# STATIONARY OPTIMAL TRANSPORT PLANS AND THE THERMODYNAMIC FORMALISM

by

Shengwen Guo

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2025

Approved by:

Dr. Kevin McGoff

Dr. Michael Grabchak

Dr. Xingjie Li

Dr. Russell Keanini

ABSTRACT

SHENGWEN GUO. Stationary Optimal Transport Plans and The Thermodynamic Formalism. (Under the direction of DR. KEVIN MCGOFF)

Optimal Transport (OT) and Thermodynamic Formalism are two famous linear optimization problems. In optimal transport problem, researchers are curious about the minimization of transportation cost and in thermodynamic formalism, problems are centered on the minimization of 'free energy' in thermodynamic physics system. In this Ph.D. project, we consider constrained versions of these two problems. That is, given $Z \subset \mathcal{C}(X)$ a closed subset and denote by $\mathcal{M}_Z(X,T)$ the set of invariant measures which equals 0 on $Z$, when probability measure $\mu$ ranges over $\mathcal{M}_Z(X,T)$, we explored the properties of optimal plan such as existence, convexity and ergodicity in the framework of optimal transport and thermodynamic formalism respectively. In addition, other topics including uniqueness of optimal plan, Lagrangian approach to optimization, optimization as zero temperature, realization and duality problem also have been studied.

In the first two chapters, project background, basic settings, and related references are introduced. Generally speaking, we are interested in the properties of optimal plans for linear optimization problems in the framework of optimal transport and thermodynamic formalism. In Chapter 3, we talk about the existence and characterization of optimal plans. When $\mathcal{M}_Z(X,T)$ is nonempty and satisfies the 'property (E)' which ensures that the extreme points in $\mathcal{M}_Z(X,T)$ are ergodic, optimal plans have some nice properties.

From Chapters 4 to 8, we studied different problems related to optimal plans. Chapter 4 studies the uniqueness property of optimal plan. In the framework of optimal transport, we have the 'generic' uniqueness property; and in the framework of thermodynamic formalism with entropy term, there is a unique optimal solution if

marginal given distributions are Bernoulli. Chapter 5 explores 'Lagrangian approach' to optimization. Such result is based on the famous Isreal[1]'s theorem. It is named as 'Lagrangian approach' because a restricted linear optimization problem can be transformed into an unrestricted problem through this approach. In Chapter 6 we studied optimization as zero temperature. In thermodynamic formalism, if potential function $\phi$ is replaced to $t\phi$ and then let $t \to \infty$, what is the behavior of optimal solution $\mu_{t\phi}$? This problem has a background in thermodynamic physics, $t$ is called an 'inverse temperature' of a system. Chapter 7 explores 'realization' problem: if a subset of measures $\mathcal{E}$ in $\mathcal{M}_Z(X,T)$ is given, can we fund a function $\phi$ such that the set of optimal plans is exactly $\mathrm{co}(\mathcal{E})$ or $\overline{\mathrm{co}(\mathcal{E})}$? We have results in both types of optimization problems. And Chapter 8 talks about duality problems. For optimization problems, its duality is always good to study. We have both Kantorovich duality and Fenchel duality with respect to the linear optimization problem in the framework of optimal transport and thermodynamic formalism.

Finally, in Chapter 9, we list several open problems. These problems are interesting and valuable enough to be explored in the future.

# DEDICATION

To my beloved Siyuan Lyu - thank you for being by my side through these years.

To my grandmothers, Jianping Lyu and Fengjuan Sun - your unconditional love made my childhood a happy and cherished time.

To my parents, Xiaozhong Yang and Wei Guo - your love and support have lifted me up and made this journey possible.

## ACKNOWLEDGEMENTS

In completing my Ph.D. project, I have received generous support from the Department of Mathematics and Statistics and from UNC Charlotte as a whole.

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Kevin McGoff. Over the past four years, his guidance has been indispensable. Without his continuous support, encouragement, and expert advice, I would not have been able to complete this project or earn my degree. I have been deeply inspired by his passion for mathematics, his rigor in research, and his patience and dedication to his students.

I would also like to sincerely thank my committee members - Dr. Michael Grabchak, Dr. Xingjie Li, and Dr. Russell Keanini - for their valuable feedback and insightful suggestions. I have learned a great deal from their perspectives and their commitment to research excellence.

Additionally, I am grateful to Dr. Shaozhong Deng and Dr. Shaoyu Li for their kind support throughout my time in Charlotte, their mentoring in my teaching, and their willingness to provide references for me.

Finally, I would like to acknowledge the financial support I received through the Graduate Assistant Support Plan (GASP) and teaching assistantship, provided by the UNC Charlotte Graduate School and the Department of Mathematics and Statistics.

# TABLE OF CONTENTS


CHAPTER 1: INTRODUCTION 1

1.1. Preliminaries 1

1.1.1. Optimal Transport 1

1.1.2. Measure-Preserving System and Invariant Measures 2

1.1.3. Ergodicity 3

1.1.4. Joinings 3

1.1.5. Measure-Theoretic Entropy 4

1.1.6. General Thermodynamic Formalism 5

1.1.7. Shift of Finite Type and Gibbs measure 6

1.2. Settings and Problems 8

1.2.1. Settings 8

1.2.2. Problems 9

1.3. Statement of main results 10

CHAPTER 2: RELATED WORK 15

2.1. Related work for problem (I) 15

2.2. Related Work for problem (II) 17

CHAPTER 3: EXISTENCE AND CHARACTERIZATION OF OPTIMAL PLANS 20

3.1. Properties of $\mathcal{M}_Z(X, T)$ 20

3.2. Existence and characterization of optimal plans for problem (I) and (II) 25

CHAPTER 4: UNIQUENESS OF OPTIMAL PLANS 31

4.1. Uniqueness results of problem (I) 31

4.2. Uniqueness results of problem (II) 34

CHAPTER 5: LAGRANGIAN APPROACHES TO OPTIMIZATION 38

5.1. Pressure function and tangent functional 38

5.2. Lagrangian approach 42

5.3. Result in uniqueness of relative equilibrium states 55

CHAPTER 6: OPTIMIZATION AS ZERO TEMPERATURE 63

6.1. An application: optimal Markov joining may not be Markovian 65

CHAPTER 7: REALIZATION PROBLEM 70

7.1. Settings 70

7.2. Necessary lemmas 72

7.3. Realization theorems and proofs 76

CHAPTER 8: DUALITY PROBLEMS 81

8.1. For problem (I) - Kantorovich duality 81

8.1.1. Geometry of Optimal Transport Plans 87

8.2. For problem (II) - Fenchel duality 89

CHAPTER 9: DISCUSSION 92

REFERENCES 95

# CHAPTER 1: INTRODUCTION

The Optimal Transport (OT) problem originates from the basic transportation problem in real life: if we want to move some goods from one or more places (warehouses) to one or more destinations, is it possible to find an 'optimal' transport plan which minimizes transportation cost? The usual setting for the OT problem consists of a fixed-cost function with some nice properties, the distribution of goods in different warehouses and the distribution of demand in different destinations. Thermodynamic Formalism is a famous topic in mathematical statistical physics, which studies the 'free energy' in certain thermodynamical systems and states that minimize system's free energy, such states are called 'equilibrium states' since systems with minimized free energy achieve an equilibrium. The key issue of general Thermodynamic Formalism is 'pressure' function, which consists of two parts: potential function that represents system's total energy and entropy that measures the disorder of a system. Both Optimal Transport and Thermodynamic Formalism are linear optimization problems. In this Ph.D. project, we focus on the linear optimization problem with restrictions under the framework of Optimal Transport and Thermodynamic Formalism. In the following, we will firstly introduce some basic concepts which are preliminaries of this Ph.D. project, and then the basic settings and problems studied in this project will be listed, finally structure of this dissertation will be given.

## 1.1 Preliminaries

### 1.1.1 Optimal Transport

The earliest Optimal Transport theory[2] only considers discrete cases that moving between finite number of places, while our work focuses on the problem under more general settings. Suppose $X$ and $Y$ are compact metrizable spaces, cost function $c : X \times Y \to \mathbb{R}$ a real valued continuous function defined on the product space $X \times Y := \{(x, y) : x \in X, y \in Y\}$, $\mathcal{M}(X)$ and $\mathcal{M}(Y)$ denote the set of Borel probability measures defined on $X$ and $Y$. The classical Optimal Transport Theory deals with the following optimization problem:

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times Y} c(x, y) d\pi(x, y)$$

This problem, initiated by Monge[3] and generalized by Kantorovich[4], is also called *Monge-Kantorovich problem*, where $\Pi(\mu, \nu)$ is the set of couplings with its marginals on $X$ and $Y$ are $\mu \in \mathcal{M}(X)$ and $\nu \in \mathcal{M}(Y)$ respectively. In other words, for any measure $\pi \in \Pi(\mu, \nu)$, $(\text{proj}_X)_\# \pi = \pi \circ \text{proj}_X^{-1} = \mu$, $(\text{proj}_Y)_\# \pi = \pi \circ \text{proj}_Y^{-1} = \nu$, where $\text{proj}_X$ and $\text{proj}_Y$ respectively stand for projection maps $(x, y) \mapsto x$ and $(x, y) \mapsto y$, the subscription '#' is read as 'push-forward'.

### 1.1.2 Measure-Preserving System and Invariant Measures

To study the stationarity of optimal transport plan, we introduce the notions of measure preserving and ergodicity. Given a probability space $(X, \mathcal{A}, \mu)$, where $X$ is sample space, $\mathcal{A}$ is the Borel $\sigma$-algebra on $X$ (in some papers it is denoted as $\mathcal{B}(X)$) and $\mu$ is a probability measure defined on $\mathcal{A}$. A transformation $T : X \to X$ is called *measure-preserving* if for any $A \in \mathcal{A}$, $T_\# \mu(A) = \mu \circ T^{-1}(A) = \mu(A)$, the measure $\mu \in \mathcal{M}(X)$ is called *$T$-invariant*. The probability space $(X, \mathcal{A}, \mu)$ together with a measure-preserving operator $T$ is called *measure-preserving system*. Let $\mathcal{M}(X, T)$ be the set of all $T$-invariant probability measures defined on $\mathcal{B}(X)$. The invariance property listed above represents stationarity, and in ergodic theory, $\mathcal{M}(X, T)$ is an important object of study. $\mathcal{M}(X, T)$ is nonempty by the following *Krylov-Bogolioubov* theorem[5]:

**Theorem 1.1** (Krylov-Bogolioubov)**.** *If $T : X \to X$ is a continuous transformation* *of a compact metric space $X$ then $\mathcal{M}(X,T)$ is nonempty.*

We have the following properties of $\mathcal{M}(X,T)$:

**Theorem 1.2** (properties of $\mathcal{M}(X,T)$[5])**.** *If $T$ is a continuous transformation of a compact metric space $X$ then*

*(i)* $\mathcal{M}(X,T)$ *is a compact subset of* $\mathcal{M}(X)$.

*(ii)* $\mathcal{M}(X,T)$ *is convex.*

*(iii)* $\mu$ *is an extreme point of* $\mathcal{M}(X,T)$ *iff* $T$ *is an ergodic measure-preserving transformation of* $(X,\mathcal{B}(X),\mu)$.

*(iv) If* $\mu,\nu \in \mathcal{M}(X,T)$ *are both ergodic and* $\mu \neq \nu$ *then they are mutually singular.*

### 1.1.3 Ergodicity

**Definition 1.1** (Ergodicity)**.** *Suppose $(X,\mathcal{A},\mu,T)$ is a measure-preserving system, a measure-preserving transformation $T$ is called ergodic if for any set $E \in \mathcal{A}$ satisfies $T^{-1}E = E$ implies $\mu(E) = 0$ or $\mu(E) = 1$.*

**Example 1.1.** *In the dynamical system $(\{0,1\}^{\mathbb{Z}},\sigma)$:*

- $\mu = \frac{1}{2}\delta_{0^\infty} + \frac{1}{2}\delta_{1^\infty}$ *is not ergodic.*
- *The Bernoulli measure $\mu$ given by $\mu([0]) = p$, $\mu([1]) = 1-p$ is ergodic.*

### 1.1.4 Joinings

The notion 'joinings' was firstly introduced by Furstenburg[6] and now widely studied in ergodic theory[7, 8]. Given two measure-preserving Borel systems $(X,\mathcal{A},\mu,T)$ and $(Y,\mathcal{B},\nu,S)$, where $T : X \to X$ and $S : Y \to Y$ are two transformations, $\mu \in \mathcal{M}(X,T)$ and $\nu \in \mathcal{M}(Y,S)$ are Borel invariant probability measures. The product operator $T \times S : X \times Y \to X \times Y$ is defined by $(T \times S)(x,y) = (T(x),S(y))$ for any $x \in X, y \in Y$. If a probability measure $\lambda$ is defined on product space $X \times Y$, is $(T \times S)$-invariant and is a coupling of $\mu$ and $\nu$ ($\lambda \in \Pi(\mu,\nu)$), then it is called the *joining* of two dynamical systems $(X, \mathcal{A}, \mu, T)$ and $(Y, \mathcal{B}, \nu, S)$. We use $J(\mu,\nu)$ to denote the set of all joinings of $(X, \mathcal{A}, \mu, T)$ and $(Y, \mathcal{B}, \nu, S)$.

### 1.1.5 Measure-Theoretic Entropy

Entropy is a well-known concept in physics, it quantifies the disorder of systems. In this project entropy quantifies the disorder of dynamical systems. Before we talk about entropy, some necessary concepts will be introduced.

Suppose that $(X, \mathcal{A}, \mu, T)$ is a measure-preserving system.

**Definition 1.2** (Partition). *A finite measurable partition of $(X, \mathcal{A}, \mu, T)$ is a collection of disjoint elements of $\mathcal{A}$. In other words, if $\mathcal{P} = \{A_1, A_2, \cdots, A_n\}$ is a finite measurable partition of $(X, \mathcal{A}, \mu, T)$, then it satisfies the following properties*

- $A_i \in \mathcal{A}$ *for any* $i = 1, 2, \cdots, n$, $A_i \cap A_j = \varnothing$ *for* $i, j \in \{1, 2, \cdots, n\}$, $i \neq j$,
- $\bigcup_{i=1}^{n} A_i = X$.

The basic definition of entropy is based on measurable partitions:

**Definition 1.3** (Entropy of a partition). *Let $\mathcal{P} = \{A_1, \cdots, A_k\}$ be a finite partition of $(X, \mathcal{A}, \mu)$, the entropy of $\mathcal{P}$ is given by*

$$H(\mathcal{P}) = -\sum_{i=1}^{k} \mu(A_i) \log \mu(A_i)$$

*where $0 \log 0 = 0$ by regulation.*

Let $T^{-1}\mathcal{P} = \{T^{-1}A_i : 1 \leq i \leq n\}$, and for any $n$ finite partitions $\mathcal{P}_1, \cdots, \mathcal{P}_n$, define their *join* as

$$\bigvee_{i=1}^{n} \mathcal{P}_i = \{C_1 \cap C_2 \cap \cdots \cap C_n : C_i \in \mathcal{P}_i\}$$

For $n \in \mathbb{N}$, let $\mathcal{P}^n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}$, now we are ready to define the measure-theoretic entropy with respect to $\mathcal{P}$:

**Definition 1.4** (Measure-theoretic entropy)**.** *Suppose* $(X, \mathcal{A}, \mu, T)$ *is a measure-preserving system,* $\mathcal{P} = \{A_1, \cdots, A_k\}$ *is a finite partition, the entropy of* $T$ *w.r.t* $\mathcal{P}$ *is given by*

$$h(T, \mathcal{P}) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{P}^n)$$

*the existence of limit above is ensured by subadditivity. Then the measure-theoretic entropy* $h(\mu) = h_\mu(T) := \sup_{\mathcal{P}} h(T, \mathcal{P})$*. The supremum is taken over all finite partitions of* $(X, \mathcal{B}, \mu)$*.*

### 1.1.6 General Thermodynamic Formalism

General Thermodynamic Formalism is centered on 'topological pressure', which consists of two parts: potential function and metric along the orbits.

The concept of *tolpological pressure*, defined by Ruelle[9] and Walters[10] firstly, relies on *Bowen's metric*: If $(X, d)$ is a compact metric space, the map $T : X \to X$ is continuous, then for $n \in \mathbb{N}$, a new metric $d_n$ on $X$ is defined by $d_n(x, y) := \max_{0\le i\le n-1} d(T^i x, T^i y)$. For $\varepsilon > 0$, a subset $E \subset X$ is called $(n, \varepsilon)$-spanning set if $\forall x \in X$, $\exists y \in E$ s.t. $d_n(x, y) < \varepsilon$. Now let $\mathcal{C}(X)$ be the space of real-valued continuous functions on $X$ and for $f \in \mathcal{C}(X)$ and $n \geq 1$, $S_n f(x) = \sum_{i=0}^{n-1} f \circ T^i(x)$, define

$$P_n(T, f, \varepsilon) := \inf\left\{\sum_{x\in E} e^{S_n f(x)} : E \text{ is a } (n, \varepsilon)\text{-spanning set for } X\right\}$$

then we put

$$P(T, f) := \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} P_n(T, f, \varepsilon)$$

the existence of limit above is again assured by subadditivity, and $P(T, f)$ is called the topological pressure for $f$. Specifically, when $f \equiv 0$, the value

$$P(T, 0) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \inf\{|E| : E \text{ is a } (n, \varepsilon)\text{-spanning set for } X\}$$

is named 'topological entropy', denoted by $h_{\mathrm{top}}(T)$.

The *variational principle* is an important result in Thermodynamic Formalism, which builds a relationship between pressure and measure-theoretic entropy:

**Theorem 1.3** (Variational Principle)**.** *Suppose that $X$ is a compact metrizable space and $T : X \to X$ is a continuous transformation, for $f \in \mathcal{C}(X)$*

$$P(T, f) = \sup_{\mu \in \mathcal{M}(X,T)} \left( \int f d\mu + h(\mu) \right)$$

where $\mathcal{M}(X,T) = \{\mu \in \mathcal{M}(X) : T_{\#}\mu = \mu\}$ and $\mathcal{M}(X)$ is the set of Borel probability measures on $X$. Proof of the variational principle please refer to[5].

The variational principle offers a way to pick out some members of $\mathcal{M}(X,T)$: If $\mu \in \mathcal{M}(X,T)$ satisfies $P(T, f) = \int f d\mu + h(\mu)$, then $\mu$ is called an *equilibrium state* for function $f$ (called *potential function*) w.r.t $T$.

Note that when $f \equiv 0$ from Theorem 1.3 we obtain a special case of the variational principle:

**Corollary 1.4.** *Suppose that $X$ is a compact metrizable space and $T : X \to X$ is a continuous transformation. Then $h_{\mathrm{top}}(T) = \sup_{\mu \in \mathcal{M}(X,T)} h(\mu)$.*

### 1.1.7 Shift of Finite Type and Gibbs measure

Gibbs measure is a special form of equilibrium state as it is based on a special dynamical system: *Shift of Finite Type* (SFT). The topics for SFTs are in the field of symbolic dynamics widely introduced by [11].

We firstly define SFT. Assume that $\{1, 2, \cdots, n\}$ is a set of possible states in a system, is it named as *alphabet* in symbolic dynamic theory. Let $\Sigma_n = \{1, 2, \cdots, n\}^{\mathbb{Z}}$ be the space of infinite sequences $\underline{x} = \{x_n\}_{n=-\infty}^{\infty}$ with $x_n \in \{1, 2, \cdots, n\}$. A transformation on $\Sigma_n$ is given by $\sigma : \Sigma_n \to \Sigma_n$, $\sigma(\underline{x})_n = x_{n+1}$, such transformation $\sigma$ is called 'left-shift' map. If $A$ is an $n \times n$ matrix with its entries either 0 or 1, define

$$\Sigma_A = \{\underline{x} \in \Sigma_n : A_{x_i x_{i+1}} = 1 \text{ for all } i \in \mathbb{Z}\}$$

That is, any sequence $\underline{x}$ in $\Sigma_A$ is an member of $\Sigma_n$ and for any $i \in \mathbb{Z}$, the $(x_i, x_j)$-th entry of matrix $A$ is 1. In other words, the matrix $A$ rules out all the sequences $\underline{x}$ in $\Sigma_n$ with $A_{x_i x_{i+1}} = 0$ for some $i \in \mathbb{Z}$. $\Sigma_A$ define above is a SFT. The set $\mathcal{F}(\Sigma_A) = \{\underline{x} \in \Sigma_A : A_{x_i x_{i+1}} = 0 \text{ if } \exists i \in \mathbb{Z}\}$ is called the set of *forbidden words* for $\Sigma_A$. The forbidden words is an important component of a SFT, if $\mathcal{F}(\Sigma_A) = \varnothing$, then $A$ is a matrix of all 1's and $\Sigma_A = \Sigma_n$ is called a *full shift* (on the alphabet $\{1, 2, \cdots, n\}$). When $\{1, 2, \cdots, n\}$ is given the discrete topology and $\Sigma_A$ (thus $\Sigma_n$) the product topology, $\Sigma_A$ is closed, compact and metrizable. The metric on $\Sigma_A$ is given by $d_n(\underline{x}, \underline{y}) = 2^{-n(\underline{x},\underline{y})}$, $n(\underline{x}, \underline{y}) = \sup\{N : x_i = y_i \text{ for } |i| < N\}$.

Mixing property of SFTs is also required there. Let $\mathcal{L}_m$ be the set of words of length $m$ (elements of $\mathcal{A}^m$) in $\Sigma_A$, $\mathcal{L} = \bigcup_{m \geq 1} \mathcal{L}_m$ is known as a *dictionary* for $\Sigma_A$. For $u \in \mathcal{L}_m$, $u = u_0 u_1 \cdots u_{m-1} := u_0^{m-1}$, define $[u] = \{\underline{x} \in \Sigma_A : \underline{x}_0^{m-1} = u\}$ a *cylinder set* in $\Sigma_A$. An SFT $\Sigma_A$ is mixing if for $u, v \in \mathcal{L}$, there exists $N > 0$ and for some $n > N$, we can find $w \in \mathcal{L}_n$ s.t. $w \in [u] \cap \sigma^{-n}[v]$. Equivalently, $\Sigma_A$ is mixing if and only if there exists an $N > 0$ s.t. $A^N$ contains all positive entries.

*Gibbs measure* is a special equilibrium states for potential function $\phi$ with proper regularity conditions, for detailed introduction please refer to [12]. Denote $\mathcal{C}(\Sigma_A)$ the set of real valued continuous functions on $\Sigma_A$, for $\phi \in \mathcal{C}(\Sigma_A)$ define

$$\mathrm{var}_n \phi := \sup\{|\phi(\underline{x}) - \phi(\underline{y})| : x_i = y_i \ \forall |i| \leq n\}$$

and

$$\mathcal{F}_A = \{\phi \in \mathcal{C}(\Sigma_A) : \mathrm{var}_n \phi \leq \beta \alpha^n \text{ for } \beta > 0, \alpha \in (0, 1)\}$$

Assume $\Sigma_A$ a mixing SFT and $\phi \in \mathcal{F}_A$, the *Gibbs measure*[12] on $\Sigma_A$ with potential $\phi$ satisfies: $\exists C_1, C_2 > 0$

$$C_1 \leq \frac{\mu_\phi([x_0^{m-1}])}{\exp(S_n\phi(\underline{x}) - mP(\phi))} \leq C_2$$

for any $\underline{x} \in \Sigma_A$ and $m \geq 1$. $P(\phi)$ is the pressure of $\phi$. $\mu_\phi$ is an equilibrium state for potential $\phi$ with respect to $\sigma$, i.e.

$$P(\phi) = \int \phi d\mu_\phi + h(\mu_\phi)$$

and $\mu_\phi$ is the unique equilibrium states for $\phi \in \mathcal{F}_A$.

## 1.2 Settings and Problems

### 1.2.1 Settings

In this project, we suppose that $X$ is a compact metric space, $T : X \to X$ is a continuous transformation. Let $\mathcal{C}(X)$ be the set of real-valued continuous functions defined on $X$ and $\mathcal{M}(X)$ be the set of Borel probability measures on $X$. $\mathcal{M}(X)$ can be seen as a subset of $\mathcal{C}^*(X)$ by defining $\mu(f) = \int f d\mu$ for any $\mu \in \mathcal{M}(X)$ and $f \in \mathcal{C}(X)$, $\mathcal{M}(X)$ is closed in weak* topology. Denote $\mathcal{M}(X,T) := \{\mu \in \mathcal{M}(X) : T_\#\mu = \mu\}$ the set of $T$-invariant Borel probability measures on $X$. For $\mathcal{A}$ the Borel $\sigma$-algebra on $X$ and given $\mu \in \mathcal{M}(X,T)$, we call $(X, \mathcal{A}, \mu, T)$ a measure-preserving dynamical system.

The following example fits for the settings above:

**Example 1.2** (Shift Space)**.** *$(X, \mathcal{A}, \mu, T)$ is a measure-preserving system if*

- *$X = \{0,1\}^{\mathbb{Z}}$, every element $\underline{x} \in X$ is a sequence consists of $0$s and $1$s.*

- *Borel $\sigma$-algebra $\mathcal{A}$ is generated by all the cylinder sets over $X$.*

- *$T = \sigma : X \to X$ is called 'left shift', for any sequence $\underline{x} := \{x_i\}_{i=-\infty}^{\infty}$ $(\sigma(\underline{x}))_i =$*

$x_{i+1}$. *It is continuous if the metric* $d$ *on* $X$ *is given by* $d(\underline{x}, \underline{y}) = 2^{-N(\underline{x},\underline{y})}$, *where*

$$N(\underline{x}, \underline{y}) := \max\{N : x_i = y_i \text{ for } |i| < N\} \text{ when } \underline{x} \neq \underline{y}$$

*and* $N(\underline{x}, \underline{y}) = 0$ *when* $\underline{x} = \underline{y}$.

- *The Bernoulli measure* $\mu$ *defined as* $\mu([0]) = p$ *and* $\mu([1]) = 1 - p$ *is measure-preserving.*

### 1.2.2 Problems

Let the 'restriction set' $Z$ be a closed subset of $\mathcal{C}(X)$ and define $\mathcal{M}_Z(X,T) := \{\mu \in \mathcal{M}(X,T) : \mu(h) = 0 \text{ for all } h \in Z\}$. In this project we have studied the following two problems:

**Problem 1.1** (Problem (I) - Linear Optimization for Dynamical Systems)**.** *Suppose that* $(X, \mathcal{A}, \mu, T)$ *is a measure-preserving systems defined as the settings,* $Z \subset \mathcal{C}(X)$ *closed such that* $\mathcal{M}_Z(X,T)$ *is nonempty. Given* $\phi \in \mathcal{C}(X)$, *the linear optimization problem for dynamical systems is*

$$\begin{aligned} &\max_{\mu \in \mathcal{M}(X,T)} \int_X \phi d\mu \\ \textit{subject to } &\int_X \psi d\mu = 0, \forall \psi \in Z \end{aligned} \tag{I}$$

*which is equivalent to*

$$\max_{\mu \in \mathcal{M}_Z(X,T)} \int_X \phi d\mu$$

**Problem 1.2** (Problem (II) - Linear Optimization for Dynamical Systems with Entropic Regularization)**.** *Suppose that* $(X, \mathcal{A}, \mu, T)$ *is a measure-preserving systems defined as the settings,* $Z \subset \mathcal{C}(X)$ *closed such that* $\mathcal{M}_Z(X,T)$ *is nonempty. Given* $\phi \in \mathcal{C}(X)$, *the linear optimization problem for dynamical systems with entropic regu-*

*larization is*

$$\begin{aligned} &\max_{\mu\in\mathcal{M}(X,T)} \int_X \phi d\mu + h(\mu) \\ \textit{subject to } &\int_X \psi d\mu = 0, \forall \psi \in Z \end{aligned} \tag{II}$$

*which is equivalent to*

$$\max_{\mu\in\mathcal{M}_Z(X,T)} \int_X \phi d\mu + h(\mu)$$

*where $h(\mu)$ is the measure-theoretic entropy for $(X, \mathcal{A}, \mu, T)$.*

## 1.3 Statement of main results

At the level of generality, problems of interests are: (i) existence and characterization of the set of optimal solutions; (ii) uniqueness of optimal solutions for problem (I) and (II); (iii) duality problems for problem (I) and (II); (iv) Lagrangian approach to optimization; (v) optimization as zero temperature and (vi) realization problem. In this section we will list our main results which may cover one or more problems of interests above.

Before the statement of main results, an important property (named 'property (E)') which ensures that $\mathcal{M}_Z(X,T)$ contains ergodic measures is given below:

(E) If for any $\mu \in \mathcal{M}_Z(X,T)$, its ergodic decomposition is

$$\mu = \int_{\mathcal{M}^e(X,T)} \nu d\tau(\nu)$$

where $\mathcal{M}^e(X,T)$ denotes the set of all ergodic measures of $\mathcal{M}(X,T)$ and $\tau$ is a probability measure defined on the Borel subsets of $\mathcal{M}(X,T)$ and supported on $\mathcal{M}^e(X,T)$. Then $\nu \in \mathcal{M}_Z(X,T)$ for $\tau$- almost everywhere of $\nu$.

**Theorem 1.5** (Lagrangian approach to optimization). *Suppose that there are $n$ dynamical systems $\{(X_i, T_i)\}_{i=1}^n$, where each $X_i$ and $T_i : X_i \to X_i$ satisfy our basic settings in section 1.2.1. Let $\phi \in \mathcal{C}(\prod_{k=1}^n X_k)$ and $\pi_i : \prod_{k=1}^n X_k \to X_i$ be the projec-*

*tion map onto the ith space. Given* $\mu_i \in \mathcal{M}^e(X_i, T_i)$ *and let*

$$Z = \left\{ \sum_{i=1}^{n} \left( f_i \circ \pi_i - \int_{X_i} f_i d\mu_i \right) : \forall f_i \in \mathcal{C}(X_i), i = 1, \cdots, n \right\}$$

*there exists* $(\hat{f}_1, \cdots, \hat{f}_n) \in \mathcal{C}(X_1) \times \cdots \times \mathcal{C}(X_n)$ *s.t.*

$$\begin{aligned} & \hat{\lambda} \in \underset{\lambda \in \mathcal{M}(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k)}{\operatorname{argmax}} \int \left[ \phi + \sum_{i=1}^{n} \left( \hat{f}_i \circ \pi_i - \int_{X_i} \hat{f}_i d\mu_i \right) \right] d\lambda + h(\lambda) \\ \Longleftrightarrow \ & \hat{\lambda} \in \underset{\lambda \in \mathcal{M}_Z(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k)}{\operatorname{argmax}} \int \phi d\lambda + h(\lambda) \end{aligned}$$

*Note that*

$$\mathcal{M}_Z(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k) = \{\mu \in \mathcal{M}(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k) : \mu \circ \pi_i^{-1} = \mu_i \ for\ i = 1, \cdots, n\}$$

The next two results are about the uniqueness of optimal solution in the framework of problem (I) and (II). In problem (I), we are interested in 'generic' uniqueness property. Denote $\mathcal{M}(X, T, f)$ the set of measures maximize $\int f d\mu$ over $\mathcal{M}_Z(X, T)$. The 'generic' uniqueness property is, given $E \subset \mathcal{C}(X)$, the set

$$\mathcal{U}(E) := \{f \in E : \mathcal{M}(X, T; f) \text{ is a singleton}\}$$

is 'topologically large' in $E$:

**Theorem 1.6** (Uniqueness of optimal solutions in problem (I))**.**

*Let* $X$ *be a compact metric space, and* $T : X \to X$ *a continuous map on* $X$*. Let* $E$ *be a topological vector space which is densely and continuously embedded in* $\mathcal{C}(X)$*. If* $Z$ *is closed in* $\mathcal{C}(X)$ *and satisfies property (E), then* $\mathcal{U}(E)$ *is a residual set. Moreover, if* $E$ *is a Baire space, then* $\mathcal{U}(E)$ *is dense in* $E$*.*

In the uniqueness result of problem (II), we will consider the problem defined on a product dynamical system with its optimal solution has specified marginal distribution, so in the following result we let

$$Z = \left\{ \varphi \circ \pi_X - \int_X \varphi d\mu + \psi \circ \pi_Y - \int_X \psi d\nu : \forall \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y) \right\}$$

where $\mu, \nu$ are given in advance.

**Theorem 1.7** (Uniqueness of optimal solutions in problem (II))**.**

*(i) Let $(X, T)$ and $(Y, S)$ be shift spaces, and let $\mu \in \mathcal{M}(X, T), \nu \in \mathcal{M}(Y, S)$ be i.i.d. invariant measures. The continuous function $\phi : X \times Y \to \mathbb{R}$ is given by $\phi(\underline{x}, \underline{y}) = \phi_0(x_0, y_0)$, and $Z$ is as above. Then problem (II) has unique optimal solution $\lambda = (\lambda_0)^{\otimes \mathbb{N}}$, where*

$$\lambda_0 = \operatorname*{argmax}_{\lambda_0' \in \Pi(\mu_0, \nu_0)} \int \phi_0 d\lambda_0' + H_{\lambda_0'}(\mathcal{P})$$

*(ii) Let $(X, \mathcal{A}, \mu, T)$ and $(Y, \mathcal{B}, \nu, S)$ are measure preserving systems. $\phi \in \mathcal{C}(X \times Y)$ and $Z$ is given as above. By Theorem 1.5, there exists a measure $\hat{\lambda} \in \mathcal{M}_Z(X \times Y, T \times S)$ and $\chi \in Z$ s.t.*

$$\hat{\lambda} \in \operatorname*{argmax}_{\lambda \in \mathcal{M}(X \times Y, T \times S)} \int (\phi + \chi) d\lambda + h(\lambda) \implies \hat{\lambda} \in \operatorname*{argmax}_{\lambda \in \mathcal{M}_Z(X \times Y, T \times S)} \int \phi d\lambda + h(\lambda)$$

*If $\phi + \chi$ is regular enough (Hölder, Lipschitz, locally constant or Walters[13]), then problem (II) has unique optimal solution.*

Define $\ell_{\tau,f}(\mu) := \int f d\mu + \tau \cdot h(\mu)$ and $L_\tau(f) := \sup_{\mu \in \mathcal{M}_Z(X,T)} \ell_{\tau,f}(\mu)$ for $\tau \in \{0, 1\}$. Denote by $\mathcal{M}(X, T; \phi) \subset \mathcal{M}_Z(X, T)$ and $\mathcal{R}(X, T; \phi) \subset \mathcal{M}_Z(X, T)$ the set of measures maximize $\ell_{\tau,\phi}(\mu)$ over $\mathcal{M}_Z(X, T)$ for $\tau = 0$ and $\tau = 1$ respectively (that is, $\mu' \in \mathcal{M}(X, T; \phi) \iff \ell_{0,\phi}(\mu') = L_0(\phi)$ and $\mu' \in \mathcal{R}(X, T; \phi) \iff \ell_{1,\phi}(\mu') = L_1(\phi)$).

Denote $\mathcal{M}_Z^e(X,T)$ the set of ergodic measures in $\mathcal{M}_Z(X,T)$, when $Z$ satisfies property (E), $\mathcal{M}_Z^e(X,T)$ is nonempty. We have the following realization result, which generalizes Jenkinson[14]:

**Theorem 1.8.** *Assume that $Z$ satisfies property (E). Let $\mathcal{E}$ be a non-empty subset of $\mathcal{M}_Z^e(X,T)$ which is weak* closed in $\mathcal{M}_Z(X,T)$. Let $\overline{co}(\mathcal{E})$ denote its closed convex hull in $\mathcal{M}_Z(X,T)$. There exists a continuous function $\phi : X \to \mathbb{R}$ such that $\mathcal{M}(X,T;\phi) = \overline{co}(\mathcal{E})$. Furthermore, if $h|_{\overline{co}(\mathcal{E})}$ is continuous, then there exists a continuous function $\psi : X \to \mathbb{R}$ such that $\mathcal{R}(X,T;\psi) = \overline{co}(\mathcal{E})$.*

For the linear optimization problems (I) and (II), we derived their duality problems. Let $Z \subset C(X)$ be a linear subspace. $\nu$ is a bounded linear functional defined on $Z$, $1 \in Z$ and $\nu(1) = 1$, and denote

$$\mathcal{M}_\nu(X,T) := \{\mu \in \mathcal{M}(X,T) : \mu|_Z = \nu\}$$

we have the Kantorovich duality problem for problem (I):

**Theorem 1.9.** *Let $W = \{g \circ T - g : g \in C(X)\}$, define*

$$\Pi_W(\nu) = \{\mu \in \mathcal{P}(X) : \int w d\mu = 0, \forall w \in W \ \ and \ \mu|_Z = \nu\}$$

*easy to see that $\Pi_W(\nu) = \mathcal{M}_\nu(X,T)$. The Kantorovich-form duality for problem (I) is:*

$$\inf_{\mu \in \Pi_W(\nu)} \int c d\mu = \sup_{\substack{f+w \le c \\ w \in W}} \nu(f)$$

*where $f \in Z$.*

The rest of this dissertation are organized as follows: In chapter 2 we organized related works of problem (I) and (II) for different selections of $Z$. Main results of this Ph.D. project were listed from chapters 3 to 8, including existence and characterization of optimal solutions, uniqueness of optimal solutions, optimization as zero temperature, realization problem, Lagrangian approaches to optimization and duality problem. In chapter 9 we summarized all the obtained outcomes and listed some future potential work.

# CHAPTER 2: RELATED WORK

For different selections of $Z$ in problems (I) and (II), there will be different famous existing works. In the following we will list related existing works associated with frameworks (I) and (II) respectively.

## 2.1 Related work for problem (I)

Firstly, when $T$ is an identity map $(Id)$, that is, $T(x) = x$ for all $x \in X$, there is no dynamic in problem (I), then

- When $X$ is finite, problem (I) is called 'Linear Programming' problem, it is a classical optimization problem and was widely introduced in [15]. The famous Sinkhorn's Algorithm in solving Optimal Transport problem is based on the techniques of Linear Programming; see [2] for details.

- When we have two compact metric spaces $X, Y$ and let $\pi_X : X \times Y \to X, \pi_Y : X \times Y \to Y$ be projection maps. Given $\mu \in \mathcal{M}(X), \nu \in \mathcal{M}(Y)$ fixed and set

$$Z = \left\{ \varphi \circ \pi_X - \int_X \varphi d\mu + \psi \circ \pi_Y - \int_Y \psi d\nu : \forall \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y) \right\}$$

  we recover the Optimal Transport (OT) problem from $(I)$. The optimal transport problem was first introduced by Monge[3] and generalized by Kantorovich[4], see Villani[16] for detailed work.

  There is a famous result for duality of the Optimal Transport problem: Kantorovich duality, which was proved by Kantorovich[4] and later generalized by Kantorovich and Rubinstein[17]. Some modern treatments ($X, Y$ are compact, Polish spaces) have been appeared in recent years, see Villani[18] for a classical duality result.

When $T : X \to X$ is a general continuous transformation, problem (I) is based on a general dynamical system, then

- When $Z = \{0\}$, there is no restriction on $\mu$, problem $(I)$ is called 'Ergodic Optimization' problem, it is a famous topic in ergodic theory and dynamical systems and has been studied in many aspects. The equivalence of ergodic optimization and time average for $\phi$ continuous was established by [19]. The generic uniqueness result was proved by Bousch [20], a similar version for $X$ Banach was given by [21]. Jenkinson[14] showed that for any ergodic measure $\mu$ in $\mathcal{M}(X,T)$, there exists a continuous function such that $\mu$ is the unique solution for problem(I). A good reference for Ergodic Optimization is Jenkinson[22].

- When $Z = \{\psi_1, \psi_2, \cdots, \psi_m\}$ is of $m$ dimensions ($m$ is any finite integer), define $\Psi = (\psi_1, \cdots, \psi_m) \in \mathcal{C}(X, \mathbb{R}^m)$, we denote by $\mathrm{Rot}(\Psi) = \{\mathrm{rv}(\mu) : \mu \in \mathcal{M}(X,T)\}$ the (generalized) rotation set of $\Psi$, where $\mathrm{rv}(\mu)$ is called 'rotation vector' of the measure $\mu \in \mathcal{M}(X,T)$ and given by

$$\mathrm{rv}(\mu) = \left( \int \psi_1 d\mu, \cdots, \int \psi_m d\mu \right)$$

  The constraint $\mu \in \mathcal{M}_Z(X,T)$ in problem (I) can be written as $\mu \in \mathcal{M}_\Psi(0)$, for $\mathcal{M}_\Psi(0) = \{\mu \in \mathcal{M}(X,T) : \mathrm{rv}(\mu) = \mathbf{0}\}$, which is non-empty when $0 \in \mathrm{Rot}(\Psi)$. In this setting, problem (I) is an optimization problem over rotation sets. The rotation set was firstly defined in the context of degree-one circle maps[23], and was studied by Geller and Misiurewicz[24] and Ziemian[25].

- In problem (I) if we have another dynamical system $(Y, S)$ where $Y$ is compact matrizable space and $S : Y \to Y$ continuous, $\nu \in \mathcal{M}(Y,S)$ is given. Moreover, there is a factor map $\pi : X \to Y$ between two dynamical systems, that is, $\pi$ is a surjection and satisfies $\pi \circ T = S \circ \pi$. When the restriction set is defined as

$$Z = \left\{ \psi \circ \pi - \int \psi d\nu : \psi \in \mathcal{C}(Y) \right\}$$

then $\mathcal{M}_Z(X,T) = \{\mu \in \mathcal{M}(X,T) : \pi_{\#}\mu = \nu\} := \mathcal{M}_\nu(X,T)$ and problem $(I)$ is called relative ergodic optimization problem. Tuncel[26] proved that if $(Y, \mathcal{B}, S, \nu)$ is a Markov shift, $(X,T)$ is a topological Markov chain and $\pi : X \to Y$ a bounded to one factor map, then $\mathcal{M}_\nu(X,T)$ contains only one point. A related problem, known as measures of maximal relative entropy, is a famous optimization problem which focuses on the maximal measure-theoretic entropy over $\mathcal{M}_\nu(X,T)$.

## 2.2 Related Work for problem (II)

When $T$ is an identity map, the measure-theoretic entropy $h(\mu) = 0$ and problem (II) is reduced to problem (I). So we only need to consider $T : X \to X$ is a general continuous transformation.

- When $Z = \{0\}$, problem (II) is called 'thermodynamic formalism', a topic of statistical mechanics. At the beginning, problem's setting was born in statistical mechanics directly, a special case that $X$ is one-dimension lattice was introduced by Dobrushin[27, 28, 29] and the optimizer was named 'Gibbs states'. Uniqueness of Gibbs states was given by Ruelle[30] and the variational principle for Gibbs states was established in the work of Lanford and Ruelle[31]. When $X$ is shift of finite type, the thermodynamic formalism and the variational principle was given by Ruelle[32], in this more general case the equilibrium states is called 'Gibbs measure'. In general thermodynamic formalism, $X$ is compact and metrizable space. For expansive $T$ the topological entropy and variational principle was given by Ruelle[9], and for general $T$, related results were proved by Goodwyn[33] and Goodman[34, 35]. Uniqueness of equilibrium states was proved by Bowen[36] firstly, in achieving uniqueness, the regularity conditions that potential function satisfies is named 'specification' condition. Bowen's result was improved by several scholars like Climengaha and Watson[37, 38], Pavlov[39] and so on.

- When $Z = \{\psi_1, \ldots, \psi_m\}$, $\Psi = (\psi_1, \cdots, \psi_m) \in \mathcal{C}(X, \mathbb{R}^m)$. Let $\mathrm{Rot}(\Psi) = \{\mathrm{rv}(\mu) : \mu \in \mathcal{M}(X, T)\}$ where $\mathrm{rv}(\mu)$ is the rotation vector for $\mu$ and is given by $(\int \psi_1 d\mu, \cdots, \int \psi_m d\mu)$. A good reference for rotation set is Jenkinson[40]. In further discussion, Kucherenko and Wolf[41] studied geometric properties for rotation set and analytical properties for entropy function over rotation sets. Assume that $\mathbf{0} \in \mathrm{Rot}(\Psi)$ (so $\mathcal{M}_Z(X, T)$ is non-empty), problem $(II)$ is called 'localized relative equilibrium states' problem. This is another restricted version of thermodynamic formalism, was studied by Kucherenko and Wolf[42]. In Kucherenko and Wolf[43], the ground states for a family of localized equilibrium states $\{\mu_t\}$ associated with potential functions $t\phi$ was studied.

- Suppose $Y$ is another compact metric space and $S : Y \to Y$ continuous, $\mathcal{C}(Y)$ and $\mathcal{M}(Y, S)$ are defined accordingly. Let $\pi : X \to Y$ be a factor map, measure $\nu \in \mathcal{M}(Y, S)$ is fixed. If the restriction set $Z = \{\psi \circ \pi - \int \psi d\nu : \psi \in \mathcal{C}(Y)\}$, then $\mathcal{M}_Z(X, T) = \{\mu \in \mathcal{M}(X, T) : \pi_{\#}\mu = \nu\}$ and problem $(II)$ is called 'relative equilibrium states', it was studied from different aspects by many researchers. Based on Bowen's definition for topological entropy[44], Ledrappier and Walters[45] gave the definitions of relative measure-theoretic entropy and relative topological pressure, and proved the relativised version of variational principle. In Walters[13], detailed analysis for relative topological pressure and relative equilibrium states is given, so as the discussion of compensation function. Uniqueness of measures of maximal relative equilibrium states was given by several scholars in different settings: when $\pi : X \to Y$ is a 1-block factor map from a 1-step SFT $X$ to a sofic shift $Y$ and $\nu$ is ergodic, Petersen, Quas and Shin[46] gave an upper bound of number of maximal relative entropy ($\phi = 0$); such upper bound was given in the form of class degree by Allahbakhshi and Quas[47]; furthermore, Allahbakhshi, Antonioli and Yoo[48] proved that number of maximal relative equilibrium states is bounded by the class degree of $\nu$, so when the class degree of $\nu$ is 1, problem (II) has a unique optimizer. The uniqueness result for a more general factor map $\pi$ was proved by Yoo[49].

# CHAPTER 3: EXISTENCE AND CHARACTERIZATION OF OPTIMAL PLANS

In this chapter, we will firstly focus on the properties of restriction set - $\mathcal{M}_Z(X,T)$ especially the nonemptyness property. Then we will introduce some facts for existence and characterization of optimal plans in Problem (I) and (II).

## 3.1 Properties of $\mathcal{M}_Z(X,T)$

By the famous result of Krylov and Bogolioubov (see [5], p.152), $\mathcal{M}(X,T)$ is nonempty under the settings in section 2.1, however, $\mathcal{M}_Z(X,T)$ may be empty for some selections of $Z$, a simple example is $Z = \mathcal{C}(X)$. The following is another example of $\mathcal{M}_Z(X,T) = \varnothing$:

**Example 3.1** (Circle Rotations)**.** *Let $S^1 = [0,1]/\sim$ be the unit circle, where $\sim$ indicates that $0$ and $1$ are identified, and $\mod 1$ makes $S^1$ an abelian group. The natural distance on $[0,1]$ induces a distance on $S^1$:*

$$d(x,y) = \min(|x-y|, 1-|x-y|)$$

*Lebesgue measure on $[0,1]$ gives a natural measure $\lambda$ on $S^1$.*

*For $\alpha \in \mathbb{R}$ irrational, let $R_\alpha$ be the rotation of $S^1$ by angle $2\pi\alpha$, i.e.*

$$R_\alpha x = x + \alpha \mod 1$$

*$(S^1, \mathcal{B}([0,1]), \lambda, R_\alpha)$ is a measure-preserving system. If $f = -(1-x)(1-\alpha-x)\mathbf{1}_{[1-\alpha,1]} \in Z$, then for any element in $\mathcal{M}_Z(X,T)$, whose support must avoid $[1-\alpha,1]$. However, such measures cannot be measure-preserving since for any interval $I = [a,b]$ with positive measure and $b-a < \alpha$, there exists an $n \in \mathbb{N}$ s.t. $R_\alpha^n I = [R_\alpha^n a, R_\alpha^n b] \subset [1-\alpha, 1]$*

*(this is because the orbit $\{R_\alpha^n\}_n$ is dense in $S^1$), thus $R^n I = 0$. So $\mathcal{M}_Z(X,T)$ is empty in this setting.*

To make problem(I) and (II) meaningful, we give a sufficient condition for the nonemptyness of $\mathcal{M}_Z(X,T)$:

**Theorem 3.1.** *Let $X$ be a compact metric space and $T : X \to X$ a continuous transformation. If there exist a Borel probability measure $\mu$ satisfies $\mu(f) = 0$ for all $f \in Z$, and for any $f \in Z$, $f \circ T \in Z$, then $\mathcal{M}_Z(X,T)$ is nonempty.*

*Proof.* The proof refers to Krylov-Bogolioubov. As $X$ is a compact and metric space, the Banach space $(\mathcal{C}(X), \|\cdot\|_{\max})$ is separable and so as its closed subset $Z$. Define

$$\psi_n(f) = \int_X \left[\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i\right] d\mu \ \forall n \in \mathbb{N} \text{ and } f \in \mathcal{C}(X)$$

It is obvious that $|\psi_n(f)| \leq \|f\|_{\max}$ for all $f \in \mathcal{C}(X), n \in \mathbb{N}$. So $\{\psi_n\}$ is a bounded sequence in $[\mathcal{C}(X)]^*$. As $\mathcal{C}(X)$ is a separable Banach space with norm $\|\cdot\|_{\max}$, by Helly's theorem, there exists a subsequence $\{\psi_{n_k}\}$ that converges, w.r.t. the weak* topology, to a bounded functional $\psi \in [\mathcal{C}(X)]^*$:

$$\lim_{k\to\infty} \psi_{n_k}(f) = \psi(f) \ \forall f \in \mathcal{C}(X)$$

Note that

$$\psi_{n_k}(f \circ T) - \psi_{n_k}(f) = \int_X \left[\frac{1}{n_k}(f \circ T^{n_k} - f)\right] d\mu$$

so if we take $k \to \infty$:

$$\psi(f \circ T) = \psi(f), \forall f \in \mathcal{C}(X)$$

which ensures the invariance of $\psi$. $\psi$ is positive, such property is inherited from $\mu$. And for any $f \in Z$, we have $\psi(f) = 0$, as $\psi_n(f) = 0$ and any $n \in \mathbb{N}$.

By Riesz-Markov Theorem, there is a Borel measure $\hat{\mu}$ for which

$$\psi(f) = \int_X f d\hat{\mu} \ \forall f \in \mathcal{C}(X)$$

$\hat{\mu}$ is invariant because

$$\int_X f \circ T d\hat{\mu} = \psi(f) = \psi(f \circ T) = \int_X f d\hat{\mu}$$

and for $f \equiv 1$, $\psi(f) = 1$ because $\psi_n(f) = 1$ for all $n \in \mathbb{N}$. For any $f \in Z$, $\hat{\mu}(f) = \psi(f) = 0$. Therefore, $\hat{\mu} \in \mathcal{M}_Z(X,T)$. $\square$

**Example 3.2.** *This example is again the rotation map dynamical system. When* $f = -(1-x)(1-\alpha-x)\mathbf{1}_{[1-\alpha,1]} \in Z$ *and* $\mu(f) = 0$. *We claim that where exist an* $n \in \mathbb{N}$ *s.t.* $f \circ R_\alpha^n \notin Z$, *or* $\mu$ *must be supported on* $[0,1]$ *but not on any* $R_\alpha^n[1-\alpha,1]$, *however, it is impossible since* $\mu$ *is an probability measure and* $\cup_{n=0}^{\infty} R_\alpha^n[1-\alpha,1] = [0,1]$.

**Example 3.3.** *This example is about the relative equilibrium states: the closed subset* $Z = \{g \circ \pi - \int_Y g d\nu : g \in \mathcal{C}(Y)\}$, *here* $(Y,S)$ *is another dynamical system and* $\nu \in \mathcal{M}(Y,S)$ *is an ergodic measure,* $\pi : X \to Y$ *is a factor map,* $\pi \circ T = S \circ \pi$. *We have already known that* $\mathcal{M}_Z(X,T)$ *is nonempty. Note that for any* $g \in \mathcal{C}(Y)$:

$$(g \circ \pi - \int_Y g d\nu) \circ T = g \circ \pi \circ T - \int_Y g d\nu = g \circ S \circ \pi - \int_Y (g \circ S) d\nu \in Z$$

**Example 3.4.** *Consider for* $X$ *a compact metrizable set, we have continuous transformation* $T_1 : X \to X$ *and* $T_2 : X \to X$.*Let*

$$Z = \{f - f \circ T_2 : f \in \mathcal{C}(X)\}$$

*Then I claim that* $M_Z(X,T_1)$ *is equal to the set of measures that are invariant under both* $T_1$ *and* $T_2$. *If* $T_1 \circ T_2 = T_2 \circ T_1$, $\mathcal{M}_Z(X,T_1)$ *is nonempty, as for* $f - f \circ T_2 \in Z$

*we have*

$$(f - f \circ T_2) \circ T_1 = f \circ T_1 - (f \circ T_1) \circ T_2 \in Z$$

In this project, the existence of ergodic measures in $\mathcal{M}_Z(X,T)$ is important. However, sometimes $\mathcal{M}_Z(X,T)$ does not contain ergodic members even if it is nonempty. The following example from Kucherenko and Wolf[42] verifies this fact:

**Example 3.5** ($\mathcal{M}_Z(X,T)$ may not contain ergodic members[42])**.** *Let $a, b, c, d \in \mathbb{R}$ with $a < b < c < d$. Let $X = [a,b] \cup [c,d]$ and $T : X \to X$ continuous with entropy map $\mu \mapsto h(\mu)$ upper-semi-continuous. In addition, assume that $T([a,b]) \subset [a,b]$, $T([c,d]) \subset [c,d]$, $T(a) = a$, $T(d) = d$, and $h_{\text{top}}(T_{[a,b]}) = h_{\text{top}}(T_{[c,d]}) \neq 0$. For $w \in (b,c)$, let $Z = \{id_X - w\}$, the set $\mathcal{M}_Z(X,T)$ does not contain any ergodic measure.*

To ensure that $\mathcal{M}_Z(X,T)$ contains ergodic measures, we may assume that $\mathcal{M}_Z(X,T)$ satisfies the following property (E):

(E) If for any $\mu \in \mathcal{M}_Z(X,T)$, its ergodic decomposition is

$$\mu = \int_{\mathcal{M}^e(X,T)} \nu d\tau(\nu)$$

where $\mathcal{M}^e(X,T)$ denotes the set of all ergodic measures of $\mathcal{M}(X,T)$ and $\tau$ is a probability measure defined on the Borel subsets of $\mathcal{M}(X,T)$ and supported on $\mathcal{M}^e(X,T)$. Then $\nu \in \mathcal{M}_Z(X,T)$ for $\tau$- almost everywhere of $\nu$.

**Remark 3.1.** *If $\mathcal{M}_Z(X,T)$ satisfies property (E), then every extreme point in $\mathcal{M}_Z(X,T)$ is ergodic.*

**Example 3.6** (Some examples of $\mathcal{M}_Z(X,T)$ satisfies property (E))**.**

*(i)* $Z = \{0\}$ *(case of ergodic optimization[22])*

*(ii)* $Z = \left\{\psi \circ \pi - \int \psi d\nu : \psi \in \mathcal{C}(Y)\right\}$ *(given $\nu$ ergodic) (case of relative ergodic optimization)*

*(iii)* $Z = \left\{ \varphi \circ \pi_X - \int_X \varphi d\mu + \psi \circ \pi_Y - \int_Y \psi d\nu : \forall \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y) \right\}$ *(given $\mu$ and $\nu$ ergodic), where $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ are projection maps. (case of ergodic optimization over product dynamical system)*

Now we give some properties of $\mathcal{M}_Z(X, T)$:

**Theorem 3.2.** *If $T$ is a continuous transformation of a compact metric space $X$, $Z \subset \mathcal{C}(X)$ is a closed subset of $\mathcal{C}(X)$ such that $\mathcal{M}_Z(X, T)$ is nonempty, then*

*(i)* $\mathcal{M}_Z(X, T)$ *is a compact subset of* $\mathcal{M}(X)$.

*(ii)* $\mathcal{M}_Z(X, T)$ *is convex.*

*(iii) If $T$ is an ergodic measure-preserving transformation of $(X, \mathcal{B}, \mu)$, then $\mu$ is an extreme point of $\mathcal{M}_Z(X, T)$.*

*(iv) If $\mu, \nu \in \mathcal{M}_Z(X, T)$ are both ergodic and $\mu \neq \nu$ then they are mutually singular.*

*Proof.*

(i) As $\mathcal{M}(X, T)$ is compact in weak* topology, we claim that $\mathcal{M}_Z(X, T)$ is also compact in weak* topology: given a sequence of measures $\{\mu_n\}_n \subset \mathcal{M}(X, T)$, we can find a sebsequence $\{n_j\}_j$ s.t. $\mu_{n_j} \to \mu \in \mathcal{M}(X, T)$, and for any $\psi \in Z$:

$$\int \psi d\mu = \lim_j \int \psi d\mu_{n_j} = 0$$

so $\mu \in \mathcal{M}_Z(X, T)$, which means $\mathcal{M}_Z(X, T)$ is compact.

(ii) Convexity is obvious.

(iii) It is clear because the extreme points in $\mathcal{M}(X, T)$ are also the extreme points in $\mathcal{M}_Z(X, T)$, and together with the fact that the extreme points in $\mathcal{M}(X, T)$ are ergodic.

(iv) This is because of Theorem 1.2(iv).

□

If $Z_1 \subset \mathcal{C}(X)$ and $Z_2 \subset \mathcal{C}(X)$, define

$$Z_1 + Z_2 = \{h_1 + h_2 : h_1 \in Z_1, h_2 \in Z_2\}$$

then we have the following results:

**Lemma 3.3** (Additivity properties of $\mathcal{M}_Z(X,T)$).

*(i) If $Z_1$ and $Z_2$ are subspaces of $\mathcal{C}(X)$ satisfying property (E), then $Z_1 + Z_2$ is a subspace of $\mathcal{C}(X)$ that satisfies property (E).*

*(ii) Let $\mathcal{C}_T(X) = \{f \in \mathcal{C}(X) : f \circ T^{-1} = f\}$. If $Z_1$ and $Z_2$ are subspaces of $\mathcal{C}_T(X)$, then $Z_1 + Z_2 \subset \mathcal{C}_T(X)$.*

*(iii) If $Z_1$ and $Z_2$ are subspaces of $\mathcal{C}(X)$, then*

$$\mathcal{M}_{Z_1}(X,T) \cap \mathcal{M}_{Z_2}(X,T) = \mathcal{M}_{Z_1+Z_2}(X,T)$$

### 3.2 Existence and characterization of optimal plans for problem (I) and (II)

The existence of optimal solutions in problem (I) is ensured by extreme value theorem of upper-semi-continuous (u.s.c) function:

**Theorem 3.4.** *If $f \in \mathcal{C}(X)$ is upper-semi-continuous, i.e. $\forall x \in E$, a convex set*

$$\limsup_{x \to x_0} f(x) \leq f(x_0)$$

*then $f$ is bounded above and the supremum of $f$ is attained.*

The next result is very classical, see Jenkinson[22]:

**Proposition 3.5.** *If $\phi$ is u.s.c., the map $\mu \mapsto \int \phi d\mu$, $\mathcal{M}(X,T) \to [-\infty, +\infty)$ is also u.s.c.: if $\mu_n \to \mu$ in $\mathcal{M}(X,T)$, then $\int \phi d\mu \geq \limsup_{n\to\infty} \int \phi d\mu_n$.*

The constraint set $\mathcal{M}_Z(X,T) := \{\mu \in \mathcal{M}(X,T) : \int \psi d\mu = 0\ \forall \psi \in Z\}$ is convex. As if $\mu_1, \mu_2 \in \mathcal{M}_Z(X,T)$, then for any $\psi \in Z$:

$$\int \psi d[\alpha\mu_1 + (1-\alpha)\mu_2] = 0 \Rightarrow \alpha\mu_1 + (1-\alpha)\mu_2 \in \mathcal{M}_Z(X,T)$$

All the facts above ensure the existence of optimal solutions. If $\mathcal{M}_Z(X,T)$ is nonempty, define

$$\mathcal{M}(X,T;\phi) := \{\mu^* \in \mathcal{M}_Z(X,T) : \int \phi d\mu^* \geq \int \phi d\mu, \forall \mu \in \mathcal{M}_Z(X,T)\}$$

the collection of optimal plans for problem (I) (measures that maximizes $\int \phi d\mu$ in problem (I)).

**Theorem 3.6** (Existence, characterization of optimal plans for problem(I))**.** *If the settings are as section 2.1 and the restriction set $\mathcal{M}_Z(X,T)$ is nonempty, then the set of optimal solutions for problem(I), denoted by $\mathcal{M}(X,T;\phi)$, satisfies*

*(i) $\mathcal{M}(X,T;\phi)$ is nonempty, compact and convex.*

*(ii) If property (E) holds, then the extreme points of $\mathcal{M}(X,T;\phi)$ are ergodic.*

*(iii) If property (E) holds, then $\mathcal{M}(X,T;\phi)$ contains an ergodic measure.*

*(iv) If $f, g \in \mathcal{C}(X)$ and if there exists $c \in \mathbb{R}$ s.t. $f - g - c$ belongs to the closure of the set $\{h \circ T - h : h \in \mathcal{C}(X)\}$, then $\mathcal{M}(X,T;f) = \mathcal{M}(X,T;g)$.*

*Proof.*

(i) • Nonemptyness. Nonemptyness is established because by Proposition 3.4, $\mu \mapsto \int_X \phi d\mu$ is u.s.c., and the restriction set $\mathcal{M}_Z(X,T)$ is convex (by

Theorem 3.2 (ii)). Thus, by Theorem 3.3, $\mathcal{M}(X,T;\phi) := \{\mu \in \mathcal{M}_Z(X,T) : \mu \text{ maximizes } (I)\}$ is nonempty.

- Compactness. If $(\mu_n)_n$ is a sequence of optimal solutions in $\mathcal{M}_{\max}(\phi)$, then by the compactness of $\mathcal{M}_Z(X,T)$, there is a subsequence $(\mu_{n_k})_k \to \mu^*$ and $\mu^* \in \mathcal{M}_Z(X,T)$. Since

$$\lim_k \int \phi d\mu_{n_k} = \int \phi d\mu^*$$

therefore $\mu^* \in \mathcal{M}(X,T;\phi)$, and we proved the compactness of $\mathcal{M}(X,T;\phi)$.

- Convexity. For $\mu, \nu \in \mathcal{M}(X,T;\phi)$ and $\alpha \in (0,1)$,

$$\int \phi d[\alpha\mu + (1-\alpha)\nu] = \alpha \int \phi d\mu + (1-\alpha)\int \phi d\nu = \int \phi d\mu = \int \phi d\nu$$

so $\alpha\mu + (1-\alpha)\nu \in \mathcal{M}(X,T;\phi)$.

(ii) Assume that $\mu \in \mathcal{M}(X,T;\phi)$ is an extreme point of $\mu \in \mathcal{M}(X,T;\phi)$ and $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$ for $\alpha \in [0,1]$, $\mu_1, \mu_2 \in \mathcal{M}_Z(X,T)$. As

$$\int \phi d\mu = \int \phi d[\alpha\mu_1 + (1-\alpha)\mu_2] = \alpha \int \phi d\mu_1 + (1-\alpha)\int \phi d\mu_2$$

and $\int \phi d\mu \geq \int \phi d\mu_1$, $\int \phi d\mu \geq \int \phi d\mu_2$, so $\int \phi d\mu = \int \phi d\mu_1 = \int \phi d\mu_2$, which implies $\mu_1, \mu_2 \in \mathcal{M}(X,T;\phi)$. Then, since $\mu$ is extreme in $\mathcal{M}(X,T;\phi)$, we have $\mu = \mu_1 = \mu_2$ and $\mu$ is an extreme point of $\mathcal{M}_Z(X,T)$, so it is ergodic.

(iii) For $\mu \in \mathcal{M}(X,T;\phi)$, let $\mu = \int_{\mathcal{M}_Z^e(X,T)} m d\tau(m)$ be the ergodic decomposition of $\mu$, where $\mathcal{M}_Z^e(X,T)$ is the set of ergodic measures of $\mathcal{M}_Z(X,T)$. Then

$$\int \phi d\mu = \int \phi d\left(\int_{\mathcal{M}_Z^e(X,T)} m d\tau(m)\right) = \int_{\mathcal{M}_Z^e(X,T)} \left(\int \phi dm\right) d\tau(m)$$

Since $\int \phi dm \leq \int \phi d\mu$, so $m \in \mathcal{M}(X,T;\phi)$ for $\tau$ almost all $m$.

(iv) Suppose that there exists $c \in \mathbb{R}$ and $h \in \mathcal{C}(X)$ such that $f = g + c + h \circ T - h$. Then for any $\mu \in \mathcal{M}(X,T;f)$ we have

$$\begin{aligned} \mu \in \operatorname*{argmax}_{m \in \mathcal{M}_Z(X,T)} \int f dm &\implies \mu \in \operatorname*{argmax}_{m \in \mathcal{M}_Z(X,T)} \int (g + c + h \circ T - h) dm \\ &\implies \mu \in \operatorname*{argmax}_{m \in \mathcal{M}_Z(X,T)} \left( c + \int g dm + \int (h \circ T - h) dm \right) \\ &\implies \mu \in \operatorname*{argmax}_{m \in \mathcal{M}_Z(X,T)} \int g dm \end{aligned}$$

so $\mu \in \mathcal{M}(X,T;g)$ and thus $\mathcal{M}(X,T;f) \subseteq \mathcal{M}(X,T;g)$. The reverse direction can be proved similarly.

□

Let $\mathcal{R}(X,T;\phi)$ be the set of optimal plans for problem(II), i.e.

$$\mathcal{R}(X,T;\phi) = \{\mu^* \in \mathcal{M}_Z(X,T) : \int \phi d\mu^* + h(\mu^*) \geq \int \phi d\mu + h(\mu), \forall \mu \in \mathcal{M}_Z(X,T)\}$$

**Theorem 3.7** (Existence, characterization of optimal solutions for problem(II)). *If the settings are as section 2.1, the restriction set $\mathcal{M}_Z(X,T)$ is nonempty. Let $\phi \in \mathcal{C}(X)$. Then*

*(i) $\mathcal{R}(X,T;\phi)$ is convex.*

*(ii) If the entropy map is u.s.c. then $\mathcal{R}(X,T;\phi)$ is compact and nonempty.*

*(iii) If $h_{\mathrm{top}}(T) < \infty$ and $\mathcal{M}_Z(X,T)$ satisfies property (E), the extreme points of $\mathcal{R}(X,T;\phi)$ are ergodic.*

*(iv) If $h_{\mathrm{top}}(T) < \infty$ and $\mathcal{M}_Z(X,T)$ satisfies property (E), then $\mathcal{R}(X,T;\phi)$ contains an ergodic measure.*

*(v) If $f, g \in \mathcal{C}(X)$ and if there exists $c \in \mathbb{R}$ s.t. $f - g - c$ belongs to the closure of the set $\{h \circ T - h : h \in \mathcal{C}(X)\}$, then $\mathcal{R}(X, T; f) = \mathcal{R}(X, T; g)$.*

Before we start to prove Theorem 3.6, two necessary lemmas will be given first: the first lemma given below shows that the entropy map $\mu \mapsto h(\mu)$ is affine in our settings (in Section 2.1):

**Lemma 3.8.** *Suppose that the settings are as in Section 2.1. The entropy map of $T$ is affine, i.e., if $\mu, \nu \in \mathcal{M}(X, T)$ and $\alpha \in [0, 1]$ then $h(\alpha\mu + (1 - \alpha)\nu) = \alpha h(\mu) + (1 - \alpha)h(\nu)$.*

For proofs of Lemma 3.7 please refer to Walters[5] p.183.

The second lemma builds a relationship between ergodic decomposition and entropy.

**Lemma 3.9** (Walters[5] Theorem 8.4)**.** *Let $T : X \to X$ be a continuous map of a compact metrizable space. If $\mu \in \mathcal{M}(X, T)$ and its ergodic decomposition is $\mu = \int_{\mathcal{M}^e(X,T)} m d\tau(m)$, then we have*

*(i) If $\mathcal{P}$ is a finite partition of $(X, \mathcal{B}(X))$ then $h(\mu, \mathcal{P}) = \int_{\mathcal{M}^e(X,T)} h(m, \mathcal{P}) d\tau(m)$.*

*(ii) $h(\mu) = \int_{\mathcal{M}^e(X,T)} h(m) d\tau(m)$*

Now we are ready to prove Theorem 3.6:

*Proof of Theorem 3.6.*

(i) It is clear. If $\mu_1, \mu_2 \in \mathcal{R}(X, T; \phi)$, then for any $\alpha \in (0, 1)$, $\alpha\mu_1 + (1 - \alpha)\mu_2 \in \mathcal{R}(X, T; \phi)$ since by Lemma 3.7 we have

$$\begin{aligned}
&\int \phi d[\alpha\mu_1 + (1 - \alpha)\mu_2] + h(\alpha\mu_1 + (1 - \alpha)\mu_2) \\
=&\alpha \left( \int \phi d\mu_1 + h(\mu_1) \right) + (1 - \alpha) \left( \int \phi d\mu_2 + h(\mu_2) \right) \\
=&\int \phi d\mu_1 + h(\mu_1) = \int \phi d\mu_2 + h(\mu_2)
\end{aligned}$$

(ii) The nonemptyness of $\mathcal{R}(X,T;\phi)$ is because the map $\mu \mapsto \int \phi d\mu + h(\mu)$ is u.s.c and $\mathcal{M}_Z(X,T)$ is convex. To prove the compactness, assume $(\mu_n)_n \in \mathcal{R}(X,T;\phi)$ a sequence of optimal solutions, then we can find a subsequence $(\mu_{n_k})_k$ s.t. $\mu_{n_k} \to \mu^* \in \mathcal{M}_Z(X,T)$ by the compactness of $\mathcal{M}_Z(X,T)$ and

$$\limsup_k \int \phi d\mu_{n_k} + h(\mu_{n_k}) \geq \int \phi d\mu + h(\mu)$$

so $\mu \in \mathcal{R}(X,T;\phi)$.

(iii) Since the map $\mu \mapsto \int \phi d\mu + h(\mu)$ is affine, proof can be followed in the same way as Theorem 3.5(ii).

(iv) By nonemptyness, we can always select a member $\mu$ of $\mathcal{R}(X,T;\phi)$. Let $\mu = \int_{\mathcal{M}_Z^e(X,T)} m d\tau(m)$ be the ergodic decomposition of $\mu$, by Lemma 3.8

$$\begin{aligned}
\int \phi d\mu + h(\mu) &= \int \phi d\left(\int_{\mathcal{M}_Z^e(X,T)} m d\tau(m)\right) + h\left(\int_{\mathcal{M}_Z^e(X,T)} m d\tau(m)\right) \\
&= \int_{\mathcal{M}_Z^e(X,T)} \left(\int \phi dm\right) d\tau(m) + \int_{\mathcal{M}_Z^e(X,T)} h(m) d\tau(m) \\
&= \int_{\mathcal{M}_Z^e(X,T)} \left(\int \phi dm + h(m)\right) d\tau(m)
\end{aligned}$$

Since $\int \phi dm + h(m) \leq \int \phi d\mu + h(\mu)$, so $m \in \mathcal{R}(X,T;\phi)$ for $\tau$ almost all $m$.

(v) Suppose that there exists $c \in \mathbb{R}$ and $h \in \mathcal{C}(X)$ such that $f = g + c + h \circ T - h$, then $\int f d\mu = \int g d\mu + c$, so

$$\int f d\mu + h(\mu) = \int g d\mu + h(\mu) + c$$

Thus $\mathcal{R}(X,T;f) = \mathcal{R}(X,T;g)$.

□

# CHAPTER 4: UNIQUENESS OF OPTIMAL PLANS

In this chapter, the uniqueness problem we studied is that under what conditions the optimal plan of problem (I) and (II) is unique? In the framework of problem (I), Jenkinson[22] got the uniqueness result for ergodic optimization problem. There are several uniqueness outcomes for different problems including general thermodynamic formalism[36, 37], Gibbs states[30, 27, 28, 29] and Gibbs measure[9, 36], relative equilibrium states[49] and localized equilibrium states problem[42] in the framework of problem (II), among them the most classical one is Bowen's uniqueness result[36] for general thermodynamic formalism.

## 4.1 Uniqueness results of problem (I)

From Theorem 3.6(i), the optimal solution set $\mathcal{M}(X,T;\phi)$ is always nonempty. Generally uniqueness cannot be ensured unless the constraint set $\mathcal{M}_Z(X,T)$ is a singleton. An extreme example is when $\phi$ is a constant and $Z=\{0\}$, then every member $\mu\in\mathcal{M}(X,T)$ maximizes $\int\phi d\mu$. So here we would like to say that a "typical" function $\phi$ *does* have a unique maximizing measure. That is, in a given function space $E\subset\mathcal{C}(X)$, can we find an open, dense subset $E'$ such that for all $\phi\in E'$, $\mathcal{M}(X,T;\phi)$ is a singleton? Or in other words, the set

$$\mathcal{U}_Z(E)=\{\phi\in E:\text{there is a unique measure in }\mathcal{M}_Z(X,T)\text{ maximizes }\int\phi d\mu\}$$

is open and dense in $E$. The next uniqueness result is generalized from ergodic optimization by Jenkinson[22].

**Theorem 4.1.** *Let $X$ be a compact metric space, and $T:X\to X$ a continuous map which has only finitely many ergodic invariant measures. Let $E$ be a topological vector* *space which is densely and continuously embedded in $\mathcal{C}(X)$. If $Z$ is closed in $\mathcal{C}(X)$ and satisfies property (E), then $\mathcal{U}_Z(E)$ is open and dense in $E$.*

*Proof.* Let $\{\mu_1, \cdots, \mu_N\}$ be the ergodic measures for $T$, and define

$$F_i = \{\phi \in E : \mu_i \text{ maximizes } \int \phi d\mu \text{ over } \mathcal{M}_Z(X,T)\}$$

for each $1 \leq i \leq N$, then

- step 1: By Theorem 3.5(iii), if problem (I) has unique optimal plan, then it must be ergodic. So $\mathcal{U}_Z(E)^c$ can be expressed as

$$\mathcal{U}_Z(E)^c = \bigcup_{i<j} F_i \cap F_j$$

- step 2: $F_i$ is closed for each $i = 1, \cdots, N$. Suppose that $\{\phi_\alpha\}$ is a net in $F_i$, with $\phi_\alpha \to \phi$ in $E$, then $\phi \in F_i$.

- step 3: each $F_i \cap F_j$ is hollow. Since $E$ is densely embedded in $\mathcal{C}(X)$, for any $i < j$ there exists $g = g_{ij} \in E$ s.t. $\int g d\mu_i \neq \int g d\mu_j$. If $\phi \in F_i \cap F_j$ then for every $\varepsilon > 0$ then function $f + \varepsilon g \notin F_i \cap F_j$. Therefore $F_i \cap F_j$ has empty interior whenever $1 \leq i < j \leq N$, so $\mathcal{U}_Z(E)$ is dense in $E$.

□

The assumption that $T$ has finite many ergodic measures is very strong. So we strive to get an uniqueness result for a more general $T$. For a specific property $\mathcal{P}$, we want to find a *residual* set $E'$ s.t. every $f \in E'$ has the property $\mathcal{P}$. A residual set is the set contains a countable intersection of open dense subsets. We say that $\mathcal{P}$ is a *generic* property if there is some residual set $E'$ s.t. every member of $E'$ has the property $\mathcal{P}$.

**Proposition 4.2.** *Let $T : X \to X$ be a continuous map on a compact metric space, $Z$ is closed in $\mathcal{C}(X)$. Let $E$ be a topological vector space which is densely and continuously embedded in $\mathcal{C}(X)$. Then $\mathcal{U}_Z(E)$ is a countable intersection of open and dense subsets of $E$.*

*If moreover $E$ is a Baire space, then $\mathcal{U}_Z(E)$ is dense in $E$.*

The sketch of Proposition 4.2's proof is given as below, for detailed proof we just need to follow Jenkinson[22]'s proof of Theorem 3.2.

*Proof.* Define $\alpha(g|\phi) := \max_{\mu \in \mathcal{M}(X,T;\phi)} \int g d\mu$ and $\mathcal{M}(X,T;g|\phi) = \{\mu \in \mathcal{M}(X,T;\phi) : \int g d\mu = \alpha(g|\phi)\}$.

- Step 1: Show that for any $\phi, g \in \mathcal{C}(X)$,

$$\left\{\int g d\mu : \mu \in \mathcal{M}(X,T;\phi+\varepsilon g)\right\} \to \alpha(g|\phi) \quad \text{as } \varepsilon \to 0$$

  The proof is, since the set $\left\{\int g d\mu : \mu \in \mathcal{M}(X,T;\phi+\varepsilon g)\right\} = [a_\varepsilon^-, a_\varepsilon^+]$, so it is enough to show that any $a_\varepsilon$ in this set we have $\lim_{\varepsilon\to 0} a_\varepsilon = \alpha(g|\phi)$. Writing $a_\varepsilon = \int g dm_\varepsilon$ for some $m_\varepsilon \in \mathcal{M}(X,T;\phi+\varepsilon g)$, it is enough to show that any weak* accumulation point of $m_\varepsilon$ belongs to $\mathcal{M}(X,T;g|\phi)$.

- Step 2: Since $\mathcal{C}(X)$ separable and $E$ is densely embedded, then there is a countable subset of $E$ which is dense in $\mathcal{C}(X)$ (we can assume $\{\phi_n : n \in \mathbb{N}\}$ a dense subset of $\mathcal{C}(X)$ and then we can choose $e_{n,i} \in E$ s.t. $e_{n,i} \to f_n$ as $i \to \infty$, $\{e_{n,i} : (n,i) \in \mathbb{N}^2\}$ is dense in $\mathcal{C}(X)$). If $\{g_i\}_{i=1}^\infty$ denotes this countable subset of $E$.

- Step 3: $\mathcal{M}(X,T;\phi)$ is a singleton if and only if $M_i(\phi) := \{\int g_i d\mu : \mu \in \mathcal{M}(X,T;\phi)\}$ is a singleton for each $i$. Define $E_{i,j} := \{\phi \in E : |M_i(\phi)| \geq 1/j\}$, then

$$\mathcal{U}_Z(E)^c = \{\phi \in E : \exists i \in \mathbb{N}, |M_i(\phi)| > 0\} = \bigcup_{j=1}^\infty \bigcup_{i=1}^\infty E_{i,j}$$

- Step 4: We need to show that each $E_{i,j}$ is closed and has empty interior, thus $\mathcal{U}_Z(E)$ is a countable intersection of open and dense subsets of $E$.

  – Closedness: Let $\{\phi_\alpha\}$ be a net in $E_{i,j}$ and $\phi_\alpha \to \phi \in E$. $M_i(\phi_\alpha) = [\int g_i d\mu_\alpha^-, \int g_i d\mu_\alpha^+]$ for $\mu_\alpha^\pm \in \mathcal{M}(X,T;\phi_\alpha) \subset \mathcal{M}_Z(X,T)$. The weak* compactness of $\mathcal{M}_Z(X,T)$ means $\exists \mu^+, \mu^-$ s.t. $\mu_\alpha^+ \to \mu^+, \mu_\alpha^- \to \mu^-$, which implies
  $$\int g_i d\mu_\alpha^+ \to \int g_i d\mu^+, \quad \int g_i d\mu_\alpha^- \to \int g_i d\mu^-$$
  Need to show $\mu^\pm \in \mathcal{M}(X,T;\phi)$: $\int \phi_\alpha d\mu_\alpha^- = \int (\phi_\alpha - \phi) d\mu_\alpha^- + \int \phi d\mu_\alpha^- \to \int \phi d\mu^-$, and for any $m \in \mathcal{M}_Z(X,T)$, $\int \phi_\alpha d\mu_\alpha^- \geq \int \phi_\alpha dm$. From these two facts, for any $m \in \mathcal{M}_Z(X,T)$, $\int \phi d\mu^- \geq \int \phi dm$. So $\mu^-, \mu^+ \in \mathcal{M}(X,T;\phi)$ and $|\int g_i d\mu^+ - \int g_i d\mu^-| \geq 1/j$, thus $\phi \in E_{i,j}$.

  – $E_{i,j}$ has empty interior: Let $\phi \in E_{i,j}$ be arbitrary, by step 1, $M_i(\phi + \varepsilon g_i) = \{\int g_i d\mu : \mu \in \mathcal{M}(X,T;\phi+\varepsilon g_i)\} \to \alpha(g_i|\phi)$ as $\varepsilon \to 0$, so for some $\varepsilon > 0$ we have $|M_i(\phi + \varepsilon g_i)| < 1/j$, which implies $\phi + \varepsilon g_i \notin E_{i,j}$ for some $\varepsilon > 0$ sufficiently small, so $\phi$ is not an interior point of $E_{i,j}$.

$\square$

## 4.2 Uniqueness results of problem (II)

Bowen's classical uniqueness result[36] for thermodynamic formalism is useful and it will be introduced firstly. A special condition given by Bowen that ensures uniqueness is called 'specification', its definition is as below:

**Definition 4.1** (Specification)**.** *A homeomorphism $T$ satisfies specification if for each $\delta > 0$ there is an integer $p(\delta)$ for which the following is true: if $I_1, \cdots, I_n$ are intervals of integers in $[a,b]$ with $d(I_i, I_j) \geq p(\delta)$ fir $i \neq j$ and $x_1, \cdots, x_n \in X$, then there is a point $x \in X$ with $T^{b-a+p(\delta)}(x) = x$ and $d(T^k(x), T^k(x_i)) < \delta$ for $k \in I_i$.*

and we define

$$V(T) := \{\phi \in \mathcal{C}(X) : \forall \varepsilon > 0, \exists K, d_n(x,y) \leq \varepsilon \Rightarrow |S_n\phi(x) - S_n\phi(y)| \leq K\}$$

Now we are ready to introduce Bowen's result:

**Theorem 4.3** (Bowen)**.** *Let $T : X \to X$ be an expansive homeomorphism of a compact metric space satisfying specification. Then each $\phi \in V(T)$ has a unique equilibrium state $\mu_\phi$.*

We studied a very special case of problem (II), in this case we consider product space $X \times Y$, $Z = \left\{\varphi \circ \pi_X - \int_X \varphi d\mu + \psi \circ \pi_Y - \int_Y \psi d\nu : \forall \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y)\right\}$ where $\phi \in \mathcal{C}(X \times Y)$, $\mu$ and $\nu$ are given invariant ergodic measures. Restriction set is actually a set of joinings between two dynamical systems $J(\mu, \nu)$.

**Theorem 4.4.** *In problem (II), let $(X,T)$ and $(Y,S)$ be shift spaces, and let $\mu \in \mathcal{M}(X,T), \nu \in \mathcal{M}(Y,S)$ be i.i.d. invariant measures. We specify*

$$Z = \left\{\varphi \circ \pi_X - \int_X \varphi d\mu + \psi \circ \pi_Y - \int_Y \psi d\nu : \forall \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y)\right\}$$

*The continuous function $\phi : X \times Y \to \mathbb{R}$ is given by $\phi(\underline{x}, \underline{y}) = \phi_0(x_0, y_0)$. That is, the function value of $\phi(\underline{x}, \underline{y})$ depends on the first digit of the sequences $\underline{x}, \underline{y}$.*

*Then problem (II) is equivalent to equilibrium states over joinings:*

$$\sup_{\lambda \in J(\mu,\nu)} \left\{\int \phi d\lambda + h(\lambda)\right\}$$

*and it has unique equilibrium states $\lambda = (\lambda_0)^{\otimes \mathbb{N}}$, where $\lambda_0$ is the optimal coupling of $\mu_0$ and $\nu_0$, i.e.*

$$\lambda_0 = \operatorname*{argmax}_{\lambda_0' \in \Pi(\mu_0,\nu_0)} \int \phi_0 d\lambda_0' + H_{\lambda_0'}(\mathcal{P})$$

*Proof.* By subadditivity of entropy, if $\mathcal{P} = \{P_{1,1}, \cdots, P_{n,m}\}$ is defined as

$$P_{i,j} = \{(x, y) \in X \times Y : (x_0, y_0) = (i, j)\}$$

$\mathcal{P}$ is the partition of the first digit, note that $\mathcal{P}$ is a generating partition, so

$$h(\lambda, \mathcal{P}) = h(\lambda) = \inf_{k \geq 1} \frac{1}{k} H(\mathcal{P}^k) \leq H(\mathcal{P})$$

where $\mathcal{P}^k = \bigvee_{i=0}^{k-1} (T \times S)^{-i} \mathcal{P}$.

So

$$\sup_{\lambda \in J(\mu,\nu)} \left\{ \int \phi d\lambda + h(\lambda) \right\} \leq \sup_{\lambda_0 \in \Pi(\mu_0,\nu_0)} \left\{ \int \phi_0 d\lambda_0 + H(\mathcal{P}) \right\}$$

that is to say, under this special setting we only need to consider the marginal on the first digit.

Note that when $\lambda = (\lambda_0)^{\otimes \mathbb{N}}$, then

$$\int \phi_0 d\lambda_0 + h(\lambda) = \int \phi_0 d\lambda_0 + H(\mathcal{P})$$

this is because of

$$h(\lambda) = \lim_{k \to \infty} (1/k) H(\mathcal{P}^k)$$

and since $|\mathcal{P}^k| = (n \times m)^k$:

$$\begin{aligned}
H(\mathcal{P}^k) &= -\sum_{A \in \mathcal{P}^k} \lambda(A) \log \lambda(A) \\
&= -\sum_{\{[i_l, j_l]\}_{l=0}^{k-1}} \lambda_0[i_0, j_0]\lambda_0[i_1, j_1] \cdots \lambda_0[i_l, j_l] \log(\lambda_0[i_0, j_0]\lambda_0[i_1, j_1] \cdots \lambda_0[i_l, j_l]) \\
&= -\sum_{[i,j] \in \mathcal{P}} \lambda_0[i, j] \log \lambda_0[i, j] = H(\mathcal{P})
\end{aligned}$$

Now we should prove the uniqueness. Assume we have another $\lambda \in J(\mu, \nu)$ s.t.

$$\int \phi_0 d\lambda_0 + h_\lambda(T \times S) = \int \phi_0 d\lambda_0 + H(\mathcal{P})$$

that is to say,

$$H(\mathcal{P}) = h(\lambda) \leq \frac{1}{n} H(\mathcal{P}^n) \leq H(\mathcal{P})$$

so we have

$$nh(\lambda) = H(\mathcal{P}^n) = nH(\mathcal{P}) \quad \forall n \in \mathbb{N}$$

which means $\lambda$'s marginal on each position is $\lambda_0$. Therefore, $\lambda = (\lambda_0)^{\otimes \mathbb{N}}$ is the unique optimal plan. $\square$

# CHAPTER 5: LAGRANGIAN APPROACHES TO OPTIMIZATION

In this chapter we will concentrate on problem (II). In the first section, we give clear definitions of the pressure function and the tangent functional in the framework of problem (II). In the second section, we will introduce our Lagrangian method and apply it to a special case of problem (II) in the last section of this chapter.

## 5.1 Pressure function and tangent functional

First of all we define a maximal restricted pressure function $P_Z : \mathcal{C}(X) \to \mathbb{R}$. For $\phi \in \mathcal{C}(X)$, let

$$P_Z(T,\phi) = \sup_{\mu \in \mathcal{M}_Z(X,T)} \int \phi d\mu + h(\mu)$$

According to Walters[5] Theorem 9.7, the topological pressure function $P(T,\cdot)$ has some nice properties. The following result shows that the map $P_Z : \mathcal{C}(X) \to \mathbb{R}$ has some similar properties:

**Lemma 5.1.** *Assume that $P_Z(T,\phi) < \infty$ for all $\phi \in \mathcal{C}(X)$,*

*(i) $f \le g$ implies $P_Z(T,f) \le P_Z(T,g)$.*

*(ii) $P_Z : \mathcal{C}(X) \to \mathbb{R}$ is 1-Lipschitz (and thus continuous).*

*(iii) $P_Z : \mathcal{C}(X) \to \mathbb{R}$ is convex.*

*(iv) $P_Z(T, f + c) = P_Z(T, f) + c$ for all $f \in \mathcal{C}(X)$ and $c \in \mathbb{R}$.*

*(v) $P_Z(f + g \circ T - g) = P_Z(f)$ for all $f, g \in \mathcal{C}(X)$.*

*Proof.*

(i) Let $P_Z(T,f) = \int f d\mu_f + h(\mu_f)$, we have

$$P_Z(T,f) \leq \int g d\mu_f + h(\mu_f) \leq \sup_{\mu \in \mathcal{M}_Z(X,T)} \int g d\mu + h(\mu) = P_Z(T,g)$$

(ii) Let $P_Z(T,f) = \int f d\mu_f + h(\mu_f)$ and $P_Z(T,g) = \int g d\mu_g + h(\mu_g)$, then

$$\begin{aligned} P_Z(T,f) - P_Z(T,g) &\geq \left( \int f d\mu_f + h(\mu_f) \right) - \left( \int g d\mu_f + h(\mu_f) \right) \\ &= \int (f-g) d\mu_f \geq -\|f-g\| \end{aligned}$$

and

$$\begin{aligned} P_Z(T,f) - P_Z(T,g) &\leq \left( \int f d\mu_g + h(\mu_g) \right) - \left( \int g d\mu_g + h(\mu_g) \right) \\ &= \int (f-g) d\mu_g \leq \|f-g\| \end{aligned}$$

So $|P_Z(T,f) - P_Z(T,g)| \leq \|f-g\|$, $P_Z(T,\cdot)$ is 1-Lipschitz.

(iii) For $f, g \in \mathcal{C}(X)$ and $\alpha \in [0,1]$:

$$\begin{aligned} P_Z(T, \alpha f + (1-\alpha) g) &= \sup_{\mu \in \mathcal{M}_Z(X,T)} \int [\alpha f + (1-\alpha) g] d\mu + h(\mu) \\ &= \sup_{\mu \in \mathcal{M}_Z(X,T)} \left[ \alpha \left( \int f d\mu + h(\mu) \right) + (1-\alpha) \left( \int g d\mu + h(\mu) \right) \right] \\ &\leq \alpha \sup_{\mu \in \mathcal{M}_Z(X,T)} \left( \int f d\mu + h(\mu) \right) + (1-\alpha) \sup_{\mu \in \mathcal{M}_Z(X,T)} \left( \int g d\mu + h(\mu) \right) \\ &= \alpha P_Z(T,f) + (1-\alpha) P_Z(T,g) \end{aligned}$$

complexity is proved.

(iv) Obviously, since $\int (f+c) d\mu = \int f d\mu + c$.

(v) Obviously, since for any invariant measure $\mu$ and any $g \in \mathcal{C}(X)$, $\int (g \circ T - g) d\mu = 0$.

□

According to Walters[5] Theorem 9.15, equilibrium states for any $\phi \in \mathcal{C}(X)$ can be connected to tangent functional to $P(T, \cdot)$ at $\phi$. Similarly, In Walters[13], any relative equilibrium state can be considered as a tangent functional to relativised pressure function at $\phi$. In the following we will give the definition of tangent functional and generalize these results.

**Definition 5.1** (Tangent functional)**.** *Let $T : X \to X$ be a continuous map of a compact metrizable space with $h_{\text{top}}(T) < \infty$ and let $f \in \mathcal{C}(X)$. Given $Z \subset \mathcal{C}(X)$ a closed subset. A tangent functional to restricted pressure function $P_Z(T, \cdot)$ at $f$ is a finite signed measure $\mu : \mathcal{B}(X) \to \mathbb{R}$ s.t. $P_Z(T, f+g) - P_Z(T, f) \geq \int g d\mu, \forall g \in \mathcal{C}(X)$.*

Let $\mathcal{W}_\phi$ denote the collection of all tangent functionals to $P_Z(T, \cdot)$ at $\phi$. The Hahn-Banach theorem implies that $\mathcal{W}_\phi$ is nonempty.

**Theorem 5.2.** *Let $Z \subset \mathcal{C}(X)$ and $\phi \in \mathcal{C}(X)$.*

*(i) All tangent functionals to $P_Z$ at $\phi$ are in $\mathcal{M}_Z(X, T)$.*

*(ii) If $\mu_0 \in \mathcal{M}_Z(X, T)$ and $\int \phi d\mu_0 + h(\mu_0) = P_Z(\phi, T)$, then $\mu_0$ is a tangent functional to $P_Z$ at $\phi$.*

*(iii) Suppose the entropy map $\mu \mapsto h(\mu)$ restricted to $\mathcal{M}_Z(X, T)$, is upper-semicontinuous at all tangent functionals to $P_Z$ at $\phi$. If $\mu$ is a signed measure on $X$, the $\mu$ is a tangent functional to $P_Z$ at $\phi$ iff $\mu \in \mathcal{M}_Z(X, T)$ and $\int \phi d\mu + h(\mu) = P_Z(\phi, T)$.*

*Proof.*

(i) Let $\mu \in \mathcal{W}_\phi$, first we show that $\mu$ takes only nonnegative values. Let $g \in \mathcal{C}(X)$ be a nonnegative function, we have

$$\begin{aligned} \int g d\mu = -\int (-g) d\mu &\geq P_Z(\phi) - P_Z(\phi - g) \\ &\geq P_Z(\phi) - [P_Z(\phi) - \inf g] \geq 0 \end{aligned}$$

so $\int gd\mu \geq 0$ for $g \geq 0$. We have $\mu(X) = 1$, because if $n \in \mathbb{Z}$ then

$$\int nd\mu \leq P_Z(\phi + n, T) - P_Z(\phi) = n$$

if $n \geq 1$, $\mu(X) \leq 1$, if $n \leq -1$, $\mu(X) \geq 1$, so $\mu(X) = 1$. Finally we need to show $\mu \circ T^{-1} = \mu$. If $n \in \mathbb{Z}$ and $g \in \mathcal{C}(X)$:

$$n \int (g \circ T - g)d\mu \leq P_Z(\phi + n(g \circ T - g), T) - P_Z(\phi, T) = 0$$

Therefore $\int g \circ T d\mu = \int gd\mu$, that is $\mu \circ T^{-1} = \mu$.

(ii) For any $\psi \in \mathcal{C}(X)$:

$$\begin{aligned}
P_Z(\phi + \psi, T) - P_Z(\phi, T) &= \sup_{\mu \in \mathcal{M}_Z(X,T)} \left( \int (\phi + \psi)d\mu + h(\mu) \right) - \left( \int \phi d\mu_0 + h(\mu_0) \right) \\
&\geq \left( \int (\phi + \psi)d\mu_0 + h(\mu_0) \right) - \left( \int \phi d\mu_0 + h(\mu_0) \right) \\
&\geq \int \psi d\mu_0
\end{aligned}$$

therefore $\mu_0 \in \mathcal{W}_\phi$.

(iii) It remains to show that if $\mu \in \mathcal{M}_Z(X, T) \cap \mathcal{W}_\phi$ then $P_Z(\phi, T) = \int \phi d\mu + h(\mu)$. As for any $g \in \mathcal{C}(X)$:

$$P_Z(\phi + g, T) - P_Z(\phi) \geq \int gd\mu$$

so

$$P_Z(\phi + g, T) - \int (\phi + g)d\mu \geq P_Z(\phi) - \int \phi d\mu$$

so by the duality result in Theorem 4.1

$$h(\mu) \geq P_Z(\phi) - \int \phi d\mu$$

then we have $P_Z(\phi, T) = \int \phi d\mu + h(\mu)$.

$\square$

## 5.2 Lagrangian approach

The Lagrangian approach to optimal plan relies on the following special cases of Isreal's Theorem:

**Lemma 5.3** (Isreal[50]). *Let $X$ be a compact metric space, and let $W : \mathcal{C}(X) \to \mathbb{R}$ be convex and continuous. Suppose $\mathcal{N}$ is a closed convex cone in $\mathcal{C}(X)$ with apex $0$. Let $f_0 \in \mathcal{C}(X)$ and let $\mu_0 \in \mathcal{M}(X)$ be $W$-bounded (i.e., $\exists c \in \mathbb{R}$ such that $\int g d\mu_0 \leq c + W(g)$, $\forall g \in \mathcal{C}(X)$). For each $\varepsilon > 0$ there exists $f \in \mathcal{C}(X)$ and $\mu \in \mathcal{C}(X)^*$ such that*

*(i) $f \in f_0 + \mathcal{N}$,*

*(ii) $\mu$ is a tangent functional to $W$ at $f$,*

*(iii) $\int h d\mu \geq \int h d\mu_0 - \varepsilon \|h\|$, $\forall h \in \mathcal{N}$.*

By Isreal[1], if $\mathcal{N}$ is a closed linear subspace of $\mathcal{C}(X)$, then for $h \in \mathcal{N}$, $(-h) \in \mathcal{N}$ and we obtain

$$\left| \int h d\mu - \int h d\mu_0 \right| \leq \varepsilon \|h\| \implies \|(\mu - \mu_0)|_{\mathcal{N}}\| \leq \varepsilon$$

The following result is the main theorem of this chapter:

**Theorem 5.4** (Lagrangian approach to optimization). *Suppose that the settings are as section 1.2.1. Let $\phi \in \mathcal{C}(X)$. If $Z \subset \mathcal{C}(X)$ is a closed linear subspace s.t. $\mathcal{M}_Z(X, T)$ is nonempty and satisfies property (E). Then there exists a measure $\mu \in \mathcal{M}(X, T)$ and $\chi \in Z$ such that $\mu$ is an equilibrium state of $\phi + \chi$ and $|\int h d\mu| \leq \varepsilon \|h\|$ for all $h \in Z$.*

*Proof.* Since $\mathcal{M}_Z(X,T)$ is nonempty so let $\mu_0$ be any element of $\mathcal{M}_Z(X,T)$. The function $W : \mathcal{C}(X) \to \mathbb{R}$ is given by

$$W(f) = \sup_{\mu \in \mathcal{M}(X,T)} \int f d\mu + h(\mu)$$

which is the pressure function so it is convex and continuous. Then $\mu_0$ is $W$-bounded because $\exists c = 0$ s.t.

$$\int g d\mu_0 \leq c + W(g) \ \forall g \in \mathcal{C}(X)$$

Let $\mathcal{N} = Z$, $f_0 = \phi$, by Isreal's theorem, there exists an element $\chi \in Z$ and there is $\mu \in \mathcal{M}(X,T)$ that is a tangent functional to $W$ defined above at $\phi + \chi$. By Theorem 5.2(iii) $\mu$ is an equilibrium state of $\phi + \chi$ for some $\chi \in Z$ and for any $\varepsilon > 0$,

$$\left| \int h d\mu - \int h d\mu_0 \right| \leq \varepsilon \|h\|, \quad \forall h \in Z$$

since $h \in Z$ and $\mu_0 \in \mathcal{M}_Z(X,T)$ so $\mu_0(h) = \int h d\mu_0 = 0$, we get

$$\left| \int h d\mu \right| \leq \varepsilon \|h\|, \quad \forall h \in Z$$

□

Now we consider a special case: if $(X,T)$ and $(Y,S)$ are topological dynamical systems ($X, Y$ are compact metrizable spaces, $T : X \to X$ and $S : Y \to Y$ are continuous transformations), we can construct a product dynamical system $(X \times Y, T \times S)$, where $T \times S : X \times Y \to X \times Y$ is given by $(T \times S)(x,y) = (T(x), S(y))$ for $(x,y) \in X \times Y$.

**Theorem 5.5.** *If $(X \times Y, T \times S)$ is a product dynamical system as above, $\phi \in \mathcal{C}(X \times$*

*$Y$). Given $\mu \in \mathcal{M}^e(X,T), \nu \in \mathcal{M}^e(Y,S)$ and let*

$$Z = \left\{ \varphi \circ \pi_X - \int_X \varphi d\mu + \psi \circ \pi_Y - \int_Y \psi d\nu : \forall \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y) \right\}$$

*Then $\mathcal{M}_Z(X \times Y, T \times S) = J(\mu,\nu)$, and there exists $(\hat{\varphi}, \hat{\psi}) \in \mathcal{C}(X) \times \mathcal{C}(Y)$ and $\hat{\lambda} \in J(\mu,\nu)$ s.t.*

$$\begin{aligned} & \hat{\lambda} \in \underset{\lambda \in \mathcal{M}(X \times Y, T \times S)}{\operatorname{argmax}} \int (\phi + \hat{\varphi} \circ \pi_X + \hat{\psi} \circ \pi_Y) d\lambda + h(\lambda) \\ \iff & \hat{\lambda} \in \underset{\lambda \in J(\mu,\nu)}{\operatorname{argmax}} \int \phi d\lambda + h(\lambda) \end{aligned}$$

Before we apply Lagrangian approach (Theorem 5.4) to prove Theorem 5.5, we have to verify that $\tilde{Z} = \{\varphi \circ \pi_X + \psi \circ \pi_Y : \forall \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y)\}$ is a closed linear subspace in $\mathcal{C}(X \times Y)$, where $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ are projection mappings. It is obvious that $Z \subset \mathcal{C}(X \times Y)$ is a linear subspace, now we will show that $Z$ is closed. A useful lemma is as follows:

**Lemma 5.6** (Kober)**.** *Let $X$ be a Banach space, $M, N$ be two closed subspaces of $X$ and $M \cap N = \{0\}$. Then $M + N$ is closed in $X$ if and only if there exists a constant $A > 0$ such that for all $x \in M$ and $y \in N$ we have $\|x\| \leq A \cdot \|x + y\|$.*

**Proposition 5.7.** *$\tilde{Z} = \{\varphi \circ \pi_X + \psi \circ \pi_Y : \varphi \in \mathcal{C}(X)$ and $\psi \in \mathcal{C}(Y)\}$ is closed.*

*Proof.* Firstly, define

$$L_X = \{\varphi \circ \pi_X : \varphi \in \mathcal{C}(X)\}, \quad L_X = \{\psi \circ \pi_Y : \psi \in \mathcal{C}(Y)\}$$

then $\mathcal{N} = L_X + L_Y$ and $L_X \cap L_Y = \{c : c \in \mathbb{R}\}$, so Lemma 4.20 cannot be applied directly.

Now we first consider the subspace: let $\tilde{Z} = \{\varphi \circ \pi_X + \psi \circ \pi_Y : \varphi \in \mathcal{C}(X)$ and $\psi \in \mathcal{C}(Y)\}$ is a subspace of $\mathcal{C}(X \times Y)$, $L_X = \{\varphi \circ \pi_X : \varphi \in \mathcal{C}(X)\}$, $D_Y = \{\psi \circ \pi_Y :$

$\psi \in \mathcal{C}(Y)\}$ and $R = \{c : c \in \mathbb{R}\}$ the set of all constant functions, which is closed in $\mathcal{C}(X \times Y)$.

Then we prove the closedness of $L_X$: assume that $h_X$ is a limit point of $L_X$, then there exists a sequence of functions $\{(h_n)_X\} \in L_X$ s.t. $(h_n)_X = \varphi_n \circ \pi_X$ and

$$\lim_{n\to\infty} (h_n)_X = \lim_{n\to\infty} \varphi_n \circ \pi_X = \left(\lim_{n\to\infty} \varphi_n\right) \circ \pi_X = h_X$$

we just need to show $\varphi_h := \lim_n \varphi_n$ is continuous: for any $\varepsilon > 0$, $\exists N$, when $n > N$ we have

$$\|\varphi_h - \varphi_N\| < \varepsilon/3$$

and for continuous function $\varphi_N$, $x \in X$, $\exists\delta$ s.t. $x_0 \in N_\delta(x) \Rightarrow |\varphi_N(x) - \varphi_N(x_0)| < \varepsilon/3$ Now

$$\begin{aligned} |\varphi_h(x) - \varphi_h(x_0)| &= |\varphi_h(x) - \varphi_N(x) + \varphi_N(x) - \varphi_N(x_0) + \varphi_N(x_0) - \varphi_h(x_0)| \\ &\leq |\varphi_h(x) - \varphi_N(x)| + |\varphi_N(x) - \varphi_N(x_0)| + |\varphi_N(x_0) - \varphi_h(x_0)| < \varepsilon \end{aligned}$$

so $\varphi_h = \lim_n \varphi_n \in \mathcal{C}(X)$, and $h_X = (\lim_n \varphi_n) \circ \pi_X \in L_X$, which means $L_X$ is closed, similarly, $L_Y$ is closed.

To apply Lemma 5.6 we consider quotient space $L_X/R$ and $L_Y/R$, we have $L_X/R \cap L_Y/R = [0]$ and $\tilde{Z}/R = L_X/R + L_Y/R$. Now we claim that if $\tilde{Z}/R$ closed in $\mathcal{C}(X \times Y)/R$ then $\tilde{Z}$ is closed in $\mathcal{C}(X \times Y)$:

**Claim 5.1.** *If $\tilde{Z}/R$ closed in $\mathcal{C}(X \times Y)/R$ then $\tilde{Z}$ is closed in $\mathcal{C}(X \times Y)$*

*Proof of Claim 5.2:* Define the quotient map $q : \mathcal{C}(X \times Y) \to \mathcal{C}(X \times Y)/R$, just need to show that $q$ is continuous: for any $\varepsilon > 0$, if $h, h' \in \mathcal{C}(X \times Y)$ and $\|h - h'\|_{\mathcal{C}(X\times Y)} < \delta = \varepsilon$,

$$\|q(h) - q(h')\|_{\mathcal{C}(X\times Y)/R} = \inf_{c\in R} \|h - h' + c\|_{\mathcal{C}(X\times Y)} \leq \|h - h'\|_{\mathcal{C}(X\times Y)} < \varepsilon.$$

□

So now we just need to show that $\tilde{Z}/R$ is closed in $\mathcal{C}(X \times Y)/R$, by lemma 5.6, we need to find $A > 0$ s.t. for any $\varphi \in \mathcal{C}(X)$ and $\psi \in \mathcal{C}(Y)$,

$$\|[\varphi \circ \pi_X]\|_{\mathcal{C}(X\times Y)/R} \leq A \cdot \|[\varphi \circ \pi_X + \psi \circ \pi_Y]\|_{\mathcal{C}(X\times Y)/R}$$

and

$$\begin{aligned}
\|[\varphi \circ \pi_X]\|_{\mathcal{C}(X\times Y)/R} &= \inf_{c\in R} \|\varphi \circ \pi_X + c\|_{\mathcal{C}(X\times Y)} \\
&= \inf_{c\in R} \sup_{(x,y)\in X\times Y} |\varphi \circ \pi_X(x,y) + c| \\
&= \inf_{c\in R} \sup_{x\in X} |\varphi(x) + c|
\end{aligned}$$

we claim that $\inf_{c\in R} \sup_{x\in X} |f(x) + c| = \frac{1}{2}(\max \varphi - \min \varphi)$.

**Claim 5.2.** $\inf_{c\in R} \sup_{x\in X} |\varphi(x) + c| = \frac{1}{2}(\max \varphi - \min \varphi)$

*Proof of Claim 5.3.* When $\varphi$ is unbounded, i.e. $\min \varphi = -\infty$ or $\max \varphi = \infty$, it is obvious that

$$\inf_{c\in R} \sup_{x\in X} |f = \varphi(x) + c| = \infty = \frac{1}{2}(\max \varphi - \min \varphi)$$

so we just need to consider the boundedness case: note that

$$\inf_{c\in R} \sup_{x\in X} |\varphi(x) + c| = \inf_{c\in R} \sup_{x\in X} |\varphi(x) - (-c)|$$

- When $-c < \min \varphi$, select $\varphi(x) = \max \varphi$ and $-c = \min \varphi$ to obtain the inf sup:

$$\inf_{c\in R} \sup_{x\in X} |\varphi(x) - (-c)| = \max \varphi - \min \varphi$$

- When $-c > \max\varphi$, select $f(x) = \min f$ and $-c = \max\varphi$ to obtain the inf sup:

$$\inf_{c\in R}\sup_{x\in X}|\varphi(x) - (-c)| = \max\varphi - \min\varphi$$

- When $\min\varphi \leq -c < \frac{1}{2}(\max\varphi + \min\varphi)$, select $\varphi(x) = \max\varphi$ and $-c = \frac{1}{2}(\max\varphi + \min\varphi)$ to obtain the inf sup:

$$\inf_{c\in R}\sup_{x\in X}|\varphi(x) - (-c)| = \frac{1}{2}(\max\varphi - \min\varphi)$$

- When $\frac{1}{2}(\max\varphi + \min\varphi) \leq -c < \max\varphi$, select $\varphi(x) = \min\varphi$ and $-c = \frac{1}{2}(\max\varphi + \min\varphi)$ to obtain the inf sup:

$$\inf_{c\in R}\sup_{x\in X}|\varphi(x) - (-c)| = \frac{1}{2}(\max\varphi - \min\varphi)$$

Therefore we proved the claim. $\square$

For any $\varphi \in \mathcal{C}(X)$ and $\psi \in \mathcal{C}(Y)$, let $F(x,y) = \varphi(x)+\psi(y)$, then $\min_{(x,y)\in X\times Y} F(x,y) = \min_{x\in X}\varphi(x)+\min_{y\in Y}\psi(y) = \min\varphi+\min\psi$, similarly, $\max F(x,y) = \max\varphi+\max\psi$, therefore

$$\begin{aligned}\|[\varphi\circ\pi_X + \psi\circ\pi_Y]\|_{\mathcal{C}(X\times Y)/R} &= \inf_{c\in R}\sup_{(x,y)\in X\times Y}|\varphi(x)+\psi(y)+c| \\ &= \frac{1}{2}(\max F - \min F) \geq \frac{1}{2}(\max\varphi - \min\varphi)\end{aligned}$$

Thus we can find $A = 1$ s.t. for any $\varphi \in \mathcal{C}(X)$ and $\psi \in \mathcal{C}(Y)$, $\varphi\circ\pi_X \in L_X/R$ and $\varphi\circ\pi_X + \psi\circ\pi_Y \in C/R$, we have

$$\|[\varphi\circ\pi_X]\|_{\mathcal{C}(X\times Y)/R} \leq A\cdot\|[\varphi\circ\pi_X+\psi\circ\pi_Y]\|_{\mathcal{C}(X\times Y)/R}$$

then by lemma 5.6, we proved that $\tilde{Z}/R$ is closed in $\mathcal{C}(X\times Y)/R$, therefore by the claim 5.2, $Z$ is closed in $\mathcal{C}(X \times Y)$. $\square$

We have proved the closedness of $\tilde{Z}$, now we want to show that $Z$ is also closed:

**Proposition 5.8.** *The subspace of* $\mathcal{C}(X \times Y)$

$$Z = \left\{ \varphi \circ \pi_X - \int_X \varphi d\mu + \psi \circ \pi_Y - \int_Y \psi d\nu : \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y) \right\}$$

*is closed.*

*Proof.* We just need to show that the set

$$Z_X = \left\{ \varphi \circ \pi_X - \int_X \varphi d\mu : \varphi \in \mathcal{C}(X) \right\}$$

is closed. We claim that $Z_X = \tilde{Z}_X := \{\tilde{\varphi} \circ \pi_X : \tilde{\varphi} \in \mathcal{C}_\mu(X)\}$, where

$$\mathcal{C}_\mu(X) = \left\{ f \in \mathcal{C}(X) : \int f d\mu = 0 \right\}$$

Since for any $F \in Z_X$, there exists $\varphi \in \mathcal{C}(X)$ s.t. $F = (\varphi - \int \varphi d\mu) \circ \pi_X$. Let $\tilde{\varphi} := \varphi - \int \varphi d\mu \in \mathcal{C}_\mu(X)$, we have $F = \tilde{\varphi} \circ \pi_X \in \tilde{Z}_X$.

In another direction, for $G \in \tilde{Z}_X$, there exists $\psi \in \mathcal{C}_\mu(X)$ s.t. $G = \psi \circ \pi_X$. Then $G = (\psi - \int \psi d\mu) \circ \pi_X \in Z_X$.

Now that $\tilde{Z}_X = Z_X$, similar to the proof of closedness of $L_X = \{\varphi \circ \pi_X : \varphi \in \mathcal{C}(X)\}$, we can show that $Z_X$ is closed. In the following steps we can define $Z_Y$ accordingly and prove the closedness of $Z_Y$. Then follow the steps of Proposition 5.7's proof to get the desired result. $\square$

Now we are ready to prove Theorem 5.5 by applying the Lagrangian approach (Theorem 5.4):

*Proof of Theorem 5.5.* In Theorem 5.4, let the topological dynamical system be $(X \times Y, T \times S)$, $\phi \in \mathcal{C}(X \times Y)$, function $W$ be the pressure function $P(T \times S, \cdot)$ that is convex and continuous on $\mathcal{C}(X \times Y)$, and the closed linear subspace $\mathcal{N} = Z$, where

$$Z = \left\{ \varphi \circ \pi_X - \int_X \varphi d\mu + \psi \circ \pi_Y - \int_Y \psi d\nu : \varphi \in \mathcal{C}(X), \psi \in \mathcal{C}(Y) \right\}$$

then by Theorem 5.4, there exists $\hat{\varphi}, \hat{\psi}$ and $\lambda^* \in \mathcal{M}(X \times Y, T \times S)$ s.t.

$$\lambda^* \in \underset{\lambda \in \mathcal{M}(X \times Y, T \times S)}{\operatorname{argmax}} \int_{X \times Y} \left( \phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu \right) d\lambda + h(\lambda) \tag{5.1}$$

and for any $\varepsilon > 0$

$$\left| \int h d\lambda^* \right| \leq \varepsilon \|h\|, \forall h \in Z \tag{5.2}$$

In the follwing we will prove that there exists a $\hat{\lambda}$ such that $\hat{\lambda} \in \mathcal{M}_Z(X \times Y, T \times S) = J(\mu, \nu)$:

Let $\lambda^* = \int_{\mathcal{M}^e(X \times Y, T \times S)} \rho d\tau(\rho)$ be the ergodic decomposition of $\lambda^*$. Then

$$\begin{cases} \lambda^* \circ \pi_X^{-1} = \int_{p \in \mathcal{M}^e(X,T)} p d(\tau \circ \tilde{\pi}_X^{-1})(p) \\ \lambda^* \circ \pi_Y^{-1} = \int_{q \in \mathcal{M}^e(Y,S)} q d(\tau \circ \tilde{\pi}_Y^{-1})(q) \end{cases}$$

are the ergodic decomposition of $\lambda^* \circ \pi_X^{-1}$ and $\lambda^* \circ \pi_Y^{-1}$ respectively, where $\tilde{\pi}_X : \mathcal{M}(X \times Y, T \times S) \to \mathcal{M}(X, T)$ and $\tilde{\pi}_Y : \mathcal{M}(X \times Y, T \times S) \to \mathcal{M}(Y, S)$ are the map induced by $\pi_X$ and $\pi_Y$.

By (5.2), when $h = \varphi \circ \pi_X - \int_X \varphi d\mu$ (let $\psi \equiv 0$) for any $\varphi \in \mathcal{C}(X)$, we have

$$\begin{aligned} \left| \int_{X \times Y} \varphi \circ \pi_X d\lambda^* - \int_X \varphi d\mu \right| &= \left| \int_X \varphi d\lambda^* \circ \pi_X^{-1} - \int_X \varphi d\mu \right| \\ &= \left| \int_X \left( \varphi - \int_X \varphi d\mu \right) d\lambda^* \circ \pi_X^{-1} - \int_X \left( \varphi - \int_X \varphi d\mu \right) d\mu \right| \\ &\leq \varepsilon \left\| \left( \varphi - \int_X \varphi d\mu \right) \right\| \end{aligned}$$

Since $\mathcal{C}(X) = \{\varphi - \int_X \varphi d\mu : \varphi \in \mathcal{C}(X)\}$, from the above inequality we have:

$$\|\lambda^* \circ \pi_X^{-1} - \mu\| = \|\tau \circ \tilde{\pi}_X^{-1} - \delta_\mu\| \leq \varepsilon$$

Similarly,

$$\|\lambda^* \circ \pi_Y^{-1} - \nu\| = \|\tau \circ \tilde{\pi}_Y^{-1} - \delta_\nu\| \leq \varepsilon$$

Since the set of equilibrium states of $\left(\phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu\right)$ is a face of $\mathcal{M}(X \times Y, T \times S)$, the set

$$\left\{\lambda \circ \pi_X^{-1} : \lambda \text{ is an equilibrium state of } \left(\phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu\right)\right\}$$

and

$$\left\{\lambda \circ \pi_Y^{-1} : \lambda \text{ is an equilibrium state of } \left(\phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu\right)\right\}$$

are faces of $\mathcal{M}(X, T)$ and $\mathcal{M}(Y, S)$ respectively.

Let

$$E_X = \{\lambda : \tilde{\pi}_X \lambda = \mu\} \qquad E_Y = \{\lambda : \tilde{\pi}_Y \lambda = \nu\}$$

and $E = \{\lambda : \lambda \, \text{is an equilibrium state of} \left(\phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu\right)\}$. $E$ is a face of $\mathcal{M}(X \times Y, T \times S)$, so $\tau(E) = 1$ as $\lambda^*$ is an equilibrium state of $\left(\phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu\right)$. Meanwhile, by the fact that $\|\tau \circ \tilde{\pi}_X^{-1} - \delta_\mu\| \leq \varepsilon$, since $[\tau \circ \tilde{\pi}_X^{-1} - \delta_\mu] : \mathcal{M}(X, T) \to \mathbb{R}$, we have

$$\|\tau \circ \tilde{\pi}_X^{-1} - \delta_\mu\| := \sup_{\substack{m \in \mathcal{C}(\mathcal{M}(X,T)) \\ \|m\| \leq 1}} |\tau \circ \tilde{\pi}_X^{-1}(m) - \delta_\mu(m)| \leq \varepsilon$$

That is, the supremum is taken over all measure $m \in \mathcal{M}(X, T)$. with its norm

$\|m\| \leq 1$, the norm of $m$ is the same of norm on $\mathcal{C}(X)^*$:

$$\|m\| = \sup_{\substack{f \in \mathcal{C}(X) \\ \|f\| \leq 1}} |m(f)| = \sup_{\substack{f \in \mathcal{C}(X) \\ \|f\| \leq 1}} \left| \int f dm \right|$$

if we set $m = \mu$, then

$$|\tau(\tilde{\pi}_X^{-1}\mu) - 1| \leq \|\tau \circ \tilde{\pi}_X^{-1} - \delta_\mu\| \leq \varepsilon \implies \tau(\tilde{\pi}_X^{-1}\mu) = \tau(E_X) \geq 1 - \varepsilon$$

Similarly, $\tau(E_Y) \geq 1 - \varepsilon$. Therefore

$$\tau(E \cap E_X \cap E_Y) = 1 - \tau(E^c \cup E_X^c \cup E_Y^c) \geq 1 - (\tau(E^c) + \tau(E_X^c) + \tau(E_Y^c)) \geq 1 - 2\varepsilon$$

As $\varepsilon$ can be any value, we can pick $\varepsilon = 1/3$, then $\tau(E \cap E_X \cap E_Y) \geq 1/3 > 0$, so some equilibrium state of $\left(\phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu\right)$, say $\hat{\lambda}$, is an element of $J(\mu, \nu)$.

Finally we will prove the 'if and only if' part: for some $\hat{\lambda} \in J(\mu, \nu)$,

$$\begin{aligned} &\hat{\lambda} \in \underset{\lambda \in \mathcal{M}(X \times Y, T \times S)}{\operatorname{argmax}} \int \left(\phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu\right) d\lambda + h(\lambda) \\ \iff &\hat{\lambda} \in \underset{\lambda \in J(\mu, \nu)}{\operatorname{argmax}} \int \phi d\lambda + h(\lambda) \end{aligned}$$

- ( $\Longrightarrow$ ): For $\hat{\lambda} \in J(\mu, \nu)$,

$$
\begin{aligned}
& \hat{\lambda} \in \operatorname*{argmax}_{\lambda \in \mathcal{M}(X \times Y, T \times S)} \int \left( \phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu \right) d\lambda + h(\lambda) \\
\Longrightarrow\ & \hat{\lambda} \in \operatorname*{argmax}_{\lambda \in \mathcal{M}(X \times Y, T \times S)} \int (\phi + \hat{\varphi} \circ \pi_X + \hat{\psi} \circ \pi_Y) d\lambda + h(\lambda) \\
\Longrightarrow\ & \hat{\lambda} \in \operatorname*{argmax}_{\lambda \in J(\mu,\nu)} \int (\phi + \hat{\varphi} \circ \pi_X + \hat{\psi} \circ \pi_Y) d\lambda + h(\lambda) \\
\Longrightarrow\ & \hat{\lambda} \in \operatorname*{argmax}_{\lambda \in J(\mu,\nu)} \int \phi d\lambda + h(\lambda) + \int_X \hat{\varphi} d\mu + \int_Y \hat{\psi} d\nu \\
\Longrightarrow\ & \hat{\lambda} \in \operatorname*{argmax}_{\lambda \in J(\mu,\nu)} \int \phi d\lambda + h(\lambda)
\end{aligned}
$$

- ( $\Longleftarrow$ ): This is directly from the fact that $\hat{\lambda}$ is an equilibrium state of $(\phi + \hat{\varphi} \circ \pi_X - \int_X \hat{\varphi} d\mu + \hat{\psi} \circ \pi_Y - \int_Y \hat{\psi} d\nu)$.

□

**Remark 5.3.**

*(i) When $Z = \{\varphi \circ \pi_X - \int_Y \varphi d\nu : \varphi \in \mathcal{C}(Y)\}$, the relativized version of the Lagrangian approach has been proved by Walters[13].*

*(ii) In Theorem 5.5 if the given $\phi \in \mathcal{C}(X \times Y)$ and $\hat{\varphi}, \hat{\psi}$ are Hölder functions or satisfy strong regularity conditions, then there exists an unique optimal plan.*

*(iii) There is another way to prove Theorem 5.5 by applying Theorem 5.4: Let $Z_X = \{\varphi \circ \pi_X - \int_X \varphi d\mu : \varphi \in \mathcal{C}(X)\}$ and $Z_Y = \{\psi \circ \pi_Y - \int_Y \pi d\nu : \psi \in \mathcal{C}(Y)\}$, and*

$$
W(f) = P_{Z_X}(T \times S, f) = \sup_{\lambda \in \mathcal{M}_{Z_X}(X \times Y, T \times S)} \int f d\lambda + h(\lambda)
$$

*by Lemma 5.1, $W$ is convex and continuous. Now, let $\mathcal{N} = Z_Y$, since $Z_Y$ is a closed subspace, we can apply Theorem 5.4 and Walters' relativized version[13],* *there exist some* $\lambda^* \in J(\mu,\nu)$ *satisfies* $\lambda^* \circ \pi_Y^{-1} = \nu$ *and* $\psi^* \in \mathcal{C}(Y)$ *s.t.*

$$\lambda^* \in \underset{\lambda \in \mathcal{M}_{Z_X}(X\times Y, T\times S)}{\operatorname{argmax}} \int (\phi + \psi^* \circ \pi_Y) d\lambda + h(\lambda) \iff \lambda^* \in \underset{\lambda \in J(\mu,\nu)}{\operatorname{argmax}} \int \phi d\lambda + h(\lambda)$$

We can generalize Theorem 5.5 if there are $n$ different dynamical systems $(X_i, T_i)$ for $i = 1, 2, \cdots, n$:

**Corollary 5.9.** *Suppose that there are* $n$ *dynamical systems* $\{(X_i, T_i)\}_{i=1}^n$*, where each* $X_i$ *and* $T_i : X_i \to X_i$ *satisfy our basic settings in section 1.2.1. Let* $\phi \in \mathcal{C}(\prod_{k=1}^n X_k)$ *and* $\pi_i : \prod_{k=1}^n X_k \to X_i$ *be the projection map onto the ith space. Given* $\mu_i \in \mathcal{M}^e(X_i, T_i)$ *and let*

$$Z = \left\{ \sum_{i=1}^n \left( f_i \circ \pi_i - \int_{X_i} f_i d\mu_i \right) : f_i \in \mathcal{C}(X_i), i = 1, \cdots, n \right\}$$

*there exists* $(\hat{f}_1, \cdots, \hat{f}_n) \in \mathcal{C}(X_1) \times \cdots \times \mathcal{C}(X_n)$ *s.t.*

$$\begin{aligned} &\hat{\lambda} \in \underset{\lambda \in \mathcal{M}(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k)}{\operatorname{argmax}} \int \left[ \phi + \sum_{i=1}^n \left( \hat{f}_i \circ \pi_i - \int_{X_i} \hat{f}_i d\mu_i \right) \right] d\lambda + h(\lambda) \\ \iff &\hat{\lambda} \in \underset{\lambda \in \mathcal{M}_Z(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k)}{\operatorname{argmax}} \int \phi d\lambda + h(\lambda) \end{aligned}$$

*Note that*

$$\mathcal{M}_Z(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k) = \{\mu \in \mathcal{M}(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k) : \mu \circ \pi_i^{-1} = \mu_i \textit{ for } i = 1, \cdots, n\}$$

*Proof.* Let

$$\tilde{Z} = \left\{ \sum_{i=1}^n f_i \circ \pi_i : f_i \in \mathcal{C}(X_i), i = 1, \cdots, n \right\}$$

by Proposition 5.7, $\tilde{Z}$ defined above is closed, and by Proposition 5.8, such $Z$ defined in Corollary 5.9 is closed. Then by the Lagrangian approach (Theorem 5.4), there exists $(\hat{f}_1, \cdots, \hat{f}_n) \in \mathcal{C}(X_1) \times \cdots \times \mathcal{C}(X_n)$ and $\lambda^* \in \mathcal{M}(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k)$ s.t.

$$\lambda^* \in \underset{\mathcal{M}(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k)}{\operatorname{argmax}} \int \left[ \phi + \sum_{i=1}^n \left( \hat{f}_i \circ \pi_i - \int_{X_i} \hat{f}_i d\mu_i \right) \right] d\lambda + h(\lambda)$$

and for any $\varepsilon > 0$ we have

$$\left| \int h d\lambda^* \right| < \varepsilon \|h\|, \ \forall h \in Z$$

For $f_i \in \mathcal{C}(X_i)$, let $h = f_i \circ \pi_i - \int_{X_i} f_i d\mu_i$, then by the proof of Theorem 5.5,

$$\left| \int h d\lambda^* \right| < \varepsilon \|h\|, \ \forall h \in Z \implies \|\lambda^* \circ \pi_i^{-1} - \mu_i\| < \varepsilon$$

Let

$$\lambda^* = \int_{\mathcal{M}^e(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k)} p d\tau(p)$$

be the ergodic decomposition of $\lambda^*$, then we have (by the proof of Theorem 5.5):

$$\|\lambda^* \circ \pi_i^{-1} - \mu_i\| = \|\tau \circ \tilde{\pi}_i^{-1} - \delta_{\mu_i}\| < \varepsilon$$

where the map $\tilde{\pi}_i : \mathcal{M}(\prod_{k=1}^n X_k, \prod_{k=1}^n T_k) \to \mathcal{M}(X_i, T_i)$ induced by $\pi_i$ and is given by

$$\tilde{\pi}_i(\lambda) = \lambda \circ \pi_i^{-1}$$

Now we define

$$E_i = \left\{ \lambda \circ \pi_i^{-1} : \lambda \text{ is an equilibrium state of } \phi + \sum_{i=1}^n \hat{f}_i \circ \pi_i \right\}$$

then $\tau(E_i) > 1 - \varepsilon$, and if $\varepsilon < 1/n$

$$\tau\left( \bigcap_{i=1}^n E_i \right) > 1 - n\varepsilon > 0$$

Since the set

$$E = \left\{ \lambda \in \mathcal{M}(\prod_{k=1}^{n} X_k, \prod_{k=1}^{n} T_k) : \lambda \text{ is an equilibrium state of } \phi + \sum_{i=1}^{n} \hat{f}_i \circ \pi_i \right\}$$

is a face of $\mathcal{M}(\prod_{k=1}^{n} X_k, \prod_{k=1}^{n} T_k)$, therefore there is a $\hat{\lambda} \in E \cap (\bigcap_{i=1}^{n} E_i)$. $\square$

## 5.3 Result in uniqueness of relative equilibrium states

Throughout this section we consider the Gibbs measures defined on Shift of Finite Types (SFTs) with Hölder continuous potential functions: Given two finite sets (alphabets) $\mathcal{H}_A$ and $\mathcal{H}_Y$. Assume that there is a one-block factor map $\pi : \mathcal{H}_A \to \mathcal{H}_Y$ between two subshifts $(\Sigma_A, \sigma)$ and $(Y, \sigma_Y)$. Here $(\Sigma_A, \sigma)$ be a topological mixing SFT and $(Y, \sigma_Y)$ be another subshift, where $\Sigma_A \subseteq \mathcal{H}_A^{\mathbb{N}}$ and $Y \subseteq \mathcal{H}_Y^{\mathbb{N}}$. Let the potential function $\phi : \Sigma_A \to \mathbb{R}$ be an member of $\mathcal{F}_A$, where

$$\mathcal{F}_A = \{\phi \in \mathcal{C}(\Sigma_A) : \mathrm{var}_n \phi \leq \beta\alpha^n, \beta > 0, \alpha \in (0,1)\}$$

and

$$\mathrm{var}_n \phi = \sup\{|\phi(\underline{x}) - \phi(\underline{y})| : \underline{x}_0^{n-1} = \underline{y}_0^{n-1}\}$$

$\sigma : \Sigma_A \to \Sigma_A$ and $\sigma_Y : Y \to Y$ are left-shift maps.

**Definition 5.2** (Fiber-wise mixing factor map)**.** *A one-block factor map $\pi$ is called fiber-wise mixing if there exists $N$ such that for any $y_0^N$ admissible in $Y$ and $x_0, x_N \in \mathcal{H}_A$ s.t. $\pi(x_0) = y_0$ and $\pi(x_N) = y_N$, then there exists $x_0 x_1^{N-1} x_N \in \Sigma_A$ with $\pi(x_0^N) = y_0^N$.*

We have the following meaningful uniqueness results in relative equilibrium states problem (by specifying $Z$ in the framework of problem (II)):

**Theorem 5.10.** *Suppose that $\phi \in \mathcal{F}_A$ and $\psi \in \mathcal{C}(Y)$ is Hölder. Let $\nu \in \mathcal{M}(Y, \sigma_Y)$ be the unique equilibrium state for $\psi$ and $\pi : \Sigma_A \to Y$ be a one-block fiber-wise* *mixing factor map from* $(\Sigma_A, \sigma)$ *onto* $(Y, \sigma_Y)$. *If* $Z = \left\{g \circ \pi - \int g d\nu : g \in \mathcal{C}(Y)\right\}$, *then we can choose a Hölder function* $g = \psi - \mathrm{Pot}(\mu_\phi \circ \pi^{-1}) \in \mathcal{C}(Y)$ *such that* $\chi = g \circ \pi - \int g d\nu$ *satisfies Lagrangian approach (where* $\mathrm{Pot}(\cdot)$ *means the potential of certain Gibbs measure). As* $g$ *is Hölder and* $\chi$ *is also Hölder, the relative equilibrium state is unique.*

Note that the theorem above gives an result for uniqueness relative equilibrium states, it is an weak version of Yoo[49]:

**Theorem 5.11** (Yoo[49])**.** *Let* $\pi : \Sigma_A \to Y$ *be a factor map from an irreducible SFT onto a sofic shift. Let* $\phi, \psi$ *be Hölder continuous functions defined on* $\Sigma_A$ *and* $Y$ *respectively. Let* $\nu$ *be the unique equilibrium state for* $\psi$. *Then there is a unique relative equilibrium state* $\mu$ *of* $\phi$ *over* $\nu$.

Before we prove Theorem 5.10, some necessary results are given as below. The first definition gives a candidate for potential functions of projection of Gibbs measures:

**Definition 5.3** ($g$-function)**.** *If* $\mu_\phi$ *is the Gibbs measure with potential* $\phi \in \mathcal{C}(\Sigma_A)$. *Denote* $g_n(\underline{x})$ *by*

$$g_n(\underline{x}) := \frac{\pi_\# \mu_\phi[x_0 \cdots x_n]}{\pi_\# \mu_\phi[x_1 \cdots x_n]}$$

*where* $\pi_\# \mu_\phi = \mu_\phi \circ \pi^{-1}$. *The* $g$*-function for* $\pi_\# \mu_\phi$ *is defined as* $g(\underline{x}) = \lim_{n \to \infty} g_n(\underline{x})$ *and* $\log g$ *is a candidate for potential functions of* $\pi_\# \mu_\phi$.

The next result given by Piraino[51] shows that projection of a Gibbs measure with Hölder potential is also Gibbsian and its potential is also Hölder:

**Proposition 5.12** (Piraino[51])**.** *Suppose that* $\phi : \Sigma_A \to \mathbb{R}$ *is continuous,* $\mu_\phi$ *the Gibbs state for* $\phi$, *and* $\pi : \Sigma_A \to Y$ *a fiber-wise mixing* 1*-block factor map. Then* $\pi_\# \mu_\phi$ *is also a Gibbs measure with Hölder continuous potential* $\log g$.

In the following we give another form of potentials candidate for $\pi_\# \mu_\phi$ when $\mu_\phi$ is a Gibbs measure with a regular potential $\phi$:

**Proposition 5.13** (Kempton[52])**.** *Suppose that $(\Sigma_A, \sigma_A)$ is a topological mixing SFT and $\pi : \Sigma_A \to Y$ is a fiber-wise mixing factor map and satisfies that there exists an integer $N$ s.t. for any $n \in \mathbb{N}$, the possible values of $x_n$ in the set of sequences $\{\underline{x} \in \Sigma_A : \pi(x_{n-N} \cdots x_{n+N}) = z_{n-N} \cdots z_{n+N}\}$ equals the possible values of $x_n$ in $\{\underline{x} \in \Sigma_A : \pi(\underline{x}) = \underline{z}\}$ (the nth position of $\underline{x}$ is locally determined). Then for $j \in \mathcal{H}_A$, $n \in \mathbb{N}$ and $\underline{w} \in \Sigma_A$ such that $j\underline{w}$ is admissible, the potential for $\pi_{\#}\mu_\phi$ is given by*

$$\psi(\underline{y}) = \lim_{n\to\infty} \log \frac{\sum_{\substack{\underline{x}=x_0^{n-1}j \\ \pi(\underline{x})=y_0^n}} \exp(S_{n+1}\phi(\underline{x\omega}))}{\sum_{\substack{\underline{x}'=x_1^{n-1}j \\ \pi(\underline{x}')=y_1^n}} \exp(S_n\phi(\underline{x}'\underline{\omega}))}$$

*for $\underline{y} \in Y$, where $x_0^{n-1} = x_0 \cdots x_{n-1}$ and $S_n\phi(\underline{x}) = \sum_{i=0}^{n-1} \phi(\sigma^i \underline{x})$.*

For completeness we give a brief proof below, for detailed proof please refer to Kempton[52].

**Lemma 5.14.** *The term*

$$\lim_{n\to\infty} \frac{\sum_{\substack{\underline{x}=x_0^{n-1}j \\ \pi(\underline{x})=y_0^n}} \exp(S_{n+1}\phi(\underline{x\omega}))}{\sum_{\substack{\underline{x}'=x_1^{n-1}j \\ \pi(\underline{x}')=y_1^n}} \exp(S_n\phi(\underline{x}'\underline{\omega}))}$$

*is well defined and independent of $j, \underline{w}$.*

**Lemma 5.15.** *There is a constant $C$ depending only on $\phi$ such that*

$$C^{-1} \le \frac{\sum_{\underline{x}=x_0^n} \exp(S_{n+1}\phi(\underline{x\omega})) \sum_{\overline{x}=x_{n+1}^s} \exp(S_{s-n}\phi(\overline{x}\underline{\omega}))}{\sum_{\underline{x}=x_0^s} \exp(S_{s+1}\phi(\underline{x\omega}))} \le C$$

Then the sketch proof of Proposition 5.13 is given below:

*Proof of Proposition 5.13.* Note that by the Gibbsian property of $\mu_\phi$ we have (for any $\underline{x} \in [x_0^n]$ and $\underline{y} \in [y_0^n]$ s.t. $\pi([x_0^n]) = [y_0^n]$)

$$C_1 \le \frac{\mu_\phi[x_0 \cdots x_n]}{\exp(S_{n+1}\phi(\underline{x}) - (n+1)P(\phi))} \le C_2$$

and

$$C_1' \leq \frac{\mu_\phi[x_1 \cdots x_n]}{\exp(S_n\phi(\sigma\underline{x}) - nP(\phi))} \leq C_2'$$

as

$$\pi_\# \mu_\phi[y_0 \cdots y_n] = \sum_{x_0^n \in \pi^{-1}(y_0^n)} \mu_\phi[x_0 \cdots x_n]$$

so

$$C_1 e^{-(n+1)P} \sum_{x_0^n \in \pi^{-1}(y_0^n)} \exp(S_{n+1}\phi(\underline{x})) \leq \pi_\#\mu_\phi[x_0 \cdots x_n] \leq C_2 e^{-(n+1)P} \sum_{x_0^n \in \pi^{-1}(y_0^n)} \exp(S_{n+1}\phi(\underline{x}))$$

$$C_1' e^{-nP} \sum_{x_1^n \in \pi^{-1}(y_1^n)} \exp(S_n\phi(\sigma\underline{x})) \leq \pi_\#\mu_\phi[x_1 \cdots x_n] \leq C_2' e^{-nP} \sum_{x_1^n \in \pi^{-1}(y_1^n)} \exp(S_n\phi(\sigma\underline{x}))$$

since $\underline{x} \in [x_0^n]$, $\sigma\underline{x} \in [x_1^n]$, and we have

$$\frac{C_1}{C_2' e^P} \cdot \frac{\sum_{x_0^n \in \pi^{-1}(y_0^n)} \exp(S_{n+1}\phi(\underline{x}))}{\sum_{x_1^n \in \pi^{-1}(y_1^n)} \exp(S_n\phi(\sigma\underline{x}))} \leq g_n(\underline{x}) \leq \frac{C_2}{C_1' e^P} \cdot \frac{\sum_{x_0^n \in \pi^{-1}(y_0^n)} \exp(S_{n+1}\phi(\underline{x}))}{\sum_{x_1^n \in \pi^{-1}(y_1^n)} \exp(S_n\phi(\sigma\underline{x}))}$$

and the potential satisfies

$$\begin{aligned} \tilde{c_1} + \lim_{n\to\infty} \log \frac{\sum_{x_0^n \in \pi^{-1}(y_0^n)} \exp(S_{n+1}\phi(\underline{x}))}{\sum_{x_1^n \in \pi^{-1}(y_1^n)} \exp(S_n\phi(\sigma\underline{x}))} &\leq \lim_{n\to\infty} \log g_n(\underline{x}) \\ \leq \tilde{c_2} + \lim_{n\to\infty} \log \frac{\sum_{x_0^n \in \pi^{-1}(y_0^n)} \exp(S_{n+1}\phi(\underline{x}))}{\sum_{x_1^n \in \pi^{-1}(y_1^n)} \exp(S_n\phi(\sigma\underline{x}))} & \end{aligned}$$

where

$$\tilde{c_1} = \log\left(\frac{C_1}{C_2' e^P}\right) \qquad \tilde{c_2} = \log\left(\frac{C_2}{C_1' e^P}\right)$$

We can choose a proper coboundary $g \circ \sigma - g + c$ where $g$ is in the same space as $\phi$ and $c$ is a constant, s.t. $P(\phi + g\circ\sigma - g + c) = 0$, and we denote $\phi := \phi + g\circ\sigma - g + c$.

Then we let $\psi_m(\underline{y}) = \log \frac{\sum_{x_0^m \in \pi^{-1}(y_0^m)} \exp(S_{m+1}\phi(\underline{x}))}{\sum_{x_1^m \in \pi^{-1}(y_1^m)} \exp(S_m\phi(\sigma\underline{x}))} = \log \frac{\sum_{\substack{\underline{x}=x_0^{m-1}j \\ \pi(\underline{x})=y_0^m}} \exp(S_{m+1}\phi(\underline{x}\underline{\omega}))}{\sum_{\substack{\underline{x}'=x_1^{m-1}j \\ \pi(\underline{x}')=y_1^m}} \exp(S_m\phi(\underline{x}'\underline{\omega}))}$

(here $\underline{\omega} \subset \Sigma_A$ and $j\underline{\omega}$ admissable) and $\psi = \lim_{m\to\infty} \psi_m$, we have the following results:

$$\begin{aligned}
S_{n+1}\psi(\underline{y}) &= \lim_{m\to\infty} \psi_m(\underline{y}) + \lim_{m\to\infty} \psi_{m-1}(\sigma\underline{y}) + \cdots + \lim_{m\to\infty} \psi_{m-n}(\sigma^n\underline{y}) \\
&= \lim_{m\to\infty} \sum_{i=0}^{n} \log \frac{\sum_{\substack{\underline{x}=x_0^{m-1}j \\ \pi(\underline{x})=y_0^n}} \exp(S_{m+1-i}\phi(\sigma^i \underline{x\omega}))}{\sum_{\substack{\underline{x}'=x_0^{m-1}j \\ \pi(\underline{x}')=y_0^n}} \exp(S_{m-i}\phi(\sigma\underline{x}'\underline{\omega}))} \\
&= \lim_{m\to\infty} \log \left( \frac{\sum_{\substack{\underline{x}=x_0^{m-1}j \\ \pi(\underline{x})=y_0^n}} \exp(S_{m+1}\phi(\underline{x\omega}))}{\sum_{\substack{\underline{x}'=x_n^{m-1}j \\ \pi(\underline{x}')=y_n^m}} \exp(S_{m-n}\phi(\sigma\underline{x}'\underline{\omega}))} \right)
\end{aligned}$$

that is because

$$\psi_{m-i}(\sigma^i\underline{y}) = \log \frac{\sum_{\substack{\underline{x}=x_i^{m-1}j \\ \pi(\underline{x})=y_i^m}} \exp(S_{m-i+1}\phi(\underline{x\omega}))}{\sum_{\substack{\underline{x}'=x_{i+1}^{m-1}j \\ \pi(\underline{x}')=y_{i+1}^m}} \exp(S_{m-i}\phi(\underline{x}'\underline{\omega}))}$$

By Lemma 5.15, we have (when $s = m$):

$$\exp(S_{n+1}\psi(\underline{y})) = \frac{\sum_{\substack{\underline{x}=x_0^m \\ \pi(\underline{x})=y_0^m}} \exp(S_{m+1}\phi(\underline{x\omega}))}{\sum_{\substack{\overline{x}=x_{n+1}^m \\ \pi(\overline{x})=y_{n+1}^m}} \exp(S_{m-n}\phi(\overline{x}\underline{\omega}))} \le C \sum_{\substack{\underline{x}=x_0^n \\ \pi(\underline{x})=y_0^n}} \exp(S_{n+1}\phi(\underline{x\omega}))$$

$$\exp(S_{n+1}\psi(\underline{y})) = \frac{\sum_{\substack{\underline{x}=x_0^m \\ \pi(\underline{x})=y_0^m}} \exp(S_{m+1}\phi(\underline{x\omega}))}{\sum_{\substack{\overline{x}=x_{n+1}^m \\ \pi(\overline{x})=y_{n+1}^m}} \exp(S_{m-n}\phi(\overline{x}\underline{\omega}))} \ge C^{-1} \sum_{\substack{\underline{x}=x_0^n \\ \pi(\underline{x})=y_0^n}} \exp(S_{n+1}\phi(\underline{x\omega}))$$

by the Gibbsian property of $\mu_\phi$:

$$C_1 \le \frac{\lambda[x_0 \cdots x_n]}{\exp(S_{n+1}\phi(\underline{x\omega}))} \le C_2$$

summing over all $\underline{x} = [x_0 \cdots x_n]$ s.t. $\pi(\underline{x}) = [y_0^n]$:

$$C_1 \sum_{\pi(\underline{x})=y_0^n} \exp(S_{n+1}\phi(\underline{x\omega})) \le \pi_\# \lambda[x_0 \cdots x_n] \le C_2 \sum_{\pi(\underline{x})=y_0^n} \exp(S_{n+1}\phi(\underline{x\omega}))$$

we have

$$\frac{C_1}{C} \leq \frac{\pi_{\#}\mu_\phi[x_0 \cdots x_n]}{\exp(S_{n+1}\psi(\underline{y}))} \leq C_2 C$$

Therefore, together with Lemma 5.13, we can pick the potential of $\pi_{\#}\mu_\phi$ as

$$\lim_{n\to\infty} \log \frac{\sum_{\substack{\underline{x}=x_0^{n-1}j \\ \pi(\underline{x})=y_0^n}} \exp(S_{n+1}\phi(\underline{x\omega}))}{\sum_{\substack{\underline{x}'=x_1^{n-1}j \\ \pi(\underline{x}')=y_1^n}} \exp(S_n\phi(\underline{x}'\underline{\omega}))}$$

□

To prove Theorem 5.10 we still need the following result:

**Proposition 5.16.** *If for $\phi \in \mathcal{C}(\Sigma_A)$ and $g \in \mathcal{C}(Y)$, then*

$$\mathrm{Pot}(\mu_{\phi+g\circ\pi} \circ \pi^{-1}) = g + \mathrm{Pot}(\mu_\phi \circ \pi^{-1})$$

*Proof.* By the result above, we have

$$\mathrm{Pot}(\mu_\phi \circ \pi^{-1})(\underline{y}) = \lim_{n\to\infty} \log \frac{\sum_{\substack{\underline{x}=x_0^{n-1}j \\ \pi(\underline{x})=y_0^n}} \exp(S_{n+1}\phi(\underline{x\omega}))}{\sum_{\substack{\underline{x}'=x_1^{n-1}j \\ \pi(\underline{x}')=y_1^n}} \exp(S_n\phi(\underline{x}'\underline{\omega}))}$$

so if the potential is $\phi + g \circ \pi$:

$$\begin{aligned}
&\mathrm{Pot}(\mu_{\phi+g\circ\pi} \circ \pi^{-1})(\underline{y}) \\
=& \lim_{n\to\infty} \log \frac{\sum_{\substack{\underline{x}=x_0^{n-1}j \\ \pi(\underline{x})=y_0^n}} \exp(S_{n+1}(\phi + g\circ\pi)(\underline{x\omega}))}{\sum_{\substack{\underline{x}'=x_1^{n-1}j \\ \pi(\underline{x}')=y_1^n}} \exp(S_n(\phi + g\circ\pi)(\underline{x}'\underline{\omega}))} \\
=& \lim_{n\to\infty} \log \frac{\exp(S_{n+1}g(\underline{y})) \sum_{\substack{\underline{x}=x_0^{n-1}j \\ \pi(\underline{x})=y_0^n}} \exp(S_{n+1}\phi(\underline{x\omega}))}{\exp(S_n g(\sigma\underline{y})) \sum_{\substack{\underline{x}'=x_1^{n-1}j \\ \pi(\underline{x}')=y_1^n}} \exp(S_n\phi(\underline{x}'\underline{\omega}))} \\
=& \lim_{n\to\infty} [S_{n+1}g(\underline{y}) - S_n g(\sigma\underline{y})] + \mathrm{Pot}(\mu_\phi \circ \pi^{-1})(\underline{y}) \\
=& g(\underline{y}) + \mathrm{Pot}(\mu_\phi \circ \pi^{-1})(\underline{y})
\end{aligned}$$

□

Now we are ready to prove Theorem 5.10:

*Proof of Theorem 5.10.* By the Lagrangian approach Theorem 5.4 and 5.5, there exists $g \in \mathcal{C}(Y)$ s.t. $\mu_{\phi+\chi} = \mu_{\phi+g\circ\pi}$, the equilibrium state of $\phi + g \circ \pi$, is a member of $\mathcal{M}_Z(\Sigma_A, \sigma)$. We claim that $g = \psi - \mathrm{Pot}(\mu_\phi \circ \pi^{-1})$ satisfies the Lagrangian approach. $g$ is a Hölder function by Proposition 5.10. We just need to show

$$\mu_{\phi+g\circ\pi} \circ \pi^{-1} = \nu$$

which is equivalent to

$$\mathrm{Pot}(\mu_{\phi+g\circ\pi} \circ \pi^{-1}) = \psi$$

As $g$ is Hölder, we know that $\mu_{\phi+g\circ\pi}$ is a Gibbs measure, and then by Proposition 5.15

$$\mathrm{Pot}(\mu_{\phi+g\circ\pi} \circ \pi^{-1}) = g + \mathrm{Pot}(\mu_\phi \circ \pi^{-1}) = [\psi - \mathrm{Pot}(\mu_\phi \circ \pi^{-1})] + \mathrm{Pot}(\mu_\phi \circ \pi^{-1}) = \psi$$

Thus we can pick $\chi = g \circ \pi - \int g d\nu \in Z$ as the Lagrange multiplier of this relativised problem. Since

$$\mu^* \in \underset{\mu \in \mathcal{M}(X,T)}{\mathrm{argmax}} \int (\phi + \chi) d\mu + h(\mu) \iff \mu^* \in \underset{\mu \in \mathcal{M}_Z(X,T)}{\mathrm{argmax}} \int \phi d\mu + h(\mu)$$

LHS problem has unique solution, so does the RHS. □

**Remark 5.4.** *Suppose that there is a factor map $\pi : \Sigma_A \to Y$, $(\Sigma_A, \sigma)$ is a topological mixing SFT and $\phi \in \mathcal{C}(\Sigma_A)$. For any $Z \subset \mathcal{C}(\Sigma_A)$, if problem (II)*

$$\sup_{\mu \in \mathcal{M}_Z(\Sigma_A,\sigma)} \int \phi d\mu + h(\mu)$$

*has unique optimal plan, then for Lagrangian multiplier* $\chi$, $\mu_{\phi+\chi}$ *is a Gibbs measure.*

# CHAPTER 6: OPTIMIZATION AS ZERO TEMPERATURE

Suppose that settings are as Section 2.1, recall that the pressure function $P(T, f)$ is defined as

$$P(f) = \sup_{\mu \in \mathcal{M}(X,T)} \left( \int \phi d\mu + h(\mu) \right)$$

If $\phi$ is replaced by $t\phi$ for $t \in \mathbb{R}$, then the entropy term loses relative importance as $t \to \infty$ (the thermodynamic interpretation of the parameter $t$ is as an *inverse temperature*, so that letting $t \to \infty$ is referred to as a *zero temperature limit*). The following result is generalized from[22]:

**Theorem 6.1** (Optimization as zero temperature)**.** *If $X$ compact metrizable, $T : X \to X$ continuous, $\phi \in \mathcal{C}(X)$, $h_{\text{top}}(T) < \infty$ and the entropy map $\mu \mapsto h(\mu)$ is u.s.c. For $t \in \mathbb{R}$, if $\mu_t$ is an equilibrium state for $t\phi$ in problem (II), then $(\mu_t)$ has at least one accumulation point $\mu^* \in \mathcal{M}_Z(X, T)$ as $t \to \infty$, and:*

*(i)* $\mu^* \in \underset{\mu \in \mathcal{M}_Z(X,T)}{\operatorname{argmax}} \int \phi d\mu,$

*(ii)* $h(\mu^*) = \max\{h(m) : m \in \mathcal{M}(X, T; \phi))\},$

*(iii)* $\lim_{t \to \infty} h(\mu_t) = h(\mu^*).$

*Proof.* The existence of $\mu_t$ for every $t$ is established by Theorem 3.6(ii). By the compactness of $\mathcal{M}_Z(X, T)$, for each sequence $(\mu_t) \subset \mathcal{M}_Z(X, T)$ there exists a convergence subsequence $(\mu_{t_k}) \to \mu^* \in \mathcal{M}_Z(X, T)$.

(i) As for each $t_k$, $\mu_{t_k}$ maximizes $\int t_k\phi d\mu + h(\mu) = t_k \int \phi d\mu + \frac{1}{t_k} h(\mu)$, when $k$ is large enough, $\frac{1}{t_k} h(\mu) < \varepsilon$ (as $h_{\text{top}}(T) < \infty$), and we denote $\mu_{\max}$ as the optimizer of problem (I). Then

$$\int \phi d\mu_{t_k} \leq \int \phi d\mu_{\max} < \int \phi d\mu_{\max} + \frac{1}{t_k} h(\mu_{t_k}) \leq \int \phi d\mu_{t_k} + \frac{1}{t_k} h(\mu_{t_k}) < \int \phi d\mu_{t_k} + \varepsilon$$

with $k \to \infty$, $\varepsilon \to 0$, and we have

$$\int \phi d\mu^* = \int \phi d\mu_{\max}$$

(ii) As for any $\phi \in \mathcal{C}(X)$ and $t \in \mathbb{R}$, there exists an optimizer $\mu_t$ for $\int t\phi d\mu + h(\mu)$. So for different $t$ we get a sequence of optimizer $(\mu_t)$, and for each $k$ and any optimal plan $\mu_{\max}$ for problem (II), i.e. for any $\mu_{\max} \in \mathcal{M}(X, T; \phi)$:

$$\int t\phi d\mu_{\max} + h(\mu_{\max}) \leq \int t\phi d\mu_t + h(\mu_t)$$

$\mu_{\max}$ maximizes $\int \phi d\mu$ over $\mathcal{M}_Z(X, T)$ so $\int t\phi d\mu_{\max} \geq \int t\phi d\mu_t$, and we can get the desired result since the entropy map is upper semi-continuous:

$$h(\mu) \leq h(\mu_t) \implies h(\mu) \leq \limsup_{t\to\infty} h(\mu_t) \leq h(\mu^*)$$

(iii) We just need to show $h(\mu^*) \leq \liminf_{t\to\infty} h(\mu_t)$: Since $\mu_t$ maximizes $\int t\phi d\mu + h(\mu)$ for $\mu \in \mathcal{M}_Z(X, T)$, then

$$\int t\phi d\mu^* + h(\mu^*) \leq \int t\phi d\mu_t + h(\mu_t)$$

from $(i)$ we know $\mu^*$ maximizes $\int \phi d\mu$, then $\int t\phi d\mu^* \geq \int t\phi d\mu_t$, so

$$h(\mu^*) \leq h(\mu_t) \implies h(\mu^*) \leq \liminf_{t\to\infty} h(\mu_t)$$

Therefore, for any sequence of optimizers $(\mu_t)$ w.r.t different values $t$:

$$\limsup_{t\to\infty} h(\mu_t) \le h(\mu^*) \le \liminf_{t\to\infty} h(\mu_t) \implies \lim_{t\to\infty} h(\mu_t) = h(\mu^*)$$

$\square$

## 6.1 An application: optimal Markov joining may not be Markovian

Markov measure is an important collection of measures defined on shift spaces. Here its definition is provided below:

**Definition 6.1** (Markov Measure). *Let $\mathcal{H} = \{0, 1, \cdots, k-1\}$ ($k \ge 2$ is an fixed integer) be an alphabet and $(\Sigma, \sigma)$ be a shift space where $\Sigma = \mathcal{H}^{\mathbb{N}}$ and the shift transformation $\sigma : \Sigma \to \Sigma$ is given by $\sigma(\underline{x})_i = x_{i+1}$ for any sequence $\underline{x} := \{x_i\}_{i=0}^{\infty} \in \Sigma$. Let $\mathcal{B}(\Sigma)$ be the $\sigma$-algebra generated by the semi-algebra of cylinder sets (i.e., $[a_0^{n-1}] := \{\underline{x} \in \Sigma : x_i = a_i \text{ for } 0 \le i \le n-1\}$). Given a probability vector $\mathbf{p} := (p_0, \cdots, p_{k-1})$ with non-zero entries ($p_i > 0$ for each $i$ and $\sum_{i=0}^{k-1} p_i = 1$) and a $k \times k$ stochastic matrix (also named 'transition matrix') $P = (p_{ij})_{i,j\in\mathcal{H}}$ ($p_{ij} \ge 0$, $\sum_{i=0}^{k-1} p_{ij} = 1$) such that $\mathbf{p}P = \mathbf{p}$. A probability measure $\mu$ defined on $(\Sigma, \mathcal{B}(\Sigma))$ is called $(\mathbf{p}, P)$ Markov measure if*

*(i)* $\mu([i]) = p_i$ *for* $0 \le i \le k-1$.

*(ii) For any cylinder set* $[a_0^n] \in \mathcal{B}(\Sigma)$,

$$\mu([a_0^n]) = p_{a_0} p_{a_0 a_1} p_{a_1 a_2} \cdots p_{a_{n-1} a_n}$$

By Walters[5], every Markov measure is ergodic and stationary. In the following we will consider Markov joinings. Given two Markov measures $\mu$ and $\nu$, where $\mu$ is $(\mathbf{p}, P)$ Markov measure, defined on the full-2 shift space $(\{A, B\}^{\mathbb{N}}, \sigma)$ and $\nu$ is $(\mathbf{q}, Q)$

Markov measure, defined on the full-2 shift space $(\{C, D\}^{\mathbb{N}}, \sigma)$. And

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix}, Q = \begin{bmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{bmatrix}, \quad \mathbf{p} = \mathbf{q} = [{}^1\!/\!_2, {}^1\!/\!_2]$$

for $0 < \alpha < 1$. Ellis[53] proved that the $\bar{d}$-metric between two Markov processes may not be attained by a Markov joining. $\bar{d}$-metric, or $\bar{d}$-distance, is a metric between two discrete time processes with finite or countable states, measures how different the two processes are. When $\mu$ and $\nu$ are stationary, we have the following useful result.

**Proposition 6.2.** *If $\mu$ and $\nu$ are stationary, then*

$$\bar{d}(\mu, \nu) = \inf_{\lambda \in J(\mu,\nu)} \int c d\lambda$$

*where* $c(x, y) = \mathbf{1}(x_0 \neq y_0) = \begin{cases} 1, & \textit{if } x_0 \neq y_0 \\ 0, & \textit{if } x_0 = y_0 \end{cases}$

Please refer to Shields[54] for detailed proof. Let $\mathcal{MJ}(\mu, \nu)$ be the set of Markov joinings of $\mu$ and $\nu$, next result is proved by Ellis[53].

**Proposition 6.3.** *Suppose that $\mu$ and $\nu$ are Markov measures as given above, for $0 < \alpha < 1/2$*

$$\bar{d}(\mu, \nu) = \frac{1}{2}(1 - 2\alpha), \quad M(\mu, \nu) = \inf_{\lambda \in \mathcal{MJ}(\mu,\nu)} \int c d\lambda = \frac{1 - 2\alpha}{2 - 2\alpha}$$

*where $M(\mu, \nu)$ is called 'Markov distance' of $\mu$ and $\nu$.*

The next theorem shows that in our framework of problem (II), if the linear optimization problem is over joinings of ergodic Markov measures, the optimal plan may not be Markovian.

**Theorem 6.4.** *Let $\mu$ and $\nu$ be Markov measures as given above. For $\phi = -c$, if $\varepsilon$ is fixed and small enough, then*

$$\sup_{\lambda \in J(\mu,\nu)} \int \phi d\lambda + \varepsilon h(\lambda) > \sup_{\lambda \in \mathcal{MJ}(\mu,\nu)} \int \phi d\lambda + \varepsilon h(\lambda)$$

To prove Theorem 6.4 we need the following properties of $\mathcal{MJ}(\mu,\nu)$:

**Lemma 6.5.** *For Markov measures $\mu$ and $\nu$ defined on $(\Sigma_\mu, \sigma)$ and $(\Sigma_\nu, \sigma)$, respectively, $\mathcal{MJ}(\mu,\nu)$ is a non-empty and compact subset in $\mathcal{M}(\Sigma_\mu \times \Sigma_\nu)$.*

*Proof.* It is obvious that $\mathcal{MJ}(\mu,\nu)$ is non-empty because $\mu \otimes \nu$ is a Markov joining of $\mu$ and $\nu$. Now we talk about the compactness, we just need to show that for a sequence of Markov joinings $\{\lambda_n\}_n \in \mathcal{MJ}(\mu,\nu)$ and $\lambda_n \to \lambda$, we have $\lambda \in \mathcal{MJ}(\mu,\nu)$. Firstly, $\lambda$ is a joining of $\mu$ and $\nu$ by Theorem 3.2(i), so we just need to show that $\lambda$ satisfies Markov property. Assume that $\mathcal{A}_\mu = \{a_0, a_1, \cdots, a_{k-1}\}$, $\Sigma_\mu = \mathcal{A}_\mu^{\mathbb{N}}$ and $\mathcal{A}_\nu = \{b_0, b_1, \cdots, b_{l-1}\}$, $\Sigma_\nu = \mathcal{A}_\nu^{\mathbb{N}}$. Denote by $T = \sigma \times \sigma$, $\lambda$ is a Markov measure on product space $(\Sigma_\mu \times \Sigma_\nu, \sigma \times \sigma)$ if and only if for every integer $m > 0$ and for every sequence $E_1, \cdots, E_m$ where $E_i \in \mathcal{A}_\mu \times \mathcal{A}_\nu$, if $I = TE_1 \cap T^2E_2 \cap \cdots \cap T^mE_m$, $c, d \in \Sigma_\mu \times \Sigma_\nu$, $F = I \cap c$

$$\lambda(TI \cap d | I) = \lambda(Tc \cap d | c)$$

Since each $\lambda_n$ satisfies the Markov property above and $\lambda_n \to \lambda$, so $\lambda$ is a Markov measure. Thus we proved the closedness and compactness of $\mathcal{MJ}(\mu,\nu)$. □

Now we are ready to prove Theorem 6.4:

*Proof of Theorem 6.4.* Assume that for each fixed $\varepsilon$, the optimizer of

$$\sup_{\lambda \in J(\mu,\nu)} \int \phi d\lambda + \varepsilon h(\lambda)$$

is denoted as $\lambda_{\varepsilon}^{J}$, correspondingly, the optimizer of

$$\sup_{\lambda \in \mathcal{MJ}(\mu,\nu)} \int \phi d\lambda + \varepsilon h(\lambda)$$

is denoted as $\lambda_{\varepsilon}^{MJ}$. Since an upper semi continuous real-valued function of a compact space attains its supremum, for each $\varepsilon_n$, optimizers $\lambda_{\varepsilon}^{J}$ and $\lambda_{\varepsilon}^{MJ}$ always exist.

We pick a sequence of numbers $\{\varepsilon_n\}_n$ with $\varepsilon_n \to 0$. For each $n$ and thus $\varepsilon_n$ there exist $\lambda_{\varepsilon_n}^{J}$ and $\lambda_{\varepsilon_n}^{MJ}$. In the following, using $\lambda_n^J$ and $\lambda_n^{MJ}$ to denote $\lambda_{\varepsilon_n}^{J}$ and $\lambda_{\varepsilon_n}^{MJ}$ respectively. By compactness, in sequence $\{\lambda_n^J\}_n$ there exists a subsequence $\{n_k\}_k$ s.t $\lambda_{n_k}^J \to \lambda^J$, and in subsequence $\{\lambda_{n_k}^{MJ}\}_k$ there exists a subsubsequence $\{n_{k_l}\}_l$ s.t $\{\lambda_{n_{k_l}}^{MJ}\}_l \to \lambda^{MJ}$. Note that $\{\lambda_{n_{k_l}}^{J}\}_l \to \lambda^J$ as well. By Theorem 6.1(i),

$$\lambda^J \in \operatorname*{argmax}_{\lambda \in \mathcal{J}(\mu,\nu)} \int \phi d\lambda, \quad \lambda^{MJ} \in \operatorname*{argmax}_{\lambda \in \mathcal{MJ}(\mu,\nu)} \int \phi d\lambda$$

By convergence result above, for any $\delta < \frac{\alpha}{8(1-\alpha)}$, there exists $l_\delta$ s.t when $l > l_\delta$, we have

$$\left| \int \phi d\lambda_{\varepsilon_{n_{k_l}}}^{J} - \int \phi d\lambda^J \right| < \delta \quad \text{and} \quad \left| \int \phi d\lambda_{\varepsilon_{n_{k_l}}}^{MJ} - \int \phi d\lambda^{MJ} \right| < \delta$$

In addition, as the topological entropy $h_{\text{top}}(\sigma \times \sigma)$ is bounded, there exists $l_h > 0$ s.t. when $l > l_h$, we have

$$\varepsilon_{n_{k_l}} < \frac{\alpha}{4(1-\alpha) h_{\text{top}}(\sigma \times \sigma)}$$

Now let $l_m = \max\{l_\delta, l_h\}$, then for $l > l_m$:

$$
\begin{aligned}
&\left[\int \phi d\lambda^J_{\varepsilon_{n_{k_l}}} + \varepsilon_{n_{k_l}} h(\lambda^J_{\varepsilon_{n_{k_l}}})\right] - \left[\int \phi d\lambda^{MJ}_{\varepsilon_{n_{k_l}}} + \varepsilon_{n_{k_l}} h(\lambda^{MJ}_{\varepsilon_{n_{k_l}}})\right] \\
>& \left[\int \phi d\lambda^J - \delta + \varepsilon_{n_{k_l}} h(\lambda^J_{\varepsilon_{n_{k_l}}})\right] - \left[\int \phi d\lambda^{MJ} + \delta + \varepsilon_{n_{k_l}} h(\lambda^J_{\varepsilon_{n_{k_l}}})\right] \\
=& -\frac{1}{2}(1-2\alpha) - \left(-\frac{1-2\alpha}{2-2\alpha}\right) - 2\delta + \varepsilon_{n_{k_l}}[h(\lambda^J_{\varepsilon_{n_{k_l}}}) - h(\lambda^{MJ}_{\varepsilon_{n_{k_l}}})] \\
>& \frac{\alpha}{2-2\alpha} - 2\delta - \varepsilon_{n_{k_l}} h_{\text{top}}(\sigma \times \sigma) \\
>& \frac{\alpha}{2-2\alpha} - \frac{\alpha}{4(1-\alpha)} - \frac{\alpha}{4(1-\alpha)} = 0
\end{aligned}
$$

Therefore, if we select $l^* > l_m$ defined above, the problem $\sup\limits_{\lambda \in J(\mu,\nu)} \int \phi d\lambda + \varepsilon_{n_{k_{l^*}}} h(\lambda)$ achieves greater maximal value than the problem $\sup\limits_{\lambda \in \mathcal{MJ}(\mu,\nu)} \int \phi d\lambda + \varepsilon_{n_{k_{l^*}}} h(\lambda)$. So if $\phi = -{}^{c}/_{\varepsilon_{n_{k_{l^*}}}}$, the optimal plan of problem (II) is not Markovian. $\square$

# CHAPTER 7: REALIZATION PROBLEM

Let $\mathcal{M}_Z(X,T)$ denote the set of $T$-invariant Borel probability measures on a compact metrizable space $X$ and $\int h d\mu = 0 \; \forall h \in Z$ for any $\mu \in \mathcal{M}_Z(X,T)$. The following realization result, generalizes Jenkinson[22], shows that if $\mathcal{E}$ is a non-empty collection of ergodic measures which is weak* closed as a subset of $\mathcal{M}_Z(X,T)$, then there is a continuous function $\phi$ such that set of optimal plans in problem (I) is the closed convex hull of $\mathcal{E}$. That is, if $\overline{\text{co}}(\mathcal{E})$ denotes the closed convex hull of $\mathcal{E}$, then $\exists \phi \in \mathcal{C}(X)$ s.t. $\mathcal{M}(X,T;\phi) = \overline{\text{co}}(\mathcal{E})$. Such realization result is also established when applied to problem (II): let $\mathcal{E}$ be a weak* closed set of ergodic measures in $\mathcal{M}_Z(X,T)$, if the entropy map $\mu \mapsto h(\mu)$ is continuous on $\overline{\text{co}}(\mathcal{E})$, there exists $\phi \in \mathcal{C}(X)$ s.t. $\mathcal{R}(X,T;\phi) = \overline{\text{co}}(\mathcal{E})$.

## 7.1 Settings

As the basic settings in section 1.2.1, $\mathcal{C}(X)$ is the space of continuous real-valued functions on $X$ (where $(X,T)$ is a compact, metrizable dynamical system). By Royden and Fitzpatrick[55], $\mathcal{C}(X)$ is a real Banach space when equipped with the supremum norm $\|\cdot\|_{\max}$. Consider the topological dual of $\mathcal{C}(X)$, denoted by $\mathcal{C}(X)^*$, which is the vector space of continuous linear functionals on $\mathcal{C}(X)$. The famous Riesz representation theorem (see Walters[5] Thm. 6.3) tells us that for each element of $J \in \mathcal{C}(X)^*$, there is a signed Borel measures $\mu$ on $X$ s.t.

$$J(\phi) = \langle \phi, \mu \rangle = \int \phi d\mu \tag{7.1}$$

Furthermore, if $J$ is a normalized positive operator (i.e., if $\phi \geq 0$ then $J(\phi) \geq 0$ and $J(1) = 1$). Then there is a Borel probability measure $\mu$ on $X$ with (7.1) satisfied.

Let $(\mathcal{C}(X)^*, w^*)$ denote the dual space $\mathcal{C}(X)^*$ equipped with the weak* topology.

By definition (see Royden and Fitzpatrick[55]), this is the weakest topology s.t. for every $\phi \in C(X)$, the linear functional $J$ on $\mathcal{C}(X)^*$ given by $\mu \mapsto \langle \phi, \mu \rangle = \int \phi d\mu$ is continuous. By Schaefer[56], weak* topology is locally convex, being generated by the family of semi-norms $\{p_\phi : \phi \in C(X)\}$, where $p_\phi(\lambda) = |\langle \phi, \lambda \rangle|$. Since $X$ is compact and metrizable, $C(X)$ is separable[55]. Therefore, from [57] Thm. 10.7 the closed unit ball $B = \{J \in \mathcal{C}(X)^* : \|J\| \leq 1\}$ in the dual space is metrizable w.r.t the weak* topology.

By the related results from [57] and [58], the dual space $\mathcal{C}(X)^*$ is also a Riesz space: it is an ordered vector space w.r.t the (convex pointed) cone $C$ of all positive Borel measures on $X$, and it is a lattice with the operations $\vee$ and $\wedge$ given by

$$(\mu \vee \nu)(C) = \sup_{D \in \mathcal{A} \otimes \mathcal{B}, D \subset C} \{\mu(D) + \nu(C \backslash D)\} \tag{7.2}$$

$$(\mu \wedge \nu)(C) = \inf_{D \in \mathcal{A} \otimes \mathcal{B}, D \subset C} \{\mu(D) + \nu(C \backslash D)\} \tag{7.3}$$

Let $E_{T,Z}$ denote the set of $Z$-restricted signed $T$-invariant measures. That is,

$$\mu \in E_{T,Z} \iff \mu \circ T^{-1} = \mu \text{ and } \mu(h) = 0 \ \forall h \in Z$$

It is obvious that $E_{T,Z}$ is a vector space. Let $B_{T,Z}(X)$ denote the closure in $\mathcal{C}(X)$ of the vector subspace generated by the set $\{f - f \circ T + h : f \in C(X), h \in Z\}$. It is easily shown that a measure $\mu \in E$ is a member of $E_{T,Z}$ iff $\int g d\mu = 0$ for all $g \in B_{T,Z}(X)$. Since $E$ is the topological dual of $C(X)$, we deduce that

$$E_{T,Z} = (C(X)/B_{T,Z}(X))^*$$

that is, $E_{T,Z}$ is the topological dual of the quotient Banach space $C(X)/B_{T,Z}(X)$. [Obviously, $E_{T,Z} \subseteq (C(X)/B_{T,Z}(X))^*$. Since $\mu \in (C(X)/B_{T,Z}(X))^*$ implies that $\forall g = f - f \circ T + h \in B_{T,Z}(X)$, $\int g d\mu = 0$, so $(C(X)/B_{T,Z}(X))^* \subseteq E_{T,Z}$.]

## 7.2 Necessary lemmas

To prove the realization result, in this section we will introduce four necessary lemmas and their proofs:

**Lemma 7.1.**

$$B_{T,Z}(X) = \{h \in C(X) : \langle h, \mu \rangle = 0 \textit{ for all } \mu \in E_{T,Z}\}$$

*Proof.* Since $E_{T,Z}$ is the topological dual of $C(X)/B_{T,Z}(X)$, it follows that the topological dual of $(E_{T,Z}, w^*) = ((C(X)/B_{T,Z}(X))^*, w^*)$ is precisely $C(X)/B_{T,Z}(X)$ (by [57] Thm. 3.16).

But there is another expression for the topological dual of $(E_{T,Z}, w^*)$, namely

$$(E_{T,Z}, w^*)^* = C(X)/\mathrm{Ann}(E_{T,Z}) \tag{7.4}$$

where $\mathrm{Ann}(E_{T,Z}) = \{h \in C(X) : \langle h, \mu \rangle = 0 \text{ for all } \mu \in E_{T,Z}\}$ denotes the annihilator of $E_{T,Z}$. To verify equation (7.4), first note that by Hahn-Banach, any continuous linear functional on $(E_{T,Z}, w^*)$ is the restriction of a continuous linear functional on $(\mathcal{C}(X)^*, w^*)$. Such a functional can therefore be identified with an element of $C(X)$, which is a topological dual of $(\mathcal{C}(X)^*, w^*)$ by [57] Thm. 3.16 and the Riesz representation theorem. But two elements of $C(X)$ yield the same functional on $E_{T,Z}$ iff their difference lies in $\mathrm{Ann}(E_{T,Z})$, so (7.4) follows. Comparison of the two expressions for $(E_{T,Z}, w^*)^*$ yields the result. $\square$

The duality of the pair $(C(X)/B_{T,Z}(X), E_{T,Z})$ will be denoted by

$$(\theta, \mu) \mapsto \langle \theta, \mu \rangle$$

which is consistent with the duality of $(C(X), E)$ in the sense that $\langle \phi, \mu \rangle = \langle \theta, \mu \rangle$ for

$\phi \in \theta \in C(X)/B_{T,Z}(X)$ and $\lambda \in E_{T,Z}$.

By our settings, $\mathcal{M}(X,T)$ is the set of $T$-invariant Borel probability measures. Clearly $\mathcal{M}(X,T)$ is convex.

**Definition 7.1** (Extreme points)**.** *Suppose that $K$ is a convex subset of a vector space $V$, a point $x \in K$ is called an extreme point of $K$ if whenever $x = \alpha x_1 + (1-\alpha)x_2$ for some $x_1, x_2 \in K$ and $0 < \alpha < 1$, then $x = x_1 = x_2$.*

If $K$ is contained in a hyperplane that does not contain the origin, it is called a *simplex* if the cone $P = \{ck : c \geq 0, k \in K\}$ defines a lattice ordering on $P - P \subseteq V$. Recall that $\mathcal{M}_Z(X,T) \subset \mathcal{M}(X,T)$ is the set of invariant Borel probability measures that is equal to 0 on $Z$.

The following lemma details some classical facts about $\mathcal{M}_Z(X,T)$.

**Lemma 7.2.** *If $\mathcal{M}_Z(X,T)$ is nonempty and satisfies property $(E)$:*

*(i) $(\mathcal{M}_Z(X,T), w^*)$ is compact and metrizable.*

*(ii) $\mathcal{M}_Z(X,T)$ is a simplex.*

*(iii) The set of extreme points of $\mathcal{M}_Z(X,T)$ is precisely $\mathcal{M}_Z^e(X,T)$, i.e., set of ergodic $Z$-restricted $T$-invariant Borel probability measures.*

*Proof.*

(i) This fact is by [7].

(ii) Since $\mathcal{M}_Z(X,T)$ lies in a hyperplane in $E$ which does not contain the origin, it suffices to show that $E_{T,Z} = C_{T,Z} - C_{T,Z}$ is a sub-lattice of $E$. (In fact, $E_{T,Z}$ is a Riesz space of $E$.) To verify that $E_{T,Z}$ is a lattice with respect to the operations $\vee$ and $\wedge$ defined by (7.2) and (7.3), it suffices to show that if $\mu \in E_{T,Z}$, then $\mu^+ = \mu \vee 0 \in E_{T,Z}$. For any $A \in \mathcal{B}$, $\mu(A) = \mu(T^{-1}A) = \mu^+(T^{-1}A) - \mu^-(T^{-1}A)$. By Rudin[59], $\mu^+(T^{-1}A) \geq \mu^+(A)$. But $A^c \in \mathcal{B}$ as well, so $\mu^+((T^{-1}A)^c) = \mu^+(T^{-1}A^c) \geq \mu^+(A^c)$, therefore $\mu^+(T^{-1}A) = \lambda^+(A)$.

(iii) Obviously since $\mathcal{M}_Z(X,T)$ satisfies property $(E)$.

□

In the following we will introduce some geometry definitions of vector spaces:

**Definition 7.2** (Convex hull)**.** *Suppose that $G$ is a non-empty subset of a convex set $K$, its convex hull* $\mathrm{co}(G)$ *is the smallest convex set containing $G$. Its closed convex hull* $\overline{\mathrm{co}}(G)$ *is the smallest closed convex set containing $G$, and it equals the closure of* $\mathrm{co}(G)$*.*

**Definition 7.3** (Face)**.** *A non-empty convex subset $F$ of $K$ is called a face of $K$ if whenever $\alpha x_1 + (1-\alpha)x_2 \in F$ for some $x_1, x_2 \in K$ and $\alpha \in (0,1)$, then $x_1, x_2 \in F$.*

**Remark 7.1.** *In the following sections we will focus on closed faces. the simplest closed faces are singletons $\{k\}$, where $k \in K$ is an extreme point.*

The following lemma summarises certain classical properties of the closed faces of $\mathcal{M}_Z(X,T)$, which follow from the fact that it is a simplex and that $\mathcal{M}_Z^e(X,T)$ is its set of extreme points.

**Lemma 7.3.** *If $\mathcal{M}_Z(X,T)$ is nonempty and satisfies property $(E)$:*

*(i) Every closed face $\mathcal{F}$ of $\mathcal{M}_Z(X,T)$ is of the form $\overline{\mathrm{co}}(\mathcal{E})$ for some non-empty subset $\mathcal{E}$ of $\mathcal{M}_Z^e(X,T)$.*

*(ii) If $\mathcal{E}$ is a non-empty subset of $\mathcal{M}_Z^e(X,T)$ which is closed in $\mathcal{M}_Z(X,T)$, then $\overline{\mathrm{co}}(\mathcal{E})$ is a face of $\mathcal{M}_Z(X,T)$.*

*(iii) If $\mathcal{M}_Z^e(X,T)$ is closed in $\mathcal{M}_Z(X,T)$, and $\mathcal{E}$ is any non-empty subset of $\mathcal{M}_Z(X,T)$, then $\overline{\mathrm{co}}(\mathcal{E})$ is a face of $\mathcal{M}_Z(X,T)$.*

*Proof.* (i) If $\mathcal{F} \subset \mathcal{M}_Z(X,T)$ is a closed face, then by Krein-Milman theorem, let $\mathcal{E}$ be the set of extreme points of $\mathcal{F}$, we have $\mathcal{F} = \overline{\mathrm{co}}(\mathcal{E})$. By [60] Prop. 2: If $F$

is a face of $K$, any point $x$ extreme in $F$ is also extreme in $K$, any point in $\mathcal{E}$ is also extreme in $\mathcal{M}_Z(X,T)$. And then by property (E), the extreme point in $\mathcal{M}_Z(X,T)$ is also ergodic. Therefore $\mathcal{E}$ must be a subset of $\mathcal{M}_Z^e(X,T)$.

(ii) By Effros[61] Thm 3.3 and Cor. 3.5 we can prove the desired result.

(iii) This result is by Alfsen[60] Prop. 4: Every closed face $F$ of $K$ can be represented in the form $F = \overline{\mathrm{co}}(\mathcal{E})$ where $\mathcal{E}$ is the subset of extreme points of $K$. If $K$ is a simplex and its set of extreme points is closed, then every set $F$ of the form is a closed face.

□

**Definition 7.4** (Affine functional)**.** *Suppose that $K$ is a convex subset of a topological vector space. A functional $\ell : K \to \mathbb{R}$ is affine if it satisfies the following*

$$\ell(\alpha x_1 + (1-\alpha)x_2) = \alpha\ell(x_1) + (1-\alpha)\ell(x_2)$$

*for all $x_1, x_2 \in K$ and $\alpha \in [0,1]$.*

In the following we will introduce a property related to functions that are continuous and affine.

**Definition 7.5** (Exposed face[62])**.** *Suppose that $K$ is a convex subset of a topological vector space. A face $F$ of $K$ is called exposed if there exists a continuous affine functional $\ell : K \to \mathbb{R}$ such that $\ell(x) = 0$ for all $x \in F$ and $\ell(x) > 0$ for all $x \in K \backslash F$. In particular, if a point $k \in K$ satisfies $\{k\}$ as an exposed face (which means that $k$ must be an extreme point), then $k$ is an exposed point.*

**Remark 7.2.** *Note that the continuity of $\ell$ means that any exposed face is necessarily closed.*

A result which is necessary in proving the main theorems of this chapter is that if $K$ is a compact metrizable simplex, then all of its closed faces are exposed:

**Lemma 7.4.** *Let $\mathcal{F}$ be a closed face of $\mathcal{M}_Z(X,T)$. There exists an affine function $\ell : \mathcal{M}_Z(X,T) \to \mathbb{R}$, continuous in weak* topology, such that $\ell(\mu) = 0$ for all $\mu \in \mathcal{F}$, and $\ell(\nu) > 0$ for all $\nu \in \mathcal{M}_Z(X,T)\backslash\mathcal{F}$.*

Such result was firstly proved by Davies[63], for a detailed proof please refer to Alfsen[62].

## 7.3 Realization theorems and proofs

Define $\ell_{\tau,f}(\mu) := \int f d\mu + \tau \cdot h(\mu)$ and $L_\tau(f) := \sup_{\mu\in\mathcal{M}_Z(X,T)} \ell_{\tau,f}(\mu)$ for $\tau \in \{0,1\}$. Denote by $\mathcal{M}(X,T;\phi) \subset \mathcal{M}_Z(X,T)$ and $\mathcal{R}(X,T;\phi) \subset \mathcal{M}_Z(X,T)$ the set of measures maximize $\ell_{\tau,\phi}(\mu)$ over $\mathcal{M}_Z(X,T)$ for $\tau = 0$ and $\tau = 1$ respectively (that is, $\mu' \in \mathcal{M}(X,T;\phi) \iff \ell_{0,\phi}(\mu') = L_0(\phi)$ and $\mu' \in \mathcal{R}(X,T;\phi) \iff \ell_{1,\phi}(\mu') = L_1(\phi)$).

The following four theorems are main results of this chapter:

**Theorem 7.5.** *Let $\mu$ be any ergodic measure on $\mathcal{M}_Z(X,T)$.*

*(i) There exists a continuous function $\phi : X \to \mathbb{R}$ such that $\mathcal{M}(X,T;\phi) = \{\mu\}$.*

*(ii) There exists a continuous function $\psi : X \to \mathbb{R}$ such that $\mathcal{R}(X,T;\psi) = \{\mu\}$.*

**Theorem 7.6.** *Let $\mathcal{E}$ be a non-empty subset of $\mathcal{M}^e_Z(X,T)$ which is weak* closed in $\mathcal{M}_Z(X,T)$. Let $\overline{\mathrm{co}}(\mathcal{E})$ denote its closed convex hull in $\mathcal{M}_Z(X,T)$. There exists a continuous function $\phi : X \to \mathbb{R}$ such that $\mathcal{M}(X,T;\phi) = \overline{\mathrm{co}}(\mathcal{E})$. Furthermore, if $h|_{\overline{\mathrm{co}}(\mathcal{E})}$ is continuous, then there exists a continuous function $\psi : X \to \mathbb{R}$ such that $\mathcal{R}(X,T;\psi) = \overline{\mathrm{co}}(\mathcal{E})$.*

If $\mathcal{M}^e_Z(X,T)$ happens to be a weak* closed subset of $\mathcal{M}_Z(X,T)$ (which in general it is not), then the conclusion of Theorem 7.6 applies if $\mathcal{E}$ is any non-empty subset of $\mathcal{M}^e_Z(X,T)$.

**Theorem 7.7.** *Suppose that $\mathcal{M}^e_Z(X,T)$ is a weak* closed subset of $\mathcal{M}_Z(X,T)$. For every non-empty subset $\mathcal{E} \subset \mathcal{M}^e_Z(X,T)$, there exists a continuous function $\phi : X \to \mathbb{R}$*

*such that* $\mathcal{M}(X,T;\phi) = \overline{\text{co}}(\mathcal{E})$. *Furthermore, if* $h|_{\overline{\text{co}}(\mathcal{E})}$ *is continuous, then there exists a continuous function* $\psi : X \to \mathbb{R}$ *such that* $\mathcal{R}(X,T;\psi) = \overline{\text{co}}(\mathcal{E})$.

We have the following characterization of those subsets of $\mathcal{M}_Z(X,T)$ which are of the form $\mathcal{M}(X,T;\phi)$ and $\mathcal{R}(X,T;\phi)$ for some $\phi \in C(X)$.

**Theorem 7.8.**

*(i) The set* $\{\mathcal{M}(X,T;\phi) : \phi \in C(X)\}$ *is precisely the set of closed faces of* $\mathcal{M}_Z(X,T)$.

*(ii) The set* $\{\mathcal{R}(X,T;\phi) : \phi \in C(X)\}$ *is precisely the set of closed faces of* $\mathcal{M}_Z(X,T)$ *on which the entropy map* $\mu \mapsto h(\mu)$ *is continuous.*

The following proposition is important in proving theorems above:

**Proposition 7.9.** *Suppose* $\ell : \mathcal{M}_Z(X,T) \to \mathbb{R}$ *is weak* continuous and affine. there exists* $\phi \in C(X)$ *such that*

$$\ell(\mu) = \langle \phi, \mu \rangle = \int \phi d\mu \quad \textit{for all } \mu \in \mathcal{M}_Z(X,T)$$

(Which means an element of $C^{**}(X)$ can be identified with an element of $C(X)$.)

*Proof of Theorems.* First, we prove Theorem 7.8. By lemma 7.4, a subset $\mathcal{F}$ of $\mathcal{M}_Z(X,T)$ is a closed face of $\mathcal{M}_Z(X,T)$ if and only if there exists a weak* continuous affine functional $\ell : \mathcal{M}_Z(X,T) \to \mathbb{R}$ such that $\ell(\mu) = 0$ when $\mu \in \mathcal{F}$, and $\ell(\nu) > 0$ when $\nu \in \mathcal{M}_Z(X,T)\backslash\mathcal{F}$. By proposition 7.9, we may write $\ell(\mu) = \int \phi d\mu$ for some $\phi \in C(X)$, so $\mathcal{F}$ is a closed face of $\mathcal{M}_Z(X,T)$ if and only if there exists a continuous function $\psi(= -\phi)$ such that $\int \psi d\mu = 0$ for all $\mu \in \mathcal{F}$ and $\int \psi d\nu < 0$ for all $\nu \in \mathcal{M}_Z(X,T)\backslash\mathcal{F}$. That is, $\mathcal{F}$ is a closed face of $\mathcal{M}_Z(X,T)$ if and only if $\mathcal{F} = \mathcal{M}(X,T;\psi)$ for some $\psi \in C(X)$, so Theorem 7.8 is proved.

If $\mathcal{E}$ is a non-empty subset of $\mathcal{M}^e_Z(X,T)$, then $\overline{\text{co}}(\mathcal{E})$ is a closed face of $\mathcal{M}_Z(X,T)$ provided either $\mathcal{E}$ is closed in $\mathcal{M}_Z(X,T)$ (by Lemma 7.3 (ii)) or $\mathcal{M}^e_Z(X,T)$ is closed in

$\mathcal{M}_Z(X,T)$ (by Lemma 7.3 (iii)). In either case, Theorem 7.8 implies the existence of some $\psi \in C(X)$ for which $\mathcal{M}(X,T;\psi) = \overline{\text{co}}(\mathcal{E})$, so Theorems 7.6 and 7.7 are proved. Theorem 7.5 follows immediately from Theorem 7.6, since the singleton $\{\mu\}$ ($\mu$ is ergodic) is a non-empty subset of $\mathcal{M}_Z^e(X,T)$ and is closed in $\mathcal{M}_Z(X,T)$. $\square$

It remains to prove Proposition 7.9. We can follow the proof of Proposition 1 in Jenkinson[14] with some replacements to finish Proposition 7.9's proof. For completeness the sketch of the proof is shown as below.

*Proof of Proposition 7.9.*

By lemma 7.1, it suffices to find $\theta = g + B_{T,Z}(X) \in C(X)/B_{T,Z}(X)$ such that $\ell(\mu) = \langle \theta, \mu \rangle$ for all $\mu \in \mathcal{M}_Z(X,T)$.

Define $C_{T,Z}$ the cone of $Z$-restricted positive invariant measures as

$$C_{T,Z} := \{c\mu : c \geq 0, \mu \in \mathcal{M}_Z(X,T)\}$$

and then we can take the following steps to prove this proposition:

(i) Define $\ell_1 : C_{T,Z} \to \mathbb{R}^+$ by setting $\ell_1(0) = 0$ and $\ell_1(m) = c\ell_1(\mu)$. $\ell_1$ is additive and weak* continuous.

(ii) Extend $\ell_1$ to a functional $\ell_2 : E_{T,Z} \to \mathbb{R}$ defined by

$$\ell_2(\mu) = \ell_1(\mu^+) - \ell_1(\mu^-)$$

$\ell_2$ is well-defined, linear and weak* continuous

(iii) Since the topological dual of $(E_{T,Z}, w^*) = ((C(X)/B_{T,Z}(X))^*, w^*)^*$ is $C(X)/B_{T,Z}(X)$. Therefore there exists a unique $\theta \in C(X)/B_{T,Z}(X)$ s.t. $\ell_2(\mu) = \langle \theta, \mu \rangle$ for all $\mu \in E_{T,Z}$. Since $\ell_2$ is an extension of $\ell : \mathcal{M}_Z(X,T) \to \mathbb{R}$, so the proposition is proved.

□

The last necessary result is a special case of Isreal and Phelps[64] Proposition 3.9:

**Proposition 7.10** ([64] Prop. 3.9)**.** *The basic settings are as section 1.2.1 and suppose that $Z \subset \mathcal{C}(X)$ closed s.t. $\mathcal{M}_Z(X,T)$ satisfies property (E). If $F \subseteq \mathcal{M}_Z(X,T)$ is a nonempty closed face of $\mathcal{M}_Z(X,T)$ and entropy map $\mu \mapsto h(\mu)$ is continuous on $F$ then there exists weak* continuous and affine functional $\ell \in \mathcal{M}_Z(X,T)^*$ such that*

$$\mu' \in F \iff \mu' \in \underset{\mu \in \mathcal{M}_Z(X,T)}{\operatorname{argmax}} \ \ell(\mu) + h(\mu)$$

Now we are ready to prove Theorem 7.5 - 7.8:

*Proof of Theorem 7.5 - 7.8.* Firstly we are going to prove Theorem 7.8: For statement (i), for any closed face $\mathcal{F}$ of $\mathcal{M}_Z(X,T)$, by Lemma 7.4, it is exposed so there exists a affine and weak* continuous functional $\ell$ s.t. $\ell|_{\mathcal{F}} \equiv 0$ and $\ell|_{\mathcal{M}_Z(X,T)\backslash\mathcal{F}} > 0$. Furthermore, from Proposition 7.9 we can find a $\phi \in \mathcal{C}(X)$ s.t. $\ell(\mu) = \int \phi d\mu$. Therefore, for $-\phi \in \mathcal{C}(X)$, we have

$$\int (-\phi) d\mu = 0, \ \forall \mu \in \mathcal{F} \quad \text{and} \quad \int (-\phi) d\mu < 0, \ \forall \mu \in \mathcal{M}_Z(X,T)\backslash\mathcal{F}$$

Thus, $\mathcal{F} = \mathcal{M}(X,T;-\phi)$. As for each $\phi \in \mathcal{C}(X)$, $\mathcal{M}(X,T;\phi)$ is a closed face (nonempty, closed, and for any $\alpha\mu_1 + (1-\alpha)\mu_2 \in \mathcal{M}(X,T;\phi)$ $(\alpha \in (0,1))$, $\mu_1, \mu_2 \in \mathcal{M}(X,T;\phi)$). So we have proved (i). The proof of statement (ii) is directly followed by the combination of Proposition 7.9 and Proposition 7.10.

In Theorem 7.6, $\mathcal{E} \subseteq \mathcal{M}_Z^e(X,T)$ and weak* closed in $\mathcal{M}(X,T)$, by Lemma 7.3(ii), $\mathrm{co}(\mathcal{E})$ is a closed face of $\mathcal{M}_Z(X,T)$. Similarly, in Theorem 7.7, $\mathcal{E} \subseteq \mathcal{M}_Z^e(X,T)$ and $\mathcal{M}_Z^e(X,T)$ is weak* closed in $\mathcal{M}(X,T)$, by Lemma 7.3(iii), $\mathrm{co}(\mathcal{E})$ is a closed face of $\mathcal{M}_Z(X,T)$. Therefore Theorem 7.6 and Theorem 7.7 are followed by Theorem 7.8.

Theorem 7.5 is a special case for Theorem 7.6 (the singleton $\mathcal{E} = \{\mu\}$ is a closed face), so it is satisfied. $\square$

# CHAPTER 8: DUALITY PROBLEMS

In this project, our problems (I) and (II) are linear optimization problems, so it is very natural to explore their duality problems. This chapter consists of two sections; in the first section we studied the duality of problem (I) - Kantorovich duality, its name origins from the duality problem of optimal transport[16]; in the second section we explored the duality of problem (II), which is called Fenchel duality.

## 8.1 For problem (I) - Kantorovich duality

Given a dynamical system $(X, T)$ with $X$ compact and metrizable and $T : X \to X$ continuous. Let $Z \subset C(X)$ be a linear subspace. $\nu$ is a bounded linear functional defined on $Z$, $1 \in Z$ and $\nu(1) = 1$, and denote

$$\mathcal{M}_\nu(X, T) := \{\mu \in \mathcal{M}(X, T) : \mu|_Z = \nu\}$$

If there is another system $(Y, S)$. Consider the product dynamical system $(X \times Y, T \times S)$, let $Z \subset C(X \times Y)$ be a linear subspace, $\rho$ is a fixed probability measure defined on $Z$, denote

$$\mathcal{M}_\rho(X \times Y, T \times S) := \{\lambda \in \mathcal{M}(X \times Y, T \times S) : \lambda|_Z = \rho\}$$

The following result is the main theorem of this chapter:

**Theorem 8.1.** *Let $W = \{g \circ T - g : g \in C(X)\}$, define*

$$\Pi_W(\nu) = \{\mu \in \mathcal{P}(X) : \int w d\mu = 0, \forall w \in W \ \textit{and}\ \mu|_Z = \nu\}$$

*we can prove that* $\Pi_W(\nu) = \mathcal{M}_\nu(X,T)$, *and the following restricted Kantorovich-form duality:*

$$\inf_{\mu\in\Pi_W(\nu)} \int c d\mu = \sup_{\substack{f+w\le c\\ w\in W}} \nu(f)$$

*where* $f \in Z$.

To prove Theorem 8.1 we need the following Kantorovich duality:

**Theorem 8.2.** *Let*

$$\Pi(\nu) = \{\mu \in \mathcal{P}(X) : \mu|_Z = \nu\}$$

*then the following duality is satisfied*

$$\inf_{\mu\in\Pi(\nu)} \int c d\mu = \sup_{f\le c} \nu(f)$$

*where* $f \in Z$.

*Proof.* As $\nu \in Z^*$ is a bounded linear functional and it is positive and continuous w.r.t $\|\cdot\|_{\max}$ norm. Let

$$U(h) = \inf_{f\in Z}\{\nu(f) : f \ge h\}$$

be a functional from $C(X)$ to $\mathbb{R}$. It can be proved that $U$ is subadditive: for $h, g \in C(X)$,

$$\begin{aligned} U(h+g) &= \inf_{f\in Z}\{\nu(f) : f \ge (g+h)\} \\ &\le \inf_{f\in Z}\{\nu(f) : f \ge g\} + \inf_{f\in Z}\{\nu(f) : f \ge h\} \\ &= U(g) + U(h) \end{aligned}$$

$U$ is positively homogeneous: for any $\alpha \in \mathbb{R}^+$

$$\begin{aligned} U(\alpha h) &= \inf_{f\in Z}\{\nu(f) : f \geq \alpha h\} \\ &= \inf_{(f/\alpha)\in Z}\{\alpha\nu(f/\alpha) : f/\alpha \geq h\} \\ &= \alpha U(h) \end{aligned}$$

Additionally, for any $t \in \mathbb{R}$, $U(th) \geq tU(h)$, we just need to show the inequality is satisfied when $t = -1$. Since $\inf E = -\sup(-E)$,

$$\begin{aligned} U(-h) &= \inf_{f\in Z}\{\nu(f) : f \geq -h\} = \inf_{f\in Z}\{\nu(f) : -f \leq h\} \\ &= \inf_{f\in Z}\{\nu(-f) : f \leq h\} = -\sup_{f\in Z}\{\nu(f) : f \leq h\} \\ &\geq -\inf_{f\in Z}\{\nu(f) : f \geq h\} = -U(h) \end{aligned}$$

the last inequality is because of

$$\sup_{f\in Z}\{\nu(f) : f \leq h\} \leq \inf_{f\in Z}\{\nu(f) : f \geq h\}$$

which follows from the positivity of the functional $\nu$.

Thus,the functional $U : C(X) \to \mathbb{R}$ is positively homogeneous and subadditive, $Z \subset C(X)$ a linear subspace, the linear functional $\nu : Z \to \mathbb{R}$ is bounded by $U$ on $Z$. Then by Hahn-Banach theorem, $\nu$ may be extended to a *linear* functional $P$ on $C(X)$ for which $P \leq U$ on $C(X)$. $P$ has the property of positivity:

Assume $P$ is not positive, then there exists $h \in C(X)$ s.t. $h \geq 0$ and $P(h) < 0$, however, the fact

$$0 < P(-h) \leq U(-h) = \inf_{f\in Z}\{\nu(f) : f \geq -h\} \leq 0$$

is contradictory. So $P$ is positive.

Let us define a new linear operator $\nu_c : \{f + tc : t \in \mathbb{R}, f \in Z\} \to \mathbb{R}$ such that

$\nu_c|_Z = \nu$ and $\nu_c(-c) = U(-c)$. By linearity of $\nu_c$ and property of $U$ we have

$$\nu_c(tc) = (-t)U(-c) \leq U(tc)$$

Thus $\nu_c$ is bounded by $U$ on its domain, so by Hahn-Banach, we can extend $\nu_c$ to the *linear* functional $P_c : C(X) \to \mathbb{R}$ such that

$$P_c|_{\{f+tc:t\in\mathbb{R},f\in Z\}} = \nu_c, P_c|_Z = \nu, P_c(-c) = U(-c), P_c \leq U$$

By the construction fo linear extensions we have

$$\sup_P P(-c) \leq U(-c) = \inf_{f\in Z}\{\nu(f) : f \geq -c\}$$

where supremum is taken over all possible linear extensions which extends $\nu$ and bounded by $U$, and

$$\begin{aligned} P_c(c) &= -P_c(-c) = -U(-c) \\ &= -\inf_{f\in Z}\{\nu(f) : f \geq -c\} = \sup_{f\in Z}\{\nu(f) : f \leq c\} \end{aligned}$$

which implies

$$\begin{aligned} -\sup_P P(-c) &\geq -\inf_{f\in Z}\{\nu(f) : f \geq -c\} \\ \implies \inf_P P(c) &\geq \sup_{f\in Z}\{\nu(f) : f \leq c\} = P_c(c) \end{aligned}$$

Since $P_c$ extends $\nu$ and dominated by $U$, we have

$$\inf_P P(c) = P_c(c) = \sup_{f\in Z}\{\nu(f) : f \leq c\}$$

As $P$ a positive linear functional on $C(X)$, by Riesz-Markov, there is a unique

Radon measure $\hat{\mu}$ on $\mathcal{B}(X)$ s.t.

$$P(c) = \int_X c d\hat{\mu} \text{ for all } c \in C(X)$$

for any $f \in Z$, since $P$ is an extension of $\nu$,

$$P(f) = \nu(f) = \int_X f d\hat{\mu} \implies \hat{\mu}|_Z = \nu$$

as $1 \in Z$, $P(1) = \nu(1) = 1$, so $\hat{\mu}$ is a probability measure. Therefore,

$$P \simeq \hat{\mu} \in \Pi(\nu)$$

Thus $P$ is a transport plan with marginal $\nu$. The desired duality problem is satisfied.

□

The next statement is a general version of minmax theorem, it is necessary to prove Theorem 8.1 and its proof can be found in [65]:

**Theorem 8.3** ([65])**.** *Let $K$ be a compact convex subset of a Hausdorff topological vector space, $Y$ be a convex subset of an arbitrary vector space, and $h$ be a real-valued function ($\leq +\infty$) on $K \times Y$, which is lower semi-continuous in $x$ for each fixed $y$, convex on $K$, and concave on $Y$. Then*

$$\min_{x \in K} \sup_{y \in Y} h(x, y) = \sup_{y \in Y} \min_{x \in K} h(x, y)$$

Now we are ready to prove Theorem 8.1:

*Proof of Theorem 8.1.* Firstly,

$$\inf_{\mu \in \Pi_W(\nu)} \int c d\mu \geq \sup_{\substack{f + w \leq c \\ w \in W, f \in Z}} \nu(f)$$

is proved below:

$$\inf_{\mu \in \Pi_W(\nu)} \int c d\mu \geq \inf_{\mu \in \Pi_W(\nu)} \sup_{f+w \leq c} \int (f+w) d\mu$$
$$= \inf_{\mu \in \Pi_W(\nu)} \sup_{f+w \leq c} \nu(f) = \sup_{f+w \leq c} \nu(f)$$

And then prove the opposite direction:

$$\sup_{\substack{f+w \leq c \\ w \in W, f \in Z}} \nu(f) = \sup_{w \in W} \sup_{\substack{f \leq c-w \\ f \in Z}} \nu(f) = \sup_{w \in W} \inf_{\mu \in \Pi(\nu)} \int (c-w) d\mu$$

the last equation is by Theorem 8.2. The next step is to apply Theorem 8.3. Let $K = \Pi(\nu)$, $Y = W$ and $h(\mu, w) = \int (c-w) d\mu$ be a real valued function on $\Pi(\nu) \times W$ satisfies:

- $h$ is l.s.c. in $\mu \in \Pi(\nu)$ for each fixed $w \in W$: for $\{\mu_k\}$ a sequence of measures in $\Pi(\nu)$ and $\mu_k$ converges to $\mu$ in weak* topology,

$$\lim_{k \to \infty} h(\mu_k, w) = \int (c-w) d\mu_k = \int (c-w) d\mu = h(\mu, w)$$

- $h$ is convex on $\Pi(\nu)$: for $\mu_1, \mu_2 \in \Pi(\nu)$, $\alpha\mu_1 + (1-\alpha)\mu_2 \in \Pi(\nu)$ and

$$h(\alpha\mu_1 + (1-\alpha)\mu_2, w) = \int (c-w) d(\alpha\mu_1 + (1-\alpha)\mu_2)$$
$$= \alpha \int (c-w) d\mu_1 + (1-\alpha) \int (c-w) d\mu_2$$
$$= \alpha h(\mu_1, w) + (1-\alpha) h(\mu_2, w)$$

- $h$ is concave on $W$: for $w_1, w_2 \in W$, any $\beta \in (0,1)$

$$\begin{aligned} h(\mu, \beta w_1 + (1-\beta)w_2) &= \int (c - (\beta w_1 + (1-\beta)w_2))d\mu \\ &= \beta \int (c - w_1)d\mu + (1-\beta) \int (c - w_2)d\mu \\ &= \beta h(\mu, w_1) + (1-\beta)h(\mu, w_2) \end{aligned}$$

Therefore, all assumptions in Theorem 8.3 are satisfied, we have

$$\sup_{w\in W} \inf_{\mu\in\Pi(\nu)} \int (c-w)d\mu = \inf_{\mu\in\Pi(\nu)} \sup_{w\in W} \int (c-w)d\mu$$

If $\mu \notin \Pi_W(\nu)$, there exists $w_1 \in W$ s.t. $\int w_1 d\mu < 0$. We can choose $w = \alpha w_1$, when $\alpha \to \infty$, $\sup_{w\in W} \int (c-w)d\mu \to \infty$, thus

$$\inf_{\mu\in\Pi(\nu)} \sup_{w\in W} \int (c-w)d\mu = \inf_{\mu\in\Pi_W(\nu)} \int c d\mu$$

□

**Example 8.1** (Relative equilibrium case)**.** *If $\pi : X \to Y$ is a factor map, we have the following duality result:*

$$\inf_{\mu\in\Pi_W(\nu)} \int c d\mu = \sup_{\substack{f\circ\pi + w \leq c \\ w\in W}} \int f d\nu$$

*where*

$$\Pi_W(\nu) = \{\mu \in \mathcal{P}(X) : \int w d\mu = 0 \ \forall w \in W, \mu \circ \pi^{-1} = \nu\}$$

This example is derived directly from Theorem 8.1 by setting the restriction set $Z = \{f \circ \pi : f \in C(Y)\}$, $\mu(f \circ \pi) = \nu(f)$. It is equivalent to say, $\mu|_Z = \nu$.

### 8.1.1 Geometry of Optimal Transport Plans

This subsection generalizes Zeav[66].

**Definition 8.1.** *For two measures $\mu_1, \mu_2$ on $X$ define the equivalence relation $\sim_W$: $\alpha \sim_W \beta$ iff*

*(i)* $\mu_1|_Z = \mu_2|_Z$.

*(ii)* $\int w d\mu_1 = \int w d\mu_2 \ \forall w \in W$

We denote by $[\mu]_W$ the equivalence class of $\mu$ w.r.t $\sim_W$. Let $S_m$ be a set of $m$ points in the space $X$, $\beta_s$ be a measure with the support $S_m$.

The following definitions and results of $(c, W)$-monotonicity are from Zaev[66]:

**Definition 8.2** ($(c, W)$-monotonicity[66])**.** *For a Borel measurable cost function $c : X \to \mathbb{R}$ and a linear subspace $W \subset C(X)$ a set $\Gamma \subset X$ is called $(c, W)$-monotone iff for any $m \in \mathbb{N}$, any $S_m \subset \Gamma$ any measure $\beta_s$, such that* $\operatorname{supp}(\beta_s) = S_m$, *and any measure $\alpha \sim_W \beta_s$:*

$$\int c d\beta_s \leq \int c d\alpha$$

**Proposition 8.4** ([66])**.** *If $W = \{0\}$, then the notion of $(c, W)$-monotonicity is equivalent to the notion of usual c-monotonicity.*

**Definition 8.3.** *A transport plan $\mu \in \Pi_W(\nu)$ is called $(c, W)$-monotone iff there is a $(c, W)$-monotone set $\Gamma$ of full $\mu$-measure: $\mu(\Gamma) = 1$.*

Now we are ready to state and prove the geometry properties of optimal plans in problem (I):

**Theorem 8.5.** *Let $X$ be Polish spaces, $\mu \in \mathcal{P}(X)$, $c \in C(X)$ is a cost function, $W \subset C(X)$ is a vector subspace, $Z \subset C(X)$ is restriction set, $\mu_* \in \Pi_W(\nu)$ is the minimizer of the primal Kantorovich problem with additional linear constraints:*

$$\inf_{\mu \in \Pi_W(\nu)} \int_X c d\mu$$

*then $\mu_*$ is a $(c, W)$-monotone transport plan.*

*Proof.* By Theorem 8.1

$$\int_X c d\mu_* = \sup_{\substack{f+w\leq c\\ f\in Z, w\in W}} \nu(f)$$

Let $(f_k, w_k)$ be a maximizing sequence in the dual problem and let $c_k = c - f_k - w_k$. Since

$$\int_X c_k d\mu_* = \int_X c d\mu_* - \nu(f_k) \to 0$$

and $c_k \geq 0$ we can find a subsequence $c_{k_j}$ and a Borel set $\Gamma$ for which $\mu_*(\Gamma) = 1$, such that $c_{k_j} \to 0$ on $\Gamma$. If $S = \{x_i\}_{i=1}^m \subset \Gamma$, $\mu_s$ is a measure with support $S$ and $\gamma \in [\mu_s]_W$ we get

$$\int c d\gamma \geq \int f_k d\gamma + \int w d\gamma$$

since $\gamma|_Z = \mu_s|_Z$, and $\int w d\mu_s = \int w d\gamma$, we have

$$\int c d\gamma \geq \int f_k d\mu_s + \int w d\mu_s = \int (c - c_k) d\mu_s$$

for any $k$. Letting $k \to \infty$ the $(c, W)$-monotonicity of $\Gamma$ follows. $\square$

However, if $\mu$ is $(c, W)$-monotone, we cannot have that $\mu$ is an optimal transport plan.

## 8.2 For problem (II) - Fenchel duality

Let $(X, T)$ as above, $Z \subset \mathcal{C}(X)$ closed. $\phi \in \mathcal{C}(X)$. $\mathcal{M}_Z(X, T) = \{\mu \in \mathcal{M}(X, T) : \int h d\mu = 0 \ \forall h \in Z\}$. Define

$$P_Z(\phi, T) = \sup_{\mu \in \mathcal{M}_Z(X,T)} \left( \int \phi d\mu + h(\mu) \right)$$

and for any $\mu_0 \in \mathcal{M}_Z(X, T)$, define

$$h_Z(\mu_0) = \inf_{\phi \in \mathcal{C}(X)} \left( P_Z(\phi, T) - \int \phi d\mu_0 \right)$$

**Theorem 8.6.** *Let $T : X \to X$ be a continuous map of a compact metrizable space with the topological entropy $h_{\text{top}}(T) < \infty$ and $Z \subset \mathcal{C}(X)$ closed with $\mathcal{M}_Z(X,T)$ nonempty. Then for $\mu_0 \in \mathcal{M}_Z(X,T)$, if the entropy map of $T$ is upper semi-continuous at $\mu_0$, we have*

$$h_Z(\mu_0) := \inf_{\phi \in \mathcal{C}(X)} \left( P_Z(\phi, T) - \int \phi d\mu_0 \right) = h(\mu_0)$$

*that is, $h_Z(\mu_0)$ coincides with the measure-theoretic entropy $h(\mu_0)$.*

We need the following fact the prove theorem 8.6:

**Lemma 8.7** ([67])**.** *If $K_1$ and $K_2$ are disjoint closed convex subsets of a locally convex linear topological space $V$ and if $K_1$ is compact there exists a continuous real-valued linear functional $F$ on $V$ such that*

$$F(x) < F(y) \quad \forall x \in K_1, y \in K_2$$

*Proof of Theorem 8.6.* As the entropy map is upper-semi continuous, by the definition of $P_Z(\phi, T)$ we have

$$\begin{aligned} h(\mu_0) &\leq P_Z(\phi, T) - \int \phi d\mu_0, \ \forall \phi \in \mathcal{C}(X) \\ \Longrightarrow h(\mu_0) &\leq \inf_{\phi \in \mathcal{C}(X)} \left( P_Z(\phi, T) - \int \phi d\mu_0 \right) \end{aligned}$$

So we just need to show $h(\mu_0) \geq \inf_{\phi \in \mathcal{C}(X)} \big( P_Z(\phi, T) - \int \phi d\mu_0 \big)$. Let $b > h(\mu_0)$ and let $C = \{(\mu, t) \in \mathcal{M}_Z(X,T) \times \mathbb{R} : 0 \leq t \leq h(\mu)\}$. $C$ is a convex set because the entropy map is affine. $C$ is a subset of $\mathcal{C}^*(X) \times \mathbb{R}$, where $\mathcal{C}^*(X)$ is equipped with weak* topology, then $(\mu_0, b) \notin \overline{C}$ by the upper-semi continuity of the entropy map at $\mu_0$. In lemma 8.7, let $V = \mathcal{C}^*(X) \times \mathbb{R}$, $K_1 = \overline{C}$, $K_2 = (\mu_0, b)$, then there is a continuous linear functional $F : \mathcal{C}^*(X) \times \mathbb{R} \to \mathbb{R}$ such that $F(\mu, t) < F(\mu_0, b)$ for all $(\mu, t) \in \overline{C}$. Since we are using the weak* topology on $\mathcal{C}^*(X)$ we know that $F$ has the form $F(\mu, t) = \int \varphi d\mu + \alpha t$ for some $\varphi \in \mathcal{C}(X)$ and some $\alpha \in \mathbb{R}$. Therefore

$$\begin{aligned} &\int \varphi d\mu + \alpha t < \int \varphi d\mu_0 + \alpha \cdot b \ \forall (\mu, t) \in \overline{C} \\ \implies &\int \varphi d\mu + \alpha h(\mu) < \int \varphi d\mu_0 + \alpha \cdot b \ \forall \mu \in \mathcal{M}_Z(X, T) \end{aligned}$$

If we set $\mu = \mu_0$, then $\alpha h(\mu_0) < \alpha \cdot b$, which implies $\alpha > 0$. So

$$h(\mu) + \int \frac{\varphi}{\alpha} d\mu < \int \frac{\varphi}{\alpha} d\mu_0 + b, \ \forall \mu \in \mathcal{M}_Z(X, T)$$

that is

$$P_Z(\varphi/\alpha, T) \leq \int \frac{\varphi}{\alpha} d\mu_0 + b$$

and which implies

$$b \geq P_Z(\varphi/\alpha, T) - \int \frac{\varphi}{\alpha} d\mu_0 \geq \inf_{\phi \in \mathcal{C}(X)} \left\{ P_Z(\phi, T) - \int \phi d\mu_0 \right\}$$

As any $b > h(\mu_0)$ satisfies the above inequality, we have

$$h(\mu_0) \geq \inf_{\phi \in \mathcal{C}(X)} \left\{ P_Z(\phi, T) - \int \phi d\mu_0 \right\}$$

$\square$

# CHAPTER 9: DISCUSSION

Throughout this paper, we explored constrained linear optimization problems on different topics. However, there are some unsolved problems during the research, which will be listed below as open problems.

The first two problems are in the framework of problem (II):

**Problem 9.1.** *Suppose that $X$ and $Y$ are irreducible shift of finite types with $T : X \to X$ and $S : Y \to Y$ define on them respectively. Let $\varphi, \psi$ be Hölder continuous functions on $X, Y$ respectively. Let $\mu \in \mathcal{M}(X)$ be the unique equilibrium states for $\varphi$ and $\nu \in \mathcal{M}(Y)$ be the unique equilibrium states for $\psi$, and let*

$$Z = \left\{ f \circ \pi_X - \int f d\mu + g \circ \pi_Y - \int g d\nu : f \in \mathcal{C}(X), g \in \mathcal{C}(Y) \right\}$$

*where $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ are projection maps and also factor maps. For $\phi \in \mathcal{C}(X \times Y)$ Hölder continuous, does the optimization problem*

$$\sup_{\lambda \in \mathcal{M}_Z(X \times Y, T \times S)} \int \phi d\lambda + h(\lambda)$$

*has unique optimal plan?*

The settings of Problem 8.1 are illuminated by Yoo[49], since under similar settings with $Z = \{f \circ \pi_X - \int f d\mu : f \in \mathcal{C}(X)\}$, problem (II) is a relative equilibrium state problem and by Theorem 5.16 (cited from[49]) there is a unique relative equilibrium state. So, it is natural to think about the uniqueness problem with two constraints in the same framework.

**Remark 9.1.** *In fact, even when $X, Y$ are 2-full shift, $\mu$ and $\nu$ are 1-step Markov* *measure determined by* $2\times 2$ *transition matrices (for example, the transition matrix for* $\mu$ *and* $\nu$ *are* $\begin{bmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{bmatrix}$ *and* $\begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix}$ *respectively with* $\alpha \in (0,1)$*) and thus the potentials* $\varphi$ *and* $\psi$ *are locally constant, the uniqueness result is hard to get.*

**Problem 9.2.** *If in the settings of Problem 9.1 there is more than one optimal plan, can you give an example or disproof the uniqueness guess?*

The next open problem is related to the 'generic' uniqueness property in the framework of problem (I). By Contreras[68], if $T: X \to X$ is expanding, there is a residual subset of the set of Lipschitz functions $\mathrm{Lip}(X)$ such that the maximizing measures are unique and supported on a single periodic orbit. For difference choices of $Z$, there may be no measure supported on a single periodic orbit in $\mathcal{M}_Z(X,T)$. Our problem is, is there a residual subset of $\mathrm{Lip}(X)$ such that the maximizing measure achieves the minimum entropy, if the maximizing measure is not supported on a single periodic orbit?

**Problem 9.3.** *Define the set*

$$\mathcal{U} = \left\{ f \in \mathrm{Lip}(X) : \mathcal{M}(X,T;f) = \{\mu^*\} \text{ and } h(\mu^*) = \inf_{\mu \in \mathcal{M}_Z(X,T)} h(\mu) \right\}$$

*Is* $\mathcal{U}$ *residual in* $\mathrm{Lip}(X)$*?*

**Remark 9.2.** *A special case of Problem 9.3 is the relative case. Suppose* $(Y,S)$ *is another dynamical system,* $\nu \in \mathcal{M}(Y)$ *is fixed and* $\pi: X \to Y$ *is the factor map. Let*

$$Z = \left\{ f \circ \pi_X - \int f d\mu : f \in \mathcal{C}(X) \right\}$$

*by the property of relative entropy, for* $\mu \in \mathcal{M}_Z(X,T) = \{\mu \in \mathcal{M}(X,T) : \mu \circ \pi^{-1} = \nu\}$*,* $h(\mu) \geq h(\nu)$*. So Problem 9.3 reduces to the following problem:*

*Given the set*

$$\mathcal{U} = \{f \in \mathrm{Lip}(X) : \mathcal{M}(X, T; f) = \{\mu\} \text{ and } h(\mu) = h(\nu)\}$$

*is* $\mathcal{U}$ *residual in* $\mathrm{Lip}(X)$*?*

# REFERENCES


[1] R. B. Israel, "Existence of phase transitions for long-range interactions," *Communications in Mathematical Physics*, vol. 43, pp. 59–68, 1975.

[2] G. Peyré, M. Cuturi, *et al.*, "Computational optimal transport," *Center for Research in Economics and Statistics Working Papers*, no. 2017-86, 2017.

[3] G. Monge, "Mémoire sur la théorie des déblais et des remblais," *Mem. Math. Phys. Acad. Royale Sci.*, pp. 666–704, 1781.

[4] L. V. Kantorovich, "On the translocation of masses," in *Dokl. Akad. Nauk. USSR (NS)*, vol. 37, pp. 199–201, 1942.

[5] P. Walters, *An introduction to ergodic theory*, vol. 79. Springer Science & Business Media, 2000.

[6] H. Furstenberg, "Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation," *Mathematical systems theory*, vol. 1, no. 1, pp. 1–49, 1967.

[7] T. De La Rue, "An introduction to joinings in ergodic theory," *arXiv preprint math/0507429*, 2005.

[8] E. Glasner, *Ergodic theory via joinings.* No. 101, American Mathematical Soc., 2003.

[9] D. Ruelle, "Statistical mechanics on a compact set with $\mathbb{Z}^\nu$ action satisfying expansiveness and specification," *Transactions of the American Mathematical Society*, vol. 185, pp. 237–251, 1973.

[10] P. Walters, "A variational principle for the pressure of continuous transformations," *American Journal of Mathematics*, vol. 97, no. 4, pp. 937–971, 1975.

[11] D. A. Lind, D. Lind, and B. Marcus, *An introduction to symbolic dynamics and coding.* Cambridge university press, 2021.

[12] R. Bowen, "Equilibrium states and the ergodic theory of anosov diffeomorphisms," *Lecture notes in mathematics*, vol. 470, pp. 1–17, 1975.

[13] P. Walters, "Relative pressure, relative equilibrium states, compensation functions and many-to-one codes between subshifts," *Transactions of the American Mathematical Society*, vol. 296, no. 1, pp. 1–31, 1986.

[14] O. Jenkinson, "Every ergodic measure is uniquely maximizing," *Discrete and Continuous Dynamical Systems*, vol. 16, no. 2, p. 383, 2006.

[15] D. Bertsimas and J. N. Tsitsiklis, *Introduction to linear optimization*, vol. 6. Athena Scientific Belmont, MA, 1997.

[16] C. Villani *et al.*, *Optimal transport: old and new*, vol. 338. Springer, 2009.

[17] L. V. Kantorovich and S. Rubinshtein, "On a space of totally additive functions," *Vestnik of the St. Petersburg University: Mathematics*, vol. 13, no. 7, pp. 52–59, 1958.

[18] C. Villani, *Topics in optimal transportation*, vol. 58. American Mathematical Soc., 2021.

[19] G. Yuan and B. R. Hunt, "Optimal orbits of hyperbolic systems," *Nonlinearity*, vol. 12, no. 4, p. 1207, 1999.

[20] T. Bousch, "La condition de walters," in *Annales scientifiques de l'Ecole normale supérieure*, vol. 34, pp. 287–311, Elsevier, 2001.

[21] G. Contreras, A. O. Lopes, and P. Thieullen, "Lyapunov minimizing measures for expanding maps of the circle," *Ergodic Theory and Dynamical Systems*, vol. 21, no. 5, pp. 1379–1409, 2001.

[22] O. Jenkinson, "Ergodic optimization," *Discrete and Continuous Dynamical Systems*, vol. 15, no. 1, p. 197, 2006.

[23] S. E. Newhouse, J. Palis, and F. Takens, "Bifurcations and stability of families of diffeomorphisms," *Publications Mathématiques de l'IHÉS*, vol. 57, pp. 5–71, 1983.

[24] W. Geller and M. Misiurewicz, "Rotation and entropy," *Transactions of the American Mathematical Society*, vol. 351, no. 7, pp. 2927–2948, 1999.

[25] K. Ziemian, "Rotation sets for subshifts of finite type," *Fundamenta Mathematicae*, vol. 146, no. 2, pp. 189–201, 1995.

[26] S. Tuncel, "Conditional pressure and coding," *Israel Journal of Mathematics*, vol. 39, pp. 101–112, 1981.

[27] P. Dobruschin, "The description of a random field by means of conditional probabilities and conditions of its regularity," *Theory of Probability & Its Applications*, vol. 13, no. 2, pp. 197–224, 1968.

[28] R. L. Dobrushin, "Gibbsian random fields for lattice systems with pairwise interactions," *Functional Analysis and its applications*, vol. 2, no. 4, pp. 292–301, 1968.

[29] R. L. Dobrushin, "The problem of uniqueness of a gibbsian random field and the problem of phase transitions," *Functional analysis and its applications*, vol. 2, no. 4, pp. 302–312, 1968.

[30] D. Ruelle, "Statistical mechanics of a one-dimensional lattice gas," *Communications in Mathematical Physics*, vol. 9, pp. 267–278, 1968.

[31] O. E. Lanford III and D. Ruelle, "Observables at infinity and states with short range correlations in statistical mechanics," *Communications in Mathematical Physics*, vol. 13, no. 3, pp. 194–215, 1969.

[32] D. Ruelle, "A measure associated with axiom-a attractors," *American Journal of Mathematics*, pp. 619–654, 1976.

[33] L. W. Goodwyn, "Topological entropy bounds measure-theoretic entropy," *Proceedings of the American Mathematical Society*, vol. 23, no. 3, pp. 679–688, 1969.

[34] T. N. Goodman, "Relating topological entropy and measure entropy," *Bulletin of the London Mathematical Society*, vol. 3, no. 2, pp. 176–180, 1971.

[35] T. N. Goodman, "Maximal measures for expansive homeomorphisms," *Journal of the London Mathematical Society*, vol. 2, no. 3, pp. 439–444, 1972.

[36] R. Bowen, "Some systems with unique equilibrium states," *Mathematical systems theory*, vol. 8, no. 3, pp. 193–202, 1974.

[37] V. Climenhaga and D. J. Thompson, "Equilibrium states beyond specification and the bowen property," *Journal of the London Mathematical Society*, vol. 87, no. 2, pp. 401–427, 2013.

[38] V. Climenhaga and D. J. Thompson, "Unique equilibrium states for flows and homeomorphisms with non-uniform structure," *Advances in Mathematics*, vol. 303, pp. 745–799, 2016.

[39] R. Pavlov, "On intrinsic ergodicity and weakenings of the specification property," *Advances in Mathematics*, vol. 295, pp. 250–270, 2016.

[40] O. Jenkinson, "Rotation, entropy, and equilibrium states," *Transactions of the American Mathematical Society*, vol. 353, no. 9, pp. 3713–3739, 2001.

[41] T. Kucherenko and C. Wolf, "Geometry and entropy of generalized rotation sets," *Israel Journal of Mathematics*, vol. 199, pp. 791–829, 2014.

[42] T. Kucherenko and C. Wolf, "Localized pressure and equilibrium states," *Journal of Statistical Physics*, vol. 160, no. 6, pp. 1529–1544, 2015.

[43] T. Kucherenko and C. Wolf, "Ground states and zero-temperature measures at the boundary of rotation sets," *Ergodic Theory and Dynamical Systems*, vol. 39, no. 1, pp. 201–224, 2019.

[44] R. Bowen, "Entropy for group endomorphisms and homogeneous spaces," *Transactions of the American Mathematical Society*, vol. 153, pp. 401–414, 1971.

[45] F. Ledrappier and P. Walters, "A relativised variational principle for continuous transformations," *Journal of the London Mathematical Society*, vol. 2, no. 3, pp. 568–576, 1977.

[46] K. Petersen, A. Quas, and S. Shin, "Measures of maximal relative entropy," *Ergodic Theory and Dynamical Systems*, vol. 23, no. 1, pp. 207–223, 2003.

[47] M. Allahbakhshi and A. Quas, "Class degree and relative maximal entropy," *Transactions of the American Mathematical Society*, vol. 365, no. 3, pp. 1347–1368, 2013.

[48] M. Allahbakhshi, J. Antonioli, and J. Yoo, "Relative equilibrium states and class degree," *Ergodic Theory and Dynamical Systems*, vol. 39, no. 4, pp. 865–888, 2019.

[49] J. Yoo, "Decomposition of infinite-to-one factor codes and uniqueness of relative equilibrium states," *arXiv preprint arXiv:1705.00448*, 2017.

[50] R. B. Israel, *Convexity in the theory of lattice gases*, vol. 62. Princeton University Press, 2015.

[51] M. Piraino, "Projections of gibbs states for hölder potentials," *Journal of Statistical Physics*, vol. 170, no. 5, pp. 952–961, 2018.

[52] T. Kempton, "Factors of gibbs measures for subshifts of finite type," *Bull. Lond. Math. Soc*, vol. 43, no. 4, pp. 751–764, 2011.

[53] M. H. Ellis, "The $\bar{d}$-distance between two markov processes cannot always be attained by a markov joining," *Israel Journal of Mathematics*, vol. 24, pp. 269–273, 1976.

[54] P. C. Shields, *The ergodic theory of discrete sample paths*, vol. 13. American Mathematical Soc., 1996.

[55] H. L. Royden and P. Fitzpatrick, *Real Analysis*, vol. 4th edition. Pearson, 2018.

[56] H. H. Schaefer, *Topological vector spaces*, vol. Graduate texts in mathematics 3. Springer-Verlag, 1966.

[57] C. D. Aliprantis and O. Burkinshaw, *Positive operators*, vol. 119. Springer Science & Business Media, 2006.

[58] C. D. Aliprantis and O. Burkinshaw, *Locally solid Riesz spaces with applications to economics*. No. 105, American Mathematical Soc., 2003.

[59] W. Rudin, *Real and Complex Analysis*. McGraw-Hill, 3rd ed., 1987.

[60] E. M. Alfsen, "On the geometry of choquet simplexes," *Mathematica Scandinavica*, vol. 15, no. 1, pp. 97–110, 1964.

[61] E. G. Effros, "Structure in simplexes. ii," *Journal of Functional Analysis*, vol. 1, no. 4, pp. 379–391, 1967.

[62] E. M. Alfsen, *Compact convex sets and boundary integrals*, vol. 57. Springer Science & Business Media, 2012.

[63] E. Davies, "A generalized theory of convexity," *Proceedings of the London Mathematical Society*, vol. 3, no. 4, pp. 644–652, 1967.

[64] R. Israel and R. Phelps, "Some convexity questions arising in statistical mechanics," *Mathematica scandinavica*, pp. 133–156, 1984.

[65] D. R. Adams and L. I. Hedberg, *Function spaces and potential theory*, vol. 314. Springer Science & Business Media, 2012.

[66] D. A. Zaev, "On the monge–kantorovich problem with additional linear constraints," *Mathematical Notes*, vol. 98, pp. 725–741, 2015.

[67] N. Dunford and J. T. Schwartz, *Linear operators, part 1: general theory*, vol. 10. John Wiley & Sons, 1988.

[68] G. Contreras, "Ground states are generically a periodic orbit," *Inventiones mathematicae*, vol. 205, no. 2, pp. 383–412, 2016.