

## OPTIMAL PATHS RELATED TO TRANSPORT PROBLEMS

QINGLAN XIA

*Rice University, Mathematics, Houston, TX, 77005, USA*  
*qlxia@rice.edu*

Received 4 December 2001

Revised 1 July 2002

In transport problems of Monge's types, the total cost of a transport map is usually an integral of some function of the distance, such as  $|x - y|^p$ . In many real applications, the actual cost may naturally be determined by a transport path. For shipping two items to one location, a "Y shaped" path may be preferable to a "V shaped" path. Here, we show that any probability measure can be transported to another probability measure through a general optimal transport path, which is given by a vector measure in our setting. Moreover, we define a new distance on the space of probability measures which in fact metrizes the weak \* topology of measures. Under this distance, the space of probability measures becomes a length space. Relations as well as related problems about transport paths and transport plans are also discussed in the end.

*Keywords:* Optimal transport problem; size minimizing current; optimal transport path.

Mathematics Subject Classification 2000: 90B06, 49Q20

### 1. Introduction

The transport problem introduced by Monge in 1781 [8] has been studied in many interesting works in the last 10 years [1, 4, 5, 7, 11]. In these works, the cost of a transport mapping or a transport plan is usually an integral of some convex (or concave [7]) function of the distance, such as  $|x - y|^p$ . However, in some real applications, the actual cost of the transport procedures is not necessarily determined by just knowing an optimal mapping from the starting position to the target position. For example in shipping two items from nearby cities to the same far away city, it may be less expensive to first bring them to a common location and put them on a single truck for most of the transport. In this case, a "Y shaped" path is preferable to a "V shaped" path. In both cases, the transport mapping is trivially the same, but the actual transport path naturally gives the total cost. We may consider the following general problem.

**Problem 1.1.** Given two general probability measures  $\mu^+$  and  $\mu^-$ , find an optimal path for transporting  $\mu^+$  to  $\mu^-$ .

To solve this problem, one needs to find a suitable category of transport paths as well as a suitable cost functional acting on these paths. Such a category should be broad enough to give existence of an optimal transport path. Also, an optimal transport path should allow the possibility that some parts overlap in a cost efficient (maybe complicated) fashion but still enjoy some nice regularity properties. If possible, one may hope to visualize such an optimal transport path using numerical analysis and computer graphics.

The family of paths we choose in this paper is a subset of the space of vector measures with divergence being the difference of the given two measures in the sense of distribution. A transport path between two atomic measures is just a directed graph with balanced weighting at interior vertices. For arbitrary measures, a transport path between them is a vector measure given by a limit of some weighted directed graphs. The cost on each transport path is a suitably modified weighted mass of the vector measure (see Sec. 2, Definition 2.2), similar to the  $H$  mass of integral currents in [3]. Unlike in [3], we work with vector measures (or 1 dimensional flat currents in the language of geometric measure theory) whose multiplicities are not necessarily integer valued. With this category and cost functional, the original optimal transport path problem becomes a Plateau-type problem as in the study of minimal surfaces. Luckily, we have the existence theorem of an optimal transport path joining any given probability measure to another.

In Sec. 4, we consider a new distance on the space of probability measures on a fixed convex set. Such a distance is different from any of the Wasserstein distances [4], but still metrizes the weak  $*$  topology of the space of probability measures.

In Sec. 5, we show that the space of probability measures with the above distance becomes a length space. The geodesic between two probability measures is an optimal transport path between them.

In Sec. 6, we use numeric analysis to visualize optimal transport paths.

In the last section, we discuss the relationship between transport paths and transport plans. A compatible pair of a transport plan and a transport path contains necessary information about the actual transportation such as how, where and when the original measure is decomposed into the targeting measure on the road.

We are presently considering regularity questions about these optimal transport paths. It seems likely that the optimal transport path is given by a countable sum of oriented segments.

In this paper, we will use the following notations:

- $X$ : a compact convex subset of a Euclidean space  $\mathbb{R}^m$ .
- $Q$ : a cube in  $\mathbb{R}^m$  contains  $X$ .
- $d$ : the edge length of  $Q$ .
- $c$ : the center of  $Q$ .
- $\alpha$ : a positive number in  $[0, 1]$ .
- $\mathcal{M}_\Lambda(X)$ : the space of Radon measure  $\mu$  on  $X$  with total mass  $\mu(X) = \Lambda$ .
- $\mathcal{M}^m(X)$ : the space of Radon vector measure  $\mu = (\mu_1, \mu_2, \dots, \mu_m)$  on  $X$ .

- $\|\mu\|$ : the total variational measure of any vector measure  $\mu \in \mathcal{M}^m(X)$ .
- $W$ : the Wasserstein 1-distance on  $\mathcal{M}_\Lambda(X)$ .

## 2. Transport Atomic Measures

Recall that a Radon measure  $a$  on  $X$  is *atomic* if  $a$  is a finite sum of Dirac measures with positive multiplicities. That is

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

for some integer  $k \geq 1$  and some points  $x_i \in X$ ,  $a_i > 0$  for each  $i = 1, \dots, k$ . For any  $\Lambda > 0$ , let

$$\mathcal{A}_\Lambda(X) \subset \mathcal{M}_\Lambda(X)$$

be the space of all atomic measures on  $X$  with total mass  $\Lambda$ .

Suppose

$$a = \sum_{i=1}^k a_i \delta_{x_i} \text{ and } b = \sum_{j=1}^l b_j \delta_{y_j} \in \mathcal{A}_\Lambda(X) \tag{2.1}$$

are two atomic measures on  $X$  of equal total mass:

$$\sum_{i=1}^k a_i = \Lambda = \sum_{j=1}^l b_j.$$

**Definition 2.1.** A **transport path from  $a$  to  $b$**  is a weighted directed graph  $G$  consists of a vertex set  $V(G)$ , a directed edge set  $E(G)$  and a weight function

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

such that

- (1)  $\{x_1, x_2, \dots, x_k\} \cup \{y_1, y_2, \dots, y_l\} \subset V(G)$ .
- (2) For each source vertex  $x_i$ ,  $i = 1, \dots, k$ ,

$$a_i = \sum_{\substack{e \in E(G) \\ e^- = x_i}} w(e),$$

where  $e^-$  denotes the starting endpoint of the edge  $e \in E(G)$ .

- (3) For each  $j = 1, \dots, l$ ,

$$b_j = \sum_{\substack{e \in E(G) \\ e^+ = y_j}} w(e),$$

where  $e^+$  denotes the ending endpoint of the edge  $e \in E(G)$ .

(4) For any interior vertex  $v \in V(G)$ ,

$$\sum_{\substack{e \in E(G) \\ e^- = v}} w(e) = \sum_{\substack{e \in E(G) \\ e^+ = v}} w(e).$$

In other words,  $G$  satisfies the Kirchoff's law at each of its interior vertices.

Denote

$$\text{Path}(a, b) = \{\text{all transport paths from } a \text{ to } b\}$$

and

$$\mathcal{G}_\Lambda(X) = \bigcup_{(a,b) \in \mathcal{A}_\Lambda(X) \times \mathcal{A}_\Lambda(X)} \text{Path}(a, b),$$

the union of all transport paths between atomic measures on  $X$  of equal total mass  $\Lambda$ .

**Remark 2.1.** Each transport path  $G \in \text{Path}(a, b)$  determines the vector measure on  $X$

$$G = \sum_{e \in E(G)} w(e)[[e]],$$

where  $[[e]]$  is the vector measure  $\mathcal{H}^1 \llcorner_e \mathbf{e}$  for each edge  $e \in E(G)$  with unit directional vector  $\mathbf{e}$ . The above conditions can be simplified to be a single divergence condition on  $G$

$$\text{div}(G) = a - b,$$

in the sense of distribution. Thus,

$$\text{Path}(a, b) = \left\{ G = \sum_{e \in E(G)} w(e)[[e]] \in \mathcal{M}^m(X) : \text{div}(G) = a - b \right\}.$$

Among all paths in  $\text{Path}(a, b)$ , we want to find an optimal path which allows the possibility that some parts overlap in a cost efficient fashion. To get such a ‘‘Y-shaped’’ optimal path, we define the following cost function on  $\mathcal{G}_\Lambda(X)$ .

**Definition 2.2.** The  $M^\alpha$  cost function on  $\mathcal{G}_\Lambda(X)$  is defined by

$$M^\alpha(G) \equiv \sum_{e \in E(G)} w(e)^\alpha \text{length}(e)$$

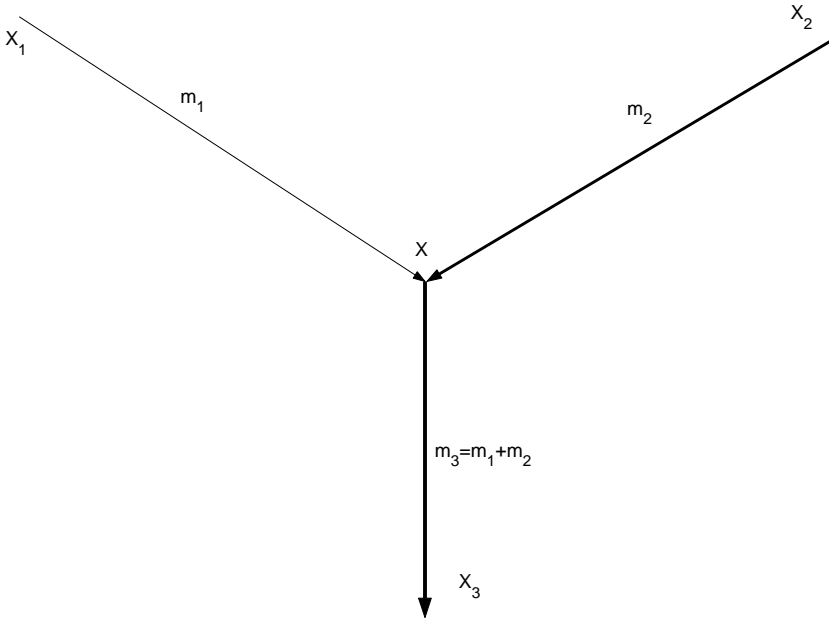
for any transport path  $G = \sum_{e \in E(G)} w(e)[[e]] \in \mathcal{G}_\Lambda(X)$ .

- This definition is motivated by the study of scans in [9] and also the  $H$  mass of integer multiplicity rectifiable currents in [3] or [12].
- Note that, when  $\alpha = 1$ , we have

$$M^1(G) = \|G\|(X),$$

where  $\|G\|$  is the total variational measure of  $G$ .

**Example 2.1.** Let  $a = m_1\delta_{x_1} + m_2\delta_{x_2}$  and  $b = m_3\delta_{x_3}$  with  $m_3 = m_1 + m_2$ . Then the optimal transport path from  $a$  to  $b$  under the  $M^\alpha$  cost looks like the following “Y shaped” graph



Here the interior vertex  $x$  is determined by a balance formula:

$$m_1^\alpha \mathbf{n}_1 + m_2^\alpha \mathbf{n}_2 = m_3^\alpha \mathbf{n}_3,$$

where  $\mathbf{n}_i = \frac{x-x_i}{|x-x_i|}$  is the unit vector from  $x$  to  $x_i$ ,  $i = 1, 2, 3$ . Let  $\theta_i$  be the angle between  $\mathbf{n}_i$  and  $-\mathbf{n}_3$  for  $i = 1, 2$  and  $k_1 = \frac{m_1}{m_1+m_2}$ ,  $k_2 = \frac{m_2}{m_1+m_2} = 1 - k_1$ . Then the above formula implies that the angles satisfy

$$\cos \theta_1 = \frac{k_1^\alpha + 1 - k_2^\alpha}{2k_1^\alpha},$$

$$\cos \theta_2 = \frac{k_2^\alpha + 1 - k_1^\alpha}{2k_2^\alpha},$$

and

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

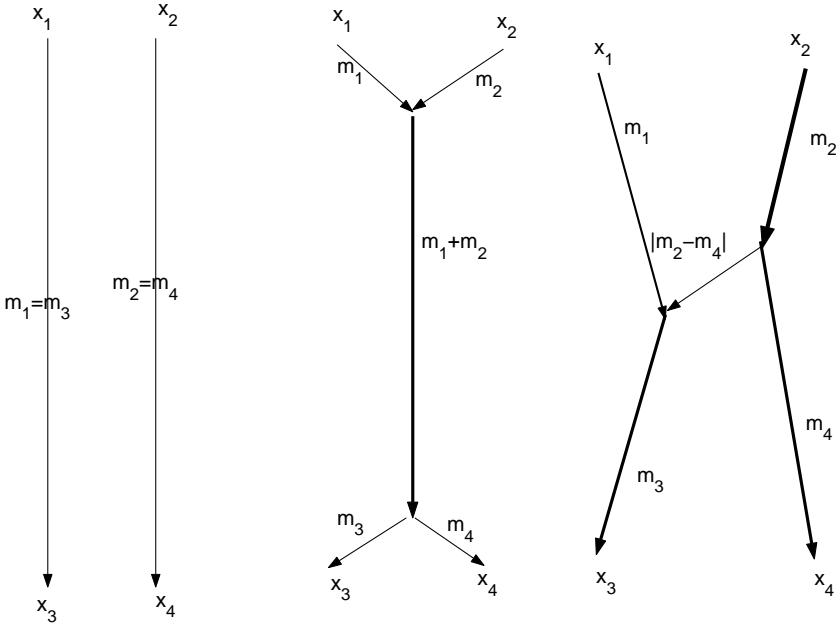
In particular, if  $m_1 = m_2$ , then

$$\theta_1 + \theta_2 = \arccos(2^{2\alpha-1} - 1).$$

Also, if  $\alpha = 1/2$ , then  $\theta_1 + \theta_2 = \pi/2$  for any  $m_1$  and  $m_2$ .

**Example 2.2.** For  $a = m_1\delta_{x_1} + m_2\delta_{x_2}$  and  $b = m_3\delta_{x_3} + m_4\delta_{x_4}$  with  $m_1 + m_2 = m_3 + m_4$ , there are three possible types of optimal transport path in  $\text{Path}(a, b)$ , depending on the positions of  $x_i$  as well as the ratios of  $m_i$ :

Three types from two points to two points



An arbitrary transport path  $G \in \text{Path}(a, b)$  is a weighted directed graph, but not necessarily a directed tree. In other words,  $G$  may contain some cycles. However, the following proposition says  $G$  can be modified to be a directed tree  $\tilde{G} \in \text{Path}(a, b)$  (i.e. contains no cycles) with less  $M^\alpha$  cost.

**Proposition 2.1.** For any  $G \in \text{Path}(a, b)$ , there exists a  $\tilde{G} \in \text{Path}(a, b)$  contains no cycles and  $M^\alpha(\tilde{G}) \leq M^\alpha(G)$ .

**Proof.** Suppose  $G$  contains some cycle  $L$ . For each edge  $e$  of  $L$ , define

$$m(e) = \frac{\alpha \text{length}(e)}{w(e)^{1-\alpha}}.$$

Arbitrarily pick an orientation for  $L$  and let

$$L_1 = \sum \{[e] : \text{edge } e \text{ of } G \text{ has the same orientation as } L\} \text{ and}$$

$$L_2 = \sum \{[e] : \text{edge } e \text{ of } G \text{ has the reverse orientation as } L\}.$$

Note that  $L_1$  or  $L_2$  is possibly empty and  $\text{div } L_1 = \text{div } L_2$ . By changing orientation on  $L$  if necessary, we may assume that

$$\sum_{e \in L_1} m(e) \leq \sum_{e \in L_2} m(e).$$

Now we set

$$G' = G + w(L_1 - L_2)$$

with  $w = \min\{w(e) : e \in L_2\}$ . Then  $G' \in \text{Path}(a, b)$  and has fewer cycles than  $G$  does. Moreover,  $M^\alpha(G') \leq M^\alpha(G)$ . To see this, we consider the function on  $[0, w]$  defined by

$$\begin{aligned} f(\lambda) &:= M^\alpha(G + \lambda(L_1 - L_2)) - M^\alpha(G) \\ &= \sum_{e \in L_1} \text{length}(e)[(w(e) + \lambda)^\alpha - w(e)^\alpha] \\ &\quad + \sum_{e \in L_2} \text{length}(e)[(w(e) - \lambda)^\alpha - w(e)^\alpha]. \end{aligned}$$

Then, since  $\alpha \leq 1$ , trivial calculations imply that

$$f''(\lambda) \leq 0, \quad f'(\lambda) \leq f'(0) = \sum_{e \in L_1} m(e) - \sum_{e \in L_2} m(e) \leq 0$$

and  $f(\lambda) \leq f(0) = 0$ . Therefore,  $M^\alpha(G') \leq M^\alpha(G)$ . Repeating the above procedure, we get the desired transport path  $\tilde{G}$  with no cycles.  $\square$

From the above proposition, we may restrict our transport paths to be the class of directed trees. For directed trees, we have the following trivial but important lemma.

**Lemma 2.1.** *Suppose  $0 < \Lambda < +\infty$  and  $G = \sum_{e \in E(G)} w(e)[[e]] \in \text{Path}(a, b)$  is a directed tree with  $a, b \in \mathcal{A}_\Lambda$  as before. Then for any edge  $e \in E(G)$ , we have*

$$0 < w(e) \leq \Lambda.$$

Thus,

$$\frac{M^\alpha(G)}{\Lambda^\alpha} \geq \frac{M^1(G)}{\Lambda}.$$

**Proof.**

$$\begin{aligned} \frac{M^\alpha(G)}{\Lambda^\alpha} &= \sum_{e \in E(G)} \left(\frac{w(e)}{\Lambda}\right)^\alpha \text{length}(e) \\ &\geq \sum_{e \in E(G)} \frac{w(e)}{\Lambda} \text{length}(e) \\ &= \frac{M^1(G)}{\Lambda}. \end{aligned}$$

$\square$

From the following elementary proposition we know that in order to transport general atomic measures, it is sufficient to transport probability atomic measures.

**Proposition 2.2.** *For any  $G = \sum_{e \in E(G)} w(e)[[e]] \in \text{Path}(a, b)$  and any positive number  $r > 0$ ,*

$$rG := \sum_{e \in E(G)} (rw(e))[[e]]$$

*is a transport path from  $ra$  to  $rb \in \mathcal{A}_{r\Lambda}(X)$ , and*

$$M^\alpha(rG) = r^\alpha M^\alpha(G).$$

*In particular,*

$$\frac{G}{\Lambda} \in \text{Path}\left(\frac{a}{\Lambda}, \frac{b}{\Lambda}\right) \text{ with } M^\alpha(G) = \Lambda^\alpha M^\alpha\left(\frac{G}{\Lambda}\right).$$

### 3. Transport General Measures

In this section, we will consider transport paths for general Radon measures in  $\mathcal{M}_\Lambda(X)$ . The idea here is to approximate those general measures by atomic measures in  $\mathcal{A}_\Lambda(X)$ , and then using the transport paths of those approximating atomic measures to approximate a transport path of general measures in  $\mathcal{M}_\Lambda(X)$ .

Recall that any Radon measure  $\mu$  on  $X$  can be approximated by atomic measures in the weak \* topology of Radon measures as we see in the following simple example.

**Example 3.1.** For any Radon measure  $\mu$  in  $\mathcal{M}_\Lambda(X)$ , we may approximate  $\mu$  by atomic measures in  $\mathcal{A}_\Lambda(X)$  as follows. For each  $i = 0, 1, 2, \dots$ , let

$$Q_i = \{Q_i^h : h \in \mathbb{Z}^m \cap [0, 2^i)^m\}$$

be a partition of  $Q$  into cubes of edge length  $\frac{d}{2^i}$ . Now, for each  $h \in \mathbb{Z}^m \cap [0, 2^i)^m$ , let  $c_i^h$  be the center of  $Q_i^h$  and

$$m_i^h = \mu(Q_i^h)$$

be the  $\mu$  mass of the cube  $Q_i^h$ . Define an atomic measure

$$A_i(\mu) = \sum_{h \in \mathbb{Z}^m \cap [0, 2^i)^m} m_i^h \delta_{c_i^h} \in \mathcal{A}_\Lambda(X).$$

Then  $\{A_i(\mu)\}$  converges to  $\mu$  weakly. This approximation  $\{A_i(\mu)\}$  is called the **dyadic approximation of  $\mu$** .

Now, we can talk about transport paths between general measures.

**Definition 3.1.** Let  $\mu^+, \mu^- \in \mathcal{M}_\Lambda(X)$  be any two Radon measures on  $X$  with equal total mass  $\Lambda$ . Extending Definition 2.1 from Sec. 1, we say a vector measure  $T \in \mathcal{M}^m(X)$  is a **transport path from  $\mu^+$  to  $\mu^-$**  if there exist two sequences  $\{a_i\}, \{b_i\}$  of atomic measures in  $\mathcal{A}_\Lambda(X)$  with a corresponding sequence of transport



paths  $G_i \in \text{Path}(a_i, b_i)$  such that

$$a_i \rightharpoonup \mu^+, b_i \rightharpoonup \mu^-, G_i \rightharpoonup T$$

weakly as Radon measures and vector measures.

The sequence of triples  $\{a_i, b_i, G_i\}$  is called **an approximating graph sequence for  $T$** . Note that for any such  $T$ ,

$$\text{div}(T) = \mu^+ - \mu^-$$

in the sense of distributions.

Let

$$\text{Path}(\mu^+, \mu^-) \subset \mathcal{M}^m(X)$$

be the space of all transport paths from  $\mu^+$  to  $\mu^-$ .

Also, given any  $\alpha \in [0, 1]$ , for any  $T \in \text{Path}(\mu^+, \mu^-)$ , we define its  $M^\alpha$  cost to be

$$M^\alpha(T) := \inf \liminf_{i \rightarrow \infty} M^\alpha(G_i),$$

where the infimum is over the set of all possible approximating graph sequence  $\{a_i, b_i, G_i\}$  of  $T$ .

**Remark 3.1.** One can easily see that  $M^\alpha$  is lower semicontinuous. Also, when  $\mu^+$  and  $\mu^-$  are atomic, the above definition about transport path and  $M^\alpha$  cost function agrees with the previous definitions in the last section.

**Remark 3.2.** In the proof of Theorem 5.1, we will see how a transport path  $T \in \text{Path}(\mu^+, \mu^-)$  with finite  $M^\alpha$  cost be represented by a Lipschitz curve in  $\mathcal{M}_1(X)$  from  $\mu^+$  to  $\mu^-$ .

For simplicity, we may assume both  $\mu^+$  and  $\mu^-$  to be probability measures, i.e.  $\mu^+(X) = \mu^-(X) = 1$ .

For any probability measure  $\mu \in \mathcal{M}_1(X)$ , we will now use the dyadic approximation  $A_i(\mu)$  of  $\mu$  to construct a transport path of finite  $M^\alpha$  cost from  $\mu$  to the Dirac measure  $\delta_c$ , where  $c$  is the center of the cube  $Q$  containing  $X$  with edge length  $d$ .

**Proposition 3.1.** *For any  $\mu \in \mathcal{M}_1(X)$ , there exists a transport path  $T \in \text{Path}(\mu, \delta_c)$  such that*

$$M^\alpha(T) \leq \frac{1}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2}$$

for any  $\alpha \in (1 - \frac{1}{m}, 1]$ .

**Proof.** Let  $\{A_i(\mu)\}$  be the dyadic approximation of  $\mu$  as in the Example 3.1.

For each  $i = 0, 1, \dots$ , and  $h \in Z^m \cap [0, 2^i)^m$ , each cube  $Q_i^h$  of level  $i$  corresponds to  $2^m$  cubes  $\{Q_{i+1}^{2^m h+h'} : h' = 0, 1, 2, \dots, 2^m - 1\}$  of level  $i + 1$  by dyadic subdivision. Then one can construct a transport path

$$G_i^h := \sum_{h'=0}^{2^m-1} m_{i+1}^{2^m h+h'} [(c_{i+1}^{2^m h+h'}, c_i^h)] \in \text{Path} \left( m_i^h \delta_{c_i^h}, \sum_{h'=0}^{2^m-1} m_{i+1}^{2^m h+h'} \delta_{c_{i+1}^{2^m h+h'}} \right),$$

which is a directed tree from the center  $c_i^h$  of  $Q_i^h$  to the centers  $c_{i+1}^{2^m h+h'}$  of  $2^m$  cubes  $\{Q_{i+1}^{2^m h+h'} : h' = 0, 1, 2, \dots, 2^m - 1\}$  with suitable weights. Now, for each  $n \geq 0$ , set

$$G_n = \sum_{i=0}^n \sum_{h \in Z^m \cap [0, 2^i)^m} G_i^h \in \text{Path} (A_0(\mu), A_{n+1}(\mu)).$$

Since

$$\begin{aligned} M^\alpha(G_n) &= \sum_{i=0}^n \sum_{h \in Z^m \cap [0, 2^i)^m} \sum_{h'=0}^{2^m-1} (m_{i+1}^{2^m h+h'})^\alpha \text{length} (c_{i+1}^{2^m h+h'}, c_i^h) \\ &= \sum_{i=0}^n \sum_{h \in Z^m \cap [0, 2^i)^m} \sum_{h'=0}^{2^m-1} (m_{i+1}^{2^m h+h'})^\alpha \frac{\sqrt{md}}{2^{i+2}} \\ &\leq \sum_{i=0}^n \sum_{h \in Z^m \cap [0, 2^i)^m} \sum_{h'=0}^{2^m-1} \left( \frac{1}{2^{m(i+1)}} \right)^\alpha \frac{\sqrt{md}}{2^{i+2}} \\ &= \sum_{i=0}^n (2^{i+1})^{m(1-\alpha)-1} \frac{\sqrt{md}}{2} \\ &\leq \frac{1}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2}, \text{ if } \alpha > 1 - \frac{1}{m}, \end{aligned}$$

where (\*) follows from the fact that the function

$$f_i(x_1, x_2, \dots, x_{2^m(i+1)}) = \sum_{h \in Z^m \cap [0, 2^i)^m} \sum_{h'=0}^{2^m-1} (x_{2^m h+h'})^\alpha \frac{\sqrt{md}}{2^{i+2}}$$

achieves its maximum at the point  $(\frac{1}{2^{m(i+1)}}, \frac{1}{2^{m(i+1)}}, \dots, \frac{1}{2^{m(i+1)}})$ .

Since  $\|G_n\|(X) = M^1(G_n) \leq M^\alpha(G_n)$  has bounded total variation, by the compactness of vector measures,  $\{G_n\}$  subconverges weakly to a vector measure  $T$  with

$$M^\alpha(T) \leq \liminf M^\alpha(G_i) \leq \frac{1}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2}.$$

Thus  $T \in \text{Path} (\mu, \delta_c)$  has finite  $M^\alpha$  cost. □

Now, we solve the Problem 1.1 in the following existence theorem:

**Theorem 3.1 (Existence theorem).** *Given two Radon measures  $\mu^+, \mu^- \in \mathcal{M}_\Lambda(X)$  on  $X \subset \mathbb{R}^m$  and  $\alpha \in (1 - \frac{1}{m}, 1]$ , there exists an optimal transport path  $S$  with least  $M^\alpha$  cost among all transport paths in the family  $\text{Path}(\mu^+, \mu^-)$ . Moreover*

$$M^\alpha(S) \leq \frac{\Lambda^\alpha}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2}.$$

**Proof.** Let  $\{T_i\}$  be an  $M^\alpha$  minimizing sequence in  $\text{Path}(\mu^+, \mu^-)$ . For each  $T_i$ , there exists a transport path  $G_i \in \text{Path}(a_i, b_i)$  such that

$$M^\alpha(G_i) \leq M^\alpha(T_i) + \frac{1}{2^i} \text{ and } W(a_i, \mu^+) + W(b_i, \mu^-) < \frac{1}{2^i},$$

where  $W$  is the Wasserstein 1-distance on  $\mathcal{M}_1(X)$  which metrizes the weak\* topology of  $\mathcal{M}_1(X)$ . By Proposition 2.1, we may assume  $G_i$  has no cycles. Thus the total variation of  $G_i$

$$\|G_i\|(X) = M^1(G_i) \leq \Lambda^{1-\alpha} M^\alpha(G_i) \leq \left( M^\alpha(T_i) + \frac{1}{2^i} \right) \Lambda^{1-\alpha}$$

is uniformly bounded. Therefore, by the compactness of vector measure,  $\{G_i\}$  is subsequently convergent to a vector measure  $S \in \text{Path}(\mu^+, \mu^-)$ . By the lower semicontinuity of  $M^\alpha$ ,  $S$  is an optimal transport path. Also, by the Proposition 3.1, we know  $M^\alpha(S) \leq \frac{\Lambda^\alpha}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2}$ . □

#### 4. A New Distance $d_\alpha$ on the Space of Probability Measures

By the existence theorem (Theorem 3.1), we give the following definition:

**Definition 4.1.** For any  $\alpha \in (1 - \frac{1}{m}, 1]$ , we define

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

for any two Radon measures  $\mu^+, \mu^- \in \mathcal{M}_\Lambda(X)$ .

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

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

Thus, we may assume  $\Lambda = 1$ .

##### 4.1. $d_\alpha$ is a distance

To show that  $d_\alpha$  is in fact a distance on  $\mathcal{M}_1(X)$ , we need the following lemma:

**Lemma 4.1.** *Given  $\mu \in \mathcal{M}_1(X)$ . Suppose  $\{a_i\}, \{b_i\} \subset \mathcal{A}_1(X)$  are two sequences of atomic probability measures on  $X$ . If  $a_i \rightarrow \mu$  and  $b_i \rightarrow \mu$ , then  $d_\alpha(a_i, b_i) \rightarrow 0$ .*

**Proof.** Given  $\epsilon > 0$ , since  $m(1 - \alpha) - 1 < 0$ , there exists a natural number  $n$  large enough so that

$$n^{m(1-\alpha)-1} \frac{1}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2} < \frac{\epsilon}{3}.$$

For any small number  $\beta > 0$ , we can find a partition  $\tilde{Q}_n = \{Q_n^h : h \in Z^m \cap [0, n]^m\}$  of  $Q$  consist of cubes of edge length between  $[(1 - \beta)\frac{d}{n}, (1 + \beta)\frac{d}{n}]$  such that for all  $i$ , the finite set  $\text{spt}(a_i) \cup \text{spt}(b_i)$  doesn't intersect the boundary of those cubes. For each  $h$ , let  $c_n^h$  be the center of  $Q_n^h$ ,  $p_i^h = a_i(Q_n^h)$  and  $q_i^h = b_i(Q_n^h)$ . Since  $a_i - b_i \rightarrow 0$ , we have  $p_i^h - q_i^h = (a_i - b_i)(\chi(\text{interior of } Q_n^h)) \rightarrow 0$  as  $i \rightarrow \infty$  for all  $h$ . Let

$$p_i = \sum_{h \in Z^m \cap [0, n]^m} p_i^h \delta_{c_n^h} \text{ and } q_i = \sum_{h \in Z^m \cap [0, n]^m} q_i^h \delta_{c_n^h}.$$

By Proposition 3.1, there exists an  $S_i^h \in \text{Path}(a_i|_{Q_n^h}, p_i^h \delta_{c_n^h})$  with

$$M^\alpha(S_i^h) \leq \frac{(p_i^h)^\alpha}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2n}.$$

Thus  $S_i = \sum_{h \in Z^m \cap [0, n]^m} S_i^h \in \text{Path}(a_i, p_i)$  and

$$\begin{aligned} M^\alpha(S_i) &\leq \sum_{h \in Z^m \cap [0, n]^m} M^\alpha(S_i^h) \\ &\leq \sum_{h \in Z^m \cap [0, n]^m} (p_i^h)^\alpha \frac{1}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2n} \\ &\leq \sum_{h \in Z^m \cap [0, n]^m} \left(\frac{1}{n^m}\right)^\alpha \frac{1}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{7n} \\ &\leq n^m \left(\frac{1}{n^m}\right)^\alpha \frac{1}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{md}}{2n} \\ &< \frac{\epsilon}{3}. \end{aligned}$$

Similarly, we may find some  $S'_i \in \text{Path}(b_i, q_i)$  with  $M^\alpha(S'_i) < \frac{\epsilon}{3}$ .

Finally, let  $G_i$  be the cone over  $p_i - q_i$  with vertex  $c$ , the center of  $Q$ . Then  $G_i \in \text{Path}(p_i, q_i)$  and

$$M^\alpha(G_i) \leq \sum_{h \in Z^m \cap [0, n]^m} (|p_i^h - q_i^h|)^\alpha \frac{\sqrt{md}}{2} < \frac{\epsilon}{3}$$

when  $i$  large enough. Therefore, we have  $T_i = S_i + G_i + S'_i \in \text{Path}(a_i, b_i)$  with  $M^\alpha(T_i) < \epsilon$  when  $i$  large enough. Thus,  $d_\alpha(a_i, b_i) \rightarrow 0$ . □

**Lemma 4.2.** For any  $\mu^+, \mu^- \in \mathcal{M}_1(X)$ , we have

$$W(\mu^+, \mu^-) \leq d_\alpha(\mu^+, \mu^-),$$

where  $W$  is the Wasserstein distance on  $\mathcal{M}_1(X)$ .

**Proof.** Let  $\{a_i, b_i, G_i\}$  be an approximating graph sequence of an optimal transport path in  $\text{Path}(\mu^+, \mu^-)$  such that

$$\lim M^\alpha(G_i) = d_\alpha(\mu^+, \mu^-).$$

For each  $i$ , one may assume that  $G_i$  contains no cycles, thus

$$W(a_i, b_i) \leq M^1(G_i) \leq M^\alpha(G_i).$$

Let  $i \rightarrow \infty$ , we have

$$W(\mu^+, \mu^-) \leq \liminf W(a_i, b_i) \leq \lim M^\alpha(G_i) = d_\alpha(\mu^+, \mu^-). \quad \square$$

**Corollary 4.1.** *If  $d_\alpha(\mu_i, \mu) \rightarrow 0$ , then  $\mu_i \rightarrow \mu$ .*

**Theorem 4.1.**  $d_\alpha$  is a distance on  $\mathcal{M}_1(X)$ .

**Proof.** From Lemma 4.2, we know that  $\mu_1 = \mu_2$  whenever  $d_\alpha(\mu_1, \mu_2) = 0$ . Thus, it's sufficient to show that

$$d_\alpha(\mu_1, \mu_3) \leq d_\alpha(\mu_1, \mu_2) + d_\alpha(\mu_2, \mu_3)$$

for any  $\mu_1, \mu_2, \mu_3 \in \mathcal{M}_1(X)$ .

In fact, given  $\epsilon > 0$ , there exist sequences of transport paths  $G_i \in \text{Path}(a_i, b_i), P_i \in \text{Path}(c_i, d_i)$  such that

$$\lim M^\alpha(G_i) \leq d_\alpha(\mu_1, \mu_2) + \frac{\epsilon}{3} \text{ and } \lim M^\alpha(P_i) \leq d_\alpha(\mu_2, \mu_3) + \frac{\epsilon}{3}$$

for some atomic approximations  $\{a_i\}, \{b_i\}, \{c_i\}, \{d_i\}$  of  $\mu_1, \mu_2, \mu_2, \mu_3$  respectively:

$$a_i \rightarrow \mu_1, b_i \rightarrow \mu_2,$$

$$c_i \rightarrow \mu_2, d_i \rightarrow \mu_3,$$

as  $i \rightarrow \infty$ . As in the proof of Lemma 4.1, we find a  $T_i \in \text{Path}(b_i, c_i)$  with  $\lim M^\alpha(T_i) < \epsilon/3$ . Thus,  $G_i + T_i + P_i \in \text{Path}(a_i, d_i)$  and

$$\begin{aligned} d_\alpha(\mu_1, \mu_3) &\leq \liminf M^\alpha(G_i + T_i + P_i) \\ &\leq d_\alpha(\mu_1, \mu_2) + d_\alpha(\mu_2, \mu_3) + \epsilon. \end{aligned}$$

Therefore,  $d_\alpha(\mu_1, \mu_3) \leq d_\alpha(\mu_1, \mu_2) + d_\alpha(\mu_2, \mu_3)$ . □

**Corollary 4.2.** *Suppose  $\{a_i, b_i, G_i\}$  is any approximating graph sequence of some transport path  $T \in \text{Path}(\mu^+, \mu^-)$ . If  $G_i$  is optimal in  $\text{Path}(a_i, b_i)$  for each  $i$ , then  $T$  is also optimal.*

**Proof.** Suppose  $S$  is an optimal transport path in  $\text{Path}(\mu^+, \mu^-)$ , and  $\{a'_i, b'_i, F_i\}$  is an approximating graph sequence of  $S$  such that

$$\lim_{i \rightarrow \infty} M^\alpha(F_i) = M^\alpha(S) = d_\alpha(\mu^+, \mu^-).$$

Then, by Lemma 4.1,

$$\begin{aligned}
 M^\alpha(T) &\leq \liminf_{i \rightarrow \infty} M^\alpha(G_i) \\
 &= \liminf_{i \rightarrow \infty} d_\alpha(a_i, b_i) \\
 &\leq \liminf_{i \rightarrow \infty} d_\alpha(a'_i, b'_i) + d_\alpha(a_i, a'_i) + d_\alpha(b_i, b'_i) \\
 &\leq \liminf_{i \rightarrow \infty} M^\alpha(F_i) = d_\alpha(\mu^+, \mu^-).
 \end{aligned}$$

Therefore,  $M^\alpha(T) = d_\alpha(\mu^+, \mu^-)$  and  $T$  is also optimal. □

**4.2. Topology on  $\mathcal{M}_1(X)$  induced by  $d_\alpha$**

In general, the  $d_\alpha$  distance is different from any of the Wasserstein distances because the optimal transport path for  $d_\alpha$  will be “Y shaped” rather than “V shaped” as in Wasserstein distances. However, we’ll show that they induce the same topology on  $\mathcal{M}_1(X)$ , namely the weak \* topology of  $\mathcal{M}_1(X)$ . We first show that atomic measures are dense in  $(\mathcal{M}_1(X), d_\alpha)$ :

**Lemma 4.3.** *For each  $\mu \in \mathcal{M}_1(X)$ , let  $\{A_n(\mu)\}$  be the dyadic approximation of  $\mu$ . Then,*

$$d_\alpha(\mu, A_n(\mu)) \leq C\beta^n$$

for some constant  $C = \frac{\sqrt{m}d/2}{2^{1-m(1-\alpha)} - 1}$  and  $0 < \beta = 2^{m(1-\alpha)-1} < 1$ . This implies atomic probability measures are dense in  $(\mathcal{M}_1(X), d_\alpha)$ .

**Proof.** As before, for any  $n \geq 1$ , we let  $\mathcal{Q}_n = \{Q_n^h : h \in Z^m \cap [0, 2^n)^m\}$  be a partition of  $[0, d]^m$  into cubes of edge length  $\frac{d}{2^n}$ , and let  $c_n^h$  be the center of  $Q_n^h$ . Then  $A_n(\mu) = \sum_h \mu(Q_n^h) \delta_{c_n^h}$  and

$$\begin{aligned}
 d_\alpha(\mu, A_n(\mu)) &\leq \sum_h d_\alpha(\mu \lfloor Q_n^h, \delta_{c_n^h}) \\
 &\leq \sum_h \frac{\mu(Q_n^h)^\alpha}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{m}d}{2^{n+1}} \\
 &\leq \sum_h \left(\frac{1}{2^{nm}}\right)^\alpha \frac{1}{2^{1-m(1-\alpha)} - 1} \frac{\sqrt{m}d}{2^{n+1}} \\
 &= \frac{\sqrt{m}d}{2^{1-m(1-\alpha)} - 1} 2^{n[m(1-\alpha)-1]-1} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad \square
 \end{aligned}$$

**Corollary 4.3.** *For any  $\mu^+, \mu^- \in \mathcal{M}_1(X)$ , let  $\{A_n(\mu^+)\}$  and  $\{A_n(\mu^-)\}$  be the dyadic approximation of  $\mu^+$  and  $\mu^-$  respectively, then*

$$|d_\alpha(\mu^+, \mu^-) - d_\alpha(A_n(\mu^+), A_n(\mu^-))| \leq 2C\beta^n.$$

**Theorem 4.2.**  $d_\alpha$  metrizes the weak\* topology of  $\mathcal{M}_1(X)$ .

**Proof.** By Lemma 4.2, it's sufficient to show that if  $\mu_i \rightharpoonup \mu$ , then  $d_\alpha(\mu_i, \mu) \rightarrow 0$ .

In fact, for each  $i$ , by Lemma 4.3, we can find an atomic probability measure  $a_i$  such that  $d_\alpha(a_i, \mu_i) \leq \frac{1}{2^i}$ . So,  $a_i \rightharpoonup \mu$  also. On the other hand, by Lemma 4.3 again, we may find a sequence of atomic probability measures  $b_i \rightharpoonup \mu$  and  $d_\alpha(b_i, \mu) \rightarrow 0$ . By Lemma 4.1,  $d_\alpha(a_i, b_i) \rightarrow 0$  and thus

$$d_\alpha(\mu_i, \mu) \leq d_\alpha(\mu_i, a_i) + d_\alpha(a_i, b_i) + d_\alpha(b_i, \mu) \rightarrow 0. \quad \square$$

### 5. Length Space Property

In this section, we will show that  $(\mathcal{M}_1(X), d_\alpha)$  is a length space, which means that for any  $\mu^+, \mu^- \in \mathcal{M}_1(X)$ , there exists a continuous map

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

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

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

We first construct such a map for atomic measures. Then, by means of such maps, we achieve a desired map for arbitrary measures.

First we need a basic lemma which will allow us to assign an appropriate “time” parameter to each vertex.

**Lemma 5.1.** Let  $G$  be any connected tree in  $X$  with vertex set

$$V(G) = \{P_1, P_2, \dots, P_{i(G)}; P_{i(G)+1}, \dots, P_{i(G)+b(G)}\}.$$

where  $\{P_1, P_2, \dots, P_{i(G)}\}$  is the set of interior vertices of  $G$  and

$$\{P_{i(G)+1}, \dots, P_{i(G)+b(G)}\}$$

is the set of boundary vertices of  $G$ . For each edge  $P_i P_j \in E(G)$ , let  $a_{ij} > 0$  be a given positive number associated to the edge  $P_i P_j$ .

Now, for any  $(c_1, c_2, \dots, c_{b(G)}) \in \mathbb{R}^{b(G)}$ , there exists a unique map

$$t: V(G) \rightarrow \mathbb{R} \tag{5.1}$$

such that for each  $j = 1, 2, \dots, b(G)$ ,

$$t(P_{i(G)+j}) = c_j \tag{5.2}$$

and for each  $i = 1, 2, \dots, i(G)$ ,

$$\sum_j \frac{a_{ij}}{t(P_j) - t(P_i)} = 0, \tag{5.3}$$

where the summation is over all  $j$  such that  $P_j$  and  $P_i$  are connected by an edge of  $G$ .

**Proof.** If  $i(G) = 1$ , then

$$V(G) = \{P_1; P_2, \dots, P_{b(G)+1}\}.$$

Since the function

$$\sum_{i=2}^{b(G)} \frac{a_{1i}}{t - c_{i-1}}$$

is a strictly increasing function with range  $(-\infty, +\infty)$ . Thus it has a unique zero  $t_1$ . Note that

$$\min\{c_1, \dots, c_{b(G)}\} \leq t_1 \leq \max\{c_1, \dots, c_{b(G)}\}$$

and when view  $t_1$  as a function of  $c_1, \dots, c_{b(G)}$ , we have

$$\frac{\partial t_1}{\partial c_{j-1}} = \frac{\frac{a_{1j}}{(t - c_{j-1})^2}}{\sum_{i=2}^{b(G)} \frac{a_{1i}}{(t - c_{i-1})^2}} \in [0, 1]$$

for each  $j = 2, \dots, b(G) + 1$ .

Suppose for each  $\bar{G}$  with  $i(\bar{G}) \leq N$ , there exists a  $\bar{t}$  satisfying (5.1)–(5.3). Then, for each interior vertex  $P$ ,

$$\min\{c_1, c_2, \dots, c_{b(\bar{G})}\} \leq \bar{t}(P) \leq \max\{c_1, c_2, \dots, c_{b(\bar{G})}\}.$$

Moreover, if one of  $P$ 's adjacent vertices is a boundary vertex  $P_{i(\bar{G})+j}$  for some  $j$ , then

$$\frac{\partial(\bar{t}(P))}{\partial c_j} \in [0, 1].$$

Now assume  $i(G) = N + 1$ . Note that there exists an interior vertex of  $G$  which has exactly one other interior vertex of  $G$  as its adjacent vertex. All other adjacent vertices are boundary vertices of  $G$ . In fact, by removing all edges with one of its endpoints being a boundary vertex of  $G$ , we get a new connected tree. Then, any boundary vertex of this new connected tree will be an interior vertex of the original  $G$  with the desired properties. Without losing generality, we may assume that the vertex is  $P_1$  and its adjacent vertices are  $P_{i(G)+1}, \dots, P_{i(G)+h}$  and  $P_2$ .

Now, removing these edges  $P_1P_{i(G)+1}, \dots, P_1P_{i(G)+h}$  from  $G$ , we get another connected tree  $\bar{G}$  with

$$V(\bar{G}) = \{P_2, \dots, P_{i(G)}; P_1, P_{i(G)+h+1}, \dots, P_{i(G)+b(G)}\}.$$

Since  $i(\bar{G}) = i(G) - 1$ , we may apply induction to  $\bar{G}$ . Thus, for any  $\lambda \in \mathbb{R}$ , there exists a solution  $\bar{t}_\lambda : V(\bar{G}) \rightarrow \mathbb{R}$  to the system (5.3) with

$$\bar{t}_\lambda(P_1) = \lambda \text{ and } \bar{t}_\lambda(P_{i(G)+j}) = c_j$$

for each  $j = h + 1, \dots, b(G)$ . Let

$$g(\lambda) = \bar{t}_\lambda(P_2).$$

By induction,

$$\min\{\lambda, c_{h+1}, c_{h+2}, \dots, c_{b(G)}\} \leq g(\lambda) \leq \max\{\lambda, c_{h+1}, c_{h+2}, \dots, c_{b(G)}\} \tag{5.4}$$

and

$$0 \leq g'(\lambda) \leq 1.$$



Now, let

$$f(\lambda) = \sum_{j=1}^h \frac{a_{1,i(G)+j}}{c_j - \lambda} + \frac{a_{12}}{g(\lambda) - \lambda}.$$

By (5.4), when  $\lambda$  is large enough, then  $g(\lambda) \leq \lambda$  and thus

$$f(\lambda) < 0$$

as  $\lambda \rightarrow +\infty$ .

Similarly, as  $\lambda \rightarrow -\infty$ ,  $g(\lambda) \geq \lambda$  and  $f(\lambda) > 0$ . Moreover,

$$f'(\lambda) = \sum_{j=1}^h \frac{a_{1,i(G)+j}}{(c_j - \lambda)^2} + \frac{a_{12}}{(g(\lambda) - \lambda)^2} (1 - g'(\lambda)) > 0,$$

thus  $f$  is strictly increasing and has a unique zero  $t_1$ . Note that, by (5.4),

$$t_1 \leq \max\{c_1, \dots, c_h, g(t_1)\} \leq \max\{c_1, \dots, c_{b(G)}\}.$$

Similarly,

$$t_1 \geq \min\{c_1, \dots, c_h, g(t_1)\} \leq \min\{c_1, \dots, c_{b(G)}\}.$$

Furthermore,

$$\frac{\partial t_1}{\partial c_j} = \frac{\frac{a_{1,i(G)+1}}{(c_1 - t_1)^2}}{\sum_{j=1}^h \frac{a_{1,i(G)+j}}{(c_j - t_1)^2} + \frac{a_{12}}{(g(t_1) - t_1)^2} (1 - g'(t_1))} \in [0, 1]$$

for each  $j = 1, \dots, h$ . Hence, when setting

$$t(P_i) = \begin{cases} \bar{t}_{t_1}(P_i), & \text{if } P_i \in V(\bar{G}), \\ t_1, & i = 1, \\ c_j, & i = i(G) + j, j = 1, \dots, h \end{cases}$$

we get a solution  $t$  to the system (5.3) for  $G$ . □

Next we consider two atomic measures

$$a = \sum_{i=1}^k a_i \delta_{x_i} \text{ and } b = \sum_{j=1}^l b_j \delta_{y_j} \in \mathcal{A}_\Lambda(X)$$

and a positive number  $H$ .

Now, for any connected tree  $G \in \text{Path}(a, b)$  which contains no cycles, there exists, by Lemma 5.1, a map

$$t: V(G) \rightarrow [0, H]$$

satisfying (5.3) such that

$$t(v) = \begin{cases} 0, & \text{if } v = x_i \text{ for some } i = 1, 2, \dots, k \\ H, & \text{if } v = y_j \text{ for some } j = 1, \dots, l \end{cases}$$

where

$$a_{ij} = w(e)^\alpha \text{ length}(e) > 0$$

for each edge  $e = P_i P_j$ . By linearly extension on every edge of  $G$ , we may extend the map  $t$  to a map on the whole graph  $G$ . Still denote the extended map by  $t$ . If  $G \setminus \{x_1, \dots, x_k, y_1, \dots, y_l\}$  is not connected, one may work with its connected components first, and then combine them to get a common map  $t$ .

Note, for any  $s \in [0, H]$ ,  $t^{-1}(s)$  is a finite subset of  $G$ . We associate a positive multiplicity  $m(a)$  for each  $a \in t^{-1}(s)$ . If  $a$  is an interior point of an edge  $e \in E(G)$ , then  $m(a) = w(e)$ . If  $a$  is a vertex of  $G$ , then

$$m(a) = \sum_{e^+=a} w(e) = \sum_{e^-=a} w(e).$$

Thus, we have a map

$$s \longmapsto \sum_{a \in t^{-1}(s)} m(a) \delta_a. \tag{5.5}$$

Since  $M^1(\sum_{a \in t^{-1}(s)} m(a) \delta_a)$  is a continuous function on  $s \in [0, H]$ , and locally constant at least outside the range of the vertices of  $t$ , we know it must be a constant, namely  $M^1(\mu^+) = \Lambda$ . Thus,  $\sum_{a \in t^{-1}(s)} m(a) \delta_a \in \mathcal{A}_\Lambda(X)$ .

**Lemma 5.2.** *For any  $G \in \text{Path}(a, b)$  as above and any  $0 \leq s_1 < s_2 \leq H$ ,*

$$M^\alpha(G \cap t^{-1}[s_1, s_2]) = \frac{s_1 - s_2}{H} M^\alpha(G).$$

**Proof.** Consider the function

$$f(s) = M^\alpha(G \cap t^{-1}[0, s])$$

on  $[0, H]$ .

Given any  $s \in [0, H]$ , since  $G$  has only finite many vertices, one may choose  $h > 0$  small enough so that  $t^{-1}((s, s + h))$  contains no vertices of  $G$ . Then, since  $t$  is linear and increasing on each edge of the graph  $G$ , we have

$$t^{-1}(s+h) = \left\{ a + h \frac{e^+ - e^-}{t(e^+) - t(e^-)} \in X : a \in t^{-1}(s), a \in [e^-, e^+) \text{ for some } e \in E(G) \right\},$$

where  $[e^-, e^+)$  denotes the line segment  $e$  including the endpoint  $e^-$ , but not the endpoint  $e^+$ . Thus,

$$G \cap t^{-1}([s, s+h]) = \cup \left\{ \left[ a, a + h \frac{e^+ - e^-}{t(e^+) - t(e^-)} \right] : a \in t^{-1}(s), a \in [e^-, e^+), e \in E(G) \right\},$$

and

$$M^\alpha(G \cap t^{-1}[s, s+h]) = \sum_{\substack{a \in t^{-1}(s) \\ a \in [e^-, e^+)}} h \frac{\text{length}(e)}{t(e^+) - t(e^-)} w(e)^\alpha.$$

Thus,

$$f'(s + 0) = \sum_{\substack{s \in [t(e^-), t(e^+)] \\ e \in E(G)}} \frac{\text{length}(e)}{t(e^+) - t(e^-)} w(e)^\alpha.$$

Similarly, for any  $s \in (0, H]$ , we have

$$f'(s - 0) = \sum_{\substack{s \in (t(e^-), t(e^+)] \\ e \in E(G)}} \frac{\text{length}(e)}{t(e^+) - t(e^-)} w(e)^\alpha.$$

Thus, if  $t^{-1}(s)$  contains no vertices of  $G$ , then  $f'(s)$  exists and its value is locally the constant

$$f'(s) = \sum_{\substack{s \in (t(e^-), t(e^+)) \\ e \in E(G)}} \frac{\text{length}(e)}{t(e^+) - t(e^-)} w(e)^\alpha.$$

On the other hand, when  $t^{-1}(s)$  contains some vertices of  $G$ , by the key Lemma 5.1,  $f'(s - 0) = f'(s + 0)$  so that  $f'(s)$  also exists. Moreover,  $f'$  becomes globally a constant. Therefore,  $f$  must be affine. Since

$$f(0) = 0 \text{ and } f(H) = M^\alpha(G),$$

we have

$$f(s) = \frac{M^\alpha(G)}{H} s. \quad \square$$

**Theorem 5.1.**  $(\mathcal{M}_1(X), d_\alpha)$  is a length space.

**Proof.** It is sufficient to show that for any  $\mu^+, \mu^- \in \mathcal{M}_1(X)$ , there exists a map

$$\psi : [0, H] \rightarrow \mathcal{M}_1(X)$$

with  $H = d_\alpha(\mu^+, \mu^-)$  such that  $\psi(0) = \mu^+, \psi(H) = \mu^-$  and

$$d_\alpha(\psi(s_1), \psi(s_2)) = s_2 - s_1 \tag{5.6}$$

for any  $0 \leq s_1 < s_2 \leq H$ .

In fact, let  $\{a_i, b_i, G_i\}$  be an approximating graph sequence for an optimal transport path  $R$  in  $\text{Path}(\mu^+, \mu^-)$ . By Lemma 5.2, there exists a map  $t_i : G_i \rightarrow [0, H]$  such that

$$M^\alpha(G_i \cap t_i^{-1}[s_1, s_2]) = \frac{s_1 - s_2}{H} M^\alpha(G_i)$$

for any  $s_1, s_2 \in [0, H]$ .

Let  $\pi$  (and  $p$ ) be the first (and second) component projection of  $X \times [0, H]$ . For each  $i$ , lift  $G_i$  into the graph of the map  $t_i$  in  $X \times [0, H]$ , we get a transport path  $\bar{G}_i \in \text{Path}(a_i \times \{0\}, b_i \times \{H\})$ . Since

$$\begin{aligned} M^1(\bar{G}_i) &\leq M^1(p_\# \bar{G}_i) + M^1(G_i) \\ &\leq H + M^\alpha(G_i) \xrightarrow{i \rightarrow \infty} 2H \end{aligned}$$

is uniformly bounded,  $\{\bar{G}_i\}$  is subseguently weakly convergent to some vector measure  $\bar{R}$  on  $X \times [0, H]$ .

Since for any  $s \in [0, H]$  and any  $i$ , we have, as in (5.5), the atomic probability measures

$$\pi_{\#}(\bar{G}_i \cap p^{-1}(s)) = \sum_{a \in \bar{t}_i^{-1}(s)} m(a)\delta_a.$$

By [6, 4.3.6] and Fatou’s lemma, one has for almost every  $s \in [0, H]$ , a subsequence of  $\{\pi_{\#}(\bar{G}_i \cap p^{-1}(s))\}$  convergent to  $\pi_{\#}(\bar{R} \cap p^{-1}(s))$ , which must also be a probability measure.

Now, for almost any  $0 \leq s_1 < s_2 \leq H$ ,  $\pi_{\#}(\bar{R} \cap p^{-1}[s_1, s_2])$  gives a path from  $\pi_{\#}(\bar{R} \cap p^{-1}(s_1))$  to  $\pi_{\#}(\bar{R} \cap p^{-1}(s_2))$ , we have the Lipschitz estimate

$$\begin{aligned} d_{\alpha}(\pi_{\#}(\bar{R} \cap p^{-1}(s_1)), \pi_{\#}(\bar{R} \cap p^{-1}(s_2))) &\leq M^{\alpha}(\pi_{\#}(\bar{R} \cap p^{-1}[s_1, s_2])) \\ &\leq \liminf M^{\alpha}(\pi_{\#}(\bar{G}_i \cap p^{-1}[s_1, s_2])) \\ &= \liminf M^{\alpha}(G_i \cap t_i^{-1}[s_1, s_2]) \\ &\leq \frac{s_2 - s_1}{H} H = s_2 - s_1. \end{aligned}$$

Thus,  $\pi_{\#}(\bar{R} \cap p^{-1}(s))$  extends to a Lipschitz 1 map  $\psi$  from all of  $[0, H]$  to the complete metric space  $(\mathcal{M}_1(X), d_{\alpha})$

$$\psi: [0, H] \rightarrow (\mathcal{M}_1(X), d_{\alpha}).$$

Note that  $\psi(0) = \mu^+$ ,  $\psi(H) = \mu^-$  and since

$$d_{\alpha}(\psi(0), \psi(H)) = d_{\alpha}(\mu^+, \mu^-) = H,$$

we readily see that  $\psi$  satisfies (5.6) and hence  $(\mathcal{M}_1(X), d_{\alpha})$  is a length space.  $\square$

**Remark 5.1.** By Theorem 5.1, one may regard  $\pi_{\#}\bar{R} \in \text{Path}(\mu^+, \mu^-)$  as a geodesic from  $\mu^+$  to  $\mu^-$ .

### 6. Computer Visualizations

In this section, we’d like to use the ideas of the previous sections to give some computer visualizations about optimal transport paths.

#### 6.1. Flows from any probability measure to a Dirac measure

Given any probability measure  $\mu$  supported in a cube  $C$  in  $\mathbb{R}^m$  with edge length  $d$ , we’d like to flow it into a Dirac measure  $\delta_p$  for some  $p \in X$  with an almost optimal transport path. Let  $H$  be a fixed positive real number.

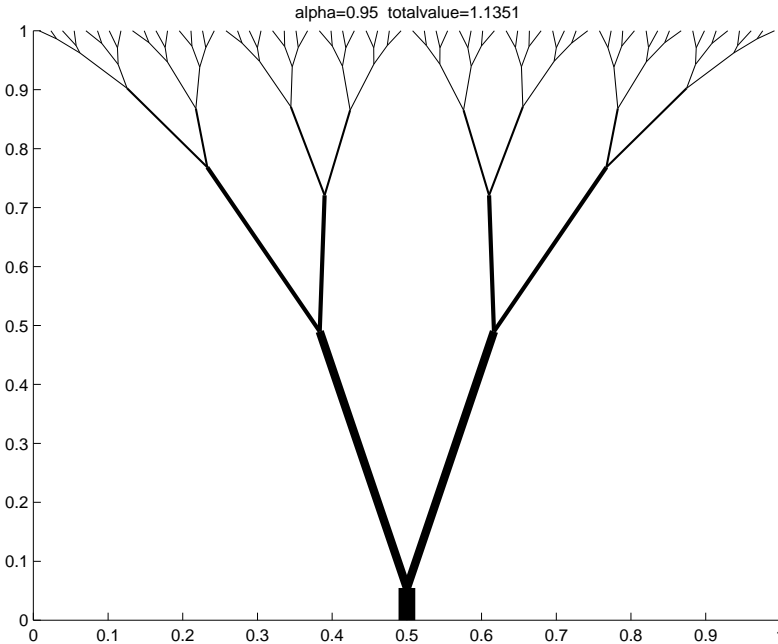
Algorithm:

- (1) Given an approximating depth  $n$ , let  $a_n = A_n(\mu)$  be the  $n$ th dyadic approximation of  $\mu$  as in Example 3.1.
- (2) For each  $h \in \mathbb{Z}^m \cap [0, 2^{n-1})^m$ , the cube  $Q_{n-1}^h$  of level  $n - 1$  consisting of  $2^m$  subcubes of level  $n$ . For any  $x \in X \times [0, H]$ , let  $G_x^h$  be the union of (the cone over  $a_n \lfloor Q_{n-1}^h$  with vertex  $x$ ) and the line segment  $\overline{xp}$  with weight  $\mu(Q_{n-1}^h)$ . Then  $G_x^h$  is a transport path in  $\text{Path}(a_n(\mu) \lfloor Q_{n-1}^h, \mu(Q_{n-1}^h)\delta_p)$ . Let  $q^h \in X \times [0, H]$  be the point at which  $M^\alpha(G_x^h)$  achieves its minimum among all  $x \in X \times [0, H]$ . Let

$$a_{n-1} = \sum_{h \in \mathbb{Z}^m \cap [0, 2^{n-1})^m} \mu(Q_{n-1}^h)\delta_{q^h} .$$

- (3) For each  $k = n - 1, \dots, 1$ , repeatedly doing step 2 to get  $a_{k-1}$ . In the end we get a transport path  $G_n \in \text{Path}(a_n, \delta_p)$  with finite  $M^\alpha$  mass.
- (4) By using Example 1, we can locally optimize the locations of the vertices of  $G$ . One may repeatedly doing upward optimization and downward optimization until the transport path converges to a fixed graph.
- (5) Increase depth  $n$  to get better approximation.

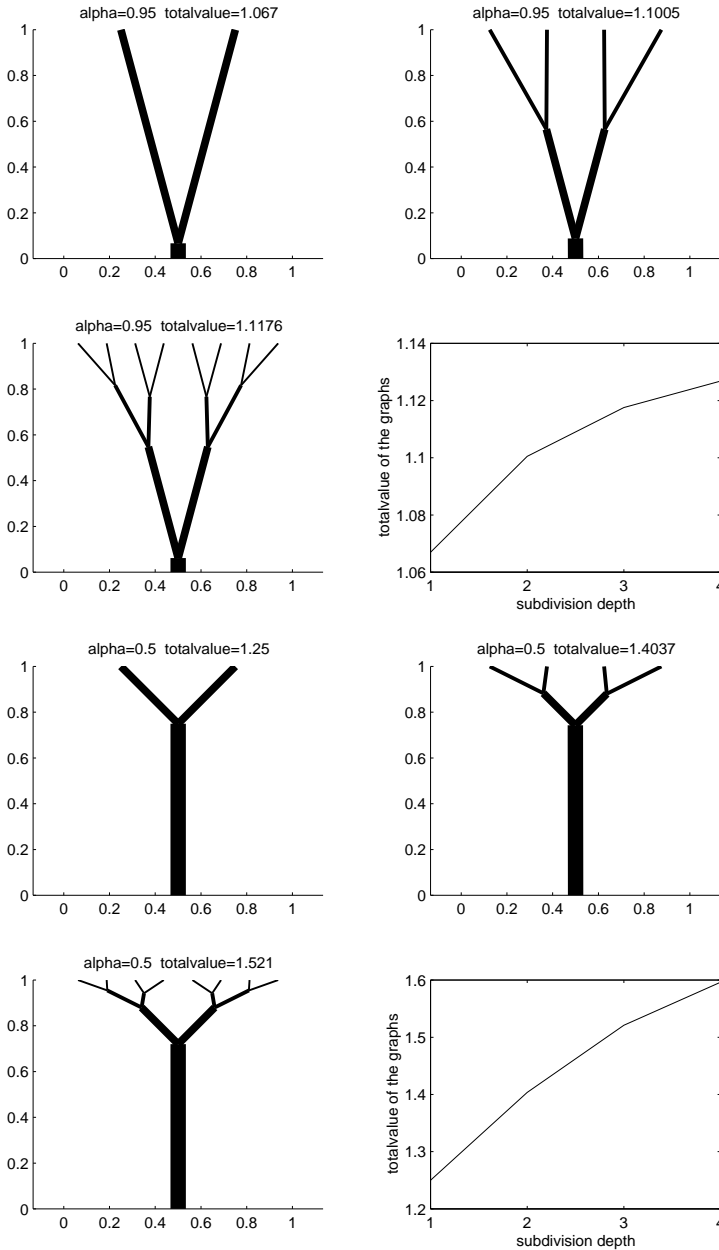
**Example 6.1.** When taking  $\mu =$  Lebesgue measure on  $[0, 1]$  and  $p = \frac{1}{2}$ ,  $\alpha = 0.95$ ,  $H = 1$  and take the depth  $n = 6$ , the above algorithm gives the following graph.

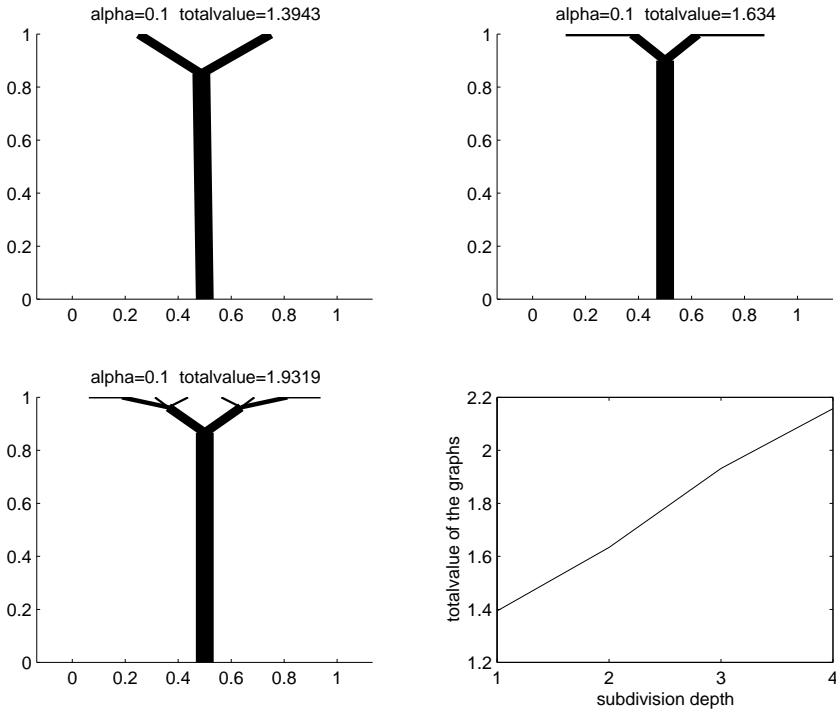


As we increase the approximating depth  $n$ , the  $M^\alpha$  mass of approximating paths may also be increasing. However, by Theorem 3.1, they will converge to a

finite number i.e. the cost of an optimal transport path. This phenomenon may be illustrated by the following example:

**Example 6.2.** Take  $\mu^+$  to be the Lebesgue measure on  $[0, 1]$  and  $\mu^- = \delta_{1/2}$  to be the Dirac measure at  $1/2$ .  $H = 1$ . Then the above algorithm gives approximating transporting flows from  $\mu^+$  to  $\mu^-$  with different  $\alpha$ :





**6.2. General measure to general measure**

Now, we want to flow a general measure  $\mu^+$  to another general measure  $\mu^-$ .

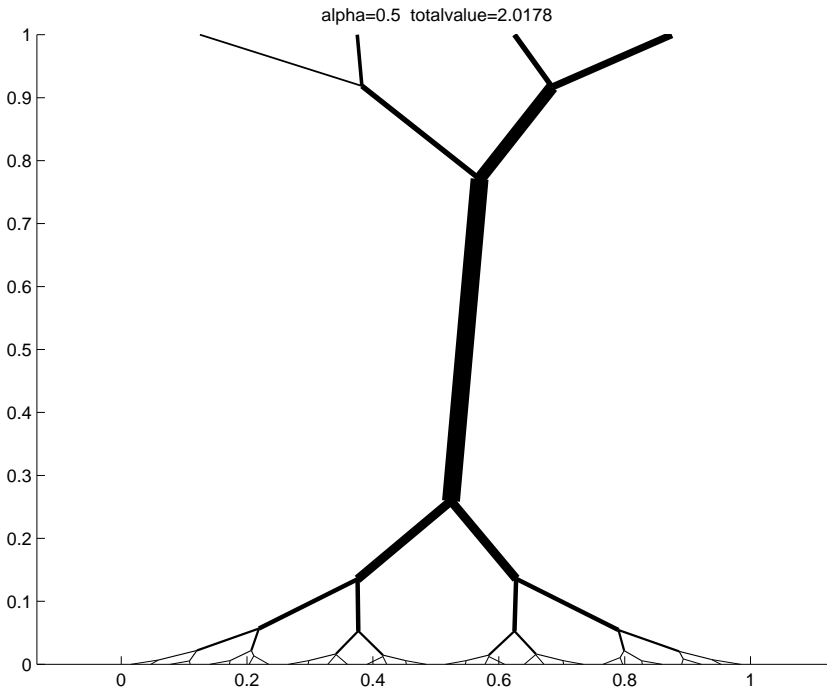
Algorithm:

- (1) Given depth  $n$ , let  $A_n(\mu^+)$  and  $A_n(\mu^-)$  be the  $n$ th dyadic approximation of  $\mu^+$  and  $\mu^-$  respectively.
- (2) Flow both  $A_n(\mu^+)$  and  $A_n(\mu^-)$  to a common Dirac measure as before.
- (3) Simplify the graphs by getting rid of unnecessary vertices (e.g. some vertex may have only one child and one parent).
- (4) Optimize the locations of each vertex as before.
- (5) If a vertex has two parents and two children (or has two parent and one child but the child has two children), use Example 2 to optimize the positions. This step may change the topology of the graphs.
- (6) Repeat steps 3-5 until it converges to an optimal path.

**Example 6.3.** Use the above algorithm, we flow the Lebesgue measure (with depth  $n = 6$ ,  $H = 1$ ) into an atomic measure

$$\frac{1}{10} \left[ \left[ \frac{1}{8} \right] \right] + \frac{2}{10} \left[ \left[ \frac{3}{8} \right] \right] + \frac{3}{10} \left[ \left[ \frac{5}{8} \right] \right] + \frac{4}{10} \left[ \left[ \frac{7}{8} \right] \right]$$

as follows:



### 7. Transport Path Versus Transport Plan

When splitting a vertex on a transport path, information about source and target may become unclear. However, we'll see very soon that those information can be traced by a transport path together with a compatible transport plan.

Recall that a transport plan for  $\mu^+, \mu^- \in \mathcal{M}_1(X)$  is a probability measure  $\gamma \in \mathcal{M}_1(X \times X)$  such that

$$\pi_{x\#}\gamma = \mu^+, \pi_{y\#}\gamma = \mu^-, \tag{7.1}$$

where  $\pi_x$  (and  $\pi_y$ ):  $X \times X \rightarrow X$  are the first (and the second) component projection. Let

$$\text{Plan}(\mu^+, \mu^-)$$

be the space of all transport plan for  $\mu^+$  and  $\mu^-$ .

#### 7.1. Atomic case

In this subsection, we fix two given atomic probability measures

$$a = \sum_{i=1}^m m_i \delta_{x_i} \text{ and } b = \sum_{j=1}^n n_j \delta_{y_j}$$



in  $\mathcal{M}_1(X)$ . Let

$$\text{Path}_0(a, b) = \{G \in \text{Path}(a, b) \text{ contains no cycles}\}.$$

Note that for any  $G \in \text{Path}_0(a, b)$ , each  $i$  and  $j$ , there exists at most one connected oriented piecewise linear curve  $g_{ij}$  from  $x_i$  to  $y_j$ , supported in  $G$ . If such curve doesn't exist, we set  $g_{ij} = 0$ . Thus, we may associate each  $G \in \text{Path}_0(a, b)$  with an  $m \times n$  1-dimensional vector measure valued matrix

$$g(G) = (g_{ij})_{m \times n} \tag{7.2}$$

with

each entry  $g_{ij}$  being either zero or an oriented piecewise linear curve

$$\text{from } x_i \text{ to } y_j \text{ and the union of support } (g_{ij}) \text{ contains no cycles.} \tag{7.3}$$

Similarly, any transport plan  $\gamma \in \text{Plan}(a, b)$  can be expressed as

$$\gamma = \sum_{i=1}^m \sum_{j=1}^n u_{ij} \delta_{(x_i, y_j)} \in \mathcal{M}_1(X \times X). \tag{7.4}$$

Thus, associated with each  $\gamma$ , there is an  $m \times n$  real matrix

$$u(\gamma) = (u_{ij})$$

with

$$u_{ij} \geq 0, \sum_{j=1}^n u_{ij} = m_i, \sum_{i=1}^m u_{ij} = n_j \text{ and } \sum_{i=1}^m \sum_{j=1}^n u_{ij} = 1. \tag{7.5}$$

**Definition 7.1.** Any pair  $(G, \gamma) \in \text{Path}_0(a, b) \times \text{Plan}(a, b)$  is said to be **compatible** if

$$G = \sum_{i=1}^m \sum_{j=1}^n u_{ij} g_{ij},$$

i.e.  $G = u(\gamma) \cdot g(G)$ , where  $u_{ij}$  and  $g_{ij}$  are given in (7.4) and (7.2) respectively.

Note that for any compatible pair  $(G, \gamma)$  and any  $i, j$

$$g_{ij} = 0 \implies u_{ij} = 0. \tag{7.6}$$

**Proposition 7.1.** Any pair  $(u, g)$  satisfying (7.5), (7.3), and (7.6) provides a compatible pair  $(G, \gamma)$  by

$$G = u \cdot g \text{ and } \gamma = u \cdot \text{bdry}(g),$$

where the matrix

$$\text{bdry}(g) = (\delta_{(x_i, y_j)}).$$

Moreover, if  $g = g(G_0)$  for some  $G_0 \in \text{Path}_0(a, b)$ , then  $G = G_0$ .

**Proof.** Suppose  $g = g(G_0)$ , then  $G - G_0$  is a cycle and is supported in the contractible 1-dimensional set support  $(G_0)$ . Thus  $G - G_0 = 0$ .  $\square$

**Corollary 7.1.** *If the matrix  $g = (g_{ij})$  has no zero entries, then  $g$  is compatible with  $u(\gamma)$  for any  $\gamma \in \text{Plan}(a, b)$ .*

**Corollary 7.2.** *There exists  $G \in \text{Path}_0(a, b)$  compatible with all  $\gamma \in \text{Plan}(a, b)$ .*

**Proof.** Pick  $c \in X$  outside the join of  $a$  and  $b$ , i.e.  $c$  is not on any line from a point of  $a$  to a point of  $b$ . Then let  $G$  be the difference of two cones:

$$G = [[c]]\#b - [[c]]\#a.$$

Since each entry of  $g(G)$  is nonzero, by the previous corollary,  $G$  is compatible with any  $\gamma \in \text{Plan}(a, b)$ .  $\square$

**Lemma 7.1.** *For any  $G \in \text{Path}_0(a, b)$ , there exists a  $\gamma \in \text{Plan}(a, b)$  compatible with  $G$ .*

**Proof.** Starting with each beginning vertex  $x_i$  of  $a$  having weight  $m_i$ , we move down vertex by vertex. At each vertex  $v \in V(G)$ , we consider the total amount received from the “ancestors” of  $v$  and decompose this amount “fairly” to the children of  $v$ , i.e. proportional to the weights on the edges connecting  $v$  and its children. For each beginning vertex  $x_i$  on  $a$  and each ending vertex  $y_j$  on  $b$ , the amount of measure  $u_{ij}$  on the “descendant”  $y_j$  inherited from the “ancestor”  $x_i$  is then well defined. Then  $\gamma = \sum_{i=1}^m \sum_{j=1}^n u_{ij} \delta_{(x_i, y_j)} \in \text{Plan}(a, b)$  is compatible with  $G$ .  $\square$

**Example 7.1.** Suppose  $G \in \text{Path}_0(a, b)$ . There exists a unique plan  $\gamma \in \text{Plan}(a, b)$  compatible with  $G$  if and only if  $G$  and  $\gamma$  have the following forms:

- Each connected component of  $G$  has either a single beginning point or a single ending point (or both).
- For each  $i, j$ , either the whole  $i$ th row or the whole  $j$ th column of the matrix representation  $u(\gamma)$  of  $\gamma$ , except the entry  $u_{ij}$ , consists of zero.

In fact, each nonzero column or row of  $u(\gamma)$  corresponds to precisely one connected component of  $G$ .

**Definition 7.2.** For any compatible pair  $(G, \gamma)$ , the pair  $(u(\gamma), g(G))$  given by (7.4), (7.2) is called the matrix representation of the pair  $(G, \gamma)$ .

**Remark 7.1.** Note that each pair  $(u, g)$  provides exactly the transporting information about source and targets. Each  $u_{ij}$  tells us the amount of transported measure from  $x_i$  to  $y_j$ , while each  $g_{ij}$  provides the actual transport path for this transportation. Moreover, the pair  $(u, g)$  also tells us how to split the measures at each possible splitting points of the transport path. Thus, each transport path,

together with a compatible transport plan, provides the necessary transporting information by its unique matrix representation  $(u(\gamma), g(G))$ .

**Remark 7.2.** If we restrict each  $g_{ij}$  in (7.3) to be the line segment  $]x_i, y_j[$  from  $x_i$  to  $y_j$ , then each transport plan  $\gamma$  may be identified with the transport path  $u(\gamma) \cdot g = \sum_{i=1}^m \sum_{j=1}^n u_{ij}]x_i, y_j[$ . In this sense, a transport path is a more general notion than a transport plan.

**7.2. General case**

**Definition 7.3.** The pair  $(T, \gamma) \in \text{Path}(\mu^+, \mu^-) \times \text{Plan}(\mu^+, \mu^-)$  is said to be **compatible** if

- (1) there exist two sequences of atomic probability measures  $\{a_i\}, \{b_i\} \in \mathcal{M}_1(X)$  such that  $a_i \rightharpoonup \mu^+$  and  $b_i \rightharpoonup \mu^-$ .
- (2) there exists a compatible pair  $(G_i, \gamma_i) \in \text{Path}(a_i, b_i) \times \text{Plan}(a_i, b_i)$  such that

$$G_i \rightharpoonup T \text{ and } \gamma_i \rightharpoonup \gamma.$$

We want to consider two optimization problems related to compatible transport paths and transport plans.

**Proposition 7.2.** *There exists a transport path  $T \in \text{Path}(\mu^+, \mu^-)$  such that  $T$  is compatible with every transport plan  $\gamma \in \text{Plan}(\mu^+, \mu^-)$ , and  $M^\alpha(T) < +\infty$ .*

**Proof.** For any  $n \geq 1$ , let  $A_n(\mu^+)$  and  $A_n(\mu^-)$  be the dyadic approximation of  $\mu^+$  and  $\mu^-$ . Let  $G_n$  be the union of two directed weighted graphs constructed as in the proof of Proposition 3.1.

$$G_n = G_n^+ + G_n^- \in \text{Path}(A_n(\mu^+), A_n(\mu^-))$$

with  $G_n^+ \in \text{Path}(A_n(\mu^+), \delta_c)$  and  $G_n^- \in \text{Path}(\delta_c, A_n(\mu^-))$ . Note that  $g(G_n)$  has no zero entries. Since  $M^\alpha(G_n)$  is uniformly bounded, it is subsequently convergent to some  $T \in \text{Path}(\mu^+, \mu^-)$  with  $M^\alpha(T) < +\infty$ .

Now, for any transport plan  $\gamma \in \text{Plan}(\mu^+, \mu^-)$ , let

$$\gamma_n = \sum_{h, h'} \gamma(Q_n^h \times Q_n^{h'}) \delta_{(c^h, c^{h'})} \rightharpoonup \gamma.$$

Since  $G_n \in \text{Path}(A_n(\mu^+), A_n(\mu^-))$  and  $g(G_n)$  has no zero entries, by Corollary 7.1,  $(G_n, \gamma_n)$  is compatible. This implies  $(T, \gamma)$  is compatible for any  $\gamma$ . □

**Proposition 7.3.** *Given a transport plan  $\gamma \in \text{Plan}(\mu^+, \mu^-)$ , there exists an optimal transport path  $T \in \text{Path}(\mu^+, \mu^-)$  with least finite  $M^\alpha$  cost among all compatible pairs  $(T, \gamma)$ .*

**Proof.** Follows from Proposition 7.2 and an analogous proof of Theorem 3.1. □

A typical realistic application for this proposition is in the mailing problem. Everyday, the headquarters of the Post Service has a given “transport plan” for sending out letters and packages, given by the addresses of the recipients. The best way to transport those items to their destinations is given by the optimal transport path compatible with the given plan as stated above.

Conversely, by Lemma 7.1 and a simple compactness argument about probability measures, one has the existence result for the following converse problem:

**Proposition 7.4.** *Given a transport path  $T \in \text{Path}(\mu^+, \mu^-)$ , there exists an optimal transport plan  $\gamma \in \text{Plan}(\mu^+, \mu^-)$  with least  $I(\gamma)$  cost among all compatible pairs  $(T, \gamma)$ , where*

$$I(\gamma) := \int_{\text{spt}(\mu^+) \times \text{spt}(\mu^-)} c(x, y) d\gamma(x, y)$$

for any given lower semicontinuous cost density function  $c : \text{spt}(\mu^+) \times \text{spt}(\mu^-) \rightarrow [0, +\infty)$  as in the Monge–Kantorovich problem.

This problem also has many applications. In shipping a given material of fixed amounts from several suppliers to be received in fixed amounts by several customers (along a fixed, possibly  $M^\alpha$  optimal transport path), there may be additional preferences of each customer  $y$  concerning the relative amounts from each supplier  $x$ . These preferences can be handled using the “cost” density function  $c(x, y)$ .

## Acknowledgment

The author wants to thank his thesis advisor Professor Robert Hardt for his infinite help and encouragements. The author also appreciates the useful suggestions of the referee and editor.

## References

- [1] L. Ambrosio, *Lecture Notes on Optimal Transport Problems*, Scuola Normale Superiore, Pisa, 2000.
- [2] L. Ambrosio and B. Kirchheim, Rectifiable sets in metric and Banach spaces, *Math. Ann.* **318**(3) (2000) 527–555.
- [3] T. De Pauw and R. Hardt, Size minimization and approximating problems, preprint.
- [4] L. C. Evans, Partial differential equations and Monge–Kantorovich mass transfer, in *Current Developments in Mathematics, 1997, Cambridge, MA*, Int. Press, Boston, MA, 1999, pp. 65–126.
- [5] L. C. Evans and W. Gangbo, Differential equations methods for the Monge–Kantorovich mass transfer problem, *Mem. Amer. Math. Soc.* **137**(653) (1999).
- [6] H. Federer, *Geometric Measure Theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag New York Inc., New York 1969.
- [7] W. Gangbo and R. J. McCann, The geometry of optimal transportation, *Acta Math.* **177**(2) (1996) 113–161.
- [8] G. Monge, Mémoire sur la théorie des déblais et de remblais, Histoire de l’Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année (1781), pp. 666–704.

- [9] R. Hardt and T. Riviere, Connecting topological Hopf singularities, preprint.
- [10] L. Kantorovich, On the translocation of masses, *C. R. (Doklady) Acad. Sci. URSS (N.S.)* **37** (1942) 199–201.
- [11] V. N. Sudakov, Geometric problems in the theory of infinite-dimensional probability distributions, (Cover to cover translation of Trudy Mat. Inst. Steklov 141 (1976)). *Proc. Steklov Inst. Math.* (1979) (2) 1–178.
- [12] B. White, Rectifiability of flat chains, *Ann. of Math. (2)* **150**(1) (1999) 165–184.