-zero entries of B_2*, there exists a transport map*

$$\varphi_2 : \{y_2, y_3, y_4, y_5, y_6, y_7, y_{10}\} \longrightarrow \{x_2, x_3, x_4, x_5, x_6, x_8\},$$

with

$$\varphi_2(y_2) = x_2,\ \varphi_2(y_3) = x_3,\ \varphi_2(y_4) = x_4,\ \varphi_2(y_5) = x_5,\ \varphi_2(y_6) = x_5,\ \varphi_2(y_7) = x_6,\ \varphi_2(y_{10}) = x_8,$$

Here, $\mu_2^- = {\varphi_2}_\# \mu_1^+$*, and* $(-T_2, \varphi_2)$ *is compatible.*

As a result, we decompose the cycle-free stair-shaped transport path $T = T_1 - T_2$ *as the difference of two map-compatible paths* T_1 *and* T_2*.*

CHAPTER 4

Transport paths under capacity constraints

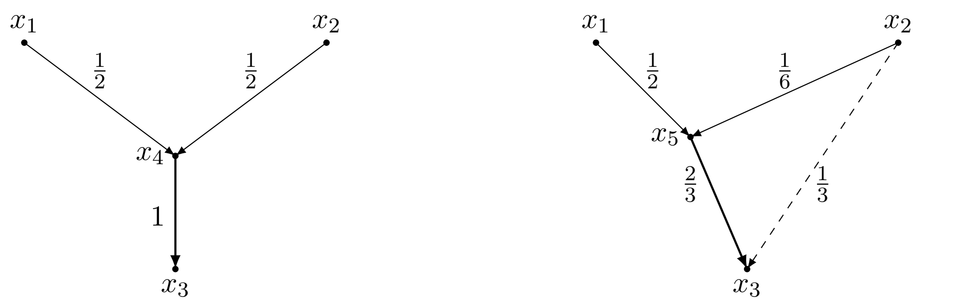
4.1. Introduction & Motivation

As illustrated in Section 2.2 and Section 3.2, transport paths between atomic measures in ramified transport system can be viewed as weighted directed graphs, as defined in Definition 2.2.1. In general, transport paths between two Radon measures can be viewed as rectifiable 1-currents, such that the value of density function equals the mass being transported at each position. Regardless of whether in atomic case or general case, the amount of mass that can be transported via any admissible transport paths has no restrictions. Hence, the phenomenon of first aggregating the total mass from the source into one place then transport through a single curve is permitted and prevalent in ramified transport paths.

As oppose to the theoretical permitted aggregation of total mass, this type of branching structure of a transport system rarely appears in real life. Transportation in reality often takes place through various kinds of medium, and most of the medium has transport capacity instantiated either as the total cumulative amount of mass transported before this medium breaks down (i.e. the life span of a product) or the maximum amount of mass this particular medium can carry all at once. In the later case, this property is often named as capacity of a medium or a particular transport path. For instance, buses, airplanes have limited seats, roads only allow a limited amount of traffic, i.e. 4 lanes, 6 lanes, etc. This brings naturally the question of ramified transport paths with capacity constraints, which can be crudely described by imposing an upper bound (called the capacity) on the weight function of a weighted directed graph or on the density function of a rectifiable 1-current. This motivates us to consider the following ramified transport problem:

Proposed problem: Given two atomic measures $\mathbf{a}$, $\mathbf{b}$ on X with equal mass, $\|\mathbf{a}\| = \|\mathbf{b}\|$, and $c > 0$. Minimize $\mathbf{M}_\alpha(G)$ among all $G \in Path(\mathbf{a}, \mathbf{b})$ with $w(e) \leq c,$ for all $e \in E(G)$.

From the description of this problem, if we assume $\|\mathbf{a}\| = \|\mathbf{b}\| \leq c$, this is equivalent to imposing no restriction on transport capacity.



(a) Y-shaped

(b) Mixture of Y-shaped and V-shaped

FIGURE 4.1. Y-shaped & Mixture of Y-shaped and V-shaped.

Note that after imposing the capacity constraint, a previously well defined transport path $G \in Path(\mathbf{a}, \mathbf{b})$, which has no capacity constraints, is not necessarily an admissible transport path anymore. This can be demonstrated in the following examples.

EXAMPLE 14. *Suppose we want to transport mass from* $\mathbf{a}$ *to* $\mathbf{b}$*, with an upper bound* c *imposed on weight functions, where*

$$\mathbf{a} = \frac{1}{2}\delta_{x_1} + \frac{1}{2}\delta_{x_2},\ \mathbf{b} = \delta_{x_3},\ c = \frac{2}{3}.$$

In this case, "Y-shaped" transport paths no longer satisfies the restriction on weight functions, since after merging at x_4 *the mass will reach 1. Changing to another kind of branching structure which is a mixture of "V-shaped" and "Y-shaped" will resolve this issue. One of the possible cases is merging* $\frac{1}{2}$ *from* x_1 *and* $\frac{1}{6}$ *from* x_2*, and let the remaining* $\frac{1}{3}$ *from* x_2 *transport directly through the dash line.*

Moreover, due to the "merging" effect, which will happen when considering a sequence of transport paths with decreasing transport cost, the **Proposed problem** may fail to have an admissible optimal solution. We may notice this "non-compactness" property of transport paths that are admissible in the **Proposed problem** from the following example.

EXAMPLE 15. *Let* $\boldsymbol{a} = \delta_x$ *and* $\boldsymbol{b} = \delta_y$ *which are atomic measures with total mass* 1 *distributed on* $\mathbb{R}^2$*. Suppose the transport capacity equals* $1/n$ *with* $n \in \mathbb{Z}^+$*, then any admissible transport paths from* $\mathbf{a}$ *to* $\mathbf{b}$ *in the* ***Proposed problem*** *need* n *curves connecting* x *to* y*, where each curve has weight* $1/n$*.*

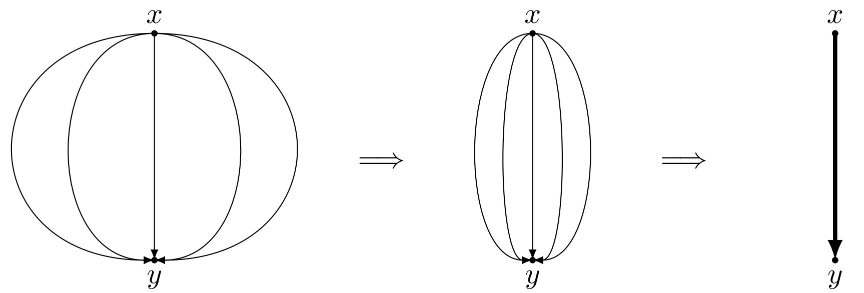


FIGURE 4.2. The above pictures give an illustration of "convergence" when $n = 5$.

Since x, $y \in \mathbb{R}^2$, the curve that has minimum distance and connects these two points is a straight line segment. Thus, when minimizing the $\mathbf{M}_\alpha$ cost of the above transport path, we have each of the n curves converges to the straight line segment connecting x and y, as illustrated from the above pictures. Hence, by taking the limit over these curves, we get a transport path that reaches the minimum $\mathbf{M}_\alpha$ cost. However, the transport path that we get after taking the limit does not satisfy the transport capacity restriction. Accumulating all the curves that have weight $1/n$ to one curve (i.e. the line segment connecting x and y) will make the weight on this curve equals 1, which is larger than the assumed transport capacity, $1/n$. Hence, the limit of a sequence of admissible transport paths in the ***Proposed problem*** *is not necessarily an admissible transport path anymore.*

4.2. Transport paths with capacity

When directly imposing upper bounds on the weight functions of transport paths or on the density functions of rectifiable 1-currents, aggregation of weights on some common curves may result in non-admissible transport paths. To overcome the "non-existence" of limit of a sequence of transport paths, we instead express transport paths into multiple components, such that each component represents a ramified transport paths with its total mass does not exceed the assumed capacity. This directs us to the following new expression of ramified transportation with capacity constraints.

PROBLEM 1 (**Ramified transportation with capacity**). *Let μ^-, μ^+ be two Radon measures on $X \subseteq \mathbb{R}^m$ with equal mass $\mu^-(X) = \mu^+(X) < \infty$, supported on compact sets, $\alpha \in (0,1)$, and $c > 0$. Minimize*

$$\boldsymbol{M}_\alpha(\vec{T}) := \sum_{k=1}^{\infty} \boldsymbol{M}_\alpha(T_k)$$

among $\vec{T} = (T_1, T_2, \ldots, T_N, \ldots)$ *satisfying*

$$T_k \in Path(\mu_k^-, \mu_k^+),\ \sum_{k=1}^{\infty} \mu_k^- = \mu^-, \sum_{k=1}^{\infty} \mu_k^+ = \mu^+, \text{and } 0 < \|\mu_k^-\| = \|\mu_k^+\| \leq c. \tag{4.2.1}$$

For simplicity of notations, denote $Path_c(\mu^-, \mu^+)$ as the set of all transport paths $\vec{T}$ satisfying conditions in (4.2.1). When $T_k = 0$ for $k \geq N+1$ (vanishing rectifiable 1-current), denote $(T_1, T_2, \ldots, T_N, \ldots)$ as $(T_1, T_2, \ldots, T_N)$ for simplicity. Note that each T_k is a rectifiable 1-current, $T_k = \underline{\underline{\tau}}(M_k, \theta_k, \xi_k)$, with $\partial T_k = \mu_k^+ - \mu_k^-$, and its $\mathbf{M}_\alpha$ cost is defined as $\mathbf{M}_\alpha(T_k) := \mathbf{M}(\underline{\underline{\tau}}(M_k, \theta_k^\alpha, \xi_k))$.

Also, note that for any $\vec{T} \in Path_c(\mu^-, \mu^+)$,

$$\sum_{i=1}^{\infty} T_k \in Path(\mu^-, \mu^+)$$

provided that the series is convergent in the following sense:

Definition 4.2.1. Let $\{T_i\}_{i=1}^{\infty}$ be any sequence of rectifiable 1-currents. We say the series $\sum_{i=1}^{\infty} T_i$ converges if the sequence $\{\sum_{i=1}^{n} T_i\}_{n=1}^{\infty}$ of partial sums converges as currents. i.e. for any differential 1-form $\omega \in \mathcal{D}^1(\mathbb{R}^m)$, the series $\sum_{i=1}^{\infty} T_i(\omega)$ of real numbers converges.

Lemma 4.2.2. *For any convergent series* $\sum_{i=1}^{\infty} T_i$ *of rectifiable 1-currents, if* $\alpha \leq 1$ *then*

$$\mathbf{M}_\alpha\left(\sum_{i=1}^{\infty} T_i\right) \leq \sum_{i=1}^{\infty} \mathbf{M}_\alpha(T_i).$$

Proof. Suppose $T_k = \underline{\underline{\tau}}(M_k, \theta_k, \xi_k)$, and let $\omega \in \mathcal{D}^1(\mathbb{R}^m)$, then

$$T_k(\omega) = \int_{M_k} \langle \omega(x), \xi_k(x) \rangle \theta_k(x)\, d\mathcal{H}^1(x),$$

and

$$\sum_{k=1}^{\infty} T_k(\omega) = \sum_{k=1}^{\infty} \int_{M_k} \langle \omega(x), \xi_k(x) \rangle\, \theta_i(x)\, d\mathcal{H}^1(x) = \int_{\bigcup_{k=1}^{\infty} M_i} \left\langle \omega(x), \sum_{k=1}^{\infty} \xi_k(x)\theta_k(x) \right\rangle d\mathcal{H}^1(x)$$

Here, we adopt the convention that for each k, $\theta_k(x) = 0$ when $x \notin M_k$. Since $\alpha \leq 1$, then for each $n \in \mathbb{N}$,

$$\left(\sum_{k=1}^{n} \theta_k(x)\right)^\alpha \leq \sum_{k=1}^{n} \theta_k(x)^\alpha \leq \sum_{k=1}^{\infty} \theta_k(x)^\alpha,$$

so that

$$\left(\sum_{k=1}^{\infty}\theta_k(x)\right)^{\alpha}=\lim_{n\to\infty}\left(\sum_{k=1}^{n}\theta_k(x)\right)^{\alpha}\leq\sum_{k=1}^{\infty}\theta_k(x)^{\alpha}.$$

Therefore,

$$\begin{aligned}\mathbf{M}_\alpha\left(\sum_{k=1}^{\infty}T_k\right) &\leq \int_{\bigcup_{k=1}^{\infty}M_k}\left(\sum_{k=1}^{\infty}\theta_k(x)\right)^{\alpha}d\mathcal{H}^1(x)\leq\int_{\bigcup_{k=1}^{\infty}M_k}\sum_{k=1}^{\infty}\theta_k(x)^{\alpha}d\mathcal{H}^1(x)\\ &= \sum_{k=1}^{\infty}\int_{\bigcup_{k=1}^{\infty}M_k}\theta_k(x)^{\alpha}d\mathcal{H}^1(x)=\sum_{k=1}^{\infty}\int_{M_k}\theta_k(x)^{\alpha}d\mathcal{H}^1(x)=\sum_{k=1}^{\infty}\mathbf{M}_\alpha(T_k).\end{aligned}$$

□

LEMMA 4.2.3. *For any transport capacity $c>0$ and any $\vec{T}\in Path_c(\mu^-,\mu^+)$, there exists a constant $N(c)\in\mathbb{N}$ with*

$$N(c)\leq\frac{2\|\mu^-\|}{c},$$

and $\vec{T'}=(T'_1,T'_2,\dots,T'_{N(c)})\in Path_c(\mu^-,\mu^+)$ with $\mathbf{M}_\alpha(\vec{T'})\leq\mathbf{M}_\alpha(\vec{T})$.

PROOF. Since $\sum_{k=1}^{\infty}\|\mu_k^-\|=\sum_{k=1}^{\infty}\|\mu_k^+\|<\infty$, there exists N such that

$$\sum_{k=N}^{\infty}\|\mu_k^-\|=\sum_{k=N}^{\infty}\|\mu_k^+\|<\frac{c}{2}.$$

For any $\vec{T}=(T_1,T_2,\dots,T_N,\dots)\in Path_c(\mu^-,\mu^+)$, denote

$$T'_N:=\sum_{k=N}^{\infty}T_k\in Path\left(\sum_{k=N}^{\infty}\mu_k^-,\sum_{k=N}^{\infty}\mu_k^+\right).$$

Then $\vec{T'}=(T_1,T_2,\dots,T_{N-1},T'_N)\in Path_c(\mu^-,\mu^+)$, and

$$\mathbf{M}_\alpha(\vec{T'})=\sum_{k=1}^{N-1}\mathbf{M}_\alpha(T_k)+\mathbf{M}_\alpha(\vec{T}'_N)=\sum_{k=1}^{N-1}\mathbf{M}_\alpha(T_k)+\mathbf{M}_\alpha\left(\sum_{k=N}^{\infty}T_k\right)\leq\sum_{k=1}^{\infty}\mathbf{M}_\alpha(T_k)=\mathbf{M}_\alpha(\vec{T}).$$

As a result, without loss of generality, we may assume that $\vec{T}$ has only finitely number of components, i.e. $\vec{T}=(T_1,T_2,\dots,T_N)$.

We may further assume that there is at most one k with $1\leq k\leq N$ satisfying $\|\mu_k^-\|\leq c/2$. Indeed, assume for some $1\leq i<j\leq N$ such that $\|\mu_i^-\|\leq c/2,\|\mu_j^-\|\leq c/2$. Let

$$\vec{T}^*:=(T_1,\dots,T_i+T_j,\dots,T_{j-1},T_{j+1},\dots,T_N),$$

then $\vec{T}^* \in Path_c(\mu^-, \mu^+)$, since $\|\mu_j^+ + \mu_j^+\| = \|\mu_j^- + \mu_j^-\| = \|\mu_j^-\| + \|\mu_j^-\| \le c$. Also,

$$\mathbf{M}_\alpha(\vec{T}^*) = \sum_{k \ne i,j} \mathbf{M}_\alpha(T_k) \ + \mathbf{M}_\alpha(T_i + T_j) \le \sum_{k \ne i,j} \mathbf{M}_\alpha(T_k) \ + \mathbf{M}_\alpha(T_i) + \mathbf{M}_\alpha(T_j) = \mathbf{M}_\alpha(\vec{T}).$$

Thus, replacing $\vec{T}$ by $\vec{T}^*$ if necessary, we may assume that there is at most one k, with $1 \le k \le N$, satisfying $\|\mu_k^-\| \le c/2$. Hence,

$$\|\mu^-\| = \sum_{k=1}^N \|\mu_k^-\| > (N-1)\frac{c}{2},$$

which implies $N < 2\|\mu^-\|/c+1$, and since N is integer valued, we have $N \le 2\|\mu^-\|/c$ as desired. □

REMARK 4.2.4. In the proof of above Lemma, it is not required to assume $\|\mu_i^-\|, \|\mu_j^-\| \le c/2$ for $1 \le i < j \le N$, and then combining these two transport paths and their corresponding source and target measures. In general, we may assume $\|\mu_i^-\|, \|\mu_j^-\| \le c/n$ for $1 \le i < j \le N$, and $n \in \mathbb{N}$, $n \ge 2$. Then similar argument gives

$$(N-(n-1)) \cdot \frac{(n-1)c}{n} < \|\mu^-\|, \text{ and this gives } N < \frac{n}{(n-1)}\frac{\|\mu^-\|}{c} + n - 1.$$

THEOREM 4.2.5. *For $\alpha \in (1-1/m, 1]$, there exists a transport path $\vec{T} \in Path_c(\mu^-, \mu^+)$ such that $\mathbf{M}_\alpha(\vec{T})$ is minimized over all admissible transport paths in Problem 1.*

PROOF. We first show that there exists a $\vec{S} \in Path_c(\mu^- \mu^+)$ satisfying (4.2.1) with $\mathbf{M}_\alpha(\vec{S}) < \infty$. Indeed, since both μ^- and μ^+ are supported on a compact set, by existence theorem [**11**] we can find $S \in Path(\mu^-, \mu^+)$ with $\mathbf{M}_\alpha(S) < +\infty$ for $\alpha \in (1-1/m, 1]$.

Pick $L \in \mathbb{N}$ large enough so that $\|\mu^+\| = \|\mu^-\| \le cL$, and let $\vec{S}$ be the L-vector of 1-rectifiable currents such that

$$\vec{S} = \left[\frac{1}{L}S, \frac{1}{L}S, \cdots, \frac{1}{L}S\right].$$

Note that $\vec{S}$ satisfies (4.2.1) with $\mu_i^\pm = \frac{1}{L}\mu^\pm$ for $i = 1, 2, \cdots, L$ and $\mu_i^\pm = 0$ for $i > L$. Moreover,

$$\mathbf{M}_\alpha(\vec{S}) = \sum_{k=1}^L \mathbf{M}_\alpha(S/L) = \sum_{k=1}^L L^{-\alpha}\mathbf{M}_\alpha(S) = L^{1-\alpha}\mathbf{M}_\alpha(S) < \infty.$$

Now, let $\{\vec{T}^{(n)}\}$ be any $\mathbf{M}_\alpha$ minimizing sequence for Problem 1 with $\mathbf{M}_\alpha(\vec{T}^{(n)}) \le \mathbf{M}_\alpha(\vec{S})$. By Lemma 4.2.3, without loss of the generality, we may assume that each $\vec{T}^{(n)} = (T_1^n, T_2^n, \cdots, T_N^n)$

with $N = N(c)$. For each $1 \le i \le N$ and $n \in \mathbb{N}$, let $T_i^n = \underline{\underline{\tau}}(M_{i,n}, \theta_{i,n}, \xi_{i,n})$ with $\theta_{i,n}(x) \le \|\mu_i^-\| \le c$, then

$$\mathbf{M}(T_i^n) = \int_{M_{i,n}} \theta_{i,n}\, \mathcal{H}^1(x) = \int_{M_{i,n}} \theta_{i,n}^{\alpha} \cdot \theta_{i,n}^{1-\alpha}\, \mathcal{H}^1(x) \le c^{1-\alpha} \int_{M_{i,n}} \theta_{i,n}^{\alpha}\, \mathcal{H}^1(x) = c^{1-\alpha}\mathbf{M}_\alpha(T_i^n).$$

Hence,

$$\mathbf{M}(T_i^n) \le c^{1-\alpha}\mathbf{M}_\alpha(T_i^n) \le c^{1-\alpha}\mathbf{M}_\alpha(\vec{T}^{(n)}) \le c^{1-\alpha}\mathbf{M}_\alpha(\vec{S}) < \infty.$$

By the weak compactness of rectifiable currents with respect to mass, each sequence $\{T_i^n\}_{n=1}^{\infty}$ sequentially converges to some rectifiable current T_i for $i = 1, 2, \ldots, N$. Since N is finite, we may assume that they have the same convergent subsequence. As a result, we have a convergent subsequence of $\{\vec{T}^{(n)} = (T_1^n, T_2^n, \cdots, T_N^n)\}$ with limit $\vec{T} = (T_1, T_2, \cdots, T_N)$. By lower-semicontinuity of mass of currents, this vector $\vec{T}$ is the desired solution for Problem 1. □

4.3. Components of transport path with capacity constraints

Given a transport N-path $(T_1, T_2, \ldots, T_N) \in Path_c(\mu^-, \mu^+)$, and note that for each k, $T_k \in Path(\mu_k^-, \mu_k^+)$, with $\partial T_k = \mu_k^+ - \mu_k^-$. Denote $\partial^- T_k := \mu_k^-$ which is the source measure, $\partial^+ T_k := \mu_k^+$ which is the target measure, and by definition of transport paths we automatically have $\|\partial^- T_k\| = \|\partial^+ T_k\|$. The conditions for transport paths in $Path_c(\mu^-, \mu^+)$ can be expressed as

$$\mu^- = \sum_{k=1}^{N} \partial^- T_k,\ \mu^+ = \sum_{k=1}^{N} \partial^+ T_k,\ \|\partial^- T_k\| = \|\partial^+ T_k\| \le c.$$

DEFINITION 4.3.1. Let $T = \underline{\underline{\tau}}(M, \theta, \xi)$ and $S = \underline{\underline{\tau}}(N, \phi, \zeta)$ be two rectifiable 1-currents. We say S is on T if $\mathcal{H}^1(N \setminus M) = 0$, and $\phi(x) \le \theta(x)$ for $\mathcal{H}^1$ almost all $x \in N$.

We now give conditions to determine whether a transport path is optimal or not.

THEOREM 4.3.2. *Given $\mu^- = \sum_{k=1}^N \mu_k^-, \mu^+ = \sum_{k=1}^N \mu_k^+, \|\mu_k^-\| = \|\mu_k^+\| \le c$, and a transport path $(T_1, T_2, \ldots, T_N) \in Path_c(\mu^-, \mu^+)$. Suppose $\vec{S} = (S_1, S_2, \ldots, S_N)$ consists of n rectifiable 1-currents, such that for each $k = 1, 2, \ldots, N$, it satisfies the following conditions:*

(1) *S_k is on T_k,*

(2) *$\partial S_k = \rho_k(x)\partial T_k$ with $|\rho_k(x)| \le 1$ and*

$$\sum_{k=1}^{N} \rho_k(x)\partial^- T_k = 0,\ \sum_{k=1}^{N} \rho_k(x)\partial^+ T_k = 0, \tag{4.3.1}$$

(3) $\|\partial^-(T_k \pm S_k)\| \le c$.

Then for $\epsilon \in [-1,1]$, $\vec{T} + \epsilon\vec{S} = (T_1 + \epsilon S_1, T_2 + \epsilon S_2, \ldots, T_N + \epsilon S_N) \in Path_c(\mu^-, \mu^+)$, *and*

$$\min\left\{\mathbf{M}_\alpha(\vec{T} + \vec{S}), \mathbf{M}_\alpha(\vec{T} - \vec{S})\right\} \le \mathbf{M}_\alpha(\vec{T}).$$

Furthermore, when $\vec{T}$ *is* α*-optimal for* $\alpha \in (0,1)$, *then* $S_k = 0$ *for all* $k = 1, 2, \cdots, N$.

Proof. Since $\partial(T_k + \epsilon S_k) = \partial T_k + \epsilon \partial S_k = (1 + \epsilon\rho_k(x))\partial T_k$ and $0 \le 1 + \epsilon\rho_k(x)$ for each k, then

$$\partial^-(T_k + \epsilon S_k) = (1 + \epsilon\rho_k(x))\partial^- T_k, \text{ and } \partial^+(T_k + \epsilon S_k) = (1 + \epsilon\rho_k(x))\partial^+ T_k.$$

By (4.3.1), this implies that

$$\sum_{k=1}^N \partial^-(T_k + \epsilon S_k) = \sum_{k=1}^N (1 + \epsilon\rho_k(x))\partial^- T_k = \sum_{k=1}^N \partial^- T_k + \epsilon \sum_{k=1}^N \rho_k(x)\partial^- T_k = \sum_{k=1}^N \partial^- T_k = \mu^-,$$

and similarly

$$\sum_{k=1}^N \partial^+(T_k + \epsilon S_k) = \sum_{k=1}^N (1 + \epsilon\rho_k(x))\partial^+ T_k = \sum_{k=1}^N \partial^+ T_k + \epsilon \sum_{k=1}^N \rho_k(x)\partial^+ T_k = \sum_{k=1}^N \partial^+ T_k = \mu^+.$$

Also,

$$\|\partial^-(T_k + \epsilon S_k)\| = \int_X 1 + \epsilon\rho_k(x)\, d(\partial^- T_k) = \int_X 1\, d(\partial^- T_k) + \epsilon \int_X \rho_k(x)\, d(\partial^- T_k),$$

which is a linear function with respect to ϵ, so that

$$\begin{aligned}
& \int_X 1\, d(\partial^- T_k) + \epsilon \int_X \rho_k(x)\, d(\partial^- T_k) \\
\le\ & \max\left\{\int_X 1\, d(\partial^- T_k) + \int_X \rho_k(x)\, d(\partial^- T_k), \int_X 1\, d(\partial^- T_k) - \int_X \rho_k(x)\, d(\partial^- T_k)\right\} \\
=\ & \max\left\{\|\partial^-(T_k + S_k)\|, \|\partial^-(T_k - S_k)\|\right\} \le c.
\end{aligned}$$

Hence, $\|\partial^-(T_k + \epsilon S_k)\| \le c$, and we can get $\|\partial^+(T_k + \epsilon S_k)\| \le c$ in a similar way. These results imply $\vec{T} + \epsilon\vec{S} \in Path_c(\mu^-, \mu^+)$.

For each $k = 1, 2, \ldots, N$, denote $T_k = \underline{\underline{\tau}}(M_k, \theta_k, \xi_k)$, $S_k = \underline{\underline{\tau}}(N_k, \phi_k, \zeta_k)$, and

$$M_k^+ = \{x \in M_k \mid \langle \xi_k(x), \zeta_k(x)\rangle = 1\},\ M_k^- = \{x \in M_k \mid \langle \xi_k(x), \zeta_k(x)\rangle = -1\}.$$

Then

$$
\begin{aligned}
\mathbf{M}_\alpha(\vec{T}+\epsilon\vec{S}) &= \sum_{k=1}^{N}\mathbf{M}_\alpha(T_k+\epsilon S_k)\\
&= \sum_{k=1}^{N}\int_{M_k^+}|\theta_k(x)+\epsilon\phi_k(x)|^\alpha d\mathcal{H}^1(x)+\int_{M_k^-}|\theta_k(x)-\epsilon\phi_k(x)|^\alpha d\mathcal{H}^1(x),
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial^2\mathbf{M}_\alpha(\vec{T}+\epsilon\vec{S})}{\partial\epsilon^2} &= \sum_{k=1}^{N}\frac{\partial^2\mathbf{M}_\alpha(T_k+\epsilon S_k)}{\partial\epsilon^2}\\
&= \alpha(\alpha-1)\Bigg(\sum_{k=1}^{N}\int_{M_k^+}\phi_k(x)^2|\theta_k(x)+\epsilon\phi_k(x)|^{\alpha-2}d\mathcal{H}^1(x)\\
&\qquad\qquad+\sum_{k=1}^{N}\int_{M_k^-}\phi_k(x)^2|\theta_k(x)-\epsilon\phi_k(x)|^{\alpha-2}d\mathcal{H}^1(x)\Bigg)\\
&\leq 0.
\end{aligned}
$$

This implies $\mathbf{M}_\alpha(\vec{T}+\epsilon\vec{S})$ is a concave function on ϵ, so that $\mathbf{M}_\alpha(\vec{T}+\epsilon\vec{S})$ reaches minimum value when ϵ reaches end points of its domain. Hence, $\min\{\mathbf{M}_\alpha(\vec{T}+\vec{S}),\mathbf{M}_\alpha(\vec{T}-\vec{S})\}\leq\mathbf{M}_\alpha(\vec{T})$.

Now assume that $\vec{T}$ is α-optimal for $\alpha\in(0,1)$ but $\vec{S}=(S_1,S_2,\ldots,S_N)$ is non-zero. i.e. there exists $k\in\{1,2,\ldots,N\}$ such that S_k is a non-vanishing current. Then,

$$
\left.\frac{\partial^2\mathbf{M}_\alpha(\vec{T}+\epsilon\vec{S})}{\partial\epsilon^2}\right|_{\epsilon=0} = \alpha(\alpha-1)\sum_{k=1}^{N}\int_{M_k^+\cup M_k^-}\phi_k(x)^2\theta_k(x)^{(\alpha-2)}d\mathcal{H}^1(x)<0,
$$

because $\vec{S}$ is nonzero and on $\vec{T}$. This says that $\mathbf{M}_\alpha(\vec{T}+\epsilon\vec{S})$ cannot achieve a local minimum at $\epsilon=0$, contradicting with $\vec{T}$ is optimal.

□

REMARK 4.3.3. Given $\partial S_k=\rho_k(x)\partial T_k$, if $\sum_k\rho_k(x)\partial^-T_k=0,\sum_k\rho_k(x)\partial^+T_k=0$, then

$$
\partial\left(\sum_k S_k\right)=\sum_k\partial S_k=\sum_k\rho_k(x)\partial T_k=\sum_k\rho_k(x)\partial^+T_k-\sum_k\rho_k(x)\partial T_k^-=0.
$$

Hence, if $\sum_k S_k$ does not form a cycle, S_k's does not satisfy criteria of the above theorem.

Also, when $\partial S_k=0$, condition (1) and (2) are automatically satisfied, since we have $\rho_k(x)=0$ and $\partial(T_k\pm S_k)=\partial S_k$.

In the following content, we will start to analyze transport paths between atomic measures, where

$$(4.3.2) \qquad \mu^- = \sum_{k=1}^{N} \mu_k^- = \sum_{i=1}^{N_1} m_i' \delta_{x_i}, \ \mu^+ = \sum_{k=1}^{N} \mu_k^+ = \sum_{j=1}^{N_2} m_j \delta_{y_j},$$

with $N_1, N_2 \in \mathbb{N} \cup \{\infty\}$, and equal total mass. Here, we also assume each $\vec{T} \in Path_c(\mu^-, \mu^+)$ consists of N components such that $\vec{T} = (T_1, T_2, \ldots, T_N)$ and $T_k \in Path(\mu_k^-, \mu_k^+)$ for $k = 1, 2, \ldots, N$. Since each $\vec{T}$ consists of N components, we may also call $\vec{T}$ as a transport N-path.

When $\vec{T} = (T_1, T_2, \ldots, T_N)$ is optimal and satisfies conditions in (4.2.1), each T_k in $\vec{T}$ is also an optimal transport path, which is acyclic. Using the definition of good decomposition and its related notations from Section 3.2, for each $k = 1, 2, \ldots, N$, there exists a good decomposition η_k of T_k, such that

$$(4.3.3) \qquad T_k = \int_\Gamma I_\gamma d\eta_k.$$

Since $T_k \in Path(\mu_k^-, \mu_k^+)$, and μ_k^-, μ_k^+ are as defined in (4.3.2), we may also write T_k as

$$T_k = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \int_{\Gamma_{x_i, y_j}} I_\gamma d\eta_k.$$

In this case,

$$\partial T_k = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \eta_k(\Gamma_{x_i, y_j})(\delta_{y_j} - \delta_{x_i}),$$

and

$$\partial^- T_k = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \eta_k(\Gamma_{x_i, y_j}) \delta_{x_i}, \ \partial^+ T_k = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \eta_k(\Gamma_{x_i, y_j}) \delta_{y_j}.$$

Using similar notation as in equation (3.4.1), for each $k = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, N_2$, denote

$$X_j(\eta_k) := \{x_i \in X : \eta_k(\Gamma_{x_i, y_j}) > 0\}.$$

Now, we would like to investigate the components of transport paths between atomic measures defined as in equation (4.3.2).

Proposition 4.3.4. *Let $\vec{T} = (T_1, T_2, \ldots, T_N) \in Path_c(\mu^-, \mu^+)$ be an optimal transport path, where μ^-, μ^+ are defined as in equation (4.3.2), and $\alpha \in (0,1)$. For each $k = 1, 2, \ldots, N$, let η_k be* *any good decomposition of T_k. Then the collection of sets*

$$\{X_j(\eta_k) : j = 1, 2, \dots, N_2\}$$

are mutually disjoint, except for at most $N_2 - 1$ many k's.

Proof. Suppose there are N_2 collections of sets

$$\{X_j(\eta_{k_\ell}) \, : \, j = 1, 2, \dots, N_2\}\,, \text{ for } \ell = 1, 2, \dots, N_2,$$

where each collection of sets are not mutually disjoint. Then for each ℓ, there exist $x_{i_\ell}, y_{j_\ell}, y_{j'_\ell}$, such that $x_{i_\ell} \in X_{j_\ell}(\eta_{k_\ell}) \cap X_{j'_\ell}(\eta_{k_\ell})$. In this case, we have

$$T_{k_\ell} = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\int_{\Gamma_{x_i,y_j}} I_\gamma d\eta_{k_\ell},\ \partial T_{k_\ell} = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\eta_{k_\ell}(\Gamma_{x_i,y_j})(\delta_{y_j} - \delta_{x_i}),$$

and

$$\eta_{k_\ell}(\Gamma_{x_{i_\ell},y_{j_\ell}}) > 0,\ \eta_{k_\ell}(\Gamma_{x_{i_\ell},y_{j'_\ell}}) > 0.$$

Our goal is to show the existence of a $\vec{S}$ such that it satisfies the conditions in Theorem 4.3.2, and reach a contradiction.

For each $\ell = 1, 2, \dots, N_2$, let $e_{j_\ell} = (0, \dots, 1, \dots, 0)$ be the vector of dimension N_2 such that it has 1 at position ℓ. Let M be the matrix

$$M := \begin{bmatrix} e_{j'_1} - e_{j_1} \\ e_{j'_2} - e_{j_2} \\ \vdots \\ e_{j'_{N_2}} - e_{j_{N_2}} \end{bmatrix},$$

then by definition of e_{j_ℓ}'s, we have

$$M \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \text{ which implies } rank\,(M) < N_2.$$

Therefore, there exists $[c_1, c_2, \ldots, c_{N_2}] \neq \vec{0}$, such that

$$\left[c_1, c_2, \ldots, c_{N_2}\right] M = \left[0, 0, \ldots, 0\right]. \tag{4.3.4}$$

Let

$$S_{k_\ell} := -\frac{c_\ell}{\eta_{k_\ell}(\Gamma_{x_{i_\ell}, y_{j_\ell}})} \int_{\Gamma_{x_{i_\ell}, y_{j_\ell}}} I_\gamma d\eta_{k_\ell} + \frac{c_\ell}{\eta_{k_\ell}(\Gamma_{x_{i_\ell}, y_{j'_\ell}})} \int_{\Gamma_{x_{i_\ell}, y_{j'_\ell}}} I_\gamma d\eta_{k_\ell},$$

for $\ell = 1, 2, \ldots, N_2$, and $S_k := 0$ for any other k's. Then $\partial S_{k_\ell} = c_\ell \delta_{y_{j'_\ell}} - c_\ell \delta_{y_{j_\ell}}$, for $\ell = 1, 2, \ldots, N_2$, and $\partial S_k = 0$ for any other k's.

For each $\ell = 1, 2, \ldots, N_2$, since

$$\eta_{k_\ell}(\Gamma_{x_{i_\ell}, y_{j_\ell}}) > 0,\ \eta_{k_\ell}(\Gamma_{x_{i_\ell}, y_{j'_\ell}}) > 0,$$

and by equation (4.3.4) we may assume that

$$0 < \max\{|c_\ell| : \ell = 1, 2, \ldots, N_2\} \leq \min\{\eta_{k_\ell}(\Gamma_{x_{i_\ell}, y_{j_\ell}}),\ \eta_{k_\ell}(\Gamma_{x_{i_\ell}, y_{j'_\ell}}) : \ell = 1, 2, \ldots, N_2\}.$$

This implies $\partial S_k = \rho_k(x) \partial T_k$, where $\rho_k(x)$ is defined as

$$\rho_{k_\ell}(x) = \begin{cases} -c_\ell / \sum_i \eta_{k_\ell}(\Gamma_{x_i, y_{j_\ell}}) & \text{if } x = y_{j_\ell} \\ c_\ell / \sum_i \eta_{k_\ell}(\Gamma_{x_i, y_{j'_\ell}}) & \text{if } x = y_{j'_\ell} \\ 0 & \text{otherwise,} \end{cases}$$

for $\ell = 1, 2, \ldots, N_2$, and $\rho_k(x) = 0$ for any other k's. Therefore, we have $|\rho_k(x)| \leq 1$, for all k and all x.

Also, since $\partial^- T_k(\{x\}) = 0$ for $x \notin \{x_1, x_2, \ldots, x_{N_1}\}$, and $\rho_k(x) = 0$ for $x \in \{x_1, x_2, \ldots, x_{N_1}\}$ then

$$\sum_k \rho_k(x) \partial^- T_k = \sum_k 0 \cdot \partial^- T_k = 0.$$

Since $\partial S_k = 0$ for $k \neq k_\ell$,

$$\sum_k \rho_k(x) \partial^+ T_k = \sum_k \rho_k(x) \left(\partial^+ T_k - \partial^- T_k\right) = \sum_k \partial S_k = \sum_{\ell=1}^{N_2} \partial S_{k_\ell} = \sum_{\ell=1}^{N_2} c_\ell \cdot (\delta_{y_{j'_\ell}} - \delta_{y_{j_\ell}}).$$

Next, by equation (4.3.4), we have

$$\begin{aligned}\sum_{\ell=1}^{N_2} c_\ell \cdot (\delta_{y_{j'_\ell}} - \delta_{y_{j_\ell}}) &= \begin{bmatrix} c_1, c_2, \ldots, c_{N_2} \end{bmatrix} M \begin{bmatrix} \delta_{y_1} \\ \delta_{y_2} \\ \vdots \\ \delta_{y_{N_2}} \end{bmatrix} = \begin{bmatrix} 0, 0, \ldots, 0 \end{bmatrix} \begin{bmatrix} \delta_{y_1} \\ \delta_{y_2} \\ \vdots \\ \delta_{y_{N_2}} \end{bmatrix} \\ &= \sum_{\ell=1}^{N_2} 0 \cdot \delta_{y_{k_\ell}} = 0.\end{aligned}$$

For $k = k_\ell$, with $\ell = 1, 2, \ldots, N_2$,

$$\|\partial^-(T_{k_\ell} \pm S_{k_\ell})\| = \|\partial^- T_{k_\ell}\| + 0 \leq c,$$

$$\|\partial^+(T_{k_\ell} \pm S_{k_\ell})\| = \mp c_\ell \pm c_\ell + \|\partial^+ T_{k_\ell}\| \leq c,$$

and for $k \neq k_\ell$, $\|\partial^-(T_k \pm S_k)\| = \|\partial^- T_k\| \leq c$ and $\|\partial^+(T_k \pm S_k)\| = \|\partial^+ T_k\| \leq c$, hold trivially.

Therefore, by Theorem 4.3.2, and $\vec{T}$ being α-optimal for $\alpha \in (0,1)$ imply that each S_k is a vanishing current, contradicting with the non-vanishing S_{k_ℓ}'s constructed above. Hence, except for at most $N_2 - 1$ transport path components (k's), the collection of sets

$$\{X_j(\eta_k) : j = 1, 2, \ldots, N_2\}$$

are mutually disjoint. □

In Chapter 3, we studied the decomposition of cycle-free transport paths such that some of the components are map-compatible. Now, we would like to introduce a similar result for optimal transport N-path $\vec{T} = (T_1, T_2, \ldots, T_N) \in Path_c(\mu^-, \mu^+)$.

Theorem 4.3.5. *Let μ^- and μ^+ be defined as in equation (4.3.2), and $\alpha \in (0,1)$. Let $\vec{T} = (T_1, T_2, \ldots, T_N) \in Path_c(\mu^-, \mu^+)$ be a solution to Problem 1. Then for each $k = 1, 2, \ldots, N$, there exists a better decomposition η_k of T_k such that*

$$|X_{j_1}(\eta_k) \cap X_{j_2}(\eta_k)| \leq 1,$$

for any $1 \leq j_1 < j_2 \leq N_2$. Moreover, except for at most $N_2 - 1$ many k's, T_k is a map-compatible transport path.

PROOF. Since $\vec{T} = (T_1, T_2, \ldots, T_N) \in Path_c(\mu^-, \mu^+)$ is a solution to Problem 1, for each k, T_k is an α-optimal transport path from μ_k^- to μ_k^+, which implies T_k is cycle-free. By Proposition 3.4.6, each T_k has a better decomposition η_k. Therefore, for any $k = 1, 2, \ldots, N$ and any $1 \le j_1 < j_2 \le N_2$, Proposition 3.4.7 gives

$$|X_{j_1}(\eta_k) \cap X_{j_2}(\eta_k)| \le 1.$$

By Definition 3.3.1, η_k is also a good decomposition for T_k.

Proposition 4.3.4 gives that, for each $k \in \{1, 2, \ldots, N\}$ with at most $N_2 - 1$ many exceptions, the collection of sets

$$\{X_j(\eta_k) : j = 1, 2, \ldots, N_2\}$$

are mutually disjoint, which implies that T_k is map-compatible by Proposition 3.5.5. □

4.4. Case study: Single target

In this section, we would like to investigate the case where μ^+ is supported on a single point, i.e. $\mu^+ = m\delta_y$. In this case, for simplicity of notation, we may denote

$$X(\eta_k) := \{x_i \in X : \eta_k(\Gamma_{x_i,y}) > 0\},$$

where η_k is a good decomposition as in equation (4.3.3). Also, in the following context, we assume $\alpha \in (0, 1)$ and $\vec{T} = (T_1, T_2, \ldots, T_N)$, consists of N components.

PROPOSITION 4.4.1. *Let $\mu^- = \sum_{i=1}^{N_1} m_i' \delta_{x_i}, \mu^+ = m\delta_y$, of equal mass, and $\vec{T} \in Path_c(\mu^-, \mu^+)$ is optimal. Then for $k_1 \neq k_2$,*

$$|X(\eta_{k_1}) \cap X(\eta_{k_2})| \le 1.$$

PROOF. For the sake of contradiction, we assume $|X(\eta_{k_1}) \cap X(\eta_{k_2})| \ge 2$ for some $k_1 \neq k_2$. Without loss of generality, assume $k_1 = 1, k_2 = 2$, and let $x_1, x_2 \in X(\eta_1) \cap X(\eta_2)$ with $x_1 \neq x_2$. Since η_1 and η_2 are good decomposition of T_1 and T_2 respectively, then

$$T_1 = \int_\Gamma I_\gamma d\eta_1, \ T_2 = \int_\Gamma I_\gamma d\eta_2,$$

and

$$\eta_1(\Gamma_{x_1,y}) > 0, \eta_1(\Gamma_{x_2,y}) > 0, \eta_2(\Gamma_{x_1,y}) > 0, \eta_2(\Gamma_{x_2,y}) > 0.$$

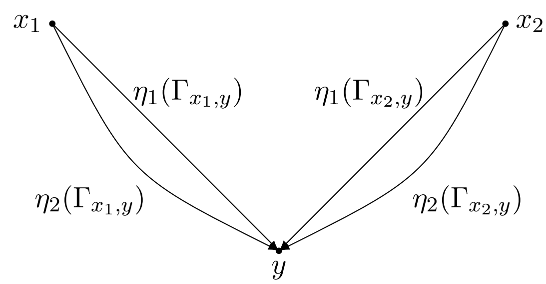


FIGURE 4.3. T_1 and T_2

Let $\epsilon_0 = \min\{\eta_1(\Gamma_{x_1,y}), \eta_1(\Gamma_{x_2,y}), \eta_2(\Gamma_{x_1,y}), \eta_2(\Gamma_{x_2,y})\} > 0$, and $\vec{S} = (S_1, S_2, S_3, \ldots, S_N)$ where

$$S_1 := \frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})}\int_{\Gamma_{x_1,y}} I_\gamma d\eta_1 - \frac{\epsilon_0}{\eta_1(\Gamma_{x_2,y})}\int_{\Gamma_{x_1,y}} I_\gamma d\eta_1$$

$$S_2 := -\frac{\epsilon_0}{\eta_2(\Gamma_{x_1,y})}\int_{\Gamma_{x_1,y}} I_\gamma d\eta_2 + \frac{\epsilon_0}{\eta_2(\Gamma_{x_2,y})}\int_{\Gamma_{x_1,y}} I_\gamma d\eta_2,$$

and $S_k := 0$ (vanishing currents) for $k \geq 3$. Suppose $T_1 = \underline{\underline{\tau}}(M, \theta, \xi)$ and $S_1 = \underline{\underline{\tau}}(N, \phi, \zeta)$, then

$$\begin{aligned}
\phi(x) &= \left|\left(\frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})}\eta_1\lfloor_{\Gamma_{x_1,y}} - \frac{\epsilon_0}{\eta_1(\Gamma_{x_2,y})}\eta_1\lfloor_{\Gamma_{x_2,y}}\right)(\Gamma_x)\right| \leq \epsilon_0\left(\frac{\eta_1(\Gamma_x \cap \Gamma_{x_1,y})}{\eta_1(\Gamma_{x_1,y})} + \frac{\eta_1(\Gamma_x \cap \Gamma_{x_2,y}}{\eta_1(\Gamma_{x_2,y})}\right) \\
&\leq \eta_1(\Gamma_x \cap \Gamma_{x_1,y}) + \eta_1(\Gamma_x \cap \Gamma_{x_2,y}) \leq \eta_1(\Gamma_x) = \theta(x).
\end{aligned}$$

As a result, S_1 is on T_1 because $\Gamma_{x_1,y} \subseteq \Gamma$.

Since

$$\partial T_1 = \int_\Gamma \partial I_\gamma d\eta_1 = \eta_1(\Gamma)\delta_y - \sum_{i=1}^{N_1}\eta_1(\Gamma_{x_i,y})\delta_{x_i} = \sum_{i=1}^{N_1}\eta_1(\Gamma_{x_i,y})\delta_y - \sum_{i=1}^{N_1}\eta_1(\Gamma_{x_i,y})\delta_{x_i},$$

and

$$\begin{aligned}
\partial S_1 &= \frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})}\int_{\Gamma_{x_1,y}} \partial I_\gamma d\eta_1 - \frac{\epsilon_0}{\eta_1(\Gamma_{x_2,y})}\int_{\Gamma_{x_1,y}} \partial I_\gamma d\eta_1 \\
&= \frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})}\int_{\Gamma_{x_1,y}} (\delta_y - \delta_{x_1}) d\eta_1 - \frac{\epsilon_0}{\eta_1(\Gamma_{x_2,y})}\int_{\Gamma_{x_1,y}} (\delta_y - \delta_{x_2}) d\eta_1 \\
&= \epsilon_0\delta_{x_2} - \epsilon_0\delta_{x_1},
\end{aligned}$$

one may express $\partial S_1 = \rho_1(x)\partial T_1$, where $\rho_1(x)$ is

$$\rho_1(x_1) = \epsilon_0/\eta_1(\Gamma_{x_1,y}),\ \rho_1(x_2) = -\epsilon_0/\eta_1(\Gamma_{x_2,y}),$$

and $\rho_1(x) = 0$ for $x \neq x_1, x_2$.

In general, by doing similar calculation as above, for $k = 1, 2, \ldots, N$, we have S_k is on T_k, $\partial S_k = \rho_k(x)\partial T_k$, and $|\rho_k(x)| \le 1$ where

	$x = x_1$	$x = x_2$	otherwise
$\rho_1(x)$	$\epsilon_0/\eta_1(\Gamma_{x_1,y})$	$-\epsilon_0/\eta_1(\Gamma_{x_2,y})$	0
$\rho_2(x)$	$-\epsilon_0/\eta_2(\Gamma_{x_1,y})$	$\epsilon_0/\eta_2(\Gamma_{x_2,y})$	0
$\rho_k(x), k \ge 3$	0	0	0

When $k = 1, 2$,

$$\|\partial^-(T_1 \pm S_1)\| = (\eta_1(\Gamma_{x_1,y}) \pm \epsilon_0) + (\eta_1(\Gamma_{x_2,y}) \mp \epsilon_0) + \sum_{i=3}^{N_1} \eta_1(\Gamma_{x_i,y}) = \sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y}) = \|\partial^- T_1\| \le c,$$

$$\|\partial^+(T_1 \pm S_1)\| = \|\partial^+ T_1\| \le c,$$

$$\|\partial^-(T_2 \pm S_2)\| = (\eta_2(\Gamma_{x_1,y}) \mp \epsilon_0) + (\eta_2(\Gamma_{x_2,y}) \pm \epsilon_0) + \sum_{i=3}^{N_1} \eta_2(\Gamma_{x_i,y}) = \sum_{i=1}^{N_1} \eta_2(\Gamma_{x_i,y}) = \|\partial^- T_2\| \le c,$$

$$\|\partial^+(T_2 \pm S_2)\| = \|\partial^+ T_2\| \le c.$$

When $k \ge 3$,

$$\|\partial^-(T_k \pm S_k)\| = \|\partial^- T_k\| \le c, \ \|\partial^+(T_k \pm S_k)\| = \|\partial^+ T_k\| \le c.$$

Also,

$$\begin{aligned} \sum_{k=1}^{N} \rho_k(x)\partial^- T_k &= \frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})} \cdot \eta_1(\Gamma_{x_1,y})\delta_{x_1} - \frac{\epsilon_0}{\eta_2(\Gamma_{x_1,y})} \cdot \eta_2(\Gamma_{x_1,y})\delta_{x_1} \\ &\quad - \frac{\epsilon_0}{\eta_1(\Gamma_{x_2,y})} \cdot \eta_1(\Gamma_{x_2,y})\delta_{x_2} + \frac{\epsilon_0}{\eta_2(\Gamma_{x_2,y})} \cdot \eta_2(\Gamma_{x_2,y})\delta_{x_2} \\ &= 0, \end{aligned}$$

and

$$\sum_{k=1}^{N} \rho_k(x)\partial^+ T_k = \sum_{k=1}^{N} 0 = 0.$$

By Theorem 4.3.2, for $\alpha \in (0, 1)$, each S_k is a vanishing current, contradicting with the non-vanishing S_1, S_2 constructed above. Hence, $|X(\eta_1) \cap X(\eta_2)| \le 1$. $\square$

Proposition 4.4.2. *Let $\mu^- = \sum_{i=1}^{N_1} m'_i \delta_{x_i}, \mu^+ = m\delta_y$, of equal mass, and $\vec{T} \in Path_c(\mu^-, \mu^+)$ is optimal. Suppose $k_1 \ne k_2$ and $|X(\eta_{k_1}) \cap X(\eta_{k_2})| = 1$, then either $\|\mu^-_{k_1}\| = c$ or $\|\mu^-_{k_2}\| = c$.*

Proof. Without loss of generality, assume $k_1 = 1, k_2 = 2$, and let $x_1 \in X(\eta_1) \cap X(\eta_2)$. Arguing by contradiction, assuming $\|\mu_1^-\|, \|\mu_2^-\| < c$. Since η_1, η_2 are good decomposition of T_1, T_2 respectively, then

$$T_1 = \int_\Gamma I_\gamma d\eta_1, \ T_2 = \int_\Gamma I_\gamma d\eta_2,$$

and

$$\eta_1(\Gamma_{x_1,y}) > 0, \ \eta_2(\Gamma_{x_1,y}) > 0.$$

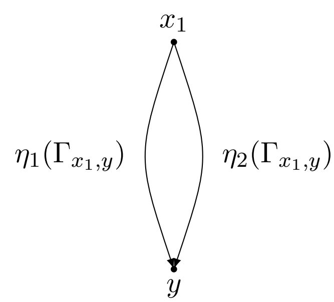


Figure 4.4. T_1 and T_2

Since $\|\mu_1^-\|, \|\mu_2^-\| < c$, the let ϵ_0 such that $0 < \epsilon_0 = \min\{\eta_1(\Gamma_{x_1,y}), \eta_2(\Gamma_{x_1,y}), c - \|\mu_1^-\|, c - \|\mu_2^-\|\}$.

Let $\vec{S} = (S_1, S_2, S_3, \ldots, S_N)$, where

$$S_1 := \frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})} \int_{\Gamma_{x_1,y}} I_\gamma d\eta_1, \ S_2 := -\frac{\epsilon_0}{\eta_2(\Gamma_{x_1,y})} \int_{\Gamma_{x_1,y}} I_\gamma d\eta_2,$$

and $S_k := 0$ for $k \geq 3$. Construction of S_k's gives S_k is on T_k, for all $k = 1, 2, \ldots, N$.

Since

$$\partial T_1 = \int_\Gamma \partial I_\gamma d\eta_1 = \eta_1(\Gamma)\delta_y - \sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y})\delta_{x_i} = \sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y})\delta_y - \sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y})\delta_{x_i},$$

and

$$\partial S_1 = \frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})} \int_{\Gamma_{x_1,y}} \partial I_\gamma d\eta_1 = \frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})} \int_{\Gamma_{x_1,y}} (\delta_y - \delta_{x_1}) d\eta_1 = \epsilon_0\delta_y - \epsilon_0\delta_{x_1},$$

then $\rho_1(x_1) = \epsilon_0/\eta_1(\Gamma_{x_1,y})$, $\rho_1(y) = \epsilon_0/\sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y})$, and $\rho_1(x) = 0$ for $x \neq x_1, y$.

By performing similar calculation as above, for $k = 1, 2, \ldots, n$, we get $\partial S_k = \rho_k(x)\partial T_k$, with $|\rho_k(x)| \leq 1$ such that

	$x = x_1$	$x = y$	otherwise
$\rho_1(x)$	$\epsilon_0/\eta_1(\Gamma_{x_1,y})$	$\epsilon_0/\sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y})$	0
$\rho_2(x)$	$-\epsilon_0/\eta_2(\Gamma_{x_1,y})$	$-\epsilon_0/\sum_{i=1}^{N_1} \eta_2(\Gamma_{x_i,y})$	0
$\rho_k(x), k \geq 3$	0	0	0

When $k = 1, 2$, since $\|\mu_k^-\| = \|\mu_k^+\|$,

$$\|\partial^-(T_1 \pm S_1)\| = \eta_1(\Gamma_{x_1,y}) \pm \epsilon_0 + \sum_{i=2}^{N_1} \eta_1(\Gamma_{x_i,y}) = \pm\epsilon_0 + \|\mu_1^-\| \leq c,$$

$$\|\partial^+(T_1 \pm S_1)\| = \pm\epsilon_0 + \sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y}) = \pm\epsilon_0 + \|\mu_1^-\| \leq c,$$

$$\|\partial^-(T_2 \pm S_2)\| = \eta_2(\Gamma_{x_1,y}) \mp \epsilon_0 + \sum_{i=2}^{N_1} \eta_2(\Gamma_{x_i,y}) = \mp\epsilon_0 + \|\mu_2^-\| \leq c,$$

$$\|\partial^+(T_2 \pm S_2)\| = \mp\epsilon_0 + \sum_{i=1}^{N_1} \eta_2(\Gamma_{x_i,y}) = \mp\epsilon_0 + \|\mu_2^-\| \leq c,$$

When $k \geq 3$,

$$\|\partial^-(T_k \pm S_k)\| = \|\partial^- T_k\| \leq c, \ \|\partial^+(T_k \pm S_k)\| = \|\partial^+ T_k\| \leq c.$$

Also,

$$\sum_{k=1}^{N} \rho_k(x)\partial^- T_k = \frac{\epsilon_0}{\eta_1(\Gamma_{x_1,y})} \cdot \eta_1(\Gamma_{x_1,y})\delta_{x_1} - \frac{\epsilon_0}{\eta_2(\Gamma_{x_1,y})} \cdot \eta_2(\Gamma_{x_1,y})\delta_{x_1} = 0,$$

and

$$\sum_{k=1}^{N} \rho_k(x)\partial^+ T_k = \frac{\epsilon_0}{\sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y})} \cdot \sum_{i=1}^{N_1} \eta_1(\Gamma_{x_i,y})\delta_y - \frac{\epsilon_0}{\sum_{i=1}^{N_1} \eta_2(\Gamma_{x_i,y})} \cdot \sum_{i=1}^{N_1} \eta_2(\Gamma_{x_i,y})\delta_y = 0.$$

Theorem 4.3.2 implies for $\alpha \in (0, 1)$, each S_k is a vanishing current, but S_1, S_2 constructed above are non-vanishing, and this leads to a contradiction. Hence, we have one of the values between $\|\mu_1^-\|$ and $\|\mu_2^-\|$ equals c. □

Corollary 4.4.3. *Let $\mu^- = \sum_{i=1}^{N_1} m_i' \delta_{x_i}, \mu^+ = m\delta_y$, of equal mass, and $\vec{T} \in Path_c(\mu^-, \mu^+)$ is optimal. Suppose*

$$\bigcap_{\ell=1}^{n} X(\eta_{k_\ell}) \neq \emptyset \text{ for some } n \geq 2.$$

Then at most one of the μ_{k_ℓ} has $\|\mu_{k_\ell}\| < c$, and any other μ_{k_ℓ}'s have mass $\|\mu_{k_\ell}\| = c$.

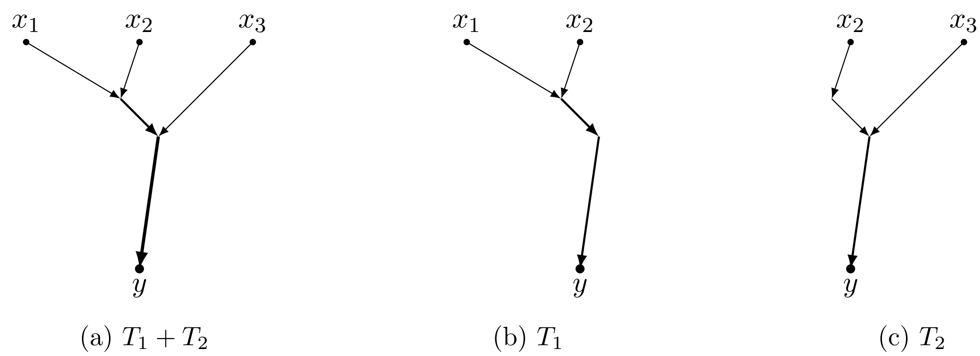


Figure 4.5. Demonstration of $X(\eta_1) \cap X(\eta_2) \neq \emptyset$, where $X(\eta_1) = \{x_1, x_2\}$, and $X(\eta_2) = \{x_2, x_3\}$.

Proof. Suppose there exist two components μ_{k_1}, μ_{k_2} with $\|\mu_{k_1}\|, \|\mu_{k_2}\| < c$. Proposition 4.4.1 implies $|X(\eta_{k_1}) \cap X(\eta_{k_2})| \leq 1$. Since $X(\eta_{k_1}) \cap X(\eta_{k_2})$ is non-empty, then $|X(\eta_{k_1}) \cap X(\eta_{k_2})| = 1$. Proposition 4.4.2 implies $\|\mu_{k_1}\| = c$ or $\|\mu_{k_2}\| = c$, which leads to contradiction. □

Results that have proved so far characterize the "support" of component measures and the weight on components of an optimal transport path. Next, we would like to apply these results to some specific cases: transport path from 1 point to 1 point and transport path from 2 points to 1 point. In the following Corollaries, denote the line segment from x to y as $\overline{xy}$. Also, denote $a[\![\gamma]\!]$ as the rectifiable 1-current, with density equals a, supported on the curve γ, and direction along this curve.

Corollary 4.4.4. *Suppose $\mu^- = m_0\delta_x$, $\mu^+ = m_0\delta_y$, and $\vec{T} \in Path_c(\mu^-, \mu^+)$ is optimal. Then up to a permutation of component indices,*

$$T_1, T_2, \ldots, T_{N-1} = c[\![\overline{xy}]\!],\ T_N = r_0[\![\overline{xy}]\!],$$

with $N = \lceil m_0/c \rceil$, $r_0 = m_0 - (N-1)c$.

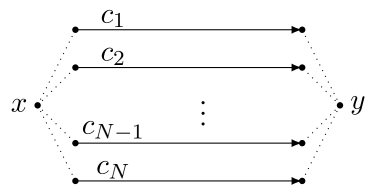


FIGURE 4.6. 1 point to 1 point.

PROOF. For $k = 1, 2, \ldots, N$, since the minimum path between two points in $\mathbb{R}^m$ is a line segment, so that $T_k = c_k[\![\overline{xy}]\!]$ for $0 < c_k \leq c$. Suppose there exist k_1, k_2 with $k_1 \neq k_2$ such that $c_{k_1}, c_{k_2} < c$, then $|X(\eta_{k_1}) \cap X(\eta_{k_2})| = |\{x\}| = 1$ and Proposition 4.4.2 implies one of the values between c_{k_1} and c_{k_2} equals c, which leads to contradiction.

Hence, for $k = 1, 2, \ldots, N$, there is at most one component k (without loss of generality assume this component index is N) such that $T_N = r_0[\![\overline{xy}]\!], r_0 \in (0, c]$, and any other components are $T_k = c[\![\overline{xy}]\!]$. The total number of components required is $N = \lceil m_0/c \rceil$, and since there is only one component has mass less or equal to c, then $r_0 = m_0 - (N-1)c$. $\square$

COROLLARY 4.4.5. *Suppose* $\mu^- = m_1\delta_{x_1} + m_2\delta_{x_2}$, $\mu^+ = (m_1 + m_2)\delta_y$, *and* $\vec{T} \in Path_c(\mu^-, \mu^+)$ *is optimal. Then there exists at most one* $k = 1, 2, \ldots, N$, *such that* $|X(\eta_k)| = 2$. *Moreover, there exist* $n_1, n_2 \in \{0\} \cup \mathbb{N}$ *with* $N = n_1 + n_2 + 2$ *such that*

(1) *if* $|X(\eta_k)| = 1$ *for each* $k = 1, 2, \ldots, N$, *then up to a permutation of component indices,*

$$T_1, T_2, \ldots, T_{n_1} = c[\![\overline{x_1y}]\!],\ T_{n_1+1}, T_{n_1+2}, \ldots, T_{N-2} = c[\![\overline{x_2y}]\!],\ T_{N-1} = \epsilon_1[\![\overline{x_1y}]\!],\ T_N = \epsilon_2[\![\overline{x_2y}]\!],$$

for $0 < \epsilon_1, \epsilon_2 \leq c$, *and* $n_1 = \lceil m_1/c \rceil - 1, n_2 = \lceil m_2/c \rceil - 1$;

(2) *if* $|X(\eta_k)| = 2$ *for some* $k = 1, 2, \ldots, N$, *then up to a permutation of component indices,* $k = N$,

$$T_1, T_2, \ldots, T_{n_1} = c[\![\overline{x_1y}]\!],\ T_{n_1+1}, T_{n_1+2}, \ldots, T_{N-2} = c[\![\overline{x_2y}]\!],\ T_{N-1} = \epsilon_3[\![\overline{x_1y}]\!] \text{ or } \epsilon_3[\![\overline{x_2y}]\!],$$

for $0 \leq \epsilon_3 \leq c$, *and* $\max\{0, \lceil (m_1 + m_2 - c)/c \rceil\} \leq N - 1 = n_1 + n_2 + 1 < \lceil (m_1 + m_2)/c \rceil$.

PROOF. For $k_1 \neq k_2$, suppose there are two transport path components, indexed as k_1, k_2, such that $|X(\eta_{k_1})| = |X(\eta_{k_2})| = 2$. Since $|supp(\mu^-)| = |\{x_1, x_2\}| = 0$, then $|X(\eta_{k_1}) \cap X(\eta_{k_2})| = |\{x_1, x_2\}| = 2$, which contradicts Proposition 4.4.1. This implies there exists at most one k such that $|X(\eta_k)| = 2$.

Case (1): Suppose $|X(\eta_k)| = 1$ for each $k = 1, 2, \ldots, N$, since the path with minimum distance from x_1, x_2 to y are line segments, this implies $T_k = c_k[\![\overline{x_1 y}]\!]$, $c_k[\![\overline{x_2 y}]\!]$.

(a) $c_k[\![\overline{x_1 y}]\!]$ (b) $c_k[\![\overline{x_2 y}]\!]$

FIGURE 4.7. Transport paths in Case 1.

Suppose there exist $T_{k_1} = c_{k_1}[\![\overline{x_1 y}]\!], T_{k_2} = c_{k_2}[\![\overline{x_1 y}]\!]$ with $k_1 \neq k_2$ such that $c_{k_1}, c_{k_2} < c$, then $|X(\eta_{k_1}) \cap X(\eta_{k_2})| = 1$. Proposition 4.4.2 gives one of values between c_{k_1} and c_{k_2} equals c, which leads to contradiction. Hence, there is at most one component, and without loss of generality this component is indexed by $N-1$, such that $T_{N-1} = \epsilon_1[\![\overline{x_1 y}]\!]$, with $0 < \epsilon_1 \leq c$. Any other components that transport mass from x_1 to y are $c[\![\overline{x_1 y}]\!]$. This gives the total number of components that transport mass from x_1 to y is $\lceil m_1/c \rceil$, $n_1 = \lceil m_1/c \rceil - 1$, and $\epsilon_1 = m_1 - (\lceil m_1/c \rceil - 1)c$.

Similarly, there is at most one component, and without loss of generality this component is indexed by N, such that $T_N = \epsilon_2[\![\overline{x_2 y}]\!]$, with $0 < \epsilon_2 \leq c$. Any other components that transport mass from x_2 to y are $c[\![\overline{x_2 y}]\!]$. This gives the total number of components that transport mass from x_2 to y is $\lceil m_2/c \rceil$, $n_2 = \lceil m_2/c \rceil - 1$, and $\epsilon_2 = m_2 - (\lceil m_2/c \rceil - 1)c$.

Case (2): Suppose $|X(\eta_k)| = 2$ for some $k = 1, 2, \ldots, N$, and without loss of generality assume this component is indexed by N. Then Proposition 4.4.1 implies all the remaining transport paths components are line segments from x_1 to y and x_2 to y.

By using similar argument as previous case, among all transport path components that transport mass from x_1 to y, there is at most one component $T_{k_1} = \epsilon_{k_1}[\![\overline{x_1 y}]\!]$ with $0 < \epsilon_{k_1} \leq c$, and among all transport path components that transport mass from x_2 to y, there is at most one component $T_{k_2} = \epsilon_{k_2}[\![\overline{x_2 y}]\!]$ with $0 < \epsilon_{k_2} \leq c$. Moreover, we claim that either $\epsilon_{k_1} = c$ or $\epsilon_{k_2} = c$. By contradiction, assume $0 < \epsilon_{k_1}, \epsilon_{k_2} < c$. Suppose η_N is a good decomposition of T_N, such that

$$T_N = \int_{\Gamma_{x_1,y}} I_\gamma d\eta_N + \int_{\Gamma_{x_2,y}} I_\gamma d\eta_N,$$

and by definition of $|X(\eta_k)| = 2$,

$$\eta_N(\Gamma_{x_1,y}) > 0, \eta_N(\Gamma_{x_2,y}) > 0.$$

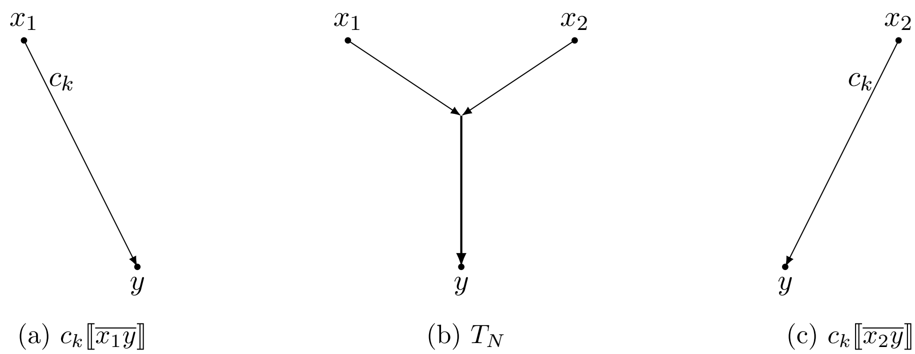


FIGURE 4.8. Transport paths in Case 2.

Let $\epsilon_0 = \min\{\epsilon_{k_1}, \epsilon_{k_2}, c - \epsilon_{k_1}, c - \epsilon_{k_2}, \eta_N(\Gamma_{x_1,y}), \eta_N(\Gamma_{x_2,y})\} > 0$, and define $\vec{S} = (S_1, S_2, \ldots, S_N)$, where

$$S_{k_1} := \epsilon_0[\![\overline{x_1 y}]\!],\ S_{k_2} := -\epsilon_0[\![\overline{x_2 y}]\!],\ S_N := -\frac{\epsilon_0}{\eta_N(\Gamma_{x_1,y})}\int_{\Gamma_{x_1,y}} I_\gamma d\eta_N + \frac{\epsilon_0}{\eta_N(\Gamma_{x_2,y})}\int_{\Gamma_{x_2,y}} I_\gamma d\eta_N,$$

and $S_k = 0$ for $k \neq k_1, k_2, N$. By construction, S_k is on T_k, for each k. The corresponding $\rho_k(x)$'s, where $\partial S_k = \rho(x)\partial T_k$ for each k, are

	$x = x_1$	$x = x_2$	$x = y$	otherwise
$\rho_{k_1}(x)$	$\epsilon_0/\epsilon_{k_1}$	0	$\epsilon_0/\epsilon_{k_1}$	0
$\rho_{k_2}(x)$	0	$-\epsilon_0/\epsilon_{k_2}$	$-\epsilon_0/\epsilon_{k_2}$	0
$\rho_N(x)$	$-\epsilon_0/\eta_N(\Gamma_{x_1,y})$	$\epsilon_0/\eta_N(\Gamma_{x_2,y})$	0	0

and $\rho_k(x) = 0$, for $k \neq k_1, k_2, N$.

Direct calculation shows that the non-vanishing S_k's constructed above satisfy conditions in Theorem 4.3.2, and when $\alpha \in (0,1)$, $\vec{T}$ is optimal, Theorem 4.3.2 gives each S_k is a vanishing current. This leads to a contradiction.

This implies one of the values between ϵ_{k_1} and ϵ_{k_2} equals c. Hence, there is only one component, and without loss of generality index it by $N-1$, $T_{N-1} = \epsilon_3[\![\overline{x_1 y}]\!]$ or $\epsilon_3[\![\overline{x_2 y}]\!]$ with $0 < \epsilon_3 \leq c$, and any other transport path components (line segments) have weight equals c. Since $0 < \|\mu_N^-\| \leq c$, then $0 \leq \sum_{k=1}^{N-1} \|\mu_k^-\| < m_1 + m_2$, and this gives

$$\max\{0, \lceil (m_1 + m_2 - c)/c \rceil\} \leq N - 1 = n_1 + n_2 + 1 < \lceil (m_1 + m_2)/c \rceil.$$

Note that because of the T_N and T_{N-1} components, we also have

$$\lceil m_1/c \rceil - 2 \leq n_1 \leq \lceil m_1/c \rceil - 1, \text{ and } \lceil m_2/c \rceil - 2 \leq n_2 \leq \lceil m_2/c \rceil - 1$$

□

Bibliography

[1] L. Ambrosio, E. Brué, and D. Semola, *Lectures on optimal transport*, Springer, Volume 130 (2021).

[2] M. Colombo, A. DeRosa, and M. A, *Improved stability of optimal traffic paths*, Calc. Var. Partial Differ. Equ., 57:28 (2018).

[3] M. Colombo, A. DeRosa, and A. Marchese, *On the well-posedness of branched transportation*, Comm. Pure Appl. Math., 74 (2021), pp. 833–864.

[4] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions revised edition*, CRC Press Taylor & Francis Group, (2015).

[5] F. Lin and X. Yang, *Geometric measure theory an introduction*, Science Press and International Press, (2002).

[6] E. Paolini and E. Stepanov, *Decomposition of acyclic normal currents in a metric space*, J Funct Anal, Vol 263, Issue 11 (2012), pp. 3358–3390.

[7] F. Santambrogio, *Optimal transport for applied mathematicians*, Springer, (2015).

[8] L. Simon, *Introduction to geometric measure theory*, https://web.stanford.edu/class/math285/ts-gmt.pdf, (2014).

[9] S. K. Smirnov, *Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows*, Algebra i Analiz, 5 (1993), pp. 206–238.

[10] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, 58 (2003).

[11] Q. Xia, *Optimal paths related to transport problems*, Commun. Contemp. Math., 5 (2003), pp. 251–279.

[12] ———, *Interior regularity of optimal transport paths*, Calc. Var. Partial Differ. Equ., 20 (2004), pp. 421–443.

[13] ———, *The formation of a tree leaf*, ESAIM Control Optim. Calc. Var., 13 (2007), pp. 359–377.

[14] ———, *Boundary regularity of optimal transport paths*, Adv. Calc. Var., 4 (2011), pp. 153–174.

[15] ———, *Motivations, ideas, and applications of ramified optimal transportation*, ESAIM: Math. Model. Numer. Anal., 49 (2015), pp. 1791 – 1832.

[16] Q. Xia and H. Sun, *Map-compatible decomposition of transport paths*, arXiv:2310.03825, (2023).

[17] Q. Xia and S. Xu, *On the ramified optimal allocation problem*, Netw. Heterog. Media, 8 (2013), pp. 591 – 624.

[18] Q. Xia and S. Xu, *Ramified optimal transportation with payoff on the boundary*, SIAM J. Math. Anal., Vol 55, No. 1 (2023), pp. 186–209.