

# Duality Theorems in Ergodic Transport

Artur O. Lopes\* and Jairo K. Mengue†

January 25, 2012

## Abstract

We analyze several problems of Optimal Transport Theory in the setting of Ergodic Theory. In a certain class of problems we consider questions in Ergodic Transport which are generalizations of the ones in Ergodic Optimization.

Another class of problems is the following: suppose  $\sigma$  is the shift acting on Bernoulli space  $X = \{0, 1\}^{\mathbb{N}}$ , and, consider a fixed continuous cost function  $c : X \times X \rightarrow \mathbb{R}$ . Denote by  $\Pi$  the set of all Borel probabilities  $\pi$  on  $X \times X$ , such that, both its  $x$  and  $y$  marginal are  $\sigma$ -invariant probabilities. We are interested in the optimal plan  $\pi$  which minimizes  $\int c d\pi$  among the probabilities on  $\Pi$ .

We show, among other things, the analogous Kantorovich Duality Theorem. We also analyze uniqueness of the optimal plan under generic assumptions on  $c$ . We investigate the existence of a dual pair of Lipschitz functions which realizes the present dual Kantorovich problem under the assumption that the cost is Lipschitz continuous. For continuous costs  $c$  the corresponding results in the Classical Transport Theory and in Ergodic Transport Theory can be, eventually, different.

We also consider the problem of approximating the optimal plan  $\pi$  by convex combinations of plans such that the support projects in periodic orbits.

## 1 Introduction

For a compact metric space  $X$ , we denote  $P(X)$  the set of probabilities acting on the Borel sigma-algebra  $\mathcal{B}(X)$ .  $C(X)$  denotes the set of continuous functions on  $X$  taking real values.

---

\*arturoscar.lopes@gmail.com, Instituto de Matemática - UFRGS - Partially supported by DynEurBraz, CNPq, PRONEX – Sistemas Dinamicos, INCT, Convenio Brasil-Franca

†jairokras@gmail.com, Instituto de Matemática - UFRGS

We denote by  $\sigma$  the shift acting on  $\{1, 2, \dots, d\}^{\mathbb{N}}$  and by  $\hat{\sigma}$  the shift acting on  $\{1, 2, \dots, d\}^{\mathbb{Z}}$ . Some of our results apply to more general cases where one can consider a continuous transformation defined on any compact metric space  $X$ . Anyway, the reader can take  $\{1, 2, \dots, d\}^{\mathbb{Z}} = X \times Y = \{1, 2, \dots, d\}^{\mathbb{N}} \times \{1, 2, \dots, d\}^{\mathbb{N}}$  as our favorite toy model.

We consider a continuous cost function  $c : X \times Y \rightarrow \mathbb{R}$ , where  $X$  and  $Y$  are compact metric spaces.

The Classical Transport Problem consider probabilities  $\pi$  on  $P(X \times Y)$  and the minimization of  $\int c(x, y) d\pi(x, y)$  under the hypothesis that the  $y$ -marginal of  $\pi$  is a fixed probability  $\nu$  and  $x$ -marginal of  $\pi$  is a fixed probability  $\mu$ . A probability  $\pi$  which minimizes such integral is called an optimal plan [27] [12].

We want to analyze a different class of problem where in some way the restriction to invariant probabilities [22] appears in some form. We present several different settings.

As a motivation one can ask: given the 2-Wasserstein metric  $W$  on the space of probabilities on  $X = \{1, 2, \dots, d\}^{\mathbb{N}}$ , and a certain fixed probability  $\mu$ , which is not invariant for the shift  $\sigma : \{1, 2, \dots, d\}^{\mathbb{N}} \rightarrow \{1, 2, \dots, d\}^{\mathbb{N}}$ , characterize the closest  $\sigma$ -invariant probability  $\nu$  to  $\mu$ . In other words, we can be interested in finding a  $\sigma$ -invariant probability  $\nu$  which minimizes the value  $W(\mu, \nu)$  for a fixed  $\mu$ . In this case we are taking  $X = Y$ . What can be said about optimal transport plans, duality, etc?

As a generalization of this problem one can consider a continuous cost  $c(x, y)$ , where  $c : \{1, 2, \dots, d\}^{\mathbb{N}} \times \{1, 2, \dots, d\}^{\mathbb{N}} \rightarrow \mathbb{R}$ , and ask about the properties of the plan  $\pi$  on  $\{1, 2, \dots, d\}^{\mathbb{N}} \times \{1, 2, \dots, d\}^{\mathbb{N}}$  which minimize  $\int c(x, y) d\pi(x, y)$  under the hypothesis that the  $y$ -marginal of  $\pi$  is a variable  $\sigma$ -invariant probability  $\nu$ , and the  $x$ -marginal of  $\pi$  is a fixed probability  $\mu$ .

We note that a plain with this marginals properties is characterized by:

$$\begin{cases} \int f(x) d\pi(x, y) = \int f(x) d\mu(x) & \text{for any } f \in C(X) \\ \int g(y) d\pi(x, y) = \int g(\sigma(y)) d\pi(x, y) & \text{for any } g \in C(Y) \end{cases} \quad (1)$$

We will show in section 2 the following:

**Theorem 1 (Kantorovich duality).** *Consider a compact metric space  $X$  and  $Y = \{1, 2, \dots, d\}^{\mathbb{N}}$ . Consider a fixed  $\mu \in P(X)$ , and a fixed continuous cost function  $c : X \times Y \rightarrow \mathbb{R}^+$ . Define  $\Pi(\mu, \sigma)$  as the set of all Borel probabilities  $\pi \in P(X \times Y)$  satisfying (1). Define  $\Phi_c$  as the set of all pair of continuous functions  $(\varphi, \psi) \in C(X) \times C(Y)$  which satisfy:*

$$\varphi(x) + \psi(y) - (\psi \circ \sigma)(y) \leq c(x, y), \quad \forall (x, y) \in X \times Y \quad (2)$$

*Then,*

I)

$$\inf_{\Pi(\mu, \sigma)} \int c d\pi = \sup_{(\varphi, \psi) \in \Phi_c} \int \varphi d\mu. \quad (3)$$

Moreover, the infimum in the left hand side is attained.

II) If  $c$  is a Lipschitz continuous function, then, there exist Lipschitz  $\varphi$  and  $\psi$  which are admissible realizers of the supremum.

Any pair  $\varphi$  and  $\psi$  satisfying (2) is called admissible. Any  $\pi$  realizing the infimum in (3) will be called an optimal plan, and its  $y$ -projection an optimal invariant probability solution for  $c$  and  $\mu$ . Moreover,  $\varphi$  and  $\psi$  are called an optimal dual pair if they realize the maximum of the right hand side expression. It is possible that does not exists an optimal dual pair (see remark bellow).

The following criteria is quite useful.

**Slackness condition** [27] [28]: suppose for all  $(x, y)$  in the support of  $\pi \in \Pi(\mu, \sigma)$  we have that

$$\varphi(x) + \psi(y) - (\psi \circ \sigma)(y) = c(x, y),$$

for some admissible  $\varphi$  and  $\psi$  satisfying  $\varphi + \psi - (\psi \circ \sigma) \leq c$ , then  $\pi$  is an optimal plan and  $\varphi$  and  $\psi$  is an optimal dual pair.

In recent years several results in the so called Ergodic Optimization Theory were obtained [15] [6] [17] [2] [23] [3] [10] [1]. We will show that the above kind of Ergodic Transport problem contains as a particular case this other theory. The subaction which possess properties of minimality described in [6] and [11] can be seen as a version of Kantorovich duality.

In the below remark we show that there are conceptual differences in the kind of analogous results we can get in the Classical and in the Ergodic Transport Theory.

**Remark 2.** We point out that if  $\mu$  is a Dirac delta in a point  $x_0$ , then the cost  $c(x_0, y)$  just depends on  $y$ . In this way if we denote  $A(y) = c(x_0, y)$  we get that the above problem is the classical one in Ergodic Optimization, where one is interested in minimizing  $\int A d\nu$  among invariant probabilities  $\nu$ . There is no big difference in this theory if one consider maximization instead of minimization. The function  $\psi$  above corresponds to the concept of subaction and the number  $\varphi(x_0)$  is equal to  $\min\{\int A(y) d\nu(y) : \nu \text{ is invariant}\}$  [6], [9], [3], [15]. It is known that for a  $C^0$ -generic continuous potential  $A$  does not exist a continuous subaction [4]. For the Classical Transport problem in compact spaces there exists continuous realizers for the dual problem when  $c$  is continuous [28]. This shows that there are non trivial differences (at least for a  $C^0$  cost function  $c$ ) between the Classical and the Ergodic transport

setting. It is known the existence of a calibrated Holder subaction for a Holder potential  $A$ . The item II on the above theorem is the correspondent result on the present setting. The expression  $\varphi(x) + \psi(\sigma(y)) - \psi(y) \leq c(x, y)$  can be for some people more natural. This can be also obtained by replacing  $\psi$  by  $-\psi$ .

The next example shows that the Ergodic Transport problem can not be derived in an easy way from Ergodic Optimization properties.

We denote by  $(a_1 \dots a_n)^\infty$  the periodic point  $(a_1 \dots a_n a_1 \dots a_n a_1 \dots)$  in  $\{0, 1\}^\mathbb{N}$ .

**Example 3.** Consider  $X = \{x_0, x_1\}$ ,  $\mu = \frac{1}{2}(\delta_{x_0} + \delta_{x_1})$ ,  $Y = \{0, 1\}^\mathbb{N}$ , and a cost function  $c$  defined on  $X \times Y$ , satisfying the following properties:

1)  $c(x_0, (01)^\infty) = 0$ ,  $c(x_0, (10)^\infty) = 1$ ,  $c(x_0, 0^\infty) = 1/4$ ,  $c(x_0, y) > 0$ , if  $y \neq (01)^\infty$ .

2)  $c(x_1, (01)^\infty) = 1$ ,  $c(x_1, (10)^\infty) = 0$ ,  $c(x_1, 1^\infty) = 1/4$ ,  $c(x_1, y) > 0$ , if  $y \neq (10)^\infty$ .

Assume  $c$  is Lipschitz continuous.

Note that, as an example, we can take

$$c(x_0, y) = d^2(y, (01)^\infty), \quad c(x_1, y) = d^2(y, (10)^\infty).$$

We observe that the measure  $\nu = \frac{1}{2}(\delta_{(01)^\infty} + \delta_{(10)^\infty})$  is not a minimizing measure for either of the potentials  $A_0(y) := c(x_0, y)$ , or,  $A_1(y) := c(x_1, y)$ . By the other hand, the unique optimal plan is given by

$$\pi = \frac{1}{2}(\delta_{(x_0, (01)^\infty)} + \delta_{(x_1, (10)^\infty)}),$$

which projects on  $\mu$  and  $\nu$ .

We will also show in section 3 that generically on  $c$  the optimal plan is unique.

In another kind of problem one can ask: given a continuous cost  $c(x, y)$ ,  $c : \{1, 2, \dots, d\}^\mathbb{N} \times \{1, 2, \dots, d\}^\mathbb{N} \rightarrow \mathbb{R}$ , what are the properties of the probability  $\pi$  on  $\{1, 2, \dots, d\}^\mathbb{N} \times \{1, 2, \dots, d\}^\mathbb{N}$  which minimize  $\int c(x, y) d\pi(x, y)$  under the hypothesis that the  $y$ -marginal of  $\pi$  is a variable invariant probability  $\nu$  and the  $x$ -marginal of  $\pi$  is a variable invariant probability  $\mu$ ? Under what assumptions on  $c$  we get that the optimal plan  $\pi$  is invariant for  $\hat{\sigma}$ ?

We will present now formal definitions of the second class of problems.

Here we fix compact metric spaces  $X$  and  $Y$  and continuous transformations

$$T_1 : X \times Y \rightarrow X, \quad T_2 : X \times Y \rightarrow Y,$$

such that,  $T : X \times Y \rightarrow X \times Y$ , given by  $T = (T_1, T_2)$ , defines a transformation of  $X \times Y$  to itself. Let  $\Pi(T)$  the set of Borel probability measures  $\pi$  in  $X \times Y$ , such that, for any  $f : X \rightarrow \mathbb{R}$ ,  $g : Y \rightarrow \mathbb{R}$ :

$$\int f(x) d\pi(x, y) = \int f(T_1(x, y)) d\pi(x, y),$$

and

$$\int g(y) d\pi(x, y) = \int g(T_2(x, y)) d\pi(x, y).$$

The set of such  $\pi$  is called the set of admissible plans.

Note that any  $T$ -invariant measure in  $X \times Y$  (which exists because  $X \times Y$  is compact) satisfies this condition. Indeed, if  $\nu$  is  $T$ -invariant, then:

$$\begin{aligned} \int f(x) d\nu(x, y) &= \int f(x, y) d\nu(x, y) = \\ &= \int f(T(x, y)) d\nu(x, y) = \int f((T_1(x, y), T_2(x, y))) d\nu(x, y) = \\ &= \int f(T_1(x, y)) d\nu(x, y). \end{aligned}$$

A similar reasoning can be applied to  $g$ .

Given a continuous function  $c : X \times Y \rightarrow [0, +\infty)$ , what can be said about

$$\alpha(c) := \inf \left\{ \int c d\pi : \pi \in \Pi(T) \right\}?$$

What are the properties of optimal plans? We are interested here in Kantorovich Duality type of results.

We will show the following:

**Theorem 4 (Kantorovich duality).**  *$\alpha(c)$  is the supremum of the numbers  $\alpha$  such that there exists continuous functions  $\varphi : X \rightarrow \mathbb{R}$ ,  $\psi : Y \rightarrow \mathbb{R}$  satisfying:*

$$\alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)) \leq c(x, y), \quad \forall (x, y) \in X \times Y.$$

We can list different interesting cases where we can apply the above result:

1) If  $T_1$  doesn't depends of  $y \in Y$  and  $T_2$  doesn't depends of  $x \in X$ , then we have the expression:

$$\alpha + \varphi(x) - \varphi(T_1(x)) + \psi(y) - \psi(T_2(y)) \leq c(x, y) \quad \forall (x, y) \in X \times Y.$$

In this case we are considering two variable invariant probabilities (one for  $T_1$  and the other for  $T_2$ ) as marginals of an admissible plan.

2) If  $X$  and  $Y$  are the Bernoulli space  $\{1, 2, \dots, d\}^{\mathbb{N}}$ ,  $T_1 = \sigma$  is the shift acting on the variable  $x$ , (doesn't depend of  $y \in Y$ ) and  $T_2 = \tau_x(y)$  (where  $\tau_j$ ,  $j = 1, 2, \dots, d$ , are the inverse branches of  $\sigma$  acting on the variable  $y$ ) we have that  $T = \hat{\sigma}$  is the shift on  $\{1, 2, \dots, d\}^{\mathbb{Z}}$  and the above expression can be written as:

$$\alpha + \varphi(x) - \varphi(\sigma(x)) + \psi(y) - \psi(\tau_x(y)) \leq c(x, y) \quad \forall (x, y) \in X \times Y.$$

In this case invariant probabilities  $\pi$  for the shift  $\hat{\sigma} : \{1, 2, \dots, d\}^{\mathbb{Z}} \rightarrow \{1, 2, \dots, d\}^{\mathbb{Z}}$  are admissible plans. But not all admissible plan is  $\hat{\sigma}$  invariant.

It is necessary to assume some special properties on  $c$  in order to get an optimal plan which is  $\hat{\sigma}$ -invariant. This is the purpose of the next example.

**Example 5.** Consider the points  $x_0 = (01)^\infty$ ,  $x_1 = (10)^\infty$ ,  $y_0 = (001)^\infty$ ,  $y_1 = (010)^\infty$ ,  $y_2 = (100)^\infty$ . Let  $c$  a Lipschitz continuous cost satisfying:

- 1)  $c(x_0, y_0) = c(x_1, y_1) = c(x_0, y_2) = c(x_1, y_2) = 0$ ,
- 2)  $c > 0$  on the other points.

*It's easy to see that the unique optimal plan is given by:*

$$\pi = \frac{1}{3}\delta_{(x_0, y_0)} + \frac{1}{3}\delta_{(x_1, y_1)} + \frac{1}{6}\delta_{(x_0, y_2)} + \frac{1}{6}\delta_{(x_1, y_2)}.$$

*The support of this plan does not contain a  $\hat{\sigma}$ -invariant probability.*

Which is the right assumption to get an optimal plan  $\pi$  which is  $\hat{\sigma}$ -invariant? In [8], [13], [18], [19], [20] some results in this direction are presented considering a cost which is dynamically defined.

Another class of examples:

3) If  $X$  and  $Y$  are the Bernoulli space  $\{1, 2, \dots, d\}^{\mathbb{N}}$ , and, for all  $(x, y) \in X \times Y$ , we have  $T_1(x, y) = x$ , and  $T_2(x, y) = \tau_{x_0}(y)$ ,  $x = (x_0, x_1, \dots) \in \{1, 2, \dots, d\}^{\mathbb{N}}$ , (where  $\tau_j$ ,  $j = 1, 2, \dots, d$ , are the inverse branches of  $\sigma$ ), then, there is no  $\varphi(x)$  in this case, and, the above expression can be written as:

$$\alpha + \psi(y) - \psi(\tau_x(y)) \leq c(x, y) \quad \forall (x, y) \in X \times Y.$$

This is the holonomic setting of [11]. A duality result is proved in section 2 in [11]. Note that the  $y$ -marginal of a holonomic probability is invariant for the shift  $\sigma$  acting on the variable  $y$  (see section 1 in [11]). In the case  $c$  is Holder it is shown the existence of Holder realizers  $\psi$  in sections 3 and 4 [11].

We will show in sections 4 and 6 here that the optimal plans can be approximated by convex combination of optimal plans (of the classical transport problem) associated to measures supported in periodic orbits. In this way one can get an approximation scheme to the optimal plan based on a finite set of conditions. Our approach here is inspired in the point of view of taking the temperature going to zero for Gibbs states at positive temperature in order to get results in Ergodic Optimization [21]. The problem of fast approximation of maximizing probabilities by measures on periodic orbits plays an important role in Ergodic Optimization [14] [5] [7].

A paper which consider Ergodic Transport problems under a continuous time setting is [16].

We would like to thanks N. Gigli for very helpful comments and advice during the preparation of this manuscript.

## 2 The case of one fixed probability and another variable invariant one

Here we will prove Theorem 1. We will adapt the main reasoning described in [27].

Given a normed Banach space  $E$  we denote by  $E^*$  the dual space containing the bounded linear functionals from  $E$  to  $\mathbb{R}$ .

We will need the following classical result [27] [28].

**Theorem 6 (Fenchel-Rockafellar duality).** *Suppose  $E$  is a vector normed space,  $E^*$  its topological dual,  $\Theta$  and  $\Xi$  two convex functions defined on  $E$  taking values in  $\mathbb{R} \cup \{+\infty\}$ . Denote  $\Theta^*$  and  $\Xi^*$ , respectively, the Legendre-Fenchel transform of  $\Theta$  and  $\Xi$ . Suppose there exists  $x_0 \in E$ , such that  $x_0 \in D(\Theta) \cap D(\Xi)$ , and, that  $\Theta$  is continuous on  $x_0$ .*

*Then,*

$$\inf_{x \in E} [\Theta(x) + \Xi(x)] = \sup_{f \in E^*} [-\Theta^*(-f) - \Xi^*(f)] \quad (4)$$

*Moreover, the supremum in (4) is attained in at least one element in  $E^*$ .*

**Proof of Theorem 1.** First we prove I).

We want to use Fenchel-Rockafellar duality in the proof.

Define

$$E = C(X \times Y) \times M(Y)$$

where  $C(X \times Y)$  is the set of all continuous functions in  $X \times Y$  taking values in  $\mathbb{R}$ , with the usual sup norm  $\| \cdot \|_\infty$ . Moreover,  $M(Y)$  is the set of bounded

linear operators in  $C(Y)$  taking values in  $\mathbb{R}$  with the total variation norm. Let  $P_\sigma(Y)$  be the set of invariant probabilities on  $Y$ .

Define  $\Theta : E \longrightarrow \mathbb{R} \cup \{+\infty\}$  by

$$\Theta(u, \nu) = \begin{cases} 0, & \text{if } u(x, y) \geq -c(x, y), \forall (x, y) \in X \times Y, \text{ and } \|\nu\| \leq 2 \\ +\infty, & \text{in the other case,} \end{cases}$$

Note that  $\Theta$  is convex.

Define  $\Xi : E \longrightarrow \mathbb{R} \cup \{+\infty\}$  by

$$\Xi(u, \nu) = \begin{cases} \int_X \varphi d\mu, & \text{if } u(x, y) = \varphi(x) + \psi(y) - (\psi \circ \sigma)(y) = \\ & \text{with } (\varphi, \psi) \in C(X) \times C(Y), \text{ and } \nu \in P_\sigma(Y) \\ +\infty, & \text{in the other case.} \end{cases}$$

Note that  $\Xi$  is well defined. Indeed, if  $u = \varphi_1(x) + \psi_1(y) - (\psi_1 \circ \sigma)(y) = \varphi_2(x) + \psi_2(y) - (\psi_2 \circ \sigma)(y)$ , then, integrating under any invariant probability  $\nu \in P_\sigma(Y)$ , we have that  $\varphi_1(x) = \varphi_2(x)$ . Also note that  $\Xi$  is convex.

Observe that if  $\nu \in P_\sigma(Y)$ , then  $(1, \nu) \in D(\Theta) \cap D(\Xi)$  and  $\Theta$  is continuous in  $(1, \nu)$ .

Observe that

$$\begin{aligned} & \inf_{(u, \nu) \in E} [\Theta(u, \nu) + \Xi(u, \nu)] \\ &= \inf \left\{ \int_X \varphi d\mu : \varphi(x) + [\psi - (\psi \circ \sigma)](y) \geq -c(x, y), (\varphi, \psi) \in C(X) \times C(Y) \right\} \\ &= \inf \left\{ - \int_X \varphi d\mu; \varphi(x) + [\psi - (\psi \circ \sigma)](y) \leq c(x, y), (\varphi, \psi) \in C(X) \times C(Y) \right\} \\ &= - \sup \left\{ \int_X \varphi d\mu; \varphi(x) + [\psi - (\psi \circ \sigma)](y) \leq c(x, y), (\varphi, \psi) \in C(X) \times C(Y) \right\} \\ &= - \sup_{(\varphi, \psi) \in \Phi_c} \int \varphi d\mu. \end{aligned}$$

Now we will compute the Legendre-Fenchel transform of  $\Theta$  and  $\Xi$ , initially, for any  $(\pi, g) \in E^*$ : by the definition of  $\Theta$  we get

$$\begin{aligned} \Theta^*((-\pi, -g)) &= \sup_{(u, \nu) \in E} \{ \langle (-\pi, -g), (u, \nu) \rangle - \Theta(u, \nu) \} \\ &= \sup_{(u, \nu) \in E} \{ -\pi(u(x, y)) - g(\nu) : -u(x, y) \leq c(x, y), \|\nu\| \leq 2 \} \\ &= \sup_{(u, \nu) \in E} \{ \pi(u(x, y)) - g(\nu) : u(x, y) \leq c(x, y), \|\nu\| \leq 2 \}. \end{aligned}$$



Following [27] we note that if  $\pi$  is not a positive functional, then, there exists a function  $v \leq 0$ ,  $v \in C(X \times Y)$ , such that,  $\pi(v) > 0$ , therefore, taking  $u = \lambda v$  (remember that  $c \geq 0$ ), and considering  $\lambda \rightarrow +\infty$ , we get that

$$\sup_{u \in E} \{\pi(u) : u(x, y) \leq c(x, y)\} = +\infty.$$

When  $\pi \in M^+(X \times Y)$  and  $c \in C(X \times Y)$  we have that the supremum of  $\pi(u)$  is, evidently,  $\pi(c)$ .

Therefore,

$$\Theta^*((-\pi, -g)) = \begin{cases} \pi(c) + \sup_{\|\nu\| \leq 2} -g(\nu), & \text{if } \pi \in M^+(X \times Y) \\ +\infty, & \text{in the other case.} \end{cases} \quad (5)$$

Analogously, by the definition of  $\Xi$  we get that

$$\begin{aligned} \Xi^*(\pi, g) &= \sup_{(u, \nu) \in E} \{ \langle (\pi, g), (u, \nu) \rangle - \Xi(u, \nu) \} \\ &= \sup_{(u, \nu) \in E} \left\{ \begin{array}{l} \pi(u(x, y)) - \int \varphi d\mu + g(\nu) : \\ u(x, y) = \varphi(x) + \psi(y) - \psi(\sigma(y)) \text{ where } (\varphi, \psi) \in C(X) \times C(Y) \\ \text{and } \nu \in P_\sigma(Y) \end{array} \right\} \\ &= \sup_{(\varphi, \psi) \in C(X) \times C(Y), \nu \in P_\sigma(Y)} \left\{ \pi(\varphi(x) + \psi(y) - \psi(\sigma(y))) - \int \varphi d\mu + g(\nu) \right\}. \end{aligned}$$

If  $\pi(\varphi(x)) \neq \int \varphi d\mu$  (we can suppose greater) for some  $\varphi$ , taking  $\lambda \cdot \varphi$  and  $\lambda \rightarrow \infty$ , the supremum will be equal to  $+\infty$ . Analogously if  $\pi(\psi(y) - \psi(\sigma(y))) \neq 0$  (we can suppose greater) for some  $\psi$ , taking  $\lambda \cdot \psi$  and  $\lambda \rightarrow \infty$ , the supremum will be  $+\infty$ .

In order to simplify the notation, define

$$\Pi^*(\mu) = \left\{ \pi \in M(X \times Y) : \begin{array}{l} \pi(\varphi(x)) = \int \varphi d\mu \text{ and } \pi(\psi(y) - \psi(\sigma(y))) = 0 \\ \forall (\varphi, \psi) \in C(X) \times C(Y) \end{array} \right\}.$$

We just show that

$$\Xi^*(\pi, g) = \begin{cases} \sup_{\nu \in P_\sigma(Y)} g(\nu), & \text{if } \pi \in \Pi^*(\mu), \\ +\infty, & \text{in the other case.} \end{cases} \quad (6)$$

We know that the left hand side (4) is equal to  $-\sup_{(\varphi, \psi) \in \Phi_c} \int \varphi d\mu$ , and also by (5) and (6), we know that the right hand side of (4) coincide with

$$\begin{aligned}
& \sup_{(\pi, g) \in E^*} \left\{ - \left( \begin{array}{l} \pi(c) + \sup_{\|\nu\| \leq 2} -g(\nu) + \sup_{\nu \in P_\sigma(Y)} g(\nu), \text{ if } \pi \in M^+(X \times Y) \cap \Pi^*(\mu) \\ +\infty \qquad \qquad \qquad \text{in the other case} \end{array} \right) \right\} \\
&= \sup_{(\pi, g) \in E^*} \left\{ \begin{array}{l} -\pi(c) + \inf_{\|\nu\| \leq 2} g(\nu) - \sup_{\nu \in P_\sigma(Y)} g(\nu), \text{ if } \pi \in M^+(X \times Y) \cap \Pi^*(\mu) \\ -\infty, \qquad \qquad \qquad \text{in the other case} \end{array} \right\} \\
&= \sup \{ -\pi(c), \pi \in M^+(X \times Y) \cap \Pi^*(\mu) \},
\end{aligned}$$

where the last equality is obtained taking  $g = 0$  because  $\|\nu\| = 1$  for any  $\nu \in P_\sigma(Y)$ .

Finally, we observe that if  $\pi \in M^+(X \times Y) \cap \Pi^*(\mu, \nu)$  then:

$$\pi(1) = \mu(1) = 1,$$

$$\pi(u) \geq 0, \text{ if } u \geq 0,$$

$\pi$  is linear.

From these properties we get that  $\pi \in P(X \times Y)$ .

Moreover, by definition of  $\Pi^*(\mu)$ , the projection of  $\pi$  in the first coordinate is  $\mu$ , and, the projection of  $\pi$  is invariant in the second coordinate. It follows that  $M^+(X \times Y) \cap \Pi^*(\mu, \nu) = \Pi(\mu, \sigma)$ .

Therefore, from this together with (4) we get

$$- \sup_{(\varphi, \psi) \in \Phi_c} \int \varphi d\mu = - \inf_{\pi \in \Pi(\mu, \sigma)} \int c d\pi$$

or,

$$\sup_{(\varphi, \psi) \in \Phi_c} \int \varphi d\mu = \inf_{\pi \in \Pi(\mu, \sigma)} \int c d\pi.$$

Note that theorem 6 claims that

$$\sup_{f \in E^*} [-\Theta^*(-f) - \Xi^*(f)] = - \inf_{\pi \in \Pi(\mu, \sigma)} \int c d\pi$$

is attained, for at least one element, and this shows the existence of an optimal plan. This shows I).

After we get the probability  $\nu$  we can consider the classical transport problem for  $\mu, \nu$  and  $c$ , and, finally, we can get some well known properties described in the classical literature (as, slackness condition,  $c$ -cyclical monotonicity, etc...).

Now, we will prove II). This will follow from the following claim.

**Claim:** Let  $X$  be a compact metric space,  $Y = \{1, \dots, d\}^{\mathbb{N}}$ ,  $c : X \times Y \rightarrow \mathbb{R}$  be a Lipschitz continuous function and  $\mu$  be a probability measure acting in  $X$ . Let  $\pi \in \Pi(\mu, \sigma)$  minimizing the integral of  $c$ . Then, there exist Lipschitz continuous functions  $\varphi(x), \psi(y)$  such that:

- i)  $\varphi(x) + \psi(\sigma(y)) - \psi(y) \leq c(x, y)$
- ii)  $\int \varphi(x) d\mu = \int c d\pi$ .

Let  $\beta$  the Lipschitz constant of  $c$ .

First note that given continuous functions  $\varphi$  and  $\psi$  satisfying

$$\varphi(x) + \psi(\sigma(y)) - \psi(y) \leq c(x, y),$$

then, there exists  $\bar{\varphi}$  and  $\bar{\psi}$ , Lipschitz functions with Lipschitz constant  $\beta$  satisfying:

$$\bar{\varphi}(x) + \bar{\psi}(\sigma(y)) - \bar{\psi}(y) \leq c(x, y),$$

and,

$$\bar{\varphi} \geq \varphi.$$

We can choose  $\bar{\psi}$  satisfying  $0 \leq \bar{\psi} \leq \beta$ .

Indeed, for any  $\sigma^n(w) = y$  and  $x_0, \dots, x_{n-1} \in X$ :

$$\psi(y) - \psi(w) \leq \sum_{i=0}^{n-1} c(x_i, \sigma^i(w)) - \varphi(x_i).$$

This shows that

$$\bar{\psi}(y) := \inf \left\{ \sum_{i=0}^{n-1} c(x_i, \sigma^i(w)) - \varphi(x_i) : n \geq 0, \sigma^n(w) = y, x_i \in X \right\}$$

is well defined.

We remark that  $\bar{\psi}$  is a Lipschitz function with the same constant  $\beta$ . Note also that

$$\varphi(x) + \bar{\psi}(\sigma(y)) - \bar{\psi}(y) \leq c(x, y).$$

Now for each  $x$  fixed, define  $\bar{\varphi}(x)$  as the greatest number such that for any  $y$ :

$$\bar{\varphi}(x) + \bar{\psi}(\sigma(y)) - \bar{\psi}(y) \leq c(x, y).$$

We note that  $\bar{\varphi} \geq \varphi$  and  $\bar{\varphi}(x) = \inf_y \{c(x, y) + \bar{\psi}(y) - \bar{\psi}(\sigma(y))\}$ . We also note that  $\bar{\varphi}$  is a Lipschitz function with same constant  $\beta$ . We remark that we can

add a constant to  $\bar{\psi}$ , and, so we can suppose without loss of generality that  $0 \leq \bar{\psi} \leq \beta$ .

Now, we prove the main claim.

By (3), there exists sequences of continuous functions  $\varphi_n$  and  $\psi_n$ ,  $n \in \mathbb{N}$ , such that

$$\varphi_n(x) + \psi_n(\sigma(y)) - \psi_n(y) \leq c(x, y)$$

and

$$\int \varphi_n d\mu \rightarrow \int c d\pi.$$

From the above reasoning we can get  $\bar{\varphi}_n, \bar{\psi}_n$  Lipschitz continuous functions such that

$$\bar{\varphi}_n(x) + \bar{\psi}_n(\sigma(y)) - \bar{\psi}_n(y) \leq c(x, y),$$

and,

$$\int \bar{\varphi}_n d\mu \rightarrow \int c d\pi.$$

For a fixed  $\epsilon > 0$ , we get

$$\int c d\pi - \epsilon < \int \bar{\varphi}_n d\mu \leq \int c d\pi,$$

for  $n$  large enough. Particularly, for a fixed  $n$  sufficiently large, there exist  $x_n, x'_n \in X$  such that

$$\int c d\pi - \epsilon \leq \bar{\varphi}_n(x_n) \quad \text{and} \quad \bar{\varphi}_n(x'_n) \leq \int c d\pi.$$

Using the fact that  $\bar{\varphi}_n$  is Lipschitz continuous, with constant  $\beta$ , and, denoting by  $D$  the diameter of  $X$ , we conclude that for large  $n$ :

$$\int c d\pi - \epsilon - D\beta \leq \bar{\varphi}_n \leq \int c d\pi + D\beta.$$

So we can apply the Arzela-Ascoli theorem and, finally, we get continuous functions  $\varphi, \psi$  satisfying:

- i)  $\varphi(x) + \psi(\sigma(y)) - \psi(y) \leq c(x, y)$
- ii)  $\int \varphi(x) d\mu = \int c d\pi.$

We know from the first reasoning that we can assume  $\varphi$  and  $\psi$  are Lipschitz continuous functions. This shows II). □

### 3 Generic properties: a unique optimal plan

**Lemma 7.** *Let  $K$  be a compact set in  $\mathbb{R}^2$ , and for each  $r > 0$ , define  $K_r$  as the set of points  $(x, y) \in K$  such that  $x + ry$  is maximal. Then the diameter of  $K_r$  converges to zero when  $r \rightarrow 0$ .*

*Proof.* See [3] page 306 for the proof. □

**Corollary 8.** *With the hypothesis of the above lemma, for each  $\epsilon > 0$ , there exists a  $r_0 > 0$ , such that, for  $r_0 > r > 0$  and  $(x_1, y_1), (x_2, y_2) \in K_r$ , we have:*

$$|y_1 - y_2| < \epsilon.$$

The bellow theorem follows from the same arguments used in proposition 9 in [3].

**Theorem 9.** *Let  $X$  be a compact metric space,  $Y = \{1, \dots, d\}^{\mathbb{N}}$  and  $\mu$  be a probability measure in  $X$ . Let  $C(X, Y)$  be the set of continuous functions from  $X \times Y$  to  $\mathbb{R}_+$  with the uniform norm. The set of functions  $c \in C(X, Y)$  with a unique Optimal Plan in  $\Pi(\mu, \sigma)$  is generic in  $C(X, Y)$ . The same is true for the Banach space  $H(X, Y)$  of the Lipschitz functions with the usual norm.*

*Proof.* On this proof, we are going to consider  $\pi$  an optimal plan if  $\int c d\pi$  is maximal (just consider the change of  $c$  by  $-c$ ).

We start studying the space  $C(X, Y)$ . Given an countable family  $(e_i)_{i \in \mathbb{N}}$  dense in  $C(X, Y)$ , the set of functions in  $C(X, Y)$  with two or more optimal plans coincides with:

$$\bigcup_{m, n \in \mathbb{N}} X_{m, n},$$

where

$$X_{m, n} := \left\{ c \in C(X, Y) : \exists \pi, \chi \text{ optimal plans, } \int e_n d(\pi - \chi) \geq \frac{1}{m} \right\}.$$

Then it's sufficiently to prove that  $X_{m, n}$  is a closed set with empty interior.

**Claim 1:**  $X_{m, n}$  is a closed set.

Indeed, we note that  $C(X, Y)$  is a normed space. Consider  $c_s$  in  $X_{m, n}$  converging to  $c$  (when  $s \rightarrow \infty$ ). Let  $(\pi_s, \chi_s)$  be the optimal ones associated with  $c_s$  in  $X_{m, n}$ . We can suppose, by taking a subsequence, that  $\pi_s \rightarrow \pi$ ,  $\chi_s \rightarrow \chi$ , where  $\pi, \chi$  are probability measures in  $X \times Y$ . So

$$\int e_n d(\pi - \chi) = \lim_{s \rightarrow \infty} \int e_n d(\pi_s - \chi_s) \geq \frac{1}{m}.$$

Clearly  $\pi, \chi \in \Pi(\mu, \sigma)$ . Also, by the above relation, they are different measures. We want to show that the limit function  $c$  is in  $X_{m,n}$ . We only need to prove that  $\pi$  and  $\chi$  are optimal plans to  $c$ . Suppose by contradiction there exists  $\zeta \in \Pi(\mu, \sigma)$  such that  $\int c d\zeta > \int c d\pi + \epsilon$ . So for  $s$  large we have:

$$\int c_s d\zeta > \int c d\zeta - \epsilon/3 > \int c d\pi + 2\epsilon/3 > \int c d\pi_s + \epsilon/3 > \int c_s d\pi_s.$$

This is impossible because  $\pi_s$  is an optimal plan for  $c_s$ . Therefore,  $\chi$  is an optimal plan for  $c$ .

**Claim 2:**  $X_{m,n}$  has empty interior.

Indeed, for a fixed  $c \in X_{m,n}$  we can show that  $c + re_n \notin X_{m,n}$  when  $r > 0$  is sufficiently small. Consider

$$K = \left\{ \left( \int c d\pi, \int e_n d\pi \right) : \pi \in \Pi(\mu, \sigma) \right\}.$$

$K$  is compact and contained in  $\mathbb{R}^2$ . Then, by the Corollary 8, when  $\epsilon = 1/2m$ , there exist a  $r_0$ , such that for  $r_0 > r > 0$  we get: if  $\int (c + r\epsilon_n) d\pi$  and  $\int (c + r\epsilon_n) d\chi$  are maximal (this means  $\pi, \chi$  optimal plans to  $f + r\epsilon_n$ ), then  $|\int (\epsilon_n) d(\pi - \chi)| < \epsilon = 1/2m$ . This show that  $c + r\epsilon_n \notin X_{m,n}$ .

In the space  $H(X, Y)$  we can get similar results. This can be obtained with the same arguments used before together with the following remarks:

- a) A dense enumerable family  $\{e_n\}$  in  $H(X, Y)$  will be a dense sequence in  $C(X, Y)$ . In this way two elements  $\pi, \chi \in \Pi(\mu, \sigma)$  will be different, if and only if,  $\int e_n d(\pi - \chi) \neq 0$ , for some  $e_n$ .
- b) Moreover, the set

$$X_{m,n} := \left\{ c \in H(X, Y) : \exists \pi, \chi \text{ optimal plans, } \int e_n d(\pi - \chi) \geq \frac{1}{m} \right\}$$

will be a closed set by the same arguments used above. We can also show, in a similar way, that it has empty interior, but we note that we need  $c + re_n \in H(X, Y)$ , and  $c + re_n \rightarrow c$  in Lipschitz norm. This is true because we can consider  $e_n \in H(X, Y)$ . □

## 4 Zeta-measures and Transport

When  $\mu, \nu$  have support in a finite number of points, then, the optimal plan  $\pi(\mu, \nu)$  for a cost  $c$  in the Classical Transport Theory can be explicitly obtained by Linear Algebra arguments [27]. Indeed, suppose

$$\mu = a_1 \delta_{x_1} + \dots + a_n \delta_{x_n},$$

and,

$$\nu = b_1\delta_{y_1} + \dots + b_m\delta_{y_m}.$$

Then any transport plan  $\pi$  have support contained in  $\{x_1, \dots, x_n\} \times \{y_1, \dots, y_m\}$ . Denoting by  $\pi_{ij}$  the mass of  $\pi$  in  $(x_i, y_j)$  we have that the variables  $\pi_{ij}$  need satisfies the linear equations:

vertical equations:

$$\begin{aligned} a_1 &= \pi_{11} + \dots + \pi_{m1} \\ &\dots \\ a_n &= \pi_{1n} + \dots + \pi_{mn} \end{aligned}$$

horizontal equations:

$$\begin{aligned} b_1 &= \pi_{11} + \dots + \pi_{1n} \\ &\dots \\ b_m &= \pi_{m1} + \dots + \pi_{mn} \end{aligned}$$

The set of solutions of this equations defines a convex in  $\mathbb{R}^{n \cdot m}$ . The conditions  $\pi_{ij} \geq 0$  will restrict the solutions to a bounded convex set with finite vertexes. So given a cost function  $c : X \times Y \rightarrow [0, \infty)$ , denoting their restricted values by  $c_{ij} := c(x_i, y_j)$ , we have that the optimal plans to  $(\mu, \nu)$  are the point of the above convex set that minimize the linear functional:

$$\sum_{i,j} c_{ij}\pi_{ij}.$$

By convexity arguments there is an optimal point in the vertexes of the underlying convex set. The conclusion is that by Linear Algebra arguments we can find a finite number of points such that at least one of these will be optimal for the integral of  $c$  with the given marginals  $\mu$  and  $\nu$ .

Note that these finite vertexes points are determined before we consider the given cost function.

It is natural in the Ergodic Theory setting to try to approximate a general invariant probability by the ones which posses the simplest behavior: the periodic probabilities. These are the ones that we can make computations more easily.

We note that to minimize the integral of the cost function  $c$  is the same that to maximize the integral of the function  $-c$ . The plan that realizes this optimal integral will the same if we add a constant to  $-c$ . Bellow we consider the problem of finding a transport plan maximizing the integral of a cost  $c$  strictly greater than zero. A transport plan from  $\mu$  to  $\nu$  maximizing the integral of  $c$  will be called a **maximizing plan**. Below we consider a compact metric space  $X$  and  $Y = \{1, \dots, d\}^{\mathbb{N}}$ .

**Definition 10.** For fixed  $\mu \in P(X)$  and a continuous function  $c : X \times Y \rightarrow \mathbb{R}$  we define a probability measure in  $X \times Y$  by the linear functional  $\zeta_{\beta,n} : C(X \times Y) \rightarrow \mathbb{R}$ , such that, for each  $w \in C(X \times Y)$  associate the number:

$$\frac{\sum_{\nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)} \int w d\pi(\mu,\nu)}{\sum_{\nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}},$$

where  $\text{Fix}_n$  denotes the set of invariant measures in  $Y$  supported in a periodic orbit of length  $n$ , and,  $\pi(\mu,\nu)$  denotes a maximizing plan from  $\mu$  to  $\nu$  with cost function  $c$  (we don't impose other conditions on the chosen the plan).

In the case  $\mu$  is supported in a unique point  $x_0$ , we can define the function  $A(y) = c(x_0, y)$ , and this measure can be written as:

$$\frac{\sum_{y \in \text{Fix}_n} e^{\beta \cdot A^n(y) \frac{\bar{w}^n(y)}{n}}}{\sum_{y \in \text{Fix}_n} e^{\beta \cdot A^n(y)}}$$

where  $\bar{w}(y) = w(x_0, y)$  and  $A^n(z) = A(z) + \dots + A(\sigma^{n-1}(z))$ . This kind of measure (also called zeta measure) is considered in Thermodynamical Formalism [24] and they can be used to approximate Gibbs states, and, also the measure that maximizes the integral of  $A$  among the invariant measures (see [21]). Therefore, in some sense, the above defined family of probabilities extend a well known concept used in Ergodic Optimization.

In the case of that  $\mu$  have finite support these zeta-measures can be determined by Linear Algebra arguments like we remarked above.

Remember that  $\Pi(\mu, \sigma)$  is the set of probabilities measures that coincides with  $\mu$  in the first coordinate and is invariant in the second coordinate. The next theorem follows the ideas used in thermodynamical limit when  $\beta, n \rightarrow \infty$  [21].

**Theorem 11.** When  $\beta, n$  goes to infinite, any limit measure  $\pi_\infty$  of convergent subsequence of  $\zeta_{\beta,n}$ , in the weak\* topology, belongs to  $\Pi(\mu, \sigma)$ . Moreover, if  $c > 0$ , then,  $\pi_\infty$  maximizes the integral of  $c$  among the measures in  $\Pi(\mu, \sigma)$ .

*Proof.* We begin by proving that for  $\beta, n$  fixed, the corresponding zeta-measure is in  $\Pi(\mu, \sigma)$ . Let  $w$  a function depending only on  $x$ . Then

$$\begin{aligned} & \frac{\sum_{\nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)} \int w d\pi(\mu,\nu)}{\sum_{\nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}} \\ &= \frac{\sum_{\nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)} \int w d\mu}{\sum_{\nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}} \end{aligned}$$



$$\begin{aligned}
&= \int w d\mu \frac{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}}{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}} \\
&= \int w d\mu.
\end{aligned}$$

Now, consider a fixed function  $w$  depending only of  $y$ . Then, we have:

$$\begin{aligned}
\zeta_{\beta,n}(w \circ \sigma) &= \frac{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)} \int w \circ \sigma d\pi(\mu,\nu)}{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}} \\
&= \frac{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)} \int w \circ \sigma d\nu}{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}} \\
&= \frac{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)} \int w d\nu}{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}} \\
&= \frac{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)} \int w d\pi(\mu,\nu)}{\sum_{\nu \in Fix_n} e^{\beta \cdot n \cdot \int c(x,y) d\pi(\mu,\nu)}} = \zeta_{\beta,n}(w).
\end{aligned}$$

This show that  $\zeta_{\beta,n} \in \Pi(\mu, \sigma)$ . So, when  $\beta, n$  goes to infinite, any limit measure  $\pi_\infty$  of convergent subsequence of  $\zeta_{\beta,n}$  is on  $\Pi(\mu, \sigma)$ .

Suppose  $\zeta_{\beta_j, n_j} \rightarrow \pi_\infty$ , when  $j \rightarrow \infty$ .

Let  $\pi^* \in \Pi(\mu, \sigma)$  maximizing the integral of  $c$ . Let  $\nu^*$  the projection of  $\pi^*$  on the second coordinate  $y$ . Then,  $\nu^*$  is a invariant measure. Let  $\nu_{n_j} \in Fix_{n_j}$  be a subsequence converging to  $\nu^*$  in the weak\* topology. If  $\pi_{n_j}$  is a maximizing plan from  $\mu$  to  $\nu_{n_j}$ , then, there exist a subsequence  $\pi_{n_i}$  converging to a maximizing plan  $\pi$  from  $\mu$  to  $\nu^*$  ([28] page 77). It is easy to see that

$$\int c d\pi = \int c d\pi^*,$$

and, therefore,  $\pi$  is maximal. In other words, it maximizes the integral of  $c$  among the measures in  $\Pi(\mu, \sigma)$ . We denote this integral by  $I(c)$ . We want to show that  $\pi_\infty(c) \geq I(c)$ , where the subsequence  $\zeta_{\beta_i, n_i}$  converges to  $\pi_\infty$  in the weak\* topology. From the above arguments we know that:

Given  $\varepsilon > 0$ , for sufficiently large  $i$  there exist  $\nu \in Fix_{n_i}$  such that:

$$\int c d\pi(\mu, \nu) > I(c) - \varepsilon.$$

Take  $\varepsilon > 0$ , such that  $(I(c) - \varepsilon) > 0$ , and define:

$$A_{n_i}(\varepsilon) = \{\nu \in \text{Fix}_{n_i} : \int c d\pi(\mu, \nu) \leq I(c) - \varepsilon\}$$

$$B_{n_i}(\varepsilon) = \{\nu \in \text{Fix}_{n_i} : \int c d\pi(\mu, \nu) > I(c) - \varepsilon\}.$$

Then, we have:

$$\begin{aligned} \sum_{\nu \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} &\leq \sum_{\nu \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon)} \\ &\leq e^{n_i \log(d) + \beta_i \cdot n_i \cdot (I(c) - \varepsilon)}, \end{aligned}$$

and

$$\sum_{\nu \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu) \leq e^{n_i \log(d) + \beta_i \cdot n_i \cdot (I(c) - \varepsilon)} (I(c) - \varepsilon).$$

By other hand, if  $n_i$  is sufficiently large,  $B_{n_i}(\varepsilon/2)$  is non empty. It follows that

$$\begin{aligned} \sum_{\nu \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} &\geq \sum_{\nu \in B_{n_i}(\varepsilon/2)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} \\ &\geq e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon/2)}, \end{aligned}$$

and,

$$\sum_{\nu \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu) \geq e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon/2)} (I(c) - \varepsilon/2).$$

Then, we get

$$\begin{aligned} 0 &\leq \lim_{i \rightarrow \infty} \frac{\sum_{\nu \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)}}{\sum_{\nu \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)}} \\ &\leq \lim_{i \rightarrow \infty} \frac{e^{n_i \log(d) + \beta_i \cdot n_i \cdot (I(c) - \varepsilon)}}{e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon/2)}} = \lim_{i \rightarrow \infty} e^{n_i \log(d) - \beta_i \cdot n_i \cdot \varepsilon/2} = 0. \end{aligned}$$

Moreover,

$$\begin{aligned} 0 &\leq \lim_{i \rightarrow \infty} \frac{\sum_{\nu \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)}{\sum_{\nu \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)} \leq \\ &\lim_{i \rightarrow \infty} \frac{e^{n_i \log(d) + \beta_i \cdot n_i \cdot (I(c) - \varepsilon)} (I(c) - \varepsilon)}{e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon/2)} (I(c) - \varepsilon/2)} = \end{aligned}$$

$$\lim_{i \rightarrow \infty} e^{n_i \log(d) - \beta_i \cdot n_i \cdot \varepsilon / 2} \frac{I(c) - \varepsilon}{I(c) - \varepsilon / 2} = 0.$$

Finally,

$$\begin{aligned} & \liminf_{i \rightarrow \infty} \frac{\sum_{\nu \in \text{Fix}_{n_i}} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)}{\sum_{\nu \in \text{Fix}_{n_i}} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)}} \\ &= \liminf_{i \rightarrow \infty} \frac{\sum_{\nu \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)}{\sum_{\nu \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)}} \\ &\geq \liminf_{i \rightarrow \infty} \frac{\sum_{\nu \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} (I(c) - \varepsilon)}{\sum_{\nu \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)}} \\ &\geq I(c) - \varepsilon. \end{aligned}$$

Taking  $\varepsilon \rightarrow 0$ , we get:

$$\liminf_{i \rightarrow \infty} \frac{\sum_{\nu \in \text{Fix}_{n_i}} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)}{\sum_{\nu \in \text{Fix}_{n_i}} e^{\beta_i \cdot n_i \cdot \int c(x,y) d\pi(\mu, \nu)}} \geq I(c).$$

Using the fact that  $\zeta_{\beta_i, n_i} \rightarrow \pi_\infty$  we conclude that  $\pi_\infty(c) \geq I(c)$ .  $\square$

## 5 Two variable invariant probabilities and other cases

We start this section by proving Theorem 4. The proof follows basically the same kind of ideas that were used in Theorem 1.

*Proof.* Define

$$E = C(X \times Y) \times M(X \times Y)$$

where  $M(X \times Y)$  is the set of bounded linear functionals from  $C(X \times Y)$  to  $\mathbb{R}$  with the norm given by the total variation.

Let  $\Theta : E \rightarrow \mathbb{R} \cup \{+\infty\}$  given by

$$\Theta(u, \nu) = \begin{cases} 0, & \text{if } u(x, y) \geq -c(x, y), \forall (x, y) \in X \times Y, \text{ and } \|\nu\| \leq 2 \\ +\infty, & \text{in the other case,} \end{cases}$$

Note that  $\Theta$  is a convex function.

Define  $\Xi : E \longrightarrow \mathbb{R} \cup \{+\infty\}$  by

$$\Xi(u, \nu) = \begin{cases} \alpha, & \text{if } u(x, y) = \alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)), \\ & \text{with } (\varphi, \psi) \in C(X) \times C(Y), \text{ and } \nu \in \Pi(T) \\ +\infty, & \text{in the other case.} \end{cases}$$

We remark that  $\Xi$  is a well defined convex function.

If  $\nu \in \Pi(T)$ , then  $(1, \nu) \in D(\Theta) \cap D(\Xi)$  and  $\Theta$  is continuous in  $(1, \nu)$ .

Note that:

$$\begin{aligned} & \inf_{(u, \nu) \in E} [\Theta(u, \nu) + \Xi(u, \nu)] \\ &= \inf \left\{ \alpha : \begin{array}{l} \exists (\varphi, \psi) \in C(X) \times C(Y), \\ \alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)) \geq -c(x, y) \end{array} \right\} \\ &= \inf \left\{ -\alpha : \begin{array}{l} \exists (\varphi, \psi) \in C(X) \times C(Y), \\ \alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)) \leq c(x, y) \end{array} \right\} \\ &= - \sup \left\{ \alpha : \begin{array}{l} \exists (\varphi, \psi) \in C(X) \times C(Y), \\ \alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)) \leq c(x, y) \end{array} \right\}. \end{aligned}$$

So the left size of (4) is:

$$- \sup \left\{ \alpha : \begin{array}{l} \exists (\varphi, \psi) \in C(X) \times C(Y), \\ \alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)) \leq c(x, y) \end{array} \right\}. \quad (7)$$

Now we will compute the Legendre-Fenchel transform of  $\Theta$  and  $\Xi$ . For each  $(\pi, g) \in E^*$ :

$$\begin{aligned} \Theta^*((-\pi, -g)) &= \sup_{(u, \nu) \in E} \{ \langle (-\pi, -g), (u, \nu) \rangle - \Theta(u, \nu) \} \\ &= \sup_{(u, \nu) \in E} \{ -\pi(u(x, y)) - g(\nu) : -u(x, y) \leq c(x, y), \|\nu\| \leq 2 \} \\ &= \sup_{(u, \nu) \in E} \{ \pi(u(x, y)) - g(\nu) : u(x, y) \leq c(x, y), \|\nu\| \leq 2 \}. \end{aligned}$$

Following [27] note that if  $\pi \notin M^+(X \times Y)$ , then there exists a function  $v \leq 0$  in  $C(X \times Y)$ , such that,  $\pi(v) > 0$ , so taking  $u = \lambda v$  (remember that  $c \geq 0$ ), when  $\lambda \rightarrow +\infty$ , we have that  $\sup_{u \in E} \{ \pi(u) : u(x, y) \leq c(x, y) \} = +\infty$ .

Moreover, if  $\pi \in M^+(X \times Y)$ , using the fact that  $c \in C(X \times Y)$ , then we have that the maximum of  $\pi(u)$  is given by  $\pi(c)$ .

Therefore,

$$\Theta^*((-\pi, -g)) = \begin{cases} \pi(c) + \sup_{\|\nu\| \leq 2} -g(\nu), & \text{if } \pi \in M^+(X \times Y) \\ +\infty, & \text{in the other case.} \end{cases} \quad (8)$$

Now we analyze  $\Xi^*$ :

$$\begin{aligned} \Xi^*(\pi, g) &= \sup_{(u, \nu) \in E} \{ \langle (\pi, g), (u, \nu) \rangle - \Xi(u, \nu) \} \\ &= \sup_{(u, \nu) \in E} \left\{ \begin{array}{l} \pi(u(x, y)) - \alpha + g(\nu) : \\ u(x, y) = \alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)), \\ \text{with } (\varphi, \psi) \in C(X) \times C(Y) \text{ and } \nu \in \Pi(T) \end{array} \right\} \\ &= \sup_{\alpha, (\varphi, \psi) \in C(X) \times C(Y), \nu \in \Pi(T)} \pi(\alpha) - \alpha + \pi(\varphi - \varphi \circ T_1) + \pi(\psi - \psi \circ T_2) + g(\nu). \end{aligned}$$

If  $\pi(\varphi(x) - \varphi(T_1(x, y))) \neq 0$  (we can suppose greater than zero) for some  $\varphi$ , by taking  $\lambda \cdot \varphi$  and  $\lambda \rightarrow \infty$ , the supremum will be equal to  $+\infty$ . Analogously, if  $\pi(\psi(y) - \psi(T_2(x, y))) \neq 0$ , the supremum will be  $+\infty$ . If  $\pi(1) \neq 1$  (we can suppose greater than one), then taking  $\alpha \rightarrow \infty$ , the supremum will be  $+\infty$ .

Define

$$\Pi^*(T) = \left\{ \pi \in M(X \times Y) : \begin{array}{l} \pi(1) = 1, \pi(\varphi - \varphi \circ T_1) = \pi(\psi - \psi \circ T_2) = 0, \\ \forall (\varphi, \psi) \in C(X) \times C(Y) \end{array} \right\}.$$

Therefore,

$$\Xi^*(\pi, g) = \begin{cases} \sup\{g(\nu) : \nu \in \Pi(T)\}, & \text{if } \pi \in \Pi^*(T), \\ +\infty, & \text{in the other case.} \end{cases} \quad (9)$$

We know that the left size of (4) is given by (7). By (8) and (9), the right size of (4) will be:

$$\begin{aligned} &\sup_{(\pi, g) \in E^*} \left\{ - \left( \begin{array}{l} (\pi(c) + \sup\{-g(\nu) : \|\nu\| \leq 2\}) + \sup\{g(\nu) : \nu \in \Pi(T)\}, \\ \text{if } \pi \in M^+(X \times Y) \cap \Pi^*(T) \\ +\infty, \text{ in the other case} \end{array} \right) \right\} \\ &= \sup_{(\pi, g) \in E^*} \left\{ \left( \begin{array}{l} (-\pi(c) + \inf\{g(\nu) : \|\nu\| \leq 2\}) - \sup\{g(\nu) : \nu \in \Pi(T)\}, \\ \text{if } \pi \in M^+(X \times Y) \cap \Pi^*(T) \\ -\infty, \text{ in the other case} \end{array} \right) \right\} \end{aligned}$$

$$= \sup\{-\pi(c), \pi \in M^+(X \times Y) \cap \Pi^*(T)\},$$

where the last equality is obtained taking  $g = 0$ .

We remark that if  $\pi \in M^+(X \times Y) \cap \Pi^*(T)$  then:

$\pi(1) = 1$ , (by definition of  $\Pi^*(T)$ )

$\pi(u) \geq 0$  if  $u \geq 0$ ,

$\pi$  is linear in  $C(X \times Y)$ ,

Therefore, we have  $\pi \in P(X \times Y)$ , and, by definition of  $\Pi^*(T)$ , we get that  $\pi \in \Pi(T)$ . The conclusion is that  $M^+(X \times Y) \cap \Pi^*(T) = \Pi(T)$ . So the right side of (4) will be:

$$\sup\{-\pi(c), \pi \in \Pi(T)\} = -\inf\{\pi(c), \pi \in \Pi(T)\}.$$

Therefore, we conclude from (4) that:

$$\begin{aligned} -\sup \left\{ \alpha : \begin{array}{l} \exists(\varphi, \psi) \in C(X) \times C(Y), \\ \alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)) \leq c(x, y) \end{array} \right\} \\ = -\inf\{\pi(c), \pi \in \Pi(T)\}, \end{aligned}$$

or, in another form that

$$\begin{aligned} \sup \left\{ \alpha : \begin{array}{l} \exists(\varphi, \psi) \in C(X) \times C(Y), \\ \alpha + \varphi(x) - \varphi(T_1(x, y)) + \psi(y) - \psi(T_2(x, y)) \leq c(x, y) \end{array} \right\} \\ = \inf\{\pi(c), \pi \in \Pi(T)\}. \end{aligned}$$

□

**Proposition 12.** *Suppose  $c$  is continuous. Denote  $\alpha = \inf_{\pi \in \Pi(T)} \int c(x, y) d\pi$ .*

*If there exist  $\varphi \in C(X)$  and  $\psi \in C(Y)$  satisfying*

$$\alpha + \varphi(x) - \varphi(T_1(x)) + \psi(y) - \psi(T_2(y)) \leq c(x, y) \quad \forall(x, y) \in X \times Y, \quad (10)$$

*then,*

$$\inf\{c(x, y) + \dots + c(T^{n-1}(x, y)) - n\alpha : n \geq 1 \text{ and } (x, y) \in X \times Y\} > -\infty.$$

*Proof.* Suppose by contradiction that

$$\inf\{c(x, y) + \dots + c(T^{n-1}(x, y)) - n\alpha : n \geq 1 \text{ and } (x, y) \in X \times Y\} = -\infty.$$

Also suppose that there exists  $\varphi$  and  $\psi$  satisfying (10). Then we have

$$\inf\{\varphi(x) - \varphi(T_1^n(x, y)) + \psi(y) - \psi(T_2^n(x, y)) : n \geq 0 \text{ and } (x, y) \in X \times Y\} = -\infty.$$

This is impossible because  $X$  and  $Y$  are compact sets and  $\varphi$  and  $\psi$  are continuous functions. □

**Proposition 13.** *Suppose  $X = Y = \{0, 1\}^{\mathbb{N}}$ ,  $T_1 = T_2 = \sigma$ , and, that  $c$  is a Lipschitz function, then there exists  $\varphi(x)$  and  $\psi(y)$  Lipschitz continuous such that*

$$\alpha + \varphi(\sigma(x)) - \varphi(x) + \psi(\sigma(y)) - \psi(y) \leq c(x, y).$$

*Proof.* Denote by  $\beta$  a Lipschitz constant for  $c$ .

By definition of  $\alpha$  there exist a increasing sequence  $\alpha_n \rightarrow \alpha$  and continuous functions  $\varphi_n, \psi_n$  such that:

$$\alpha_n + \varphi_n(\sigma(x)) - \varphi_n(x) + \psi_n(\sigma(y)) - \psi_n(y) \leq c(x, y).$$

From this relation we have that if  $\sigma^m(z) = x$  and  $y_0, \dots, y_{m-1}$  belongs to  $Y$ , then

$$\varphi_n(x) - \varphi_n(z) \leq \sum_{i=0}^{m-1} c(\sigma^i z, y_i) + \psi_n(y_i) - \psi_n(\sigma(y_i)) - \alpha_n.$$

Therefore,

$$\inf \left\{ \sum_{i=0}^{m-1} c(\sigma^i z, y_i) + \psi_n(y_i) - \psi_n(\sigma(y_i)) - \alpha_n : m \geq 0, \sigma^m(z) = x, y_i \in Y \right\} > -\infty.$$

Denote by  $\bar{\varphi}_n(x)$  this infimum. Note that this function satisfies:

$$\alpha_n + \bar{\varphi}_n(\sigma(x)) - \bar{\varphi}_n(x) + \psi_n(\sigma(y)) - \psi_n(y) \leq c(x, y).$$

It is easy to see that  $\bar{\varphi}_n$  is Lipschitz continuous with the same Lipschitz constant  $\beta$  of  $c$ . Using this last inequality and the same arguments used before (applied to  $\psi$ ) we can construct a Lipschitz continuous function  $\bar{\psi}_n$  with the same Lipschitz constant  $\beta$  satisfying:

$$\alpha_n + \bar{\varphi}_n(\sigma(x)) - \bar{\varphi}_n(x) + \bar{\psi}_n(\sigma(y)) - \bar{\psi}_n(y) \leq c(x, y).$$

Note that we can add constants on  $\bar{\varphi}_n$  and  $\bar{\psi}_n$  and the conclusions are the same. Then, we get: there exists Lipschitz continuous functions  $\bar{\varphi}_n$  and  $\bar{\psi}_n$  with Lipschitz constant  $\beta$ , which bounded bellow and above, respectively, by 0 and  $\beta$ , such that:

$$\alpha_n + \bar{\varphi}_n(\sigma(x)) - \bar{\varphi}_n(x) + \bar{\psi}_n(\sigma(y)) - \bar{\psi}_n(y) \leq c(x, y).$$

Now using Arzela-Ascoli theorem we obtain continuous functions  $\varphi$  and  $\psi$  satisfying:

$$\alpha + \varphi(\sigma(x)) - \varphi(x) + \psi(\sigma(y)) - \psi(y) \leq c(x, y).$$

Applying the same reasoning of the previous arguments we can also construct Lipschitz functions satisfying this inequality.  $\square$

**Proposition 14.** *Let  $C(X, Y)$  be the set of continuous functions from  $X \times Y$  to  $\mathbb{R}_+$  with the uniform norm. The set of functions  $c \in C(X, Y)$  with a unique Optimal Plan in  $\Pi(T)$  is generic in  $C(X, Y)$ . The same is true for the Banach space  $H(X, Y)$  of the Lipschitz functions with the usual norm.*

*Proof.* The result follows from adapting the proof of Theorem 9.  $\square$

## 6 Zeta-measures for the second class of problems

In this section we suppose  $X = Y = \{0, 1\}^{\mathbb{N}}$  and  $T_1 = T_2 = \sigma$  is the shift. On this case we have  $\Pi(T) = \Pi(\sigma)$  is the set of probabilities  $\pi$  in  $X \times Y$  such that project on  $\sigma$ -invariant measures in  $X$  and  $Y$ .

Bellow we consider the problem of finding a transport plan in  $\Pi(\sigma)$  maximizing the integral of a cost  $c$  strictly greater than zero. A transport plan maximizing this integral will be called a **maximizing plan**. By changing the signal of the cost we can get from this the analysis of usual minimization problem of Transport Theory.

**Definition 15.** *For a fixed cost  $c$  we define a probability measure in  $X \times Y$  by the linear functional  $\zeta_{\beta, n} : C(X \times Y) \rightarrow \mathbb{R}$ , such that, to each  $w \in C(X \times Y)$  we associate the number:*

$$\frac{\sum_{\nu, \mu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)} \int w d\pi(\mu, \nu)}{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)}},$$

where  $\text{Fix}_n$  denote the set of invariant measures in  $X = Y$  supported in a periodic orbit of length  $n$ , and,  $\pi(\mu, \nu)$  denote a maximizing plan from  $\mu$  to  $\nu$  with cost function  $c$  (we don't impose other conditions on the plan).

This zeta-measures can be determined by Linear Algebra arguments. Indeed, note that if  $\mu, \nu \in \text{Fix}_n$ , then the plan  $\pi(\mu, \nu)$  can be determined by the study of certain permutations (see page 5 in [27]).

**Theorem 16.** *When  $\beta, n$  goes to infinite, any limit measure  $\pi_\infty$  of convergent subsequence of  $\zeta_{\beta, n}$ , in the weak\* topology, is on  $\Pi(\sigma)$ . Also, if  $c > 0$ , then,  $\pi_\infty$  maximizes the integral of  $c$  between the measures in  $\Pi(\sigma)$ .*

*Proof.* We begin by proving that for  $\beta, n$  fixed, the zeta-measure is in  $\Pi(\sigma)$ . Let  $w$  be a function depending only on  $y$ . Then

$$\zeta_{\beta, n}(w \circ \sigma) = \frac{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)} \int w \circ \sigma d\pi(\mu, \nu)}{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)}}$$



$$\begin{aligned}
&= \frac{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)} \int w \circ \sigma d\nu}{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)}} \\
&= \frac{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)} \int w d\nu}{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)}} \\
&= \frac{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)} \int w d\pi(\mu, \nu)}{\sum_{\mu, \nu \in \text{Fix}_n} e^{\beta \cdot n \cdot \int c(x, y) d\pi(\mu, \nu)}} = \zeta_{\beta, n}(w).
\end{aligned}$$

If  $w$  depends only on  $x$  the argument is similar. This shows that  $\zeta_{\beta, n} \in \Pi(\sigma)$ . Then, when  $\beta, n$  goes to infinite, via a convergent subsequence, any limit measure  $\pi_\infty$  of  $\zeta_{\beta, n}$ , in the weak\* topology, will be on  $\Pi(\sigma)$ .

Suppose  $\zeta_{\beta_j, n_j} \rightarrow \pi_\infty$ , when  $j \rightarrow \infty$ .

Let  $\pi^* \in \Pi(\sigma)$  be a probability maximizing the integral of  $c$ . Let  $\mu^*, \nu^*$ , respectively, the projection of  $\pi^*$  in the first and second coordinates. Then,  $\mu^*, \nu^*$  are invariant measures. Let  $\mu_{n_j}, \nu_{n_j} \in \text{Fix}_{n_j}$  subsequences converging to  $\mu^*, \nu^*$  in the weak\* topology. If  $\pi_{n_j}$  is a maximizing plan from  $\mu_{n_j}$  to  $\nu_{n_j}$ , then, there exist a subsequence  $\pi_{n_i}$  converging to a maximizing plan  $\pi$  from  $\mu^*$  to  $\nu^*$  ([28] page 77). Clearly

$$\int c d\pi = \int c d\pi^*,$$

and, therefore,  $\pi$  is maximal. This means that  $\pi$  maximizes the integral of  $c$  among the measures in  $\Pi(\sigma)$ . We denote this integral by  $I(c)$ . We want to show that  $\pi_\infty(c) \geq I(c)$ . We note that the subsequence  $\zeta_{\beta_i, n_i}$  converges to  $\pi_\infty$  in the weak\* topology. From the above arguments we get:

given  $\varepsilon > 0$ , for sufficiently large  $i$  there exists  $\mu, \nu \in \text{Fix}_{n_i}$ , such that,

$$\int c d\pi(\mu, \nu) > I(c) - \varepsilon.$$

Consider a fixed  $\varepsilon > 0$ , such that,  $(I(c) - \varepsilon) > 0$ , and, define:

$$A_{n_i}(\varepsilon) = \{(\mu, \nu) \in \text{Fix}_{n_i} : \int c d\pi(\mu, \nu) \leq I(c) - \varepsilon\},$$

$$B_{n_i}(\varepsilon) = \{(\mu, \nu) \in \text{Fix}_{n_i} : \int c d\pi(\mu, \nu) > I(c) - \varepsilon\}.$$

Then, we have that

$$\sum_{(\mu, \nu) \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \leq \sum_{(\mu, \nu) \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon)}$$

$$\leq e^{2n_i \log(2) + \beta_i \cdot n_i \cdot (I(c) - \varepsilon)},$$

and,

$$\sum_{(\mu, \nu) \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu) \leq e^{2n_i \log(2) + \beta_i \cdot n_i \cdot (I(c) - \varepsilon)} (I(c) - \varepsilon).$$

By other hand, if  $n_i$  is sufficiently large,  $B_{n_i}(\varepsilon/2)$  is not empty. Moreover,

$$\begin{aligned} \sum_{(\mu, \nu) \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} &\geq \sum_{(\mu, \nu) \in B_{n_i}(\varepsilon/2)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \\ &\geq e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon/2)}, \end{aligned}$$

and,

$$\sum_{(\mu, \nu) \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu) \geq e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon/2)} (I(c) - \varepsilon/2).$$

Then,

$$\begin{aligned} 0 &\leq \lim_{i \rightarrow \infty} \frac{\sum_{(\mu, \nu) \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)}}{\sum_{(\mu, \nu) \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)}} \\ &\leq \lim_{i \rightarrow \infty} \frac{e^{2n_i \log(2) + \beta_i \cdot n_i \cdot (I(c) - \varepsilon)}}{e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon/2)}} = \lim_{i \rightarrow \infty} e^{2n_i \log(2) - \beta_i \cdot n_i \cdot \varepsilon/2} = 0, \end{aligned}$$

and,

$$\begin{aligned} 0 &\leq \lim_{i \rightarrow \infty} \frac{\sum_{(\mu, \nu) \in A_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)}{\sum_{(\mu, \nu) \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)} \\ &\leq \lim_{i \rightarrow \infty} \frac{e^{2n_i \log(2) + \beta_i \cdot n_i \cdot (I(c) - \varepsilon)} (I(c) - \varepsilon)}{e^{\beta_i \cdot n_i \cdot (I(c) - \varepsilon/2)} (I(c) - \varepsilon/2)} \\ &= \lim_{i \rightarrow \infty} e^{2n_i \log(2) - \beta_i \cdot n_i \cdot \varepsilon/2} \frac{I(c) - \varepsilon}{I(c) - \varepsilon/2} = 0. \end{aligned}$$

Therefore,

$$\begin{aligned} &\liminf_{i \rightarrow \infty} \frac{\sum_{(\mu, \nu) \in Fix_{n_i}} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)}{\sum_{(\mu, \nu) \in Fix_{n_i}} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)}} \\ &= \liminf_{i \rightarrow \infty} \frac{\sum_{(\mu, \nu) \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)}{\sum_{(\mu, \nu) \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)}} \end{aligned}$$

$$\begin{aligned} &\geq \liminf_{i \rightarrow \infty} \frac{\sum_{(\mu, \nu) \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} (I(c) - \varepsilon)}{\sum_{(\mu, \nu) \in B_{n_i}(\varepsilon)} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)}} \\ &\geq I(c) - \varepsilon. \end{aligned}$$

Taking  $\varepsilon \rightarrow 0$  we get

$$\liminf_{i \rightarrow \infty} \frac{\sum_{(\mu, \nu) \in Fix_{n_i}} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)} \int c d\pi(\mu, \nu)}{\sum_{(\mu, \nu) \in Fix_{n_i}} e^{\beta_i \cdot n_i \cdot \int c(x, y) d\pi(\mu, \nu)}} \geq I(c).$$

Then, using the fact that  $\zeta_{\beta_i, n_i}$  converges to  $\pi_\infty$ , we finally get  $\pi_\infty(c) \geq I(c)$ .  $\square$

## References

- [1] R. Bissacot and E. Garibaldi, Weak KAM methods and ergodic optimal problems for countable Markov shifts, *Bull. Braz. Math. Soc.* 41, N.3, 321-338, (2010).
- [2] T. Bousch, Le poisson n'a pas d'arêtes, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, **36**, 489-508, (2000).
- [3] T. Bousch, La condition de Walters, *Ann. Sci. ENS*, 34, pp. 287–311, (2001).
- [4] T. Bousch and O. Jenkinson, Cohomology classes of dynamically non-negative  $C^k$  functions, *Inventiones Mathematicae*, 148 (2002), 207–217
- [5] X. Bressaud and A. Quas, Rate of approximation of minimizing measures, *Nonlinearity*, 20, no. 4, 845-853, (2007).
- [6] G. Contreras, A. O. Lopes and Ph. Thieullen, Lyapunov minimizing measures for expanding maps of the circle, *Ergodic Theory and Dynamical Systems*, **21**, (2001), pp. 1379–1409.
- [7] D. Collier and I. D. Morris, Approximating the maximum ergodic average via periodic orbits. *Ergodic Theory Dynam. Systems*, 28, no. 4, 1081-1090, (2008)
- [8] G. Contreras, A. O. Lopes and E. R. Oliveira, Ergodic Transport Theory, periodic maximizing probabilities and the twist condition, preprint (2011), Arxiv

- [9] J.-P. Conze & Y. Guivarc, *Croissance des sommes ergodiques*, manuscript, circa 1993.
- [10] E. Garibaldi and Ph. Thieullen, Minimizing orbits in the discrete Aubry-Mather model, *Nonlinearity* 24 (2011), no. 2, 563-611
- [11] E. Garibaldi and A. O. Lopes, On the Aubry-Mather Theory for Symbolic Dynamics, *Erg Theo and Dyn Systems*, Vol 28 , Issue 3, 791-815 (2008)
- [12] N. Gigli, Introduction to Optimal Transport: Theory and Applications, XXVIII Coloquio Brasileiro de Matematica, 2011, IMPA, Rio de Janeiro
- [13] E. Garibaldi and A. O. Lopes, The effective potential and transshipment in Thermodynamic Formalism at temperature zero, to appear in *Stoch. Dyn.*
- [14] B. R. Hunt and G. C. Yuan, Optimal orbits of hyperbolic systems. *Nonlinearity* 12, (1999), 1207-1224.
- [15] O. Jenkinson, Ergodic optimization, *Discrete and Continuous Dynamical Systems, Series A* 15 (2006), 197-224.
- [16] B. Kloeckner, Optimal transport and dynamics of expanding circle maps acting on measures, preprint (2010), Arxiv
- [17] R. Leplaideur, A dynamical proof for the convergence of Gibbs measures at temperature zero. *Nonlinearity* 18, no. 6, (2005), 2847-2880.
- [18] A. O. Lopes, E. R. Oliveira and Ph. Thieullen, The dual potential, the involution kernel and transport in ergodic optimization, *preprint*, (2008).
- [19] A. O. Lopes, E. R. Oliveira and D. Smania, Ergodic Transport Theory and Piecewise Analytic Subactions for Analytic Dynamics. preprint (2011)
- [20] A. O. Lopes and E. R. Oliveira, On the thin boundary of the fat attractor, preprint (2012)
- [21] A. O. Lopes and J. Mengue, Zeta measures and Thermodynamic Formalism for temperature zero, *Bulletin of the Brazilian Mathematical Society* 41 (3), 449-480, (2010)
- [22] R. Mane, *Ergodic Theory and Differentiable Dynamics*, Springer Verlag (1987)

- [23] I. D. Morris, A sufficient condition for the subordination principle in ergodic optimization, *Bull. Lond. Math. Soc.* **39**, no. 2, (2007), 214-220.
- [24] W. Parry and M. Pollicott, Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque* N. **187-188** (1990).
- [25] S. T. Rachev and L. Rüschendorf, *Mass transportation problems – Volume I: Theory, Volume II: Applications*, Springer-Verlag, New York, 1998.
- [26] C. Robinson, *Dynamical Systems*, CRC Press, (1995)
- [27] C. Villani, *Topics in optimal transportation*, AMS, Providence, (2003).
- [28] C. Villani, *Optimal transport: old and new*, Springer-Verlag, Berlin, (2009).