UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Programa de Pós-Graduação em Matemática

Thermodynamic formalism for general IFS and quantum channels

Jader Eckert Brasil

Agradecimentos

Gostaria de agradecer ao professor Rafael Rigão Souza e ao professor Artur Oscar Lopes pelos valiosos ensinamentos, conselhos e compreensão durante todos esses anos.

Aos meus colegas, Josué Knorst, Guilherme Feltes, Hugo Ibanez, Marcus Vinícius da Silva, William Braucks, entre outros que sempre estiveram dispostos a discutir uma boa questão e me apresentar ideias extraordinárias.

A minha mãe, Helga Eckert, e ao meu irmão, Marx Eckert Brasil, que sempre estiveram ao meu lado me dando suporte.

Ao meu falecido pai, Claudiomar da Silva Brasil, que me ensinou a importancia do estudo e do conhecimento, nunca esquecerei de vários de seus ensinamentos entre eles não esquecerei que: "A caneta é mais leve que a pá".

Agradeço também a minha amada Júlia Carolyne, que esteve ao meu lado nos momentos mais difíceis da minha vida e também nos mais felizes.

Agradeço aos membros banca, pelas sugestões e pelo tempo que dedicaram para conhecer meu trabalho e participar do fim dessa etapa.

Obrigado à CAPES pelo suporte financeiro durante todo o trabalho.

Por fim gostaria de dizer que nada seria possível sem todos vocês. Sozinho eu nunca chegaria tão longe. Muito obrigado a todos!

Tese submetida por Jader Eckert Brasil[1], como requisito parcial para a obtenção do grau de Doutor em Ciência Matemática, pelo Programa de Pós-Graduação em Matemática, do Instituto de Matemática e Estatística da Universidade Federal do Rio Grande do Sul.

Professor Orientador: Prof. Dr. Rafael Rigão Souza
Professor Coorientador: Prof. Dr. Artur Oscar Lopes

Banca Examinadora:
Prof. Dr. Rafael Rigão Souza (PPGMAT/UFRGS, Orientador)
Prof. Dr. Artur Oscar Lopes (PPGMAT/UFRGS, Coorientador)
Prof. Dr. Elismar da Rosa Oliveira (PPGMAT/UFRGS)
Prof. Dr. Leandro Cioletti (UnB)
Prof. Dr. Leonardo Fernandes Guidi (UFRGS)

[1]Bolsista CAPES - Coordenação de Aperfeiçoamento Pessoal

Contents

Chapter 1

Introduction

This work is composed of two independent parts with a common motivation. Inspired by the general theory of Thermodynamic Formalism described in [LMMS15], we develop concepts like pressure and entropy for general IFS and quantum channels.

In Chapter 2, we talk about Iterated Function Systems with measures (IFSm). In this setting, we have a compact metric space $\mathbf{X}$, a family $(\tau_\theta)_{\theta\in\Theta}$ of maps on $\mathbf{X}$, and finite positive probabilities $(\mathrm{q}_x))_{x\in\mathbf{X}}$ on Θ. Starting on $x_0 \in \mathbf{X}$, in the next step we will be on $\tau_\theta(x_0)$ with probability $d\mathrm{q}_x(\theta)$.

While in classical Thermodynamic Formalism we study invariant measures, when we analyze the ergodic properties of IFS, it is natural to consider the holonomic measures to play this role. This concept was adapted from the corresponding one, initially introduced in Aubry-Mather Theory, and first appeared in thermodynamic formalism papers in [GL08] and [LO09].

We show that the Thermodynamic Formalism for IFSm, in some sense, is a generalization of the Thermodynamic Formalism for a certain class of dynamical systems (see section 2.8). This corresponds to consider the inverse branches of the dynamical system to define an IFSm.

For such perspective, we can verify that, with some regularity in the potential, the pressure is the same for the IFS and the dynamical system that inspired the IFS. Furthermore, the transfer operator for the IFS is equal to the Ruelle Operator of that dynamical system.

It is natural to consider that, under the perspective of IFS, the Ruelle Operator depends on the inverse branches (the backwards dynamics and not the forward dynamics). The material of this chapter is part of [BOS22], and a future question is whether we can obtain ergodic optimization results in this setting.

Chapter 3 is part of [BKL21b] and is related to another work described in [BKL21a]. Both works are inspired by the Benoist paper [BFPP19], which introduces the idea to consider a quantum channel as an integral and elaborates about (ϕ-Erg) and (Pur) properties, which in [BKL21b, BKL21a] were shown generic.

At first glance, it might seem that there is no relation between the previous chapter and this setting. But if we consider the space D_k of density matrices with complex entries, a function $L : M_k \to M_k$, and a measure μ on M_k, with certain hypotheses described better in section 3.6, we have a quantum trajectory that if we start on $\rho_0 \in D_k$ state, we go to $\frac{L(v)\rho_0 L(v)^\dagger}{\operatorname{tr}(L(v)\rho_0 L(v)^\dagger}$ state with probability $\operatorname{tr}(L(v)\rho_0 L(v)^\dagger)\,\mathrm{d}\mu(v)$.

Looking at this quantum trajectory, we can see similarities with the IFSm setting. By writing this quantum trajectory with the notation of IFSm, we will see a relation between the linear operator $\phi(\rho) = \int_{M_k} L(v)\rho L(v)^\dagger\,\mathrm{d}\mu(v)$ and the Ruelle Operator.

In this thesis, we observe that irreducible quantum channels have similar properties to the dual of the Ruelle Operator. First, it maps densities in densities, and, as it is known, the dual of the Ruelle Operator maps probabilities in probabilities. Furthermore, the Theorem 3.3.5 is similar to a Ruelle-Perron-Frobenius in this setting, and we can define, for a class of potentials, a notion of pressure (Definition 3.4.3) of this same potential that we prove to be equal to the log of the spectral radius of the quantum channel defined from the potential in Theorem 3.4.8.

In section 3.8 we present some interesting examples, one of them showing that our entropy, somehow, generalizes the concept of entropy for a classical Markov chain. This shows that our definition is quite natural.

In [BKL21a] we consider Lyapunov exponents associated with quantum channels, and show a quantum version of Pesin Theorem relating the Lyapunov exponents with entropy. This reaffirms the claim that our concept of entropy is natural in this quantum setting.

The Benoist paper [BFPP19] describes quantum channels under a dynamical and ergodic perspective, but does not present the concept of entropy.

The entropy we defined for this quantum setting is of dynamical nature, which differs from the Von Neumann entropy which is not. The idea to define this entropy comes from the paper [BLLC10], where the authors were inspired by certain results in [Sło03].

Chapter 2

Thermodynamic formalism for general IFS

Abstract

This chapter is part of [BOS22] and introduces a theory of Thermodynamic Formalism for Iterated Function Systems with Measures (IFSm). We study the spectral properties of the Transfer and Markov operators associated to a IFSm. We introduce variational formulations for the topological entropy of holonomic measures and the topological pressure of IFSm given by a potential. A definition of equilibrium state is then natural and we prove its existence for any continuous potential. We show, in this setting, a uniqueness result for the equilibrium state requiring only the Gâteaux differentiability of the pressure functional.

2.1 Introduction

The modern study of Iterated Function Systems (IFS for short) come back to the early 80's with the works of J. Hutchinson [Hut81] and M. Barnsley [BD85] where the theory was unified both in the geometric and the analytical point of view, generating what we call today the Hutchinson-Barnsley theory for IFS, meaning that each IFS, which is a family of maps acting from a set to itself, having good contraction hypothesis has an invariant compact set called the fractal attractor and, if we add weights having good continuity hypothesis to each function, the IFS acts on probabilities having an invariant probability whose support is the fractal attractor set. Although, several works on geometric features of fractals were done in the previous decades by Mandelbrot and others, but after the 80's the IFS assumed the central role in the generation and study of fractals and its applications.

For a typical dynamical system $T : X \to X$, an initial point $x_0 \in X$ is iterated by T producing the orbit $\{x_0, T(x_0), T^2(x_0), ...\}$, whose limit or the cluster points are the objects of main interest, from a dynamical point of view. On the other hand, for an IFS $(X, \tau_\theta)_{\theta\in\Theta}$, we iterate the initial point by choosing at each step a possibly different map $\tau_\theta : X \to X$, indexed by the generally finite set Θ, producing multiple orbits $\{Z_j, j \geqslant 0\} = \{Z_0 = x_0, Z_1 = \tau_{\theta_0}(x_0), Z_2 = \tau_{\theta_1}(\tau_{\theta_0}(x_0)), ...\}$. We notice that the orbit is now a set of orbits controlled by the sequence $\{\theta_0, \theta_1, ...\} \in \Theta^{\mathbb{N}}$. To avoid this complication Hutchinson defined the fractal operator $F : K(X) \to K(X)$ by

$$F(B) = \bigcup_{\theta\in\Theta} \tau_\theta(B)$$

for $B \in K(X)$, the family of nonempty compact sets of X. This operator is called the Hutchinson-Barnsley operator and a compact set is invariant or fractal if $F(\Omega) = \Omega$. Additionally Ω is a fractal attractor if the orbit of B by F, given by $\{B, F(B), F^2(B), ...\}$ converge, w.r.t. the Hausdorff-Pompeiu metric to Ω, for any $B \in K(X)$ (see [BP13] for details on the Hausdorff-Pompeiu metric).

Other possible point of view to understand the dynamics of an IFS is the probabilistic one. In this case we consider that, in each step the function to be iterated is chosen according to some probability, thus we are actually studying a stochastic process $X_0, X_1, X_2, ... \in X$ where each X_{j+1} is a random variable whose distribution is obtained from the previous X_j by a transition kernel using the IFS law. In other words, given an initial distribution μ_0 we iterate it by the Markov operator $M : \mathcal{P}(X) \to \mathcal{P}(X)$, defined as the transfer operator's dual, obtaining the distributions $\mu_0, \mu_1 = M(\mu_0), \mu_2 = M^2(\mu_0), ... \in \mathcal{P}(X)$. Analogously to the fractal attractor, we say that $\mu \in \mathcal{P}(X)$ is an invariant measure if $M(\mu) = \mu$ and that $\mu \in \mathcal{P}(X)$ is an attracting invariant measure (or Hutchinson-Barnsley measure) if $M^j(\mu_0)$ converge to μ w.r.t. the Monge-Kantorovich metric(see [Hut81]), for any $\mu_0 \in \mathcal{P}(X)$. It is possible to prove that the support of the invariant attracting measure is the fractal attractor (see [Hut81]). Since the set of measures over X is the dual of $C(X)$ the Markov operator is often defined by duality w.r.t. the transference operator $L : C(X) \to C(X)$.

To illustrate that we consider the classical case of IFS with constant probabilities studied by Hutchinson, Barnsley and many others in the beginnings of the 80's. We consider $\Theta = \{1, 2, ..., n\}$, meaning that we have a finite number of maps, and each one is chosen according to a probability $p_j > 0$ where $p_1 + \cdots + p_n = 1$, constituting an IFS with probabilities (IFSp for short). Under this conditions the classic transfer operator (also called Ruelle operator, see Ruelle [Rue67, Rue68], Walters [Wal75] and Fan [FL99]) is given by

$$L(f)(x) = \sum_{j=1}^{n} p_j\, f(\theta_j(x)),$$

for any $f \in C(X)$, and the Markov operator acting on μ, $M(\mu)$, is implicitly defined by the property

$$\int_X f dM(\mu) = \int_X L(f)(x) d\mu,$$

for any $f \in C(X)$.

The final feature of IFS dynamics we need to understand is the connection between IFS orbits and the invariant measures. The first one is the celebrated result due to M. Barnsley, the Chaos Game Theorem (CGT for short) claiming that, from the initial probabilities p_j's, we can built a probability $\mathbb{P}$ over the space $\Theta^{\mathbb{N}}$ such that $\mathbb{P}$-a.e. $(\theta_0, \theta_1, ...) \subset \Theta^{\mathbb{N}}$ the correspondent orbit $\{x_0, \tau_{\theta_0}(x_0), \tau_{\theta_1}(\tau_{\theta_0}(x_0)), ...\}$ approximate the fractal attractor Ω, for any initial point x_0. The second one is the Elton's Ergodic Theorem (EET for short) [Elt87] claiming that, from the initial probabilities p_j's, we can built a probability $\mathbb{P}$ over the space $\Theta^{\mathbb{N}}$ such that $\mathbb{P}$-a.e. $(\theta_0, \theta_1, ...) \subset \Theta^{\mathbb{N}}$ the correspondent average of visits of the orbit $\{x_0, \tau_{\theta_0}(x_0), \tau_{\theta_1}(\tau_{\theta_0}(x_0)), ...\}$ to a measurable set $B \subset X$ is equal to $\mu(B)$, if $\mu(\partial(B)) = 0$, analogously to the usual Birkhoff ergodic theorem for a single map, where μ is the invariant measure of the IFS in consideration. For continuous functions it means that

$$\frac{1}{N}\left(f(x_0) + f(\tau_{\theta_0}(x_0)) + \cdots + f(\tau_{\theta_{N-1}}(\cdots \tau_{\theta_0}(x_0)))\right) \to \int_X f d\mu,$$

for any $f \in C(X)$, as $N \to \infty$. In other words

$$\frac{1}{N}\left(\delta_{x_0} + \delta_{\tau_{\theta_0}(x_0)} + \cdots + \delta_{\tau_{\theta_{N-1}}(\cdots \tau_{\theta_0}(x_0))}\right) \to \mu,$$

as a distribution. In synthesis, the CGT and the EET are random procedures to approximate the fractal attractor and the invariant measure, respectively.

The study of the conditions under what we have, for a given IFS, a fractal attractor which is the support of an invariant measure is called the Hutchinson-Barnsley theory. Such conditions has been extremely relaxed and generalized in several ways in the past forty years. A first generalization was for IFSp where the probability $p_j > 0$ where $p_1 + \cdots + p_n = 1$, were replaced by variable probabilities $p_j(x) > 0$ where $p_1(x) + \cdots + p_n(x) = 1$ for all $x \in X$. Now, the transfer operator is defined by

$$L(f)(x) = \sum_{j=1}^{n} p_j(x)\, f(\theta_j(x)),\ \forall x \in X,$$

for any $f \in C(X)$. Very general conditions for the existence of the invariant measure for such IFS are given in [BDEG88]. We point out that the EET was also proved for variable probabilities and finite functions in [Elt87].

In Fan [FL99], 1999, the condition $p_1(x) + \cdots + p_n(x) = 1$ is finally dropped assuming only that each $p_\theta(x) \geqslant 0$ for $\theta \in \{1, ..., n\}$. In this work Fan study a contractive system which is a triplet $(X, \tau_\theta, p_\theta)_{\theta \in \{1,...,n\}}$, where each τ_θ is a contractive map and each $p_\theta(x) \geqslant 0$ for $\theta = 1, ..., n$, generalizing the notion of IFS with probabilities. In this setting, Fan proves a Ruelle-Perron-Frobenius theorem (RPF theorem, for short), meaning the existence of a positive eigenfunction for the operator L and an eigenmeasure for the operator M with the same eigenvalue which is the spectral radius of L.

The next key improvement was given by Stenflo [Ste02], where random iterations are used to represent the iterations of a so called IFS with probabilities, $(X, \tau_\theta, \mu)_{\theta \in \Theta}$ for an arbitrary measurable space Θ. The approach here is slightly different from the previous works on IFS with probabilities, instead considering weights, the iterations from $Z_0 \in X$ are $Z_{j+1} = \tau_{I_j}(Z_j)$ governed by a sequence of i.i.d variables $\{I_j \in \Theta\}_{j \in \mathbb{N}}$, with distribution μ, generating a Markov chain $\{Z_j,\ j \geqslant 0\}$ with transfer operator given by

$$L(f)(x) = \int_\Theta f(\tau_\theta(x)) d\mu(\theta),$$

for any $f \in C(X)$. The main goal of Stenflo [Ste02] is to establish, when L is Feller, the existence of an unique attracting invariant measure π, for this Markov chain.

In our work we will extend the variational results in [LMMS15, Lop11] and, more recently, the preprint Cioletti and Oliveira [CO17], to a general IFS called IFS with measures (IFSm), $(X, \tau_\theta, q)_{\theta \in \Theta}$ for an arbitrary compact space Θ (see Dumitru [Dum13] for the Hutchinson-Barnsley theory for such infinite systems or Lukawska [GLJ05] for infinite countable ones). The approach here consist in a generalization of Stenflo [Ste02]. We take a family $q_x(\cdot) \in \mathcal{M}(\Theta)$, indexed by $x \in X$, generating a Markov chain with transfer operator given by

$$B_q(f)(x) = \int_\Theta f(\tau_\theta(x)) dq_x(\theta),$$

for any $f \in C(X)$. The meaning of the distribution $q_x(\cdot) \in \mathcal{M}(\Theta)$ is such that, the position x of a previous iteration of the IFS determine the distribution $q_x(\cdot)$ of θ used to choose the function τ_θ and produce the new point $\tau_\theta(x)$. When $q_x(\cdot) = \mu$, for any $x \in X$, is a constant distribution we recover the setting from Stenflo [Ste02].

In our setting the IFSm $(X, \tau_\theta, q)_{\theta \in \Theta}$ can be studied as the sample paths of the Markov process $\{Z_j,\ j \geqslant 0\}$ with initial distribution $\mu_0 = \mu \in \mathcal{M}(X)$

and $\mu_{j+1} = \mathcal{L}_q(\mu_j)$, where for any $\nu \in \mathcal{M}(X)$,

$$\int_X f(x) d\mathcal{L}_q(\nu)(x) = \int_X B_q(f)(x) d\nu(x),$$

for any $f \in C(X)$. Such degree of generality is necessary to enlarge the range of application for the IFS theory, specially the thermodynamic formalism. In Section 2.9 we present a situation where we believe the tools developed in the previous sections can be applied when analysing an interesting problem in economics.

Our goal is to present a complete theory of thermodynamical formalism for these IFS with measures, that is, good definitions for transfer operators, invariant measures, entropy, pressure, equilibrium measures and a variational principle. Finally, we want to use these tools to characterize the solutions of the ergodic optimization problem.

For sake of completeness we would like to point out that we do not prove a RPF theorem for those systems, only the existence of positive eigenfunctions, but we establish all the results that can be derived if we have assumed such a property. To the best of our knowledge the RPF theorem for IFSm has not been established an it is a very hard problem. There are several works on the matter of finding IFS for which the RPF theorem holds, those IFS are said to have the RPF property. In 2009 Lopes and Oliveira [Lop11] studied those systems renaming it as weighted systems or IFS with weights, having the RPF property, producing a self contained notion of entropy and topological pressure through a variational principle for holonomic measures allowing to establish a thermodynamical formalism for IFS. Other approaches for IFS thermodynamic formalism were developed by Urbański [SSU01, MU00, HMU02] and many others.

It's worth to mention that, in Urbański et al. [HMU02] a thermodynamic formalism for conformal infinite (countable) iterated function systems is presented using the conformal structure via partition functions. Also in Käenmäki [Käe04] a thermodynamical formalism for IFS is studied with the help of cylinder functions, where general IFS means that $(X, \tau_\theta)_{\theta \in \Theta}$ and Θ is the increasing union of finite alphabets. In Lopes et al. [LMMS15] a thermodynamic formalism for shift spaces, taking values on a compact metric space is presented, although this problem is closely related to thermodynamic formalism for IFS when we associate the pre images of the shift map with a respective maps producing an infinite IFS. Also in [ACR18] a variational principle for the specific entropy on the context of symbolic dynamics of compact metric space alphabets was developed generalizing somehow the results in [Lop11].

The structure of the paper is the following: in Section 2.2, we present the basic definitions on IFS with measures (IFSm) and a fundamental result about the eigenspace associated to the maximal eigenvalue of the transfer operator. In Section 2.3 we define the Markov Operator, which in the case of a normalized IFSm gives the evolution of the distribution of the associated Markov Process, and show that the set of eigenmeasures for it is non-empty. In Section 2.4, we introduce holonomic measures, which play the role of invariant measures in the IFS setting. In Section 2.5 we define entropy for a IFSm, the topological pressure of a given potential function, as well as the concept of equilibrium states. In Section 2.6 a uniqueness result for the equilibrium states is obtained. Section 2.7 prove the existence of a positive eigenfunction for the transfer operator associated to the spectral radius and give a constructive proof of the existence of equilibrium states. In Section 2.8 we show how the classical thermodynamical formalism for a dynamical system is a particular case of the IFSm Thermodynamic Formalism. Finally, in Section 2.9 we present a possible application in economic theory of the theory developed in the previous sections.

2.2 IFS with measures

In this section we set up the basic notation and present a fundamental result about the eigenspace associated to the maximal eigenvalue (or spectral radius) of the transfer operator.

In this paper $\mathbf{X}, \Theta$ are compact metric spaces, equipped with $\mathscr{B}(\mathbf{X})$ and $\mathscr{B}(\Theta)$ respectively the Borel σ-algebra for $\mathbf{X}$ and Θ.

The Banach space of all real continuous functions equipped with supremum norm is denoted by $C(\mathbf{X}, \mathbb{R})$. Its topological dual, as usual, is identified with $\mathscr{M}_s(\mathbf{X})$, the space of all finite Borel signed measures endowed with total variation norm. We use the notation $\mathscr{M}_1(X)$ for the set of all Borel probability measures over X supplied with the weak-$*$ topology. Since we are assuming that X is compact metric space then we have that the topological space $\mathscr{M}_1(X)$ is compact and metrizable.

Take $\mathrm{q} = (\mathrm{q}_x)_{x\in\mathbf{X}}$ a collection of measures on $\mathscr{B}(\Theta)$, such that

(q1) $\mathrm{q}_x(\Theta) < \infty$ for all $x \in \mathbf{X}$,

(q2) $\inf \mathrm{q} \equiv \inf_{x\in\mathbf{X}} \mathrm{q}_x(\Theta) > 0$,

(q3) $x \mapsto \mathrm{q}_x(A)$ is a Borel map, i.e, is $\mathscr{B}(\mathbf{X})$-measurable for all fixed $A \in \mathscr{B}(\Theta)$.

(q4) $x \mapsto \mathrm{q}_x$ is weak*-continuous.

An Iterated Function System with measures q, IFSm for short, is a triple $\mathcal{R}_{\mathrm{q}} = (\mathbf{X}, \tau, \mathrm{q})$, where $\tau = (\tau_\theta)_{\theta\in\Theta}$ is a collection of functions from $\mathbf{X}$ to itself with the following

(τ1) $\tau : (\Theta, \mathbf{X}) \mapsto \mathbf{X}$, where $\tau(\theta, x) = \tau_\theta(x)$ is continuous.

The $\mathcal{R}_{\mathrm{q}}$ is said to be **normalized** if for all $x \in \mathbf{X}$, q_x is a probability measure.

Definition 2.2.1. Let $\mathcal{R}_{\mathrm{q}} = (\mathbf{X}, \tau, \mathrm{q})$ be an IFSm. The **Transfer Operator** $\mathrm{B}_{\mathrm{q}} : C(\mathbf{X}, \mathbb{R}) \circlearrowleft$ associated to $\mathcal{R}_{\mathrm{q}}$ is defined by:

$$\mathrm{B}_{\mathrm{q}}(f)(x) = \int_\Theta f(\tau_\theta(x))\, \mathrm{d}\mathrm{q}_x(\theta), \qquad \forall x \in X.$$

B_{q} is well defined. In fact, B_{q} is continuous once that

$$\|\mathrm{B}_{\mathrm{q}}(f)\|_\infty = \sup_x \left| \int f(\tau_\theta(x))\, \mathrm{d}\mathrm{q}_x(\theta) \right| \leqslant \sup \mathrm{q}\, \|f\|_\infty < \infty.$$

Futhermore, for a fixed $f \in C(\mathbf{X}, \mathbb{R})$ and $x \in \mathbf{X}$, given $\varepsilon > 0$, take $\delta > 0$ s.t.

$$\sup_{\theta\in\Theta} d(f(\tau_\theta(x)), f(\tau_\theta(y))) < \frac{\varepsilon}{2 \sup \mathrm{q}},$$

and

$$\left| \int_\Theta f(\tau_\theta(x))\mathrm{q}_x(\theta) - \int_\Theta f(\tau_\theta(x))\mathrm{q}_y(\theta) \right| < \frac{\varepsilon}{2},$$

for all $y \in \mathbf{X}$ with $d(x, y) < \delta$. Then,

$$\begin{aligned}
|\mathrm{B}_{\mathrm{q}}(f)(x) - \mathrm{B}_{\mathrm{q}}(f)(y)| &= \left| \int_\Theta f(\tau_\theta(x))\, \mathrm{d}\mathrm{q}_x(\theta) - \int_\Theta f(\tau_\theta(y))\, \mathrm{d}\mathrm{q}_y(\theta) \right| \\
&\leqslant \int_\Theta |f(\tau_\theta(x)) - f(\tau_\theta(y))|\, \mathrm{d}\mathrm{q}_x(\theta) + \left| \int_\Theta f(\tau_\theta(y))\, \mathrm{d}\mathrm{q}_x(\theta) - \int_\Theta f(\tau_\theta(y))\, \mathrm{d}\mathrm{q}_y(\theta) \right| \\
&< \frac{\varepsilon}{2 \sup \mathrm{q}} \int_\Theta \mathrm{d}\mathrm{q}_x(\theta) + \frac{\varepsilon}{2} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

This shows that, for $f \in C(\mathbf{X}, \mathbb{R})$ and $x \in \mathbf{X}$, given $\varepsilon > 0$, there is $\delta > 0$ such that for every $d(x, y) < \delta$, $|\mathrm{B}_{\mathrm{q}}(f)(x) - \mathrm{B}_{\mathrm{q}}(f)(y)| < \epsilon$, therefore $\mathrm{B}_{\mathrm{q}}(f)(x)$ is continuous.

Proposition 2.2.2. *Let $\mathcal{R}_q = (\boldsymbol{X}, \tau, q)$ be a continuous IFSm. Then for the N-th iteration of B_q we have*

$$B_q^N(1)(x) = \int_{\Theta^N} dP_x^q(\theta_0, \dots, \theta_{N-1})$$

where,

$$dP_x^q(\theta_0, \dots, \theta_{N-1}) \equiv \prod_{j=1}^{N} dq_{x_{N-j}}(\theta_{N-j}),\ x_0 = x \text{ and } x_{j+1} = \tau_{\theta_j} x_j.$$

Proof. This expression can be obtained by proceeding a formal induction on N. For $N = 2$ and $x = x_0$, we have

$$\begin{aligned}
\mathrm{B}_\mathrm{q}^2(1)(x_0) &= \int_\Theta \mathrm{B}_\mathrm{q}(1)(\tau_{\theta_0}(x_0))\, \mathrm{dq}_{x_0}(\theta_0) \\
&= \int_\Theta \int_\Theta \mathrm{dq}_{x_1}(\theta_1)\, \mathrm{dq}_{x_0}(\theta_0) \\
&= \int_{\Theta^2} \mathrm{dP}_x^\mathrm{q}(\theta_0, \theta_1).
\end{aligned}$$

And, if

$$\mathrm{B}_\mathrm{q}^N(1)(x) = \int_{\Theta^N} \mathrm{dP}_x^\mathrm{q}(\theta_0, \dots, \theta_{N-1}),$$

then

$$\begin{aligned}
\mathrm{B}_\mathrm{q}^{N+1}(1)(x) &= \int_\Theta \mathrm{B}_\mathrm{q}^N(1)(x_1)\, \mathrm{dq}(\theta_0) \\
&= \int_\Theta \int_{\Theta^N} \mathrm{dP}_{x_1}^\mathrm{q}(\theta_1, \dots, \theta_N)\, \mathrm{dq}(\theta_0) \\
&= \int_\Theta \cdots \int_\Theta \left(\prod_{j=0}^{N-1} \mathrm{dq}_{x_{N-j}}(\theta_{N-j}) \right) \mathrm{dq}(\theta_0) \\
&= \int_\Theta \cdots \int_\Theta \left(\prod_{j=1}^{N+1} \mathrm{dq}_{x_{N+1-j}}(\theta_{N+1-j}) \right) \\
&= \int_{\Theta^N} \mathrm{dP}_x^\mathrm{q}(\theta_0, \dots \theta_N).
\end{aligned}$$

■

Remark 2.2.3. The formal notation used for $\mathrm{P}_x^{\mathrm{q}}$, in fact, means that $\mathrm{P}_x^{\mathrm{q}}$ is a measure in Θ^N defined by,

$$\mathrm{P}_x^{\mathrm{q}}(\Theta_0 \times \cdots \times \Theta_{N-1}) = \int_{\Theta_0} \cdots \int_{\Theta_{N-1}} \mathrm{dq}_{x_{N-1}}(\theta_{N-1}) \cdots \mathrm{dq}_{x_0}(\theta_0).$$

In the case $N = 2$ for instance,

$$\begin{aligned} \mathrm{P}_x^{\mathrm{q}}(\Theta_0 \times \Theta_1) &= \int_{\Theta_0} \int_{\Theta_1} \mathrm{dq}_{\tau_{\theta_0} x}(\theta_1)\, \mathrm{dq}_x(\theta_0) \\ &= \int_{\Theta_0} \mathrm{q}_{\tau_{\theta_0} x}(\Theta_1)\, \mathrm{dq}_x(\theta_0). \end{aligned}$$

Note that $\mathrm{q}_{\tau_{\theta_0}(x_0)}(\Theta_1)$, with fixed Θ_1 and x_0, is a function of θ_0 that is measurable: indeed, if $A \in \mathscr{B}(\Theta)$, $f_A : \mathbf{X} \to \mathbb{R}$ defined by $f_A(x) = q_x(A)$ is measurable by (*q3*) and by (*τ1*) implies τ is measurable. Thus, $F_A := f_A \circ \tau$ is measurable.

Proposition 2.2.4. *If $f : \boldsymbol{X} \to \mathbb{R}$ is a measurable nonnegative function, then*

$$H(x) := \int_{\Theta_0} f \circ \tau(\theta, x)\, dq_x(\theta),$$

is measurable.

Using Proposition 2.2.4, it is a simple induction to prove that

$$x \mapsto \mathrm{P}_x^{\mathrm{q}}(\Theta_0 \times \cdots \times \Theta_{N-1}) = \int_{\Theta_0} \mathrm{P}^{\mathrm{q}}_{\tau_{\theta_0} x_0}(\Theta_1 \times \cdots \times \Theta_{N-1})\, \mathrm{dq}_{x_0}(\theta_0)$$

is measurable for any $\Theta_i \in \mathcal{B}(\Theta)$.

In this way we conclude that $\mathrm{P}_x^{\mathrm{q}}$ is well defined for each space Θ^N.

Theorem 2.2.5. *Let $\mathcal{R}_q = (\boldsymbol{X}, \tau, q)$ be a continuous IFSm and suppose that there are a positive number ρ and a strictly positive continuous function $h : \boldsymbol{X} \to \mathbb{R}$ such that $B_q(h) = \rho h$. Then the following limit exits*

$$\lim_{N \to \infty} \frac{1}{N} \ln \left(B_q^N(1)(x) \right) = \ln \rho(B_q) \tag{2.1}$$

the convergence is uniform in x and $\rho = \rho(B_q)$ is the spectral radius of B_q acting on $C(X, \mathbb{R})$.

Remark 2.2.6. We will adress the question of existence of positive eigenfunctions in section 2.7.

Proof. From the hypothesis we can build a normalized IFSm $\mathcal{R}_{\mathrm{p}} = (\mathbf{X}, \tau, \mathrm{p})$ where

$$\mathrm{dp}_x(\theta) = \frac{h(\tau_\theta(x))}{\rho h(x)}\, \mathrm{dq}_x(\theta).$$

Note that $\mathrm{dP}_x^{\mathrm{q}}$ and $\mathrm{dP}_x^{\mathrm{p}}$ are related in the following way

$$\begin{aligned}
\mathrm{dP}_x^{\mathrm{q}}(\theta_0, \ldots, \theta_{N-1}) &= \prod_{j=1}^{N} \mathrm{dq}_{x_{N-j}}(\theta_{N-j}) \\
&= \prod_{j=1}^{N} \frac{\rho h(x_{N-j})}{h(\tau_\theta(x_{N-j}))}\, \mathrm{dp}_{x_{N-j}}(\theta_{N-j}) \\
&= \rho^N \prod_{j=1}^{N} \frac{h(x_{N-j})}{h(x_{N-j+1})}\, \mathrm{dp}_{x_{N-j}}(\theta_{N-j}) \\
&= \rho^N \frac{h(x_0)}{h(x_N)} \prod_{j=1}^{N} \mathrm{dp}_{x_{N-j}}(\theta_{N-j}) \\
&= \rho^N \frac{h(x_0)}{h(x_N)}\, \mathrm{dP}_x^{\mathrm{p}}(\theta_0, \ldots, \theta_{n-1})
\end{aligned}$$

Since $\mathbf{X}$ is compact and h is a strictly positive continuous function, we have for some positive constant a and b the following inequalities

$$0 < a \leqslant h(x_0)/h(x_N) \leqslant b.$$

Using the Proposition 2.2.2 and the above inequalities, we obtain for any fixed $N \in \mathbb{N}$ the following expression

$$\begin{aligned}
\frac{1}{N} \ln(\mathrm{B}_{\mathrm{q}}^N(1)(x)) &= \frac{1}{N} \ln\left(\int_{\Theta^N} \mathrm{dP}_x^{\mathrm{q}}(\theta_0, \ldots, \theta_{N-1}) \right) \\
&= \frac{1}{N} \ln\left(\int_{\Theta^N} \rho^N \frac{h(x_0)}{h(x_N)}\, \mathrm{dP}_x^{\mathrm{p}}(\theta_0, \ldots, \theta_{N-1}) \right) \\
&= \ln \rho + \frac{1}{N} \ln\left(\int_{\Theta^N} \frac{h(x_0)}{h(x_N)}\, \mathrm{dP}_x^{\mathrm{p}}(\theta_0, \ldots, \theta_{N-1}) \right).
\end{aligned}$$

Futhermore,

$$\begin{aligned}\frac{1}{N}\ln\left(\int_{\Theta^N}\frac{h(x_0)}{h(x_N)}\,\mathrm{dP}^{\mathrm{p}}_x(\theta_0,\ldots,\theta_{N-1})\right) &\geqslant \frac{1}{N}\ln\left(\int_{\Theta^N} a\,\mathrm{dP}^{\mathrm{p}}_x(\theta_0,\ldots,\theta_{N-1})\right)\\ &= \frac{1}{N}\ln a + \frac{1}{N}\ln\int_{\Theta_N}\mathrm{dP}^{\mathrm{p}}_x\\ &= \frac{1}{N}\ln a \xrightarrow{N\to\infty} 0\end{aligned}$$

and

$$\begin{aligned}\frac{1}{N}\ln\left(\int_{\Theta^N}\frac{h(x_0)}{h(x_N)}\,\mathrm{dP}^{\mathrm{p}}_x(\theta_0,\ldots,\theta_{N-1})\right) &\leqslant \frac{1}{N}\ln\left(\int_{\Theta^N} b\,\mathrm{dP}^{\mathrm{p}}_x(\theta_0,\ldots,\theta_{N-1})\right)\\ &= \frac{1}{N}\ln b + \frac{1}{N}\ln\int_{\Theta_N}\mathrm{dP}^{\mathrm{p}}_x\\ &= \frac{1}{N}\ln b \xrightarrow{N\to\infty} 0.\end{aligned}$$

Therefore, for every $N \geqslant 1$ we have

$$\sup_{x\in\mathbf{X}}\left|\frac{1}{N}\ln\left(\mathrm{B}^N_{\mathrm{q}}(1)(x)\right)-\ln\rho\right| = O(1/N),$$

where $O(1/N)$ is independent of x, wich proves (2.1). From the above inequality and Gelfand's formula for the spectral radius we have

$$\begin{aligned}|\ln\rho(\mathrm{B}_{\mathrm{q}})-\ln\rho| &= \left|\ln\left(\lim_{N\to\infty}\left\|\mathrm{B}^N_{\mathrm{q}}\right\|^{\frac{1}{N}}\right)-\ln\rho\right| = \lim_{N\to\infty}\left|\frac{1}{N}\ln\left\|\mathrm{B}^N_{\mathrm{q}}\right\|-\ln\rho\right|\\ &\leqslant \limsup_{N\to\infty}\ \sup_{x\in\mathbf{X}}\left|\frac{1}{N}\ln\left(\mathrm{B}^N_{\mathrm{q}}(1)(x)\right)-\ln\rho\right|\\ &\leqslant \limsup_{N\to\infty}\frac{C}{N} = 0.\end{aligned}$$

■

2.3 Markov Operator and its Eigenmeasures

In this section we define the Markov Operator, which in the case of a normalized IFSm gives the evolution of the distribution of the associated Markov Process, and show that the set of eigenmeasures for it is non-empty.

Definition 2.3.1. The **Markov Operator** $\mathcal{L}_{\mathrm{q}} : \mathscr{M}_s(X) \circlearrowleft$ is the unique bounded linear operator satisfying

$$\int_{\mathbf{X}} f \, \mathrm{d}[\mathcal{L}_q(\mu)] = \int_{\mathbf{X}} \mathrm{B}_{\mathrm{q}}(f) \, \mathrm{d}\mu,$$

for all $\mu \in \mathscr{M}_s(X)$ and $f \in C(\mathbf{X}, \mathbb{R})$.

In the case of a normalized IFSm, we can consider the Markov Process $\{Z_j, j \geqslant 0\}$ with initial distribution $Z_0 \sim \mu_0$, where $\mu_0 \in \mathscr{M}_1(X)$, and $Z_{j+1} = \tau_{\theta_j}(Z_j)$ for $j \geqslant 0$, where $\theta_j \sim q_{Z_j}$. Then, if $Z_j \sim \mu_j$, we have $\mu_{j+1} = \mathcal{L}_{\mathrm{q}}(\mu_j)$.

Theorem 2.3.2. *Let $\mathcal{R}_q = (\boldsymbol{X}, \tau, q)$ be a continuous IFSm. Then there exists a positive number $\rho \leqslant \rho(B_q)$ such that the set $\mathcal{G}^*(q) = \{\nu \in \mathscr{M}_1(X) : \mathcal{L}_q\nu = \rho\nu\}$ is not empty.*

Proof. Notice that the mapping

$$\mathscr{M}_1(X) \ni \gamma \mapsto \frac{\mathcal{L}_q(\gamma)}{\mathcal{L}_q(\gamma)(X)}$$

sends $\mathscr{M}_1(X)$ to itself. From its convexity and compactness, in the weak topology which is Hausdorff when X is metric and compact, it follows from the continuity of $\mathcal{L}_q$ and the Tychonov-Schauder Theorem that there is at least one probability measure ν satisfying $\mathcal{L}_q(\nu) = (\mathcal{L}_q(\nu)(X))\,\nu$.

We claim that

$$\inf_{x \in \mathbf{X}} \mathrm{q}_x(\Theta) \leqslant \mathcal{L}_q(\gamma)(\mathbf{X}) \leqslant \sup_{x \in \mathbf{X}} \mathrm{q}_x(\Theta) \tag{2.2}$$

for every $\gamma \in \mathscr{M}_1(X)$.

Indeed,

$$\mathrm{B}_{\mathrm{q}}(1)(x) = \int_{\Theta} 1 \, \mathrm{d}\mathrm{q}_x(\theta) = \mathrm{q}_x(\Theta),$$

$$\mathcal{L}_{\mathrm{q}}(\gamma)(\mathbf{X}) = \int_{\mathbf{X}} 1 \, \mathrm{d}[\mathcal{L}_{\mathrm{q}}\gamma] = \int_{\mathbf{X}} \mathrm{B}_{\mathrm{q}}(1) \, \mathrm{d}\gamma = \int_{\mathbf{X}} \mathrm{q}_x(\Theta) \, \mathrm{d}\gamma(x),$$

$$0 < \inf_{x \in \mathbf{X}} \mathrm{q}_x(\Theta) \leqslant \int_{\mathbf{X}} \mathrm{q}_x(\Theta) \, \mathrm{d}\gamma(x) \leqslant \sup_{x \in \mathbf{X}} \mathrm{q}_x(\Theta) < \infty.$$

From the inequality (2.2) follows that

$$0 < \rho \equiv \sup\{\mathcal{L}_{\mathrm{q}}(\nu)(X) : \mathcal{L}_{\mathrm{q}}(\nu) = (\mathcal{L}_{\mathrm{q}}(\nu)(X))\,\nu\} < +\infty.$$

By a compactness argument one can show the existence of $\nu \in \mathscr{M}_1(\mathbf{X})$ so that $\mathcal{L}_{\mathrm{q}}\nu = \rho\nu$. Indeed, let $(\nu_n)_{n \in \mathbb{N}}$ be a sequence such that $\mathcal{L}_{\mathrm{q}}(\nu_n)(\mathbf{X}) \uparrow \rho$, when n goes to infinity. Since $\mathscr{M}_1(X)$ is compact metric space in the weak topology we can assume, up to subsequence, that $\nu_n \rightharpoonup \nu$. This convergence together with the continuity of $\mathcal{L}_{\mathrm{q}}$ provides

$$\mathcal{L}_{\mathrm{q}}\nu = \lim_{n\to\infty} \mathcal{L}_{\mathrm{q}}\nu_n = \lim_{n\to\infty} \mathcal{L}_{\mathrm{q}}(\nu_n)(X)\nu_n = \rho\,\nu,$$

thus showing that the set $\mathcal{G}^*(q) \equiv \{\nu \in \mathscr{M}_1(X) : \mathcal{L}_{\mathrm{q}}\nu = \rho\,\nu\} \neq \varnothing$.

To finish the proof we observe that by using any $\nu \in \mathcal{G}^*(\mathrm{q})$, we get the following inequality

$$\rho^N = \int_X \mathrm{B}_{\mathrm{q}}^N(1)\,\mathrm{d}\nu \leqslant \left\|\mathrm{B}_{\mathrm{q}}^N\right\|.$$

From this inequality and Gelfand's Formula follows that $\rho \leqslant \rho(B_q)$. ■

2.4 Holonomic Measure and Disintegrations

Now we introduce holonomic measures, which play the role of invariant measures in the IFS setting.

An invariant measure for a classical dynamical system $T : \mathbf{X} \circlearrowleft$ on a compact space is a measure μ satisfying for all $f \in C(X, \mathbb{R})$

$$\int_{\mathbf{X}} f(T(x))\,\mathrm{d}\mu = \int_{\mathbf{X}} f(x)\,\mathrm{d}\mu, \quad \text{equivalently} \quad \int_{\mathbf{X}} f(T(x)) - f(x)\,\mathrm{d}\mu = 0.$$

From the Ergodic Theory point of view the natural generalization of this concept for an IFS $\mathcal{R} = (\mathbf{X}, \tau)$ is the concept of holonomy.

Consider the cartesian product space $\Omega \equiv \mathbf{X} \times \Theta$ and for each $f \in C(X, \mathbb{R})$ its "Θ-differential" $\mathrm{d}f : \Omega \to \mathbb{R}$ defined by $[\mathrm{d}_x f](\theta) \equiv f(\tau_\theta(x)) - f(x)$.

Definition 2.4.1. A measure $\hat{\mu}$ over Ω is said holonomic, with respect to an IFS $\mathcal{R}$ if for all $f \in C(X, \mathbb{R})$ we have

$$\int_\Omega [\mathrm{d}_x f](\theta)\, d\hat{\mu}(x, \theta) = 0.$$

Notation,

$$\mathcal{H}(\mathcal{R}) \equiv \{\hat{\mu} \,|\, \hat{\mu} \text{ is a holonomic probability measure with respect to } \mathcal{R}\}.$$

Since Ω is compact the set of all holonomic probability measures is obviously convex and compact. It is also not empty because Ω is compact and any average

$$\hat{\mu}_N \equiv \frac{1}{N}\sum_{j=0}^{N-1}\delta_{(x_j,\theta_j)},$$

where $x_{j+1} = \tau_{\theta_j}(x_j)$ and $x_0 \in \mathbf{X}$ is fixed, will have their cluster points in $\mathcal{H}(\mathcal{R})$. Indeed, for all $N \geqslant 1$ we have the following identity

$$\int_\Omega [\mathrm{d}_x f](\theta)\, d\hat{\mu}_N(x,\theta) = \frac{1}{N}\sum_{j=0}^{N-1}[\mathrm{d}_{x_j} f](\theta_j) = \frac{1}{N}(f(\tau_{\theta_{N-1}}(x_{N-1})) - f(x_0)).$$

From the above expression is easy to see that if $\hat{\mu}$ is a cluster point of the sequence $(\hat{\mu}_N)_{N\geqslant 1}$, then there is a subsequence $(N_k)_{k\to\infty}$ such that

$$\begin{aligned}\int_\Omega [\mathrm{d}_x f](\theta)\, d\hat{\mu}(x,\theta) &= \lim_{k\to\infty}\int_\Omega [\mathrm{d}_x f](\theta)\, d\hat{\mu}_{N_k}(x,\theta)\\ &= \lim_{k\to\infty}\frac{1}{N_k}(f(x_{N_k}) - f(x_0)) = 0.\end{aligned}$$

Theorem 2.4.2 (Disintegration)**.** *Let X and Y be compact metric spaces, $\hat{\mu} : \mathscr{B}(Y) \to [0,1]$ a Borel probability measure, $T : Y \to X$ a Borel mensurable function and for each $A \in \mathscr{B}(X)$ define a probability measure $\mu(A) \equiv \hat{\mu}(T^{-1}(A))$. Then there exists a family of Borel probability measures $(\mu_x)_{x\in X}$ on Y, uniquely determined μ-a.e, such that*

1. *$\mu_x(Y\backslash T^{-1}(x)) = 0$, μ-a.e;*
2. *$\displaystyle\int_Y f\, d\hat{\mu} = \int_X \left[\int_{T^{-1}(x)} f(y)\ d\mu_x(y)\right] d\mu(x)$.*

This decomposition is called the disintegration of $\hat{\mu}$, with respect to T.

Proof. For a proof of this theorem, see [DM78] p.78 or [AGS05], Theorem 5.3.1. ■

In this paper we are interested in disintegrations in cases where Y is the cartesian product $\Omega \equiv \mathbf{X}\times\Theta$ and $T : \Omega \to \mathbf{X}$ is the projection on the first coordinate. In such cases if $\hat{\mu}$ is any Borel probability measure on Ω, then follows from the first conclusion of Theorem 2.4.2 that the disintegration of $\hat{\mu}$ provides for each $x \in \mathbf{X}$ a unique probability measure μ_x (μ-a.e.) supported on $\{x\}\times\Theta$. So we can write the disintegration of $\hat{\mu}$ as $\mathrm{d}\hat{\mu}(x,\theta) = \mathrm{d}\mu_x(\theta)\mathrm{d}\mu(x)$, where here we are abusing notation identifying $\mu_x(\{x\}\times A)$ with $\mu_x(A)$.

Now we take $\hat{\mu} \in \mathcal{H}(\mathcal{R})$ and $f : \Omega \to \mathbb{R}$ as being any bounded continuous function, depending only on its first coordinate. From the very definition of holonomic measures we have the following equations

$$\int_{\Omega} [\mathrm{d}_x f](\theta)\, \mathrm{d}\hat{\mu}(x,\theta) = 0 \Longleftrightarrow \int_{\Omega} f(\tau_\theta(x))\, \mathrm{d}\hat{\mu}(x,\theta) = \int_{\Omega} f(x)\, \mathrm{d}\hat{\mu}(x,\theta)$$

by disintegrating both sides of the second equality above we get that

$$\int_{\mathbf{X}} \int_{\Theta} f(\tau_\theta(x))\, \mathrm{d}\mu_x(\theta) \mathrm{d}\mu(x) = \int_{\mathbf{X}} \int_{\Theta} f(x)\, \mathrm{d}\mu_x(\theta) \mathrm{d}\mu(x).$$

The above equation establish a natural link between holonomic measures for an IFS $\mathcal{R}$ and disintegrations. Given an IFS $\mathcal{R} = (\mathbf{X}, \tau)$ and $\hat{\mu} \in \mathcal{H}(\mathcal{R})$ we can use the previous equation to define an IFSm $\mathcal{R}_{\mathrm{q}} = (\mathbf{X}, \tau, \mathrm{q})$, where q_x is a probability defined by $\mathrm{q}_x(A) \equiv \mu_x(A)$ for every $A \in \mathcal{B}(\Theta)$. If B_{q} denotes the transfer operator q_x. If B_{q} denotes the transfer operator associated to $\mathcal{R}_{\mathrm{q}}$, we have from the last equation the following identity

$$\int_{\mathbf{X}} \mathrm{B}_{\mathrm{q}}(f)\, \mathrm{d}\mu = \int_{\mathbf{X}} f\, d\mu.$$

Since in the last equation f is an arbitrary bounded measurable function, depending only on the first coordinate, follows that the Markov operator associated to the IFSm $\mathcal{R}_{\mathrm{q}}$ satisfies

$$\mathcal{L}_{\mathrm{q}}(\mu) = \mu.$$

In other words the "second marginal" μ of a holonomic measure $\hat{\mu}$ is always an eingemeasure for the Markov operator associated to the IFSm $\mathcal{R}_{\mathrm{q}} = (\mathbf{X}, \tau, q)$ above defined.

Reciprocally, since the last five equations are equivalent, given an IFSm $\mathcal{R}_{\mathrm{q}} = (\mathbf{X}, \tau, q)$ such that the associated Markov operator has at least one fixed point, i.e., $\mathcal{L}_{\mathrm{q}}(\mu) = \mu$, then it is possible to define a holonomic probability measure $\hat{\mu} \in \mathcal{H}(\mathcal{R})$ given by $\mathrm{d}\hat{\mu}(x,\theta) = \mathrm{d}\mu_x(\theta)\, \mathrm{d}\mu(x)$, where $\mathrm{d}\mu_x(\theta) \equiv \mathrm{dq}_x(\theta)$. This Borel probablity measure on Ω will be called the **holonomic lifting** of μ, with respect to $\mathcal{R}_{\mathrm{q}}$.

2.5 Entropy and Pressure for IFSm

We now define two concepts of entropy, compare then, show sufficient conditions for them to be equal and introduce the topological pressure of a given potential, as well as the concept of equilibrium states. We show in this section a first result on the existence of equilibrium states. In all that follows, the a priori measure has a special role (see [LMMS15]).

As in the previous section the mapping $T : \Omega \to X$ denotes the projection on the first coordinate. Even when not explicitly mentioned, any disintegrations of a probability measure $\hat{\nu}$, defined over Ω, will be from now on considered with respect to T.

Definition 2.5.1 (Variational Entropy)**.** Let $\mathcal{R}$ be an IFS, $\hat{\nu} \in \mathcal{H}(\mathcal{R})$, μ a probability on Θ and $d\hat{\nu}(x,\theta) = d\nu_x(\theta)d\nu(x)$ a disintegration of $\hat{\nu}$, with respect to T. The variational entropy of $\hat{\nu}$ with a priori probability μ is defined by

$$h^{\mu}_{\mathrm{v}}(\hat{\nu}) \equiv \inf_{\substack{g \in C(\mathbf{X},\mathbb{R}) \\ g>0}} \left\{ \int_{\mathbf{X}} \ln \frac{\mathrm{B}_{\mu}(g)}{g} \, \mathrm{d}\nu \right\}.$$

Definition 2.5.2. When $\mathrm{q} = (\mathrm{q}_x)_{x \in \mathbf{X}}$ is a family of measures on Θ and μ a probability on Θ, and ν is a probability on $\mathbf{X}$, we say that the family q is ν-almost everywhere absolutely continuous with respect to μ when $\mathrm{q}_x \ll \mu$ for ν-almost everywhere x on $\mathbf{X}$ and write $\mathrm{q} \ll_{\nu} \mu$.

If $\mathrm{q} \ll_{\nu} \mu$, we define $J_x(\theta)$ such that $J_x = \frac{\mathrm{d}\mathrm{q}_x}{\mathrm{d}\mu}$ when $\mathrm{q}_x \ll \mu$ and $J_x(\theta) = 0$ otherwise.

Definition 2.5.3 (Average entropy)**.** Let $\mathcal{R}$ be an IFS, $\hat{\nu} \in \mathcal{H}(\mathcal{R})$, $d\hat{\nu}(x,\theta) = d\nu_x(\theta)d\nu(x)$ a disintegration of $\hat{\nu}$ with respect to T and μ a probability on Θ such that $(\nu_x) \ll_{\nu} \mu$. The average entropy of $\hat{\nu}$ with respect to μ is defined by

$$h^{\mu}_{\mathrm{a}}(\hat{\nu}) \equiv -\int_{\Omega} \ln J_x(\theta) \, \mathrm{d}\nu_x(\theta) \, \mathrm{d}\nu(x).$$

Definition 2.5.4 (Optimal Function)**.** Let $\mathcal{R}$ be an IFS, $\hat{\nu} \in \mathcal{H}(\mathcal{R})$, $\mathrm{d}\hat{\nu}(x,\theta) = \mathrm{d}\nu_x(\theta)\,\mathrm{d}\nu(x)$ a desintegration of $\hat{\nu}$ with respect to T and $q_x = \nu_x$, for all $x \in \mathbf{X}$. If $\mathrm{q} \ll_{\nu} \mu$, we say that a positive function $g \in C(\mathbf{X}, \mathbb{R})$ is optimal, with respect to the IFSm $\mathcal{R}_{\mathrm{q}}$, if we have

$$J_x(\theta) = \frac{g(\tau_\theta(x))}{\mathrm{B}_{\mu}(g)(x)}.$$

Proposition 2.5.5. *If* $dq_x(\theta) = Q_x(\theta)\, d\mu(\theta)$ *and* $dp_x(\theta) = P_x(\theta)\, d\mu(\theta)$ *are probabilities, then*

$$-\int_{\boldsymbol{X}}\int_{\Theta} \log(Q_x(\theta))\, dq_x(\theta)\, d\nu(x) \leqslant -\int_{\boldsymbol{X}}\int_{\Theta} \log(P_x(\theta))\, dq_x(\theta)\, d\nu(x).$$

Proof. Using Jensen's Inequality on $f(x) = -x\log(x)$ concave function

$$\begin{aligned}\int_{\mathbf{X}}\int_{\Theta} -\log\left(\frac{Q_x(\theta)}{P_x(\theta)}\right)\frac{Q_x(\theta)}{P_x(\theta)}\,\mathrm{dp}_x(\theta)\,\mathrm{d}\nu(x)\\ = \int_{\Omega} f\left(\frac{Q_x(\theta)}{P_x(\theta)}\right)\,\mathrm{dp}_x(\theta)\,\mathrm{d}\nu(x)\\ \leqslant f\left(\int_{\Omega}\mathrm{dp}_x(\theta)\,\mathrm{d}\nu(x)\right) = f(1) = 0.\end{aligned}$$

Then,

$$\begin{aligned}\int_{\mathbf{X}}\int_{\Theta} -\log\left(\frac{Q_x(\theta)}{P_x(\theta)}\right)\frac{Q_x(\theta)}{P_x(\theta)}P_x(\theta)\,\mathrm{d}\mu(\theta)\,\mathrm{d}\nu(x)\\ = \int_{\mathbf{X}}\int_{\Theta} -\log\left(\frac{Q_x(\theta)}{P_x(\theta)}\right)\,\mathrm{dq}_x(\theta)\,\mathrm{d}\nu(x) \leqslant 0\end{aligned}$$

Therefore,

$$-\int_{\mathbf{X}}\int_{\Theta}\log\left(Q_x(\theta)\right)\,\mathrm{dq}_x(\theta)\,\mathrm{d}\nu(x) \leqslant -\int_{\mathbf{X}}\int_{\Theta}\log\left(P_x(\theta)\right)\,\mathrm{dq}_x(\theta)\,\mathrm{d}\nu(x)$$

∎

Theorem 2.5.6. *Let $\mathcal{R}$ be an IFS, $\hat{\nu} \in \mathcal{H}(\mathcal{R})$, $d\hat{\nu}(x,\theta) = d\nu_x(\theta)\,d\nu(x)$ a disintegration of $\hat{\nu}$ with respect to T, $\mathcal{R}_q = (X,\tau,q)$ the IFSm with $q_x = \nu_x$ for all $x \in X$ and μ a probability on Θ such that $q \ll_{\nu} \mu$. Then*

1. $h_a^{\mu}(\hat{\nu}) \leqslant h_v^{\mu}(\hat{\nu}) \leqslant 0$;
2. *if there exists some optimal function ϕ, with respect to $\mathcal{R}_q$, then*

$$h_a^{\mu}(\hat{\nu}) = h_v^{\mu}(\hat{\nu}) = \int_{\boldsymbol{X}} \ln\frac{B_{\mu}(\phi)}{\phi}\,d\nu.$$

Proof. We first prove item *1*. Since q_x is a probability measure follows from the definition of average entropy that $h_{\mathrm{a}}^{\mu}(\hat{\nu}) \geqslant 0$.

From the definition of variational entropy we obtain

$$h_{\mathrm{v}}^{\mu}(\hat{\nu}) = \inf_{\substack{g\in C(\mathbf{X},\mathbb{R})\\ g>0}}\left\{\int_{\mathbf{X}}\ln\frac{\mathrm{B}_{\mu}(g)}{g}\,\mathrm{d}\nu\right\} \leqslant \int_{\mathbf{X}}\ln\frac{\mathrm{B}_{\mu}(1)}{1}\,\mathrm{d}\nu = 0.$$

To finish the proof of item 1 it remains to show that $h^\mu_{\mathrm{a}}(\hat\nu) \leqslant h^\mu_{\mathrm{v}}(\hat\nu)$. Let $g : X \to \mathbb{R}$ be continuous positive function and define for each $x \in X$ a probability where $\mathrm{dp}_x(\theta) = g(\tau_\theta(x))/\mathrm{B}_\mu(g)(x)\,\mathrm{d}\mu(\theta)$. From Proposition 2.5.5 and the properties of the holonomic measures we get the following inequalities for any continuous and positive function g:

$$\begin{aligned}
h^\mu_{\mathrm{a}}(\hat\nu) &= -\int_{\mathbf{X}}\int_\Theta \ln J_x(\theta)\,\mathrm{dq}_x(\theta)\,\mathrm{d}\nu(x)\\
&\leqslant -\int_{\mathbf{X}}\int_\Theta \ln\left(\frac{g\circ\tau_\theta}{\mathrm{B}_\mu(g)}\right)\mathrm{dq}_x(\theta)\,\mathrm{d}\nu(x)\\
&= -\int_{\mathbf{X}}\left[\int_\Theta \ln(g\circ\tau_\theta)\,\mathrm{dq}_x(\theta) - \int_\Theta \ln(\mathrm{B}_\mu(g))\,\mathrm{dq}_x(\theta)\right]\mathrm{d}\nu(x)\\
&= -\int_{\mathbf{X}} \mathrm{B}_{\mathrm{q}}(\ln g)\,\mathrm{d}\nu + \int_{\mathbf{X}} \ln(\mathrm{B}_\mu(g))\,\mathrm{d}\nu\\
&= -\int_{\mathbf{X}} \ln g\,\mathrm{d}\nu + \int_{\mathbf{X}} \ln(\mathrm{B}_\mu(g))\,\mathrm{d}\nu\\
&= \int_{\mathbf{X}} \ln\frac{\mathrm{B}_\mu(g)}{g}\,\mathrm{d}\nu,
\end{aligned}$$

Therefore, $h^\mu_{\mathrm{a}}(\hat\nu) \leqslant h^\mu_{\mathrm{v}}(\hat\nu)$. Futhermore, if $J_x(\theta) = \phi\circ\tau_\theta(x)/\mathrm{B}_\mu(\phi)(x)$ for some $\phi > 0$ continuous function, then $h^\mu_{\mathrm{a}}(\hat\nu) = \int_{\mathbf{X}} \log\frac{\mathrm{B}_\mu(\phi)}{\phi} = h^\mu_{\mathrm{v}}(\hat\nu)$. ■

Definition 2.5.7. Let $\psi : \mathbf{X} \to \mathbb{R}$ be a positive continuous function, μ a probability on Θ, and $\mathcal{R}_\psi(\mathbf{X}, \tau, \mathrm{q})$ an IFSm, where $\mathrm{dq}_x(\theta) = \psi\circ\tau_\theta(x)\,\mathrm{d}\mu(\theta)$. The topological pressure of ψ, with respect to $\mathcal{R}_\psi$, is defined by

$$P(\psi) = \sup_{\hat\nu\in\mathcal{H}(\mathcal{R})}\ \inf_{g\in C(\mathbf{X};\mathbb{R})g>0}\left\{\int_{\mathbf{X}} \ln\frac{\mathrm{B}_{\mathrm{q}}(g)}{g}\,\mathrm{d}\nu\right\}, \tag{2.3}$$

where $\nu := T_*\hat\nu$ for $T : \Omega \to \mathbf{X}$ the $\mathbf{X}$ projection.

Observe that, for the potential $\psi = 1$, the pressure $P(\psi) = \sup_{\hat\nu\in\mathcal{H}(\mathcal{R})} h^\mu_{\mathrm{v}}(\hat\nu)$. We also can obtain the following alternative forma for pressure.

Lemma 2.5.8. *Let $\psi : \boldsymbol{X} \to \mathbb{R}$ be a positive continuous function and $\mathcal{R}_\psi = (\boldsymbol{X}, \tau, q)$ be the $IFSm$ defined above, where $dq_x(\theta) = \psi\circ\tau_\theta(x)\,d\mu(\theta)$. Then, the topological pressure of ψ is alternatively given by*

$$P(\psi) = \sup_{\hat\nu\in\mathcal{H}(\mathcal{R})}\left\{h^\mu_v(\hat\nu) + \int_{\boldsymbol{X}} \log\psi\,d\nu\right\}.$$

Proof. First, note that $\mathrm{B_q}(g) = \mathrm{B}_\mu(g \cdot \psi)$ where $(g \cdot \psi)(x) = g(x)\psi(x)$. In fact

$$\mathrm{B_q}(g)(x) = \int_\Theta g \circ \tau_\theta(x) \psi \circ \tau_\theta(x) \, \mathrm{d}\mu(\theta) = \int_\Theta (g \cdot \psi) \circ \tau_\theta(x) \, \mathrm{d}\mu(\theta) = \mathrm{B}_\mu(g \cdot \psi)(x).$$

To finish the proof, we only need to use the pressure's definition and some basic properties as follows:

$$\begin{aligned}
P(\psi) &= \sup_{\hat{\nu} \in \mathcal{H}(\mathcal{R})} \inf_{g>0} \left\{ \int_{\mathbf{X}} \ln \frac{\mathrm{B_q}(g)}{g} \, \mathrm{d}\nu \right\}, \\
&= \sup_{\hat{\nu} \in \mathcal{H}(\mathcal{R})} \inf_{g>0} \left\{ \int_{\mathbf{X}} \log \psi \, \mathrm{d}\nu - \int_{\mathbf{X}} \log \psi \, \mathrm{d}\nu + \int_{\mathbf{X}} \ln \frac{\mathrm{B_q}(g)}{g} \, \mathrm{d}\nu \right\} \\
&= \sup_{\hat{\nu} \in \mathcal{H}(\mathcal{R})} \inf_{g>0} \left\{ \int_{\mathbf{X}} \log \psi \, \mathrm{d}\nu + \int_{\mathbf{X}} \ln \frac{\mathrm{B}_\mu(g \cdot \psi)}{g \cdot \psi} \, \mathrm{d}\nu \right\} \\
&= \sup_{\hat{\nu} \in \mathcal{H}(\mathcal{R})} \left\{ \int_{\mathbf{X}} \log \psi \, \mathrm{d}\nu + \inf_{\tilde{g}>0} \int_{\mathbf{X}} \ln \frac{\mathrm{B}_\mu(\tilde{g})}{\tilde{g}} \, \mathrm{d}\nu \right\} \\
&= \sup_{\hat{\nu} \in \mathcal{H}(\mathcal{R})} \left\{ h_{\mathrm{v}}^{\mu}(\hat{\nu}) + \int_{\mathbf{X}} \log \psi \, \mathrm{d}\nu \right\}.
\end{aligned}$$

■

Remark 2.5.9. Note that, if $\mathrm{dq}_x(\theta) = \frac{\psi \circ \tau_\theta(x)}{\mathrm{B}_\mu(\psi)(x)} \, \mathrm{d}\mu(\theta)$, by the Theorem 2.3.2 there exists $\rho > 0$ and ν s.t. $\mathcal{L}_\mathrm{q}(\nu) = \rho\nu$.

But,

$$\begin{aligned}
\rho = \int_{\mathbf{X}} \mathrm{d}\mathcal{L}_\mathrm{q}(\nu) &= \int_{\mathbf{X}} \mathrm{B_q}(1)(x) \, \mathrm{d}\nu(x) = \int_\Omega \frac{\psi \circ \tau_\theta(x)}{\mathrm{B}_\mu(\psi)(x)} \, \mathrm{d}\mu(\theta) \, \mathrm{d}\nu(x) \\
&= \int_{\mathbf{X}} \mathrm{B}_\mu(\psi)(x)^{-1} \int_\Theta \psi \circ \tau_\theta(x) \, \mathrm{d}\mu(\theta) \, \mathrm{d}\nu(x) = \int_{\mathbf{X}} \mathrm{d}\nu = 1.
\end{aligned}$$

Therefore we have

$$P(\psi) \geqslant \sup_{\nu \in \{\mathcal{L}_\mathrm{q}(\nu) = \nu\}} \int_{\mathbf{X}} \ln B_\mu(\psi) \, \mathrm{d}\nu.$$

Definition 2.5.10 (Equilibrium States)**.** Let $\mathcal{R}$ be an IFS, $\hat{\nu} \in \mathcal{H}(\mathcal{R})$ and μ a probability on Θ. Let $\psi : \mathbf{X} \to \mathbb{R}$ be a positive continuous function. We say that the holonomic measure $\hat{\nu}$ is an equilibrium state for (ψ, μ) if

$$h_{\mathrm{v}}^{\mu}(\hat{\nu}) + \int_{\mathbf{X}} \log \psi \, \mathrm{d}\nu = P(\psi).$$

Lemma 2.5.11. *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be compact separable metric spaces and $T : \boldsymbol{Y} \to \boldsymbol{X}$ be a continuous mapping. Then the push-forward mapping $\Phi_T \equiv \Phi : \mathscr{M}_1(\boldsymbol{Y}) \to \mathscr{M}_1(\boldsymbol{X})$ given by*

$$\Phi(\hat{\mu})(A) = \hat{\mu}(T^{-1}(A)), \quad \textit{where } \hat{\mu} \in \mathscr{M}_1(\boldsymbol{Y}) \textit{ and } A \in \mathscr{B}(\boldsymbol{X})$$

is weak-$*$ to weak-$*$ continuous.

Proof. Since we are assuming that $\mathbf{X}$ and $\mathbf{Y}$ are separable compact metric spaces then we can ensure that the weak-$*$ topology of both $\mathscr{M}_1(\mathbf{Y})$ and $\mathscr{M}_1(\mathbf{X})$ are metrizable. Therefore is enough to prove that Φ is sequentially continuous. Let $(\hat{\mu}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{M}_1(\mathbf{Y})$ so that $\hat{\mu}_n \rightharpoonup \hat{\mu}$. For any continuous real function $f : \mathbf{X} \to \mathbb{R}$ we have from change of variables theorem that

$$\int_{\mathbf{X}} f \, d[\Phi(\hat{\mu}_n)] = \int_{\mathbf{Y}} f \circ T \, d\hat{\mu}_n,$$

for any $n \in \mathbb{N}$. From the definition of the weak-$*$ topology follows that the rhs above converges when $n \to \infty$ and we have

$$\lim_{n\to\infty} \int_{\mathbf{X}} f \, d[\Phi(\hat{\mu}_n)] = \lim_{n\to\infty} \int_{\mathbf{Y}} f \circ T \, d\hat{\mu}_n = \int_{\mathbf{Y}} f \circ T \, d\hat{\mu} = \int_{\mathbf{X}} f \, d[\Phi(\hat{\mu})].$$

The last equality shows that $\Phi(\hat{\mu}_n) \rightharpoonup \Phi(\hat{\mu})$ and consequently the weak-$*$ to weak-$*$ continuity of Φ. ∎

For any $\hat{\nu} \in \mathcal{H}(\mathcal{R})$ it is always possible to disintegrate it as $d\hat{\nu}(x,i) = d\nu_x(i) d[\Phi(\hat{\nu})](x)$, where $\Phi(\hat{\nu}) \equiv \nu$ is the probability measure on $\mathscr{B}(\mathbf{X})$, defined for any $A \in \mathscr{B}(\mathbf{X})$ by

$$\nu(A) \equiv \Phi(\hat{\nu})(A) \equiv \hat{\nu}(T^{-1}(A)), \tag{2.4}$$

where $T : \Omega \to \mathbf{X}$ is the canonical projection of the first coordinate. This observation together with the previous lemma allow us to define a continuous mapping from $\mathcal{H}(\mathcal{R})$ to $\mathscr{M}_1(\mathbf{X})$ given by $\hat{\nu} \longmapsto \Phi(\hat{\nu}) \equiv \nu$.

We now prove a theorem ensuring the existence of equilibrium states for any continuous positive function ψ. Although this theorem has clear and elegant proof and works in great generality it has the disadvantage of providing no description of the set of equilibrium states.

Theorem 2.5.12 (Existence of Equilibrium States)**.** *Let $\mathcal{R}$ be an IFS, $\psi : X \to \mathbb{R}$ a positive continuous function and μ a probability on Θ. Then the set of equilibrium states for (ψ, μ) is not empty.*

Proof. As we observed above we can define a weak-$*$ to weak-$*$ continuous mapping

$$\mathcal{H}(\mathcal{R}) \ni \hat{\nu} \longmapsto \nu \in \mathscr{M}_1(\mathbf{X}),$$

where $d\hat{\nu}(x,i) = d\nu_x(i)d\nu(x)$ is the above constructed disintegration of $\hat{\nu}$. From this observation follows that for any fixed positive continuous g we have that the mapping $\mathcal{H}(\mathcal{R}) \ni \hat{\nu} \longmapsto \int_{\mathbf{X}} \ln(B_1(g)/g)\, d\nu$ is continuous with respect to the weak-$*$ topology. Therefore the mapping

$$\mathcal{H}(\mathcal{R}) \ni \hat{\nu} \longmapsto \inf_{\substack{g\in C(\mathbf{X},\mathbb{R})\\ g>0}} \left\{ \int_{\mathbf{X}} \ln \frac{B_1(g)}{g}\, d\nu \right\} \equiv h_v(\hat{\nu}).$$

is upper semi-continuous (USC) which implies by standard results that the following mapping is also USC

$$\mathcal{H}(\mathcal{R}) \ni \hat{\nu} \longmapsto h_v(\hat{\nu}) + \int_{\mathbf{X}} \ln(\psi(x))\, d\nu(x).$$

Since $\mathcal{H}(\mathcal{R})$ is compact in the weak-$*$ topology and the above mapping is USC then follows that this mapping attains its supremum at some $\hat{\mu} \in \mathcal{H}(\mathcal{R})$, i.e.,

$$\sup_{\hat{\nu}\in\mathcal{H}(\mathcal{R})} \left\{ \int_{\mathbf{X}} \ln\psi\, d\nu + h_v(\hat{\nu}) \right\} = \int_{\mathbf{X}} \ln\psi\, d\mu + h_v(\hat{\mu})$$

thus proving the existence of at least one equilibrium state. ■

2.6 Pressure Differentiability and Equilibrium States

We show in this section a uniqueness result for the equilibrium state introduced in the last section. In order to do that we will need to consider the functional $p : C(\mathbf{X}, \mathbb{R}) \to \mathbb{R}$ given by

$$p(\varphi) = P(\exp(\varphi)). \tag{2.5}$$

It is immediate to verify that p is a convex and finite valued functional. We say that a Borel signed measure $\nu \in \mathscr{M}_s(X)$ is a **subgradient** of p at φ if it satisfies the following subgradient inequality

$$p(\eta) \geqslant p(\varphi) + \nu(\eta - \varphi) \ \text{ for any } \ \eta \in \mathscr{M}_s(X).$$

The set of all subgradients at φ is called **subdifferential** of p at φ and denoted by $\partial p(\varphi)$. It is well-known that if p is a continuous mapping then $\partial p(\varphi) \neq \varnothing$ for any $\varphi \in C(\mathbf{X}, \mathbb{R})$.

We observe that for any pair $\varphi, \eta \in C(\mathbf{X}, \mathbb{R})$ and $0 < t < s$, follows from the convexity of p the following inequality

$$s(p(\varphi + t\eta) - p(\varphi)) \leqslant t(p(\varphi + s\eta) - p(\varphi)).$$

In particular, the one-sided directional derivative $d^+p(\varphi) : C(\mathbf{X}, \mathbb{R}) \to \mathbb{R}$ given by

$$d^+p(\varphi)(\eta) = \lim_{t \downarrow 0} \frac{p(\varphi + t\eta) - p(\varphi)}{t}$$

is well-defined for any $\varphi \in C(\mathbf{X}, \mathbb{R})$.

Theorem 2.6.1. *For any fixed $\varphi \in C(\boldsymbol{X}, \mathbb{R})$ we have*

1. *the signed measure $\nu \in \partial p(\varphi)$ iff $\nu(\eta) \leqslant d^+p(\varphi)(\eta)$ for all $\eta \in C(X, \mathbb{R})$;*
2. *the set $\partial p(\varphi)$ is a singleton iff $d^+p(\varphi)$ is the Gâteaux derivative of p at φ.*

Proof. This theorem is a consequence of Theorem 7.16 and Corollary 7.17 of the reference [CA06]. ■

Theorem 2.6.2. *Let $\mathcal{R}$ be an IFS, $\psi : \boldsymbol{X} \to \mathbb{R}$ a positive continuous function, μ a probability on Θ and $\Phi(\hat{\nu}) = \nu$ for $\hat{\nu} \in \mathcal{H}(\mathcal{R})$ where ν is given by disintegration with respect to T. If the functional p defined on* (2.5) *is Gâteaux differentiable at $\varphi \equiv \log \psi$, then*

$$\#\{\Phi(\hat{\mu}) : \ \hat{\mu} \text{ is an equilibrium state for } \psi\} = 1.$$

Proof. Suppose that $\hat{\mu}$ is an equilibrium state for ψ. Then we have from the definition of the pressure that

$$\begin{aligned} p(\varphi + t\eta) - p(\varphi) &= P(\psi \exp(t\eta)) - P(\psi) \\ &\geqslant h_v(\hat{\mu}) + \int_{\mathbf{X}} \ln \psi \; d\mu + \int_{\mathbf{X}} t\eta \, d\mu - h_v(\hat{\mu}) - \int_{\mathbf{X}} \ln \psi \; d\mu \\ &= t \int_{\mathbf{X}} \eta \, d\mu. \end{aligned}$$

Since we are assuming that p is Gâteaux differentiable at φ follows from the above inequality that $\mu(\eta) \leqslant d^+p(\varphi)(\eta)$ for all $\eta \in C(\mathbf{X}, \mathbb{R})$. From this inequality and Theorem 2.6.1 we can conclude that $\partial p(\varphi) = \{\mu\}$. Therefore for all equilibrium state $\hat{\mu}$ for ψ we have $\Phi(\hat{\mu}) = \partial p(\varphi)$, thus finishing the proof. ■

2.7 A Constructive Approach to Equilibrium States

In this section we prove the existence of a positive eigenfunction for the transfer operator associated to the spectral radius, and give a constructive proof of the existence of equilibrium states.

Let $\mathcal{R}_{\mathrm{q}} = (\mathbf{X}, \tau, \mathrm{q})$ assuming that there is μ a probability on Θ s.t. $\forall x \in \mathbf{X}$, $\mathrm{q}_x \ll \mu$ and $J : \Theta \times \mathbf{X} \to \mathbb{R}$, defined by $J(x, \theta) := \frac{\mathrm{dq}_x}{\mathrm{d}\mu}(\theta)$, a continuous function. Define $u(x, \theta) = \log J(\theta, x)$, and consider a parametric family of variable discount functions $\delta_n : [0, +\infty) \to \mathbb{R}$, where $\delta_n(t) \to I(t) = t$, when $n \to \infty$, pointwise and the normalized limits $\lim_n w_n(x) - \max w_n$ of the fixed points

$$w_n(x) := \log \int_\Theta e^{u(\theta,x)+\delta_n(w_n(\tau(\theta,x)))} \,\mathrm{d}\mu = \log \int_\Theta e^{\delta_n(w_n(\tau(\theta,x)))} \,\mathrm{dq}_x(\theta)$$

of a variable discount decision-making process, as defined in [CO19], $S_n := (\mathbf{X}, \Theta, \Psi, \tau, u, \delta_n)$ where $\Psi(x) = \Theta$ for all $x \in \mathbf{X}$ and the sequence (δ_n) satisfies the admissibility conditions:

1. the contration modulus γ_n of δ_n is also a variable discount function;
2. $\delta_n(0) = 0$ and $\delta_n(t) \leqslant t$ for any $t \in (0, +\infty)$;
3. for any fixed $\alpha > 0$ we have $\delta_n(t + \alpha) - \delta_n(t) \to \alpha$ when $n \to \infty$, uniformly in $t > 0$.

Theorem 2.7.1. *Let $\mathcal{R}_q$ and (δ_n) in above conditions such that the above defined u satisfy:*

1. *u is uniformly δ-bounded for (δ_n);*
2. *u is uniformly δ-dominated for (δ_n).*

Then there exists a positive and continuous eigenfunction h such that $B_q(h) = \rho(B_q)h$.

Proof. Theorem 3.28 of [CO19] implies that there exists $k \in [0, \|u\|_\infty]$ and $\varphi(x) := e^{h(x)}$ continuous and positive function with

$$e^k \varphi(x) = \int_\Theta \varphi \circ \tau(\theta, x)\, e^{u(x,\theta)} \,\mathrm{d}\mu(\theta) = \mathrm{B}_{\mathrm{q}}(\varphi)(x),$$

for all $x \in \mathbf{X}$. Now use the Theorem 2.2.5 and the theorem is proven. ■

Let $\psi : \mathbf{X} \to \mathbb{R}$ be a positive continuous function, μ a probability on Ω, and $\mathcal{R}_\psi(\mathbf{X}, \tau, \mathrm{q})$ an IFSm, where $\mathrm{dq}_x(\theta) = \psi \circ \tau_\theta(x)\, \mathrm{d}\mu(\theta)$. Suppose that there is h a positive continuous function such that $\mathrm{B}_\mathrm{q}(h) = \mathrm{B}_\mu(h \cdot \psi) = \rho(\mathrm{B}_\mathrm{q})h$. Then we can define, following Definition 2.5.4, $\mathcal{R}_\mathrm{p} = (\mathbf{X}, \tau, \mathrm{p})$ where

$$\frac{\mathrm{dp}_x}{\mathrm{d}\mu}(\theta) := \frac{(h \cdot \psi) \circ \tau_\theta(x)}{\mathrm{B}_\mu(h \cdot \psi)} = \frac{(h \cdot \psi) \circ \tau_\theta(x)}{\rho(\mathrm{B}_\mathrm{q})h} = \frac{h \circ \tau_\theta(x)}{\rho(\mathrm{B}_\mathrm{q})h} \cdot \frac{dq_x}{\mathrm{d}\mu}(\theta).$$

The IFSm $\mathcal{R}_\mathrm{p}$ is called the normalization of $\mathcal{R}_\mathrm{q}$. Take $\mathcal{L}_\mathrm{p}(\nu) = \nu$ and let $\hat{\nu}$ be the holonomic lifting of ν. Then by the Theorem 2.5.6 we know that

$$h^\mu_\mathrm{a}(\hat{\nu}) = h^\mu_\mathrm{v}(\hat{\nu}) = \int_\mathbf{X} \log \frac{\mathrm{B}_\mu(h \cdot \psi)}{h \cdot \psi}\, \mathrm{d}\nu = \log \rho(\mathrm{B}_\mathrm{q}) - \int_\mathbf{X} \log \psi\, \mathrm{d}\nu.$$

Then, choosing this $\hat{\nu}$ as particular in supremum given in Lema 2.5.8, $P(\psi) \geqslant h^\mu_\mathrm{v}(\hat{\nu}) + \int_\mathbf{X} \log \psi\, \mathrm{d}\nu = \log \rho(\mathrm{B}_\mathrm{q})$.

But, remember that the pressure, defined in expression (2.3), is

$$P(\psi) = \sup_{\hat{\nu} \in \mathcal{H}(\mathcal{R})} \inf_{g \in C(\mathbf{X};\mathbb{R}) g>0} \left\{ \int_\mathbf{X} \ln \frac{\mathrm{B}_\mathrm{q}(g)}{g}\, \mathrm{d}\nu \right\},$$

then,

$$\inf_{g \in C(\mathbf{X};\mathbb{R}) g>0} \left\{ \int_\mathbf{X} \ln \frac{\mathrm{B}_\mathrm{q}(g)}{g}\, \mathrm{d}\nu \right\} \leqslant \int_\mathbf{X} \ln \frac{\mathrm{B}_\mathrm{q}(h)}{h}\, \mathrm{d}\nu = \log \rho(\mathrm{B}_\mathrm{q}).$$

Taking the supremum over $\mathcal{H}(\mathcal{R})$ in both sides of above inequality, we have $P(\psi) \leqslant \log \rho(\mathrm{B}_\mathrm{q})$. Since the reverse inequality is already shown, we prove that

$$P(\psi) = \log \rho(\mathrm{B}_\mathrm{q}).$$

2.8 Example: Thermodynamic Formalism for right shift

Now we show that the $IFSm$ Thermodynamic Formalism generalizes the Thermodynamical Formalism for a dynamical system.

Let Θ be a compact metric space and $\mathbf{X} = \Theta^\mathbb{N}$. For each $\theta \in \Theta$ define $\sigma_\theta(x_1, x_2, \ldots) = (\theta, x_1, x_2, \ldots)$ the inverse branch of the right shift σ. Take μ a a-priori probability on Θ. Let $\psi : \Omega \to \mathbb{R}$ be a positive potential and $A = \log \psi$.

Now we define $\mathrm{dq}_x(\theta) = e^{A\circ\sigma_\theta(x)}\,\mathrm{d}\mu(\theta)$.

Then,

$$\mathrm{B}_{\mathrm{q}}(\phi) = \int_S e^{A\circ\sigma_\theta(x)}\phi\circ\sigma_\theta(x)\,\mathrm{d}\mu(\theta) = L_A(\phi)(x),$$

where L_A is the Ruelle Operator for the right shift σ and the potential A (see [LMMS15] for more details).

By Definition 2.5.7,

$$\begin{aligned} P(\psi) &= \sup_{\hat{\nu}\in\mathcal{H}(\mathcal{R})}\inf_{g>0}\left\{\int_{\mathbf{X}} \ln\frac{\mathrm{B}_{\mathrm{q}}(g)}{g}\,\mathrm{d}\nu\right\} \\ &= \sup_{\nu\in\mathcal{M}_\sigma}\inf_{g>0}\left\{\int_{\mathbf{X}} \ln\frac{L_A(g)}{g}\,\mathrm{d}\nu\right\}. \end{aligned}$$

The last expression is exactly the pressure of the potential A in Thermodynamical Formalism. Suppose that there is ϕ_A a positive continuous function such that $\mathrm{B}_{\mathrm{q}}(\phi_A) = L_A(\phi_A) = \rho(\mathcal{R}_A)\phi_A = \lambda_A\phi_A$. Then $P(e^A) = \log\lambda_A$. For instance, we know that if A is Hölder, then there exists such $\phi_A > 0$ function.

From this example, we can see that the $IFSm$ Thermodynamic Formalism, in certain sense, generalizes the Thermodynamical Formalism for a dynamical system. When we look at the family $\{\sigma_\theta\}_{\theta\in\Theta}$ of functions, we are looking at the inverse branches of the dynamical system.

2.9 Example: IFSm and a Possible Application in Economics

In Gupta et al. [GS+21] a chaos game is used to represent a time series as a PC plot and compare similarities and dissimilarities in different time frame such as the global pandemic of COVID-19. More precisely, the author consider the set $X = [0,1]^2$ as the base space and the four linear contractions

$$\begin{cases} \tau_A(x,y) = (0.5x, 0.5y) \\ \tau_B(x,y) = (0.5x+0.5, 0.5y) \\ \tau_C(x,y) = (0.5x, 0.5y+0.5) \\ \tau_D(x,y) = (0.5x+0.5, 0.5y+0.5) \end{cases} \tag{2.6}$$

so $(X, \tau_\theta)_{\theta\in\Theta}, \Theta = \{A, B, C, D\}$, is a classic contractive system whose attractor is X itself. Consider the identification:

A – if the market falls more than 0.01% of the previous value,

B – if the market falls less than 0.01% of the previous value,

C – if the market gains less than 0.01% of the previous value and
D – if the market gains more than 0.01% of the previous value,
in this way the time series of length N associated to a certain economic indicator is translated in to a genetic sequence

$$\gamma = (DACCADCDACDC...AACCBADD) \in \Theta^N.$$

Fixed an arbitrary initial point $Z_0 = (x_0, y_0) = (0.5, 0.5)$ the chaos game consist in iterating (x_0, y_0) by each map $Z_1 = (x_1, y_1) = \tau_D(x_0, y_0), Z_2 = (x_2, y_2) = \tau_A(x_1, y_1), Z_3 = (x_3, y_3) = \tau_C(x_2, y_2), ...$ according to γ. Considering $M \geqslant 2$ and the diadic partition of X given by

$$\bigcup_{\gamma' \in \Theta^M} \tau_{\gamma'_M}(\cdots(\tau_{\gamma'_1}(X))),$$

the PC plot W is a grey scale picture where the color of the each individual part $\Lambda = \tau_{\gamma'_M}(\cdots(\tau_{\gamma'_1}(X)))$ is the frequency of visits of the chaos game orbit $\{Z_j, j \geqslant 0\}$ to Λ that is,

$$W(\Lambda) = \frac{1}{N} \sharp\{j = 0, ..., N-1 \,|\, Z_j \in \Lambda\} \sim \mu(\Lambda).$$

Obviously, $\nu_N = \sum_\Lambda W(\Lambda)\delta_{(x_\Lambda, y_\Lambda)}$, where $(x_\Lambda,\, y_\Lambda) \in \Lambda$ is arbitrary, is a discrete probability and, if $\mu(\partial(\Lambda)) = 0$ then by the EET ([Elt87], Corollary 2), when $N \to \infty$, ν_N converge in distribution to the invariant measure μ for the IFS with probabilities $(X, \tau_\theta, p_\theta)_{\theta \in \Theta}$, where p_A, p_B, p_C, p_D are the relative frequency of each symbol A, B, C, D in γ, respectively.
For instance, if $N = 100$ and if a certain time series produces the genetic sequence

$$\gamma = (A, A, D, D, A, ..., A, D, B, C, C, B, A, D) \in \{A, B, C, D\}^{100},$$

we obtain the frequencies $[p_A, p_B, p_C, p_C] = [0.39,\ 0.17,\ 0.15,\ 0.29]$, and considering $M = 4$ we obtain the following PC plot which is an approximation for the invariant measure μ of the associated IFS with probabilities $[0.39,\ 0.17,\ 0.15,\ 0.29]$.

In order to generalize this idea we need to consider an infinite compact continuous range of values of the economic indicator, such as $\Theta = [0\%, 100\%]$, instead of taking only four values $\Theta = \{A, B, C, D\}$. Also, it is not reasonable to suppose that the probability of a change of $\theta\%$ in the indicator is independent of the current state of the indicator: the distribution of the occurrence of $\theta \in [0\%, 100\%]$, given the current state $Z \in X$ must be a measure of probability $q_Z(\cdot)$ over $[0\%, 100\%]$. Therefore, we believe the theory developed in the previous sections should be used when making this generalization.

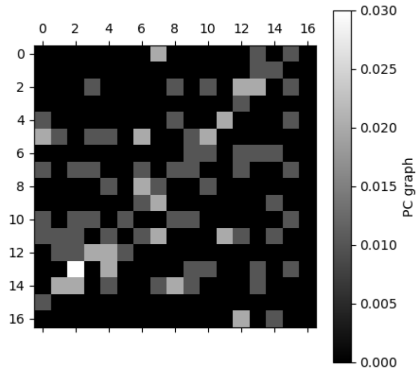

Figure 2.1: PC plot where each square represents one element Λ of the diadic partition and grey scale value $0 \leqslant \frac{1}{N} \sharp\{j = 0, ..., N-1 \mid Z_j \in \Lambda\} \leqslant 1$.

Aknowledgement: *We would like to thanks Prof. Leandro Cioletti whose contribution on the related preprint [CO17] was fundamental to establish the tools and ideas that we now generalize in our work.*

Chapter 3

Thermodynamic Formalism for Quantum Channels

Abstract

This entire chapter is part of the article [BKL21b]. Denote M_k the set of complex k by k matrices. We will analyze here quantum channels ϕ_L of the following kind: given a measurable function $L : M_k \to M_k$ and a measure μ on M_k we define the linear operator $\phi_L : M_k \to M_k$, via the expression $\rho \to \phi_L(\rho) = \int_{M_k} L(v)\rho L(v)^\dagger \, d\mu(v)$, where $L(v)^\dagger$ is the adjunt matrix of $L(v)$.

This paper [BFPP19] is our starting point. They considered the case where L was the identity.

Under some mild assumptions on the quantum channel ϕ_L we analyze the eigenvalue property for ϕ_L and we define entropy for such channel. For a fixed μ (the *a priori* measure) and for a given a Hamiltonian $H : M_k \to M_k$ we present a version of the Ruelle Theorem: a variational principle of pressure (associated to such H) related to an eigenvalue problem for the Ruelle operator. We introduce the concept of Gibbs channel.

We also show that for a fixed μ (with more than one point in the support) the set of L such that it is ϕ-Erg (also irreducible) for μ is a generic set.

We describe a related process X_n, $n \in \mathbb{N}$, taking values on the projective space $P(\mathbb{C}^k)$ and analyze the question of the existence of invariant probabilities. We also consider an associated process ρ_n, $n \in \mathbb{N}$, with values on $\mathcal{D}_k$ ($\mathcal{D}_k$ is the set of density operators). Via the barycenter, we associate the invariant probability mentioned above with the density operator fixed for ϕ_L.

3.1 Introduction

There are many different definitions and meanings for the concept of quantum dynamical entropy. We mention first the more well-known concepts due to Connes-Narnhofer-Thirring (see [CNT87]), Alicki-Fannes (see [AF94]), Accardi-Ohya-Watanabe (see [ASS20]), Stormer [Stø02] and Kossakowski-Ohya-Watanabe (see [KOW99]). In this case, the entropy can be exactly computed for several examples of quantum dynamical systems.

A different approach appears in [SS17] and [SŻ94] where the authors present their definition of quantum dynamical entropy (see also [AW19]).

Classical texts on quantum entropy are [AF01], [Ben09], [Ben03] and [OP04], and for quantum channels we also mention [JP19], [Lid19], and [Wol12].

We present here a certain concept of dynamical quantum entropy. A confirmation that this entropy is in fact a concept that describes valuable information from a dynamic point of view is its relationship with Lyapunov exponents as presented in [BKL21a] by the same authors. Lyapunov exponents are quite important tools that are used in Physics, Dynamics, and Fractals. Moreover, in [BKL21a] we will show that the purification property is C^0-generic.

One of the most challenging open problems in quantum information theory is to introduce a good definition capable of quantifying how *entanglement* behaves when part of an entangled state is sent through a quantum channel. Therefore the understanding of quantum channels is a problem of central importance.

Denote M_k the set of complex k by k matrices. We will analyze here quantum channels ϕ_L of the following kind: given a measurable function $L : M_k \to M_k$ and the measure μ on M_k we define the linear operator $\phi_L : M_k \to M_k$, via the expression $\rho \to \phi_L(\rho) = \int_{M_k} L(v)\rho L(v)^\dagger \, d\mu(v)$.

The probability μ will play the role of an *a priori* probability for defining entropy (in the spirit of [LMMS15]) as described in section 3.4.

In [BFPP19] the authors present interesting results for the case $L = I$. This paper is our starting point and we follow its notation as much as possible. Given L (as above) one can consider in the setting of [BFPP19] a new probability $\mu_L = \mu \circ L^{-1}$ and part of the results presented here can be recovered from there (using μ_L instead of μ).

We will present all the proofs here using L and μ as above (and not via μ_L) because this will be more natural for our future reasoning (for instance when analyzing generic properties).

In the Thermodynamic Formalism version of Quantum Information, the L will help on the one hand to express the *analogous concept of function* *(even the analog of a Hamiltonian)* and on the other hand, a certain class of L - together with the *a priori* probability μ on M_k - will help to describe the analogous concept of *invariant probability.* Later we will elaborate on that.

This paper is self-contained.

For a fixed μ and a general L we present a natural concept of entropy for a channel in order to develop a version of Gibbs formalism which seems natural to us. Example 3.8.5 in Section 3.8 (the Markov model in quantum information) will show that our definition is a natural extension of the classical concept of entropy. We point out that the definition of entropy we will consider here is a generalization of the concept described on the papers [BLLC10], [BLLC11a] and [BLLC11b]. This particular way of defining entropy is inspired by the results of [Sło03] which consider iterated function systems.

For a given $H : M_k \to M_k$ (which plays the role of a Hamiltonian) we present a version of the Ruelle Theorem for ϕ_H: a variational principle of pressure related to an eigenvalue problem for a kind of Ruelle operator (see Theorem 3.4.8).

A question of terminology: the operator H (mentioned above as Hamiltonian) could also be naturally called Liouvillian; it would make perfect sense taking into account that M_k is an algebra of quantum observables where the operator acts (Heisenberg picture of QM). The notation L used by the authors in [BFPP19] was probably inspired by their understanding that L plays the role of a Liouvillian operator.

We say that $E \subset \mathbb{C}^k$ is (L, μ)-invariant if $L(v)(E) \subset E$, for all v in the support of μ. Given $L : M_k \to M_k$ and μ on M_k, we say that L is ϕ-Erg for μ, if there exists an unique minimal non-trivial space E, such that, E is (L, μ)-invariant. We will show in Section 3.7 that for a fixed μ (with more than one point in the support) the set of L such that it is ϕ-Erg for μ is generic. In fact, the set of L which are irreducible is dense according to Theorem 3.7.5.

The introduction of this variable L allows us to consider questions of a generic nature in this type of problem.

We point out that here we explore the point of view that the (discrete-time dynamical) classical Kolmogorov-Shannon entropy of an invariant probability is in some way attached to an *a priori* probability (even if this is not transparent on the classical definition). This point of view becomes more clear when someone tries to analyze the generalized XY model (the symbolic space $M^{\mathbb{N}}$ where the alphabet M is a compact metric space) which is a case with the property that each point has an uncountable number of preimages (see [LMMS15] and [BCL$^+$11] for discussion). In the dynamical setting of [LMMS15] to define entropy it is necessary first to introduce the transfer (Ruelle) operator (which we claim - in some sense - is a more fundamental concept than entropy) which requires an *a priori* probability (not a general measure). Our results correspond to the case where the alphabet (that in some sense corresponds to the support of the *a priori* probability μ) can be uncountable.

The point of view of defining entropy via the limit of dynamical partitions is not suitable for the generalized XY model. We are just saying that in any case the concept of entropy can be recovered via the Ruelle operator.

We point out, as a curiosity, that for the computation of the classical Kolmogorov-Shannon entropy of a shift invariant probability on $\{1, 2, \ldots, d\}^{\mathbb{N}}$ one should take as the *a priori* measure (not a probability) the counting measure on $\{1, 2, \ldots, d\}$ (see discussion in [LMMS15]). In the case, we take as *a priori probability* μ the uniform normalized probability on $\{1, 2, \ldots, d\}$ the entropy will be negative (it will be Kolmogorov-Shannon entropy - $\log d$). In this case the independent $1/d$ probability on $\{1, 2, \ldots, d\}^{\mathbb{N}}$ will have maximal entropy equal 0.

A general reference for Thermodynamic Formalism is [PP90] and [Lop11].

We point out that we consider here Quantum Channels but the associated discrete-time process is associated with a Classical Stochastic Process (a probability on the infinite product of an uncountable state space) and not to a quantum spin-lattice, where it is required the use of the tensor product (see [LMMM18] and [BLMM21]).

After some initial sections describing basic properties which will be required later we analyze in Section 3.3 the eigenvalue property for ϕ_L.

Under some mild assumptions on ϕ_L, we define the entropy of the channel ϕ_L in Section 3.4. For a fixed μ (the *a priori* measure) and a given Hamiltonian $H : M_k \to M_k$ we present a variational principle of pressure and we associate with all this an eigenvalue problem on Section 3.3. In Definition 3.4.4 we introduce the concept of Gibbs channel for the Hamiltonian H (or, for the channel ϕ_H).

In Section 3.5 we describe (adapting [BFPP19] to the present setting) a process X_n, $n \in \mathbb{N}$, taking values on the projective space $P(\mathbb{C}^k)$. We also analyze the existence of an initial invariant probability for this process (see Theorem 3.5.2).

In Section 3.6 we consider a process ρ_n, $n \in \mathbb{N}$ (called quantum trajectory by T. Benoist, M. Fraas, Y. Pautrat, and C. Pellegrini) taking values on $\mathcal{D}_k$, where $\mathcal{D}_k$ is the set of density operators on M_k. Using the definition of barycenter taken from [Sło03] we relate in proposition 3.6.2 the invariant probabilities of Section 3.5 with the fixed point of Section 3.3.

In Section 3.7, for a fixed measure μ, we show that ϕ-Erg (and also irreducible) is a generic property for L (see Corollary 3.7.10).

In Section 3.8, we present several examples that will help the reader in understanding the theory. Example 3.8.5 shows that the definition of entropy for Quantum Channels described here is the natural generalization of the classical concept of entropy. In another example in this section, we consider the case where μ is a probability with support on a linear space of M_2 (see Example 3.8.6), and among other things we estimate the entropy of the channel.

In the final section 3.9 we will present some clarifications on which directions our work is related to relevant issues in the area connected to quantum entropy.

3.2 General properties

We present some basic definitions.

We denote by M_k, $k \in \mathbb{N}$, the set of complex k by k matrices. We consider $\mathcal{M}$ the standard Borel sigma-algebra over M_k and on $\mathbb{C}^k$, we consider the canonical Euclidean inner product.

We denote by Id_k the identity matrix on M_k.

According to our notation, $\dagger$ denotes the operation of taking the dual of a matrix with respect to the canonical inner product on $\mathbb{C}^k$.

Here tr denotes the trace of a matrix.

Given two matrices A and B we define the Hilbert-Schmidt product

$$\langle A , B \rangle = \operatorname{tr}(A B^\dagger).$$

This induces a norm $\|A\| = \sqrt{\langle A , A \rangle}$ on the Hilbert space M_k which will be called the Hilbert-Schmidt norm.

Given a linear operator ϕ on M_k we denote by $\phi^* : M_k \to M_k$ the dual linear operator in the sense of Hilbert-Schmidt, that is, if for all X, Y we get

$$\langle \phi(X) , Y \rangle = \langle X , \phi^*(Y) \rangle.$$

Now, consider a measure μ on $\mathcal{M}$.

For an integrable transformation $F : M_k \to M_k$:

$$\int_{M_k} F(v)\, d\mu(v) = \left(\int_{M_k} F(v)_{i,j}\, d\mu(v) \right)_{i,j},$$

where $F(v)_{i,j}$ is the entry (i,j) of the matrix $F(v)$.

We will list a sequence of trivial results (without proof) that will be used next.

Lemma 3.2.1. *For an integrable transformation $F : M_k \to M_k$*

$$\operatorname{tr} \int_{M_k} F(v)\ d\mu(v) = \int_{M_k} \operatorname{tr} F(v)\ d\mu(v).$$

Lemma 3.2.2. *Given a matrix $B \in M_k$ and an integrable transformation $F : M_k \to M_k$, then,*

$$B \int_{M_k} F(v)\ d\mu(v) = \int_{M_k} BF(v)\ d\mu(v).$$

Proposition 3.2.3. *If $l : M_k \to \mathbb{C}$ is a linear functional and $F : M_k \to M_k$ is integrable, then,*

$$l\left(\int_{M_k} F(v)\ d\mu(v)\right) = \int_{M_k} l(F(v))\ d\mu(v).$$

Definition 3.2.4. Given a measure μ on M_k and a measurable funtion $L : M_k \to M_k$, we say that μ is L-square integrable, if

$$\int_{M_k} \|L(v)\|^2\ d\mu(v) < \infty.$$

For a fixed L we denote by $\mathcal{M}(L)$ the set of L-square-integrable measures. We also denote $\mathcal{P}(L)$ the set of L-square-integrable probabilities.

Definition 3.2.5. Given a measurable function $L : M_k \to M_k$ and a L-square-integrable measure μ we define the linear operator $\phi_L : M_k \to M_k$ via the expression

$$\rho \to \phi_L(\rho) = \int_{M_k} L(v)\rho L(v)^\dagger\ d\mu(v).$$

For a given $H : M_k \to M_k$ (which plays the role of a Hamiltonian) we present a version of the Ruelle Theorem: a variational principle of pressure related to an eigenvalue problem for a kind of Ruelle operator (see Theorem 3.4.8).

Remember that if $A, B \in M_k$ with $A, B \geqslant 0$, then $\operatorname{tr}(AB) \leqslant \operatorname{tr}(A)\operatorname{tr}(B)$.

Therefore, if $\rho \geqslant 0$, we have

$$\begin{aligned}
\|\phi_L(\rho)\|^2 &= \operatorname{tr}\left(\phi_L(\rho)\phi_L(\rho)^\dagger\right) \\
&= \int_{M_k}\int_{M_k} \operatorname{tr}\left(L(v)\rho L(v)^\dagger L(w)^\dagger \rho L(w)\right) d\mu(v)\, d\mu(w) \\
&= \int_{M_k}\int_{M_k} \operatorname{tr}\left(\rho L(v)^\dagger L(w)^\dagger \rho L(w) L(v)\right)\ d\mu(v)\, d\mu(w) \\
&\leqslant \operatorname{tr}(\rho) \int_{M_k}\int_{M_k} \operatorname{tr}\left(\rho L(w) L(v) L(v)^\dagger L(w)^\dagger\right)\ d\mu(v)\, d\mu(w) \\
&\leqslant \operatorname{tr}(\rho)^2 \int_{M_k}\int_{M_k} \operatorname{tr}\left(L(w) L(v) L(v)^\dagger L(w)^\dagger\right)\ d\mu(v)\, d\mu(w) \\
&\leqslant \operatorname{tr}(\rho)^2 \int_{M_k} \|L(v)\|^2\ d\mu(v) \int_{M_k} \|L(w)\|^2\ d\mu(w) < \infty.
\end{aligned}$$

For a general $\rho \in M_k$, we write $\rho = \rho_+ - \rho_-$ where $\rho_+ = |\rho|$ and $\rho_- = |\rho| - \rho$ are both positive semidefinite matrices. By linearity of ϕ_L, we have

$$\phi_L(\rho) = \phi_L(\rho_+) - \phi_L(\rho_-),$$

hence, ϕ_L is well defined.

Proposition 3.2.6. *Given a measurable function $L : M_k \to M_k$ and a L-square integrable measure μ, then, the dual transformation ϕ_L^* is given by*

$$\phi_L^*(\rho) = \int_{M_k} L(v)^\dagger \rho L(v)\ d\mu(v).$$

Definition 3.2.7. Given a measurable function $L : M_k \to M_k$ and a L-square integrable measure μ over M_k, then, the transformation ϕ_L is called **stochastic** if

$$\phi_L^*(\mathrm{Id}_k) = \int_{M_k} L(v)^\dagger L(v)\ d\mu(v) = \mathrm{Id}_k.$$

By abuse of language, we sometimes say L stochastic to mean that ϕ_L is stochastic.

We will be able to define the concept of entropy when the ϕ_L is stochastic.

Definition 3.2.8. A linear map $\phi : M_k \to M_k$ is called **positive** if takes positive matrices to positive matrices.

Definition 3.2.9. A positive linear map $\phi : M_k \to M_k$ is called **completely positive**, if for any m, the linear map $\phi_m = \phi \otimes I_m : M_k \otimes M_m \to M_k \otimes M_m$ is positive, where I_m is the identity operator acting on the matrices in M_m.

Definition 3.2.10. If $\phi : M_k \to M_k$ is a linear map and satisfies

1. ϕ is completely positive;
2. ϕ preserves trace.

Then, we say that ϕ is a quantum channel.

Theorem 3.2.11. *Given $L : M_k \to M_k$ and μ a L-square measure. Then the associated transformation ϕ_L is completely positive. Moreover, if ϕ_L is stochastic then preserves trace.*

Proof. *1. ϕ_L is completely positive:* suppose $A \otimes B \in M_n \otimes M_k$ satisfies $A \otimes B \geqslant 0$ and $\psi \in \mathbb{C}^n \otimes \mathbb{C}^k$. Then, if $\psi_L(v) = (Id_n \otimes L(v)^\dagger)\psi$ we get

$$\begin{aligned}
\langle \psi | A \otimes \phi_L(B) | \psi \rangle &= \langle \psi | A \otimes \int_{M_k} L(v) B L(v)^\dagger \, d\mu(v) | \psi \rangle \\
&= \int_{M_k} \langle \psi | A \otimes (L(v) B L(v)^\dagger) | \psi \rangle \, d\mu(v) \\
&= \int_{M_k} \langle \psi | (\mathrm{Id}_n \otimes L(v))(A \otimes B)(Id_n \otimes L(v)^\dagger) | \psi \rangle \, d\mu(v) \\
&= \int_{M_k} \langle (\mathrm{Id}_n \otimes L(v)^\dagger)\psi | (A \otimes B) | (Id_n \otimes L(v)^\dagger)\psi \rangle \, d\mu(v) \\
&= \int_{M_k} \langle \psi_L(v) | (A \otimes B) | \psi_L(v) \rangle \, d\mu(v) \geqslant 0.
\end{aligned}$$

Above we use the positivity of $A \otimes B$ in order to get $\langle \psi_L(v) | (A \otimes B) | \psi_L(v) \rangle \geqslant 0$. We also used in some of the equalities the fact that $l(X) := \langle \psi | A \otimes X | \psi \rangle$ is a linear functional and therefore we can apply proposition 3.2.3.

2. Under our assumption ϕ_L preserves trace: given $B \in M_k$

$$\begin{aligned}
\operatorname{tr} \phi_L(B) &= \operatorname{tr} \left(\int_{M_k} L(v) B L(v)^\dagger \, d\mu(v) \right) \\
&= \int_{M_k} \operatorname{tr} \left(L(v) B L(v)^\dagger \right) \, d\mu(v) \\
&= \int_{M_k} \operatorname{tr} \left(B L(v)^\dagger L(v) \right) \, d\mu(v) \\
&= \operatorname{tr} \left(B \int_{M_k} L(v)^\dagger L(v) \, d\mu(v) \right) \\
&= \operatorname{tr}(B).
\end{aligned}$$

■

Remark 3.2.12 (ϕ_L^* is completely positive)**.** When L is measurable, then, using the same reasoning as above one can show that ϕ_L^* is completely positive.

We say that ϕ_L preserves unity if $\phi_L(\mathrm{Id}) = \mathrm{Id}$. In this case, ϕ_L^* preserves trace. When ϕ_L^* preserves the identity then ϕ_L preserves trace.

3.3 The eigenvalue property for ϕ_L

In this section, we will investigate questions related to the existence of eigenvalues and eigenmatrices for the setting of Quantum Information. An important role will be played by a result about positive maps on C^*-algebras described in [Eva78] which presents a noncommutative version of the Perron Theorem.

Definition 3.3.1 (Irreducibility)**.** We say that $\phi : M_k \to M_k$ is irreducible if one of the equivalent properties is true

- Does not exists $\lambda > 0$ and a projection p on a proper non-trivial subspace of $\mathbb{C}^k$, such that, $\phi(p) \leqslant \lambda p$;
- For all non null $A \geqslant 0$, $(\mathrm{Id} + \phi)^{k-1}(A) > 0$;
- For all non null $A \geqslant 0$ there exists $t_A > 0$, such that, $(e^{t_A \phi})(A) > 0$;
- If $P \in M_k$ is a hermitian projector such that $\phi(PM_kP) \subset PM_kP$, then $P \in \{0, \mathrm{Id}\}$;
- For all pair of non null positive matrices $A, B \in M_k$ there exists a natural number $n \in \{1, \ldots, k-1\}$, such that, $\mathrm{tr}\,[B\phi^n(A)] > 0$.

The proof of the equivalence of the two first items appears in [Eva78].

The equivalence of the two middle ones appears in [Sch00] where also one can find the proof of the improved positivity (to be defined below) which implies irreducibility. For the proof that the last two items, we refer to [Wol12].

Definition 3.3.2 (Irreducibility)**.** Given μ we will say (by abuse of language) that L is irreducible for μ or μ-irreducible if the associated ϕ_L is irreducible.

Lemma 3.3.3. *Given $L : M_k \to M_k$ and μ a L-square measure, the following statements are equivalent:*

1. *ϕ_L is irreducible;*

2. *If $E \subset \mathbb{C}^k$ is a subspace such that $L(v)E \subset E$ for all $v \in \operatorname{supp}\mu$, then $E \in \{\{0\}, \mathbb{C}^k\}$.*

Proof. 1. → 2.: If ϕ_L is irreducible and $E \subset \mathbb{C}^k$ is a subspace such that $L(v)E \subset E$ for all $v \in \operatorname{supp}\mu$, take P the orthogonal projection on E. Then $PL(v)P = L(v)P$ for all $v \in \operatorname{supp}\mu$. Moreover, for every $A \in M_k$

$$
\begin{aligned}
\phi_L(PAP) &= \int_{M_k} L(v)PAPL(v)^\dagger \, d\mu(v) \\
&= \int_{\operatorname{supp}\mu} PL(v)PAPL(v)^\dagger P \, d\mu(v) \\
&= P \int_{\operatorname{supp}\mu} L(v)PAPL(v)^\dagger \, d\mu(v) P \in PM_kP,
\end{aligned}
$$

and by the fourth equivalence of 3.3.1, $P \in \{0, \mathrm{Id}\}$. Therefore $E = \{0\}$ or $E = \mathbb{C}^k$.

2. → 1.: If there is $P \in M_k$ Hermitian projection such that $\phi_L(PM_kP) \in PM_kP$, take $E = \operatorname{Im} P$, $x \in E$ and $A = |\, x \,\rangle\langle\, x \,|$. Then we have

$$
\begin{aligned}
0 &= \operatorname{tr}\left(\phi_L(PAP) - P\phi_L(PAP)P\right) = \\
&= \int_{M_k} \operatorname{tr}\left(L(v)AL(v)^\dagger - PL(v)AL(v)^\dagger P\right) d\mu(v) \\
&= \int_{B_x} \operatorname{tr}\left(L(v)AL(v)^\dagger - PL(v)AL(v)^\dagger P\right) d\mu(v) \\
&= \int_{B_x} \operatorname{tr}\left(L(v)AL(v)^\dagger (I - P)\right) d\mu(v),
\end{aligned}
$$

where $B_x := \{v \,|\, L(v)x \notin E\}$. Suppose $P \notin \{0, \mathrm{Id}\}$, then we have

$$
\operatorname{tr}\left(L(v)AL(v)^\dagger\right) - \operatorname{tr}\left(PL(v)AL(v)^\dagger P\right) > 0,
$$

and, since the integral is zero, $\mu(B_x) = 0$. Thus, for all $x \in E$, $\operatorname{supp}\mu \subset B_x^c$ and so $L(v)E \subset E$ for $v \in \operatorname{supp}\mu$. By hypothesis, we have that $E \in \{\{0\}, \mathbb{C}^k\}$ which brings us to an absurd.

■

Definition 3.3.4 (Improving positivity)**.** We say that ϕ_L is positivity improving, if $\phi_L(A) > 0$, for any non-null $A \geqslant 0$. Note that improving positivity implies irreducibility.

For any μ and square-integrable L the Theorem 3.2.11 assures that ϕ_L is completely positive. In the case ϕ_L is irreducible we can use the Theorem 2.3 and 2.4 of [Eva78] in order to get λ and $\rho > 0$, such that, $\phi_L(\rho) = \lambda\rho$ (ρ is unique up to multiplication by scalar). For what comes next, we will choose ρ such that $\operatorname{tr}\rho = 1$. Moreover, in the same work the authors show that ϕ_L is irreducible, if and only if, ϕ_L^* also is completely positive, and therefore we get:

Theorem 3.3.5 (The spectral radius is a simple eigenvalue)**.** *Given a square integrable $L : M_k \to M_k$ assume that the associated ϕ_L is irreducible. On a Hilbert space, the spectral radius $\lambda_L > 0$ of ϕ_L and ϕ_L^* is the same. In this case it is also an eigenvalue and it is simple. We denote, respectively, by $\rho_L > 0$ and $\sigma_L > 0$, the eigenmatrices, such that, $\phi_L(\rho_L) = \lambda_L\rho_L$ and $\phi_L^*(\sigma_L) = \lambda_L\sigma_L$, where ρ_L and σ_L are the unique non null eigenmatrices (up to multiplication by scalar).*

The above theorem is the natural version of the Perron-Frobenius Theorem for the present setting. It is natural to think that ϕ_L acts on density matrices and ϕ_L^* acts in selfadjoint matrices.

Remark 3.3.6. We choose ρ_L in such way that $\operatorname{tr}\rho_L = 1$ and after that, we take σ_L such that $\operatorname{tr}(\sigma_L\rho_L) = 1$. By doing that, we have chosen the precise scalar multiples that makes both ρ_L and $\sigma_L\rho_L$ densities. Notice that, as eigendensity, ρ_L is unique. We point out that at this moment it is natural to make an analogy with Thermodynamic Formalism: ϕ_L^* corresponds to the Ruelle operator (acting on functions) and ϕ_L to the dual of the Ruelle operator (acting on probabilities). We refer the reader to [PP90] for details. In this sense, the density operator $\sigma_L\rho_L$ plays the role of an equilibrium probability. The paper [Spi72] by Spitzer describes this formalism in a simple way in the case the potential depends on two coordinates.

Remark 3.3.7. If L is irreducible and stochastic (resp. ϕ_L is unital, i.e., $\phi_L(\mathrm{Id}) = \mathrm{Id}$) then $\lambda_L = 1$ and $\sigma_L = \mathrm{Id}_k$ (resp. $\rho_L = \mathrm{Id}_k$) by Proposition 6.1 on [Wol12] page 91.

3.3.1 Normalization

We consider in this section a fixed measure μ over M_k which plays the role of the *a priori* probability.

In this section we will introduce the concept of normalized transformation L (see definition 3.3.9). If L is not normalized we will be able to find an associated $\hat{L}$ which is normalized (see (3.1)).

Given a continuous L **(variable)** we assume in this section that ϕ_L is irreducible (we do not assume that preserves trace).

We will associate to this square integrable transformation $L : M_k \to M_k$ (and the associated ϕ_L) another transformation $\hat{L} : M_k \to M_k$ which will correspond to a normalization of L. This will define another quantum channel $\phi_{\hat{L}} : M_k \to M_k$.

Results of this section have a large intersection with some material in [Wol12]. For completeness, we describe here what we will need later.

Consider σ_L e λ_L as described above. As σ_L is positive we consider ${\sigma_L}^{1/2} > 0$ and ${\sigma_L}^{-1/2} > 0$.

In this way we define

$$\hat{L}(v) = \frac{1}{\sqrt{\lambda_L}} {\sigma_L}^{1/2} L(v) {\sigma_L}^{-1/2}. \tag{3.1}$$

Using the measure μ we can define the associated $\phi_{\hat{L}}$.

Therefore,

$$\begin{aligned}
\phi_{\hat{L}}^*(\mathrm{Id}) &= \frac{1}{\lambda_L} \int_{M_k} {\sigma_L}^{-1/2} L(v)^\dagger {\sigma_L}^{1/2} {\sigma_L}^{1/2} L(v) \sigma_L^{-1/2}\, d\mu \\
&= \frac{1}{\lambda_L} \sigma_L^{-1/2} \int_{M_k} L(v)^\dagger \sigma_L L(v)\, d\mu\ \sigma_L^{-1/2} \\
&= \frac{1}{\lambda_L} \sigma_L^{-1/2} \phi_L^*(\sigma_L)\, \sigma_L^{-1/2} \\
&= \frac{1}{\lambda_L} \sigma_L^{-1/2} \lambda_L \sigma_L\, \sigma_L^{-1/2} \\
&= \sigma_L^{-1/2} \sigma_L \sigma_L^{-1/2} \\
&= \mathrm{Id}.
\end{aligned}$$

Note that $\hat{L}(v)^\dagger = \frac{1}{\sqrt{\lambda_L}} \sigma_L^{-1/2} L(v)^\dagger \sigma_L^{1/2}$. From this we get easily that $\phi_{\hat{L}}$ is completely positive and preserves trace (is stochastic).

We will show that $\phi_{\hat{L}}$ is irreducible. Given $A \in M_k$ we have

$$\phi_{\hat{L}}(A) = \frac{1}{\lambda_L} \sigma_L^{1/2} \phi_L(\sigma_L^{-1/2} A \sigma_L^{-1/2}) \sigma_L^{1/2}.$$

Then,

$$\begin{aligned}\phi_{\hat{L}}^2(A) &= \frac{1}{\lambda_L}\sigma_L^{1/2}\phi_L(\sigma_L^{-1/2}\frac{1}{\lambda_L}\sigma_L^{1/2}\phi_L(\sigma_L^{-1/2}A\sigma_L^{-1/2})\sigma_L^{1/2}\sigma_L^{-1/2})\sigma_L^{1/2}\\ &= \frac{1}{\lambda_L^2}\sigma_L^{1/2}\phi_L^2(\sigma_L^{-1/2}A\sigma_L^{-1/2})\sigma_L^{1/2}.\end{aligned}$$

By induction we get

$$\phi_{\hat{L}}^n(A) = \frac{1}{\lambda_L^n}\sigma_L^{1/2}\phi_L^n(\sigma_L^{-1/2}A\sigma_L^{-1/2})\sigma_L^{1/2}.$$

Given $A, B \geqslant 0$, note that $\sigma_L^{-1/2}A\sigma_L^{-1/2} \geqslant 0$ and $\sigma_L^{1/2}B\sigma_L^{1/2} \geqslant 0$. Therefore, using irreducibility of ϕ_L, there exists an integer $n \in \{1, \ldots, k-1\}$, such that,

$$\begin{aligned}0 &< \lambda_L^{-n}\mathrm{tr}\,[\sigma_L^{1/2}B\sigma_L^{1/2}\phi_L^n(\sigma_L^{-1/2}A\sigma_L^{-1/2})]\\ &= \lambda_L^{-n}\mathrm{tr}\,[B\sigma_L^{1/2}\phi_L^n(\sigma_L^{-1/2}A\sigma_L^{-1/2})\sigma_L^{1/2}]\\ &= \mathrm{tr}\,[B\phi_{\hat{L}}^n(A)].\end{aligned}$$

Therefore, $\phi_{\hat{L}}$ is irreducible and completely positive and preserves trace. In this way, to the given L we can associate $\hat{L}$ which will be called the normalization of L. The transformation $\phi_{\hat{L}}$ is a quantum channel.

Definition 3.3.8. Given the measure μ over M_k we denote by $\mathfrak{L}(\mu)$ the set of all integrable L such that the associated ϕ_L is irreducible.

Definition 3.3.9. Suppose L is in $\mathfrak{L}(\mu)$. We say that L is normalized if ϕ_L has spectral radius 1 and preserves trace. We denote by $\mathfrak{N}(\mu)$ the set of all normalized L.

Note that the transformation $\hat{L}$ defined above in (3.1) is normalized.

If $L \in \mathfrak{N}(\mu)$, then, we get from Theorem 3.3.5 and the fact that $\phi_L^*(\mathrm{Id}) = \mathrm{Id}$, that the spectral radius, which is also a simple eigenvalue, is $\lambda_L = 1$. According to Remark 3.3.6, there is a unique eigendensity ρ_L such that $\phi_L(\rho_L) = \rho_L$. These properties will be important for what will come next.

Theorem 3.3.10 (Ergodicity and temporal means)**.** *Suppose $L \in \mathfrak{N}(\mu)$. Then, for all density matrix $\rho \in M_k$ it is true that*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\phi_L^n(\rho) = \rho_L,$$

where ρ_L is the density matrix associated to L.

Proof. The proof follows from Theorem 3.3.5 and Corollary 6.3 in [Wol12]. ■

The above result connects irreducibility and ergodicity (the temporal means have a unique limit).

3.4 Entropy

In this section, we will define entropy for ϕ_L, when the associated L is irreducible and stochastic (see Definition 3.4.2). After that, it will be possible to give a meaning for a certain variational principle of pressure in Definition 3.4.3 (this is similar to the setting in Thermodynamic Formalism which is described in [PP90], for instance).

Remember the classical entropy is defined just for invariant (stationary) probabilities. Something of this sort is required for defining the entropy of a quantum channel ϕ_L: L has to be stochastic. These ϕ_L will play in some sense the role of the different possible invariant probabilities.

We will explore some ideas which were already present on the paper [BLLC10] (which explores some previous nice results on [ŁŻS03] and [Sło03]) which considers a certain *a priori* probability.

Hereafter, we consider **fixed a measure** μ over M_k which plays the role of the *a priori* probability. Given $L \in \mathfrak{L}(\mu)$ we will associate in a natural way the transformation $\phi_L : M_k \to M_k$.

Definition 3.4.1. We denote by $\phi = \phi_\mu$ the set of all L such that the associated $\phi_L : M_k \to M_k$ is irreducible and stochastic.

We will describe a discrete-time process that takes values on M_k.

Suppose L is irreducible and stochastic. We will associate to such L a kind of "transition probability kernel" P_L (to be defined soon) acting on matrices. Given the matrices v and w the value $P_L(v, w)$ will describe the probability of going in the next step to w if the process is on v.

Given L, suppose that the discrete-time process is given in such a way that the initial state is described by the density matrix ρ_L which is invariant for ϕ_L (see Theorem 3.3.5).

The reasoning here is that such process should be in "some sense stationary" because ρ_L is invariant by ϕ_L. As we said before in ergodic theory the concept of Shanon-Kolmogorov entropy has a meaning just for invariant (for a discrete-time dynamical system) probabilities. Therefore, something of this order is required.

In our reasoning given that the state is described by ρ, then, in the next step of the process we get $\frac{L(v)\rho L(v)^\dagger}{\operatorname{tr}(L(v)\rho L(v)^\dagger)}$ with probability $\operatorname{tr}(L(v)\rho L(v)^\dagger)\; d\mu(v)$.

This discrete-time process takes values on density operators in M_k.

Definition 3.4.2. We define entropy for L (or, for ϕ_L) by the expression (when finite):

$$h(L) = h_\mu(L) := -\int_{M_k \times M_k} \operatorname{tr}(L(v)\rho_L L(v)^\dagger) P_L(v,w) \log P_L(v,w)\;\; d\mu(v)\, d\mu(w),$$

where

$$P_L(v,w) := \frac{\operatorname{tr}(L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger)}{\operatorname{tr}(L(v)\rho_L L(v)^\dagger)}.$$

This definition is a generalization of the analogous concept presented on the papers [BLLC10], [BLLC11a] and [BLLC11b].

Note that $\operatorname{tr}(L(v)\rho_L L(v)^\dagger)$ is the probability of being in state $\frac{L(v)\rho_L L(v)^\dagger}{\operatorname{tr}(L(v)\rho_L L(v)^\dagger)}$. Moreover, $P_L(v,w)$ describes the probability of going from v to w, being in state

$$\frac{L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger}{\operatorname{tr}(L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger)}.$$

In this way $h_\mu(L)$ in some way resembles the analogous expression of entropy for the case of Markov chains.

We will show in Example 3.8.5 that the above definition of entropy is indeed a natural generalization of the classical one in Ergodic Theory.

Suppose $H : M_k \to M_k$ is square integrable, irreducible and $H(v) \neq 0$, for μ-*a.e.* v. For such H, consider the corresponding ρ_H, σ_H and λ_H which are given by Theorem 3.3.5, where $\operatorname{tr}\rho_H = 1$ and $\operatorname{tr}\sigma_H\rho_H = 1$.

This H describes the action of a potential.

Then, we define

$$U_H(v) := \log\left(\operatorname{tr}(\sigma_H H(v)\rho_H H(v)^\dagger)\right).$$

Definition 3.4.3. We define the pressure of H by

$$P_\mu(H) = P(H) := \sup_{L\in\phi}\left\{h_\mu(L) + \int U_H(v)\; \operatorname{tr}(L(v)\rho_L L(v)^\dagger)\;\; d\mu(v)\right\}.$$

Remember that ϕ_μ is the set of all $L : M_k \to M_k$ which are square-integrable, irreducible, and stochastic.

Definition 3.4.4. Given μ and H as above we say that ϕ_L, for some $L \in \phi_\mu$, is a Gibbs channel, if

$$P_\mu(H) = h_\mu(L) + \int U_H(v)\ \mathrm{tr}\,(L(v)\rho_L L(v)^\dagger)\, d\mu(v).$$

We will need soon the following well-known result (see [PP90]).

Proposition 3.4.5. *Suppose $p, q : M_k \to \mathbb{R}_+$ are such that $p, q > 0$, μ-almost everywhere, $\int_{M_k} p\ \ d\mu = 1$ and $\int_{M_k} q\ \ d\mu = 1$. Then,*

$$-\int p \log p\ \ d\mu + \int p \log q\ \ d\mu \leqslant 0.$$

Moreover, the above inequality is an equality just when $p = q$, μ-almost everywhere.

Theorem 3.4.6. *Assume that $H : M_k \to M_k$ is continuous, irreducible and $H(v) \neq 0$ for μ-a.e. v, then,*

$$P(H) := \sup_{L \in \phi} \left\{ h_\mu(L) + \int U_H(v)\ \mathrm{tr}\,(L(v)\rho_L L(v)^\dagger)\ \ d\mu(v) \right\} \leqslant \log(\lambda_H),$$

The supremum is attained only if

$$\frac{\mathrm{tr}\,(L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger)}{\mathrm{tr}\,(L(v)\rho_L L(v)^\dagger)} = \frac{1}{\lambda_H}\mathrm{tr}\,(\sigma_H H(w)\rho_H H(w)^\dagger), \quad \textit{for}\, \mu\textit{-a.e.}\, v, w.$$

In this case, $P(H) = \log(\lambda_H)$.

Proof. We define $q(w) := \frac{1}{\lambda_H}\mathrm{tr}\,(\sigma_H H(w)\rho_H H(w)^\dagger)$. Note that

$$\begin{aligned}
\int q\ \ d\mu &= \frac{1}{\lambda_H}\int \mathrm{tr}\,(\sigma_H H(w)\rho_H H(w)^\dagger)\ \ d\mu(w) \\
&= \frac{1}{\lambda_H}\mathrm{tr}\,\left(\sigma_H \int H(w)\rho_H H(v)^\dagger\ d\mu(w)\right) \\
&= \frac{1}{\lambda_H}\mathrm{tr}\,(\sigma_H \lambda_H \rho_H) \\
&= \mathrm{tr}\,(\sigma_H \rho_H) = 1.
\end{aligned}$$

For fixed v and irreducible and stochastic L take

$$p_v(w) = P_L(v, w) = \frac{\mathrm{tr}\,(L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger)}{\mathrm{tr}\,(L(v)\rho_L L(v)^\dagger)},$$

If $\operatorname{tr}(L(v)\rho_L L(v)^\dagger) \neq 0$ and $p_v(w) = 0$ otherwise. It follows that

$$\int p_v(w)\ d\mu(w) = \int \frac{\operatorname{tr}(L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger)}{\operatorname{tr}(L(v)\rho_L L(v)^\dagger)}\ d\mu(w)$$
$$= \frac{\operatorname{tr}(L(v)\rho_L L(v)^\dagger \int L(w)^\dagger L(w))\ d\mu(w)}{\operatorname{tr}(L(v)\rho_L L(v)^\dagger)} = 1.$$

From Proposition 3.4.5 we get that for each v

$$-\int p_v(w)\log(p_v(w))\ d\mu(w) + \int p_v(w)\log(q(w))\ d\mu(w) \leqslant 0. \tag{3.2}$$

Equality will happen when

$$\frac{\operatorname{tr}(L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger)}{\operatorname{tr}(L(v)\rho_L L(v)^\dagger)} = \frac{1}{\lambda_H}\operatorname{tr}(\sigma_H H(w)\rho_H H(w)^\dagger),$$

for μ-almost everywhere w.
Note that from (3.2) it follows that

$$\int -P_L(v,w)\log P_L(v,w) + P_L(v,w)\log\big(\operatorname{tr}(\sigma_H H(w)\rho_H H(w)^\dagger)\big)\ d\mu(w)$$
$$\leqslant \int P_L(v,w)\log(\lambda_H)\ d\mu(w) = \log(\lambda_H).$$

Now we multiply both sides of the above inequality by $\operatorname{tr}(L(v)\rho_L L(v)^\dagger)$, integrate with respect to v (remember that $\int \operatorname{tr}(L(v))\rho_L L(v)^\dagger) = 1$) and we get

$$\begin{aligned}
&h_\mu(L) + \\
&\quad\int \operatorname{tr}(L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger)\log\big(\operatorname{tr}(\sigma_H H(w)\rho_H H(w)^\dagger)\big)\ \ d\mu(w)\, d\mu(v)\\
&\quad= h_\mu(L) + \int \operatorname{tr}(L(w)\phi_L(\rho_L)L(w)^\dagger)\log\big(\operatorname{tr}(\sigma_H H(w)\rho_H H(w)^\dagger)\big)\ \ d\mu(w)\\
&\quad= h_\mu(L) + \int \log\big(\operatorname{tr}(\sigma_H H(v)\rho_H H(v)^\dagger)\big)\ \operatorname{tr}(L(v)\rho_L L(v)^\dagger)\ \ d\mu(v)\\
&\quad\leqslant \log(\lambda_H).
\end{aligned}$$

As this is true for any $L \in \phi$, we take the sup over all such L to finally get:

$$P(H) \leqslant \log(\lambda_H).$$

■

A natural question: is there a $L \in \phi$ such that the supremum is attained? This kind of result would correspond in our setting to the Ruelle Theorem of Thermodynamic Formalism (see [PP90]). In this direction, we are able to get Theorem 3.4.8.

Before trying to address this question we point out that given H as above one can get the associated normalized $\hat{H}$ by the expression $\hat{H} = \frac{1}{\sqrt{\lambda_H}}\sigma_H^{1/2} H \sigma_H^{-1/2}$.

Note that $\sigma_{\hat{H}} = \mathrm{Id}$, $\rho_{\hat{H}} = \sigma_H^{1/2}\rho_H\sigma_H^{1/2}$ and $\lambda_{\hat{H}} = 1$. Therefore,

$$\begin{aligned}
&\int \log\left(\mathrm{tr}\,(\sigma_{\hat{H}}\hat{H}(v)\rho_{\hat{H}}\hat{H}(v)^\dagger)\right)\,\mathrm{tr}\,(L(v)\rho_L L(v)^\dagger)\;d\mu(v)\\
&= \int \log\left(\mathrm{tr}\,(\frac{1}{\lambda_H}\sigma_H^{1/2}H(v)\sigma_H^{-1/2}\sigma_H^{1/2}\rho_H\sigma_H^{1/2}\sigma_H^{-1/2}H(v)^\dagger\sigma_H^{1/2})\right)\,\mathrm{tr}\,(L(v)\rho_L L(v)^\dagger)\;d\mu(v)\\
&= \int \log\left(\mathrm{tr}\,(\frac{1}{\lambda_H}\sigma_H H(v)\rho_H H(v)^\dagger)\right)\,\mathrm{tr}\,(L(v)\rho_L L(v)^\dagger)\;d\mu(v)\\
&= \int \log\left(\mathrm{tr}\,(\sigma_H H(v)\rho_H H(v)^\dagger)\right)\,\mathrm{tr}\,(L(v)\rho_L L(v)^\dagger)\;d\mu(v) - \log(\lambda_H).
\end{aligned}$$

From the above reasoning we get:

Theorem 3.4.7. *Assume that $H : M_k \to M_k$ is irreducible, square integrable and $H(v) \neq 0$, for μ-a.e. v. If $\hat{H}$ denotes the associated normalization, then,*

$$P(\hat{H}) = P(H) - \log(\lambda_H).$$

Note that $\hat{H} \in \phi_\mu$.

Theorem 3.4.8. *If H is irreducible, square integrable and $H(v) \neq 0$, for μ-a.e. v, then,*

$$P(H) = \log \lambda_H.$$

Proof. We already know that $P(H) \leqslant \log \lambda_H$. We will show that there exists an irreducible and stochastic L which attains the supremum. In order to do that we take an orthonormal basis $\{|\, i\,\rangle\}_{i=1,2,\ldots,k}$ of $\mathbb{C}^k$. Then, we define an operator P such that $P|\, i+1\,\rangle = |\, i\,\rangle$ (for instance, $P = \sum_{i=1}^k |\, i\,\rangle\,\langle\, i+1\,|$ and by convention $|\, 1\,\rangle = |\, k+1\,\rangle$).

Note that the dual of P is $P^\dagger = \sum_i | \, i+1 \, \rangle\langle \, i \, |$. This is so because given $u, v \in \mathbb{C}^k$, we get that

$$\langle \, u \, , \, Pv \, \rangle = \sum_i \langle \, u \, , \, | \, i \, \rangle\langle \, i+1 \, | v \, \rangle = \sum_i \langle \, | \, i+1 \, \rangle\langle \, i \, | u \, , \, v \, \rangle = \langle \, P^\dagger u \, , \, v \, \rangle.$$

Moreover, $P^\dagger P = \mathrm{Id}$. Indeed,

$$\sum_{i,j} | \, j+1 \, \rangle\langle \, j \, || \, i \, \rangle\langle \, i+1 \, | = \sum_i | \, i \, \rangle\langle \, i \, | = \mathrm{Id}.$$

Now, take $Q = (q_{ij})$ the matrix with $q_{kk} = -1$, $q_{ii} = 1$, for $i = 1, ..., k-1$, and $q_{ij} = 0$ otherwise. Note that $Q^\dagger Q = \mathrm{Id}$.

Consider $\rho_H, \sigma_H, \lambda_H$ given by Theorem 3.3.5, where $\mathrm{tr}\,(\sigma_H \rho_H) = 1$ and $\mathrm{tr}\,(\rho_H) = 1$ and let $\varphi(v) = \sqrt{\frac{1}{\lambda_H} \mathrm{tr}\,(\sigma_H H(v) \rho_H H(v)^\dagger)}$.

Note that if $\# \operatorname{supp} \mu = 1$, H can't be irreducible because any eingenvector of $H(v)$ for $v \in \operatorname{supp} \mu$ generates an invariant subspace.

There exist $v_1, v_2 \in \operatorname{supp} \mu$ with $\varphi(v_i) \neq 0$ by hypothesis. Take O an open set with $v_1 \in O$ and $d(\overline{O}, v_2) > 0$. Now we can define L by $L(v) = \varphi(v) P$, for $v \notin O$, and $L(v) = \varphi(v) Q$, for $v \in O$.

Observe that $L(v)^\dagger L(v) = |\varphi(v)|^2 \, \mathrm{Id}$, for all v, and $\int |\varphi(v)|^2 \; d\mu(v) = 1$. This implies that $\phi_L^*(\mathrm{Id}) = \mathrm{Id}$.

Suppose that E is an invariant subspace of $\mathbb{C}^k$ for all $L(v)$ with $v \in \operatorname{supp} \mu$. Of course, as $\varphi(v_i) \neq 0$, E is invariant for P and Q. In this sense, taking $x = (x_1, \ldots, x_k) \in E$, we get $Qx = (x_1, \ldots, -x_k) \in E$. As E is a linear subspace this implies that $(x_1, \ldots, x_{k-1}, 0) \in E$, and $(0, \ldots, 0, x_k) \in E$. Taking $P^n(0, \ldots, 0, x_k)$, for $n = 0, \ldots, k-1$, if $x_k \neq 0$, we get a base of $\mathbb{C}^k$ in E. Therefore, if $x_k \neq 0$, we have $E = \mathbb{C}^k$. On the other hand, if initially $x_k = 0$, we take $P^n x$, where $(P^n x)_k \neq 0$, and we use the previous argument. If there is no $x \in E$ and n such that $(P^n x)_k \neq 0$, then $E = \{0\}$. Therefore, ϕ_L is irreducible by Lemma 3.3.3.

To show that L satisfy the supremum for pressure, from the inequality give by Theorem 3.4.5, it is enough to show that

$$\frac{\mathrm{tr}\,(L(w) L(v) \rho_L L(v)^\dagger L(w)^\dagger)}{\mathrm{tr}\,(L(v) \rho_L L(v)^\dagger)} = \frac{1}{\lambda_H} \mathrm{tr}\,(\sigma_H H(w) \rho_H H(w)^\dagger).$$

In order to get this, observe that

$$\begin{aligned}
&\operatorname{tr}(L(w)L(v)\rho_L L(v)^\dagger L(w)^\dagger) \\
&\quad = \operatorname{tr}(L(v)\rho_L L(v)^\dagger L(w)^\dagger L(w)) \\
&\quad = \operatorname{tr}(L(v)\rho_L L(v)^\dagger)\, |\varphi(w)|^2 .
\end{aligned}$$

Thus, the required equation holds. ■

3.5 Process X_n, $n \in \mathbb{N}$, taking values on $P(\mathbb{C}^k)$

Consider a fixed measure μ on M_k and a fixed $L : M_k \to M_k$, such that, $\int_{M_k} \|L(v)\|^2 \; d\mu(v) < \infty$, and, also that ϕ_L is irreducible and stochastic.

Note that if, for example, μ is a probability and the the function $v \to \|L(v)\|$ is bounded we get that $\int_{M_k} \|L(v)\|^2 \; d\mu(v) < \infty$.

Denote by $P(\mathbb{C}^k)$ the projective space on $\mathbb{C}^k$ with the metric $d(\hat{x}, \hat{y}) = (1 - |\langle x\,,\, y\rangle|^2)^{1/2}$, where x, y are representatives with norm 1 and $\langle \cdot\,,\,\cdot \rangle$ is the canonical inner product.

We choose representatives and from now on for generic $\hat{x}, \hat{y}$ the associated ones are denoted by x, y. We assume "continuity" on these choices.

Take $\hat{x} \in P(\mathbb{C}^k)$ and $S \subset P(\mathbb{C}^k)$. For a stochastic ϕ_L we consider the kernel

$$\Pi_L(\hat{x}, S) = \int_{M_k} \mathbf{1}_S(L(v) \cdot \hat{x})\, \|L(v)x\|^2 \; d\mu(v), \tag{3.3}$$

where the norm above is the euclidean one.

Above $L(v) \cdot \hat{x}$ denotes the projectivized element in $P(\mathbb{C}^k)$.

As ϕ_L is stochastic we get that $\Pi_L(\hat{x}, P(\mathbb{C}^k)) = 1$. $\Pi_L(\hat{x}, S)$ describes the probability of getting in the next step a state in S, if the system is presently at the state $\hat{x}$.

Remember that $\operatorname{tr}(L(v)\pi_{\hat{x}} L(v)^\dagger) = \|L(v)x\|^2$, where $\pi_{\hat{x}} = |\, x \rangle\langle x \,|$ and x are representatives of norm 1 in the class of $\hat{x}$.

This discrete-time process (described by the kernel) taking values on $P(\mathbb{C}^k)$ is determined by μ and L. If ν is a probability on the Borel σ-algebra $\mathcal{B}$ of $P(\mathbb{C}^k)$ define

$$
\begin{aligned}
\nu\Pi_L(S) &= \int_{P(\mathbb{C}^k)} \Pi_L(\hat{x}, S)\ d\nu(\hat{x}) \\
&= \int_{P(\mathbb{C}^k)\times M_k} \mathbf{1}_S(L(v)\cdot\hat{x})\,\|L(v)x\|^2\ d\nu(\hat{x})\,d\mu(v).
\end{aligned}
$$

$\nu\Pi_L$ is a new probability on $P(\mathbb{C}^k)$ and Π_L is a Markov operator. The above definition of $\nu \to \nu\Pi_L$ is a simple generalization of the one in [BFPP19], where the authors take the L considered here as the identity transformation.

The map $\nu \to \nu\,\Pi_L$ (acting on probabilities ν) is called the Markov operator obtained from ϕ_L in the paper [ŁŻS03]. There the *a priori* measure μ is a sum of Dirac probabilities. Here we consider a more general setting.

Definition 3.5.1. We say that the probability ν over $P(\mathbb{C}^k)$ is invariant for Π_L, if $\nu\Pi_L = \nu$.

The natural question is: does exist such invariant probability for Π_L?

About the question of existence, we are going to prove that the kernel defined above is a continuous Markov operator (in the weak-star topology). So, leaving the compact set of probabilities over $P(\mathbb{C}^k)$ invariant, by the Markov-Kakutani theorem there exists a fixed point, which means that there exists an invariant probability. In order to do that we only need to find a linear operator $U : C_0(P(\mathbb{C}^k), \mathbb{C}) \to C_0(P(\mathbb{C}^k), \mathbb{C})$ such that $\langle Uf, \nu\rangle = \langle f, \nu\Pi_L\rangle$. Here, $C_0(P(\mathbb{C}^k), \mathbb{C})$ stands for continuous functions from $P(\mathbb{C}^k)$ to $\mathbb{C}$ with the C_0 norm which we denote by $\|\cdot\|_\infty$. When such U exists we say that the Markov operator Π_L is Feller.

According to Proposition 2.10 in [Sło03] if such U exists, then, Π_L is continuous in weak-star topology and by Markov-Kakutani theorem, there is a fixed probability in $P(\mathbb{C}^k)$.

In Example 3.8.5 we calculate the explicit expression of the invariant probability ν.

Theorem 3.5.2. *Suppose that L is such that $\int_{M_k} \|L(v)\|^2\ d\mu(v) < \infty$. Then, there exists at least one invariant probability ν for the Markov operator Π_L.*

Proof. Define $U : C_0(P(\mathbb{C}^k), \mathbb{C}) \to C_0(P(\mathbb{C}^k), \mathbb{C})$ by

$$
Uf(\hat{x}) = \int_{M_k} f(L(v)\cdot\hat{x})\,\|L(v)x\|^2\ d\mu(v).
$$

Notice that

$$\begin{aligned}\langle Uf, \nu\rangle &= \int_{P(\mathbb{C}^k)} Uf(\hat{x})\, d\nu(\hat{x}) \\ &= \int_{P(\mathbb{C}^k)\times M_k} f(L(v)\cdot \hat{x})\, \|L(v)x\|^2\,,\, d\mu(v)d\nu(\hat{x}) \\ &= \int_{P(\mathbb{C}^k)} f(\hat{x})\, d(\nu\Pi_L)(\hat{x}) = \langle \nu\Pi_L\rangle.\end{aligned}$$

Therefore, $\langle Uf, \nu\rangle = \langle f, \nu\Pi_L\rangle$.

Then, we only need to prove that Uf is a continuous function of $P(\mathbb{C}^k)$.

Consider a sequence $(\hat{x_n}) \in P(\mathbb{C}^k)$, such that, $\hat{x_n} \longrightarrow \hat{x} \in P(\mathbb{C}^k)$. We are going to show that $Uf(\hat{x_n}) \longrightarrow Uf(\hat{x})$. Define $F, F_n : M_k \to \mathbb{C}$ by

$$F_n(v) = f(L(v)\cdot \hat{x_n})\, \|L(v)x_n\|^2$$

and

$$F(v) = f(L(v)\cdot \hat{x})\, \|L(v)x\|^2$$

This way, $Uf(\hat{x_n}) = \int F_n(v)\, d\mu(v)$ and $Uf(\hat{x}) = \int F(v)\, d\mu(v)$. Since the function f and the norm are continuous, we have $F_n(v) \longrightarrow F(v)$, for all $v \in \mathcal{M}_k$.

Also,

$$|F_n(v)| = |f(L(v)\cdot \hat{x_n})| \cdot \|L(v)x_n\|^2 \leqslant \|f\|_\infty\ \mathrm{tr}\,(L(v)|\, x_n \rangle\langle\, x_n\,|L(v)^\dagger)$$

$$= \|f\|_\infty\ \mathrm{tr}\,(|\, x_n \rangle\langle\, x_n\,|L(v)L(v)^\dagger) \leqslant \|f\|_\infty\ \mathrm{tr}\,(L(v)L(v)^\dagger) = \|f\|_\infty\, \|L(v)\|^2\,.$$

As $\int \|L(v)\|^2\ d\mu(v) < \infty$, we can apply Lebesgue Dominated Convergence Theorem to conclude that

$$Uf(\hat{x_n}) = \int F_n(v)\, d\mu(v) \longrightarrow \int F(v)\, d\mu(v) = Uf(\hat{x}).$$

So we have that Uf is continuous and this is the end of the proof. ■

3.6 Process ρ_n, $n \in \mathbb{N}$, taking values on $\mathcal{D}_k$

For a fixed μ over M_k and L such that ϕ_L is irreducible and stochastic, one can naturally define a process (ρ_n) on $\mathcal{D}_k = \{\rho \in M_k : \mathrm{tr}\,\rho = 1 \text{ and } \rho \geqslant 0\}$ which is called *quantum trajectory* by T. Benoist, M. Fraas, Y. Pautrat, and C. Pellegrini in [BFPP19]. Given a ρ_0 initial state, we get

$$\rho_n = \frac{L(v)\rho_{n-1}L(v)^\dagger}{\mathrm{tr}\,(L(v)\rho_{n-1}L(v)^\dagger)}$$

with probability

$$\operatorname{tr}\left(L(v)\rho_{n-1}L(v)^{\dagger}\right) d\mu(v), \quad n \in \mathbb{N}.$$

This process has similarities with the previous one in $P(\mathbb{C}^k)$ and we explore some relations between them. In this section, we follow closely the notation of [BFPP19].

We want to relate the invariant probabilities of the last section with the fixed point $\rho_{inv} = \rho_{inv}^L$ of ϕ_L.

First, denote $\Omega := M_k^{\mathbb{N}}$, and for $\omega = (\omega_i)_{i\in\mathbb{N}}$, take $\pi_n(\omega) = (\omega_1, \ldots, \omega_n)$. Recall that $\mathcal{M}$ is the Borel sigma-algebra on M_k. For all, $n \in \mathbb{N}$, consider $\mathcal{O}_n$ the sigma algebra on Ω generated by the cylinder sets of size n, that is, $\mathcal{O}_n := \pi_n^{-1}(\mathcal{M}^n)$. We equip Ω with the smallest sigma algebra $\mathcal{O}$ which contains all $\mathcal{O}_n$, $n \in \mathbb{N}$.

Denote $\mathcal{J}_n := \mathcal{B} \times \mathcal{O}_n$ and $\mathcal{J} := \mathcal{B} \times \mathcal{O}$. In this way, $(P(\mathbb{C}^k) \times \Omega, \mathcal{J})$ is an measurable space. By abuse of language we consider $V_i : \Omega \to M_k$ as a random variable $V_i(\omega) = \omega_i$. We also introduce another random variable $W_n := L(V_n) \cdots L(V_1)$, where $W_n(\omega) = L(\omega_n) \cdots L(\omega_1)$.

For a given a probability ν on $P(\mathbb{C}^k)$, we define for $S \in \mathcal{B}$ and $O_n \in \mathcal{O}_n$ another probability on $P(\mathbb{C}^k \times \Omega)$ by

$$\mathbb{P}_{\nu,n}(S \times O_n) := \int_{S\times O_n} \|W_n(\omega)x\|^2 \; d\nu(\hat{x})\, d\mu^{\otimes n}(\omega). \tag{3.4}$$

Remark 3.6.1. We can extend the above probability $\mathbb{P}_\nu$ over $\mathcal{B} \times \mathcal{O}$. We claim that $\mathbb{P}_{\nu,n}$, $n \in \mathbb{N}$, is a consistent family over the cylinders of size n (then, we can use the Caratheodory-Kolmogorov extension theorem).

Indeed, note that $W_{n+1}(\omega) = L_{n+1}(\omega)W_n(\omega)$. Then

$$\begin{aligned}
\mathbb{P}_{\nu,n+1}(S \times O_n \times M_k) &= \int_{S\times O_n\times M_k} \|W_{n+1}(\omega)x\|^2 \; d\nu(\hat{x})\, d\mu^{\otimes n+1}(\omega) \\
&= \int_{S\times O_n\times M_k} \operatorname{tr}\left(L(\omega_{n+1})W_n(\omega)\pi_{\hat{x}}W_n(\omega)^{\dagger}L(\omega_{n+1})^{\dagger}\right) d\nu(\hat{x})\, d\mu^{\otimes n+1}(\omega) \\
&= \int_{S\times O_n} \operatorname{tr}\left(W_n(\omega)\pi_{\hat{x}}W_n(\omega)^{\dagger}\int_{M_k} L(\omega_{n+1})^{\dagger}L(\omega_{n+1})\, d\mu(\omega_{n+1})\right) d\nu(\hat{x})\, d\mu^{\otimes n}(\omega) \\
&= \int_{S\times O_n} \|W_n(\omega)x\|^2 \; d\nu(\hat{x})\, d\mu^{\otimes n}(\omega) \\
&= \mathbb{P}_{\nu,n}(S \times O_n).
\end{aligned}$$

Since the set $\{W_n x = 0\}$ leads to a null integrating term in (3.4), we have $\mathbb{P}_\nu(W_n x = 0) = 0$. Therefore, we define the expression for each n and then extend it. In this way $W_n(\omega)x \neq 0$. Remember that $W_n(\omega) \cdot \hat{x}$ is the representative of the class $W_n(\omega)x$, when $W_n(\omega)x \neq 0$.

Denote $\mathbb{E}_\nu$ the expected value with respect to $\mathbb{P}_\nu$. Now observe that for a ν probability on $P(\mathbb{C}^k)$, if π_{X_0} is an orthogonal projection on subspace generated by X_0 on $\mathbb{C}^k$, we have

$$\rho_\nu := \mathbb{E}_\nu(\pi_{X_0}) = \int_{P(\mathbb{C}^k)} \pi_{x_0}\ d\nu(x_0).$$

We call ρ_ν barycenter of ν, and it is easy to see that $\rho_\nu \in \mathcal{D}_k$.

Note that for each $\rho \in \mathcal{D}_k$, exists (v_n) an orthonormal basis of eigenvectors with eigenvalues a_i such that $\rho = \sum_i a_i \pi_{v_i}$. Therefore, exists $\nu = \sum a_i \delta_{v_i}$ such that $\rho_\nu = \rho$.

We collect the above results in the next proposition (which was previously stated as Proposition 2.1 in [BFPP19] for the case $L = I$).

Proposition 3.6.2. *If ν is invariant for Π_L, then*

$$\rho_\nu = \mathbb{E}_\nu(\pi_{\hat{X}_0}) = \mathbb{E}_\nu(\pi_{\hat{X}_1}) = \phi_L(\rho_\nu).$$

Therefore, for an irreducible L, every invariant measure ν for Π_L has the same barycenter.

We point out that in this way we can recover ρ_{inv}, the fixed point of ϕ_L, by taking the barycenter of any invariant probability (the quantum channel ϕ_L admits only one fixed point). That is, for any invariant probability ν for Π_L, we get that $\rho_\nu = \rho_{inv}$.

Note that the previous process can be seen as $\rho_n : \Omega \to \mathcal{D}_k$, such that, $\rho_0(\hat{x}, \omega) = \rho_\nu$ and, and $n \in \mathbb{N}$

$$\rho_n(\omega) = \frac{W_n(\omega)\rho_0 W_n(\omega)^\dagger}{\operatorname{tr}(W_n(\omega)\rho_0 W_n(\omega)^\dagger)}.$$

Using an invariant ρ we can define a Stationary Stochastic Process taking values on M_k. That is, we will define a probability $\mathbb{P}$ over $\Omega = (M_k)^{\mathbb{N}}$.

Take $O_n \in \mathcal{O}_n$ and define

$$\mathbb{P}^\rho(O_n) = \int_{O_n} \operatorname{tr}(W_n(\omega)\rho W_n(\omega)^\dagger)\ d\mu^{\otimes n}(\omega).$$

The probability $\mathbb{P}$ on Ω defines a Stationary Stochastic Process.

3.7 ϕ-Erg and irreducible is Generic

Definition 3.7.1. Given $L : M_k \to M_k$, μ on M_k and E subspace of $\mathbb{C}^k$, we say that E is (L, μ)-invariant, if $L(v)E \subset E$, for all $v \in \operatorname{supp} \mu$.

Definition 3.7.2. Given $L : M_k \to M_k$, μ on M_k, we say that L is ϕ-Erg for μ, if there exists an unique minimal non-trivial space E, such that, E is (L, μ)-invariant.

In the case the space E is equal to $\mathbb{C}^k$, as shown in Lemma 3.3.3, we have L irreducible for μ (or μ-irreducible) in the sense of Definition 3.3.2.

Consider $\mathcal{B}(M_k) = \{L : M_k \to M_k \,|\, L \text{ is continuous and bounded}\}$ where $\|L\| = \sup_{v \in M_k} \|L(v)\|$. We write $\mathcal{B} = \mathcal{B}(M_k)$ when k is implicit.

Proposition 3.7.3. *Given $L \in \mathcal{B}(M_k)$, μ over M_k, $v_1 \in$ supp μ and $\varepsilon > 0$, there exists $L_\varepsilon \in \mathcal{B}(M_k)$ such that $\|L - L_\varepsilon\| < \frac{\varepsilon}{2}$ and $L_\varepsilon(v_1)$ has k distinct eigenvalues.*

Proof. Take $v_1 \in \operatorname{supp} \mu$. Denote by J the Jordan canonical form for the complex matrix $L(v_1)$ and take B such that $L(v_1) = B^{-1}JB$. Define $D_n = (d_{i,j})_{i,j} \in M_k$, where

$$d_{i,j} = \begin{cases} 1 & \text{if } i = n \text{ and } j = n \\ 0 & \text{otherwise.} \end{cases}$$

Now, we look for each diagonal element of J. If the first, i.e., the element $(1, 1)$ is zero, we sum $\frac{\delta}{4} D_1$. If the second element is not different from the first or is not different of zero, then, we sum $\frac{\delta}{2^i} D_2$, where $i > 2$ is chosen to satisfy both. We repeat this process until all the elements of diagonal are considered. After that, we get that all diagonal elements of $J + \sum_j \frac{\delta}{2^{i_j}} D_j$ are different and none is zero. Moreover, $\left\| \sum_j \frac{\delta}{2^{i_j}} D_j \right\| \leqslant \sum_j \frac{\delta}{2^{i_j}} \leqslant \frac{\delta}{2}$.

We define $D^\delta = \sum_j \frac{\delta}{2^{i_j}} D_j$ and $L_\varepsilon = L + B^{-1} D^\delta B$. Therefore, $\|L_\varepsilon - L\| = \left\| B^{-1} D^\delta B \right\| \leqslant \frac{\delta}{2} \left\| B^{-1} \right\| \|B\|$. Choosing $\delta < \frac{\varepsilon}{\|B^{-1}\| \|B\|}$ we get

$$\|L_\varepsilon - L\| < \frac{\varepsilon}{2}.$$

Therefore, as $J + D^\delta$ has the same eigenvalues of $L_\varepsilon(v_1)$, we finished the proof. ■

Lemma 3.7.4. *Consider eigenvectors $v_i \in \mathbb{C}^k, 1 \leqslant i \leqslant n$ of a linear transformation A with respective eigenvalues λ_i, where $\lambda_i \neq \lambda_j$, for $i \neq j$. If a subspace $F \subseteq \mathbb{C}^k$ is invariant for A and satisfies for some non-null constants $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$*

$$\alpha_1 v_1 + \cdots + \alpha_n v_n \in F,$$

then, $v_i \in F$ for all $1 \leqslant i \leqslant n$.

Proof. We proceed by induction. Suppose $n = 2$. Since $A(\alpha_1 v_1 + \alpha_2 v_2) \in F$ and $\lambda_1(\alpha_1 v_1 + \alpha_2 v_2) \in F$, we have

$$\lambda_1(\alpha_1 v_1 + \alpha_2 v_2) - A(\alpha_1 v_1 + \alpha_2 v_2)$$

$$= \lambda_1(\alpha_1 v_1 + \alpha_2 v_2) - (\lambda_1 \alpha_1 v_1 + \lambda_2 \alpha_2 v_2)$$
$$= (\lambda_1 - \lambda_2)\alpha_2 v_2 \in F.$$

Therefore, $v_1, v_2 \in F$. Now, assuming that the claim is true for every $n \leqslant k$, we get

$$\lambda_{k+1}(\alpha_1 v_1 + \cdots + \alpha_{k+1} v_{k+1}) - A(\alpha_1 v_1 + \cdots + \alpha_{k+1} v_{k+1}) \in F.$$

Which means $(\lambda_{k+1} - \lambda_1)\alpha_1 v_1 + \cdots + (\lambda_{k+1} - \lambda_k)\alpha_k v_k \in F$. From the hypothesis, this implies $v_1, \cdots, v_k \in F$. It follows that $v_{k+1} \in F$.

Theorem 3.7.5. *Given $L \in \mathcal{B}(M_k)$, μ over M_k with $\# \operatorname{supp} \mu > 1$ and $\varepsilon > 0$, there exists $M_\delta \in \mathcal{B}(M_k)$, such that, $\|L - M_\delta\| < \varepsilon$ and M_δ is ϕ-Erg and irreducible for μ.*

Proof. Given an $\varepsilon > 0$, take $v_1 \in \operatorname{supp} \mu$ such that $v_1 \neq 0$, the respective L_ε from Proposition 3.7.3 and moreover $\{x_1, \ldots, x_k\}$ such that they are a base of eigenvectors of $L_\varepsilon(v_1)$, with corresponding eigenvalues λ_i. If L_ε is irreducible for μ, we are done. Otherwise, there exists a decomposition in $E_1, \ldots, E_n$ minimal non-trivial subspaces that are invariant for all $L_\varepsilon(v)$, with v in $\operatorname{supp} \mu$ and $k > \dim E_1 \geqslant \dim E_i$, for all i.

Remember that $E_i \cap E_j = \{0\}$ and since all E_i are invariant for $L_\varepsilon(v_1)$, they are generated by some of its eigenvectors.

Relabel $x_1, \ldots, x_k$ in such way that we get:

$E_1 = \langle x_1, \ldots, x_{d_1} \rangle$, $E_2 = \langle x_{d_1+1}, \ldots, x_{d_2} \rangle, \ldots, E_n = \langle x_{d_{n-1}+1}, \ldots, x_{d_n} \rangle$ and $K = \langle x_{d_n+1}, \ldots, x_k \rangle$, with $\mathbb{C}^k = E_1 \oplus \cdots \oplus E_n \oplus K$, where K is either $\{0\}$ or is not invariant for all $L_\varepsilon(v)$.

Now, define the linear transformation $A : \mathbb{C}^k \to \mathbb{C}^k$ by $A(x_j) = x_{j+1}$. By abuse of notation, we assume that $x_{k+1} = x_1$. Consider, for a $\delta > 0$, the operator $M_\delta(v) = L_\varepsilon(v) + \frac{\delta\varphi(v)}{2\|A\|}A$, where $\varphi(v) = \frac{\|v-v_1\|}{\|v\|+\|v_1\|} \leqslant 1$. Denote $c(v) = \frac{\delta\varphi(v)}{2\|A\|} \geqslant 0$. Note that $c(v) > 0$, for all $v \neq v_1$. Notice that $M_\delta(v_1) = L_\varepsilon(v_1)$. The idea here is to make an element x_i move to all of the other subspaces, making it impossible to have an invariant and proper subspace for all $M_\delta(v)$. This combined with the proximity of the original L will give us the result.

Claim: There exists a $\delta > 0$, such that the only non-trivial (and therefore minimal) subspace invariant for all $M_\delta(v)$, with $v \in \textit{supp } \mu$, is $\mathbb{C}^k$.

Suppose $F \subseteq \mathbb{C}^k$ is such a subspace. There exists a non-trivial element $\alpha_1 x_1 + \cdots + \alpha_k x_k \in F \cap E_i$, for some constants $a_l \in \mathbb{C}^k$ and some i. This is so because if K is $\{0\}$ or not invariant for $M_\delta(v_1) = L_\varepsilon(v_1)$, then $F \nsubseteq K$. Since not all a_i can be zero, we have by the above lemma that some $x_j \in F$.

We take a matrix $v_2 \in \textit{supp } \mu$, $v_2 \neq v_1$. Now,

$$M_\delta(v_2)x_j = L_\varepsilon(v_2)x_j + c(v_2)Ax_j = L_\varepsilon(v_2)x_j + c(v_2)x_{j+1} \in F.$$

As E_i is invariant for $L_\varepsilon(v_2)$, we get that

$$L_\varepsilon(v_2)x_j = \sum_{m=d_{i-1}+1}^{d_i} \alpha_m x_m.$$

Now, again, F is invariant for $M_\delta(v_1) = L_\varepsilon(v_1)$, and then

$$L_\varepsilon(v_1)M_\delta(v_2)(x_j) = L_\varepsilon(v_1)\left(\sum_{m=d_{i-1}+1}^{d_i} \alpha_m x_m + c(v_2)x_{j+1}\right)$$

$$= \sum_{m=d_{i-1}+1}^{d_i} \lambda_m \alpha_m x_m + c(v_2)\lambda_{j+1}\ x_{j+1} \in F.$$

Moving on, $L_\varepsilon(v_1)M_\delta(v_2)x_j - \lambda_{j+1} \cdot M_\delta(v_2)x_j \in F$. This means

$$\sum_{m=d_{i-1}+1}^{d_i} (\lambda_m - \lambda_{j+1})\alpha_m x_m \in F.$$

By the lemma, $x_m \in F$, for all m which are not $j+1$ and the corresponding α_m is not zero. Now, suppose that $x_{j+1} \notin E_i$ (this excludes the possibility of $m = j+1$ above). In this way, $\alpha_m x_m \in F$, for all $m \in \{d_{i-1}+1, \ldots, d_i\}$, with no exceptions. It follows that $\sum_m \alpha_m x_m \in F$ and

$$M_\delta(v_2)x_j - \sum_{m=d_{i-1}+1}^{d_i} \alpha_m x_m \in F$$

$$= \sum_{m=d_{i-1}+1}^{d_i} \alpha_m x_m + c(v_2)x_{j+1} - \sum_{m=d_{i-1}+1}^{d_i} \alpha_m x_m$$
$$= c(v_2)x_{j+1} \in F.$$

As $c(v_2) \neq 0$, we get $x_{j+1} \in F$. Now suppose $x_{j+1} \in E_i$. Then

$$M_\delta(v_2)x_j - \sum_{\substack{m=d_{i-1}+1 \\ m \neq j+1}}^{d_i} \alpha_m x_m \in F.$$

This means $c(v_2)x_{j+1}+\alpha_{j+1}x_{j+1} \in F$. If $c(v_2)+\alpha_{j+1} = 0$ we get a problem. In order to fix this, we need that $\frac{\delta\varphi(v_2)}{2\|A\|} \neq -\alpha_{j+1} \iff \delta \neq \frac{-2\alpha_{j+1}\|A\|}{\varphi(v_2)}$. But, note that α_{j+1} does not depend on δ. In fact, it appears only in the decomposition

$$L_\varepsilon(v_2)x_j = \sum_{m=d_{i-1}+1}^{d_i} \alpha_m x_m.$$

Since we can do this decomposition for all j, we only have to check that

$$\delta \notin \left\{ \frac{-2\alpha_{j+1}\|A\|}{\varphi(v_2)}; 1 \leqslant j \leqslant d_n \right\}.$$

Taking δ small enough, we accomplish this and also we get $\delta < \varepsilon$. Now, we get the claim in the same way: $x_{j+1} \in F$ and $F = \mathbb{C}^k$. So, for this δ we get that M_δ is irreducible. Finally,

$$\|L - M_\delta\| \leqslant \|L - L_\varepsilon\| + \|L_\varepsilon - M_\delta\| < \varepsilon/2 + \left\| \frac{\delta\varphi(v)A}{2\|A\|} \right\| < \varepsilon.$$

■

Definition 3.7.6. For a fixed measure μ over M_k, define

$$\mathcal{B}_\mu(M_k) = \{L \in \mathcal{B} \,|\, \text{L irreducible for}\mu\},$$

and

$$\mathcal{B}_\mu^\phi(M_k) = \{L \in \mathcal{B} \,|\, \text{L is } \phi\text{-Erg for } \mu\}.$$

Corollary 3.7.7. *Given μ over M_k with $\#\operatorname{supp}\mu > 1$, $\mathcal{B}_\mu(M_k)$ is dense on $\mathcal{B}(M_k)$.*

Proof. It follows from the above. ■

Proposition 3.7.8. *$\mathcal{B}_\mu(M_k)$ is open for a fixed μ on M_k.*

Proof. We will prove that the complement of $\mathcal{B}_\mu(M_k)$ is closed in $\mathcal{B}(M_k)$. Let L_n be a sequence outside $\mathcal{B}_\mu(M_k)$ converging to some $L \in \mathcal{B}(M_k)$. For each n, consider E_n a non-trivial (L_n, μ)-invariant subspace and P_n the projection on E_n.

The (L_n, μ)-invariance is equivalent to say that $L_n(v)P_n = P_n L_n(v)P_n$, for all $v \in \operatorname{supp}\mu$. Therefore, there is a subsequence such that $P_{n_i} \to P$, where P is a projection. Rename $P_n \to P$. Furthermore, $L_n \to L$, thus $P_n L_n(v)P_n = L_n(v)P_n \to PL(v)P = L(v)P$, for all $v \in \operatorname{supp}\mu$. This implies that $E := \Im(P)$ is (L, μ)-invariant for L. Of course, E is not the trivial space because $\|P\| \geqslant 1$. Moreover, we know that $\ker(P_n)$ is non-trivial for all n, once L_n is not μ-irreducible. So, take $x_n \in \ker(P_n)$ with $\|x_n\| = 1$, and rename it in order to get a subsequence such that $x_n \to x$. Observe that $P_n x_n = 0$, for all n and $P_n x_n \to Px$. This implies that $Px = 0$ and, of course, $\ker(P)$ is non-trivial. Hence, $E \neq \mathbb{C}^k$ and L is not μ-irreducible.

■

Proposition 3.7.9. *$\mathcal{B}^\phi_\mu(M_k)$ is open for a fixed μ on M_k.*

Proof. Take $L_n \to L$ such that L_n is not ϕ-Erg. Therefore, there exists $E_{1,n} \oplus E_{2,n} \oplus E_{0,n} = \mathbb{C}^k$, with $E_{i,n}$ minimal (L_n, μ)-invariant for L_n, where $i = 1, 2$ and $E_{0,n}$ is not necessarily (L_n, μ)-invariant. Take $P_{i,n}$ the projection on $E_{i,n}$. Rename them in order to get a subsequence such that $P_{i,n} \to P_i$, for all $i = 1, 2, 0$. By using the same argument as the one used in Proposition 3.7.8, we observe that $E_i = \Im(P_i)$ is (L, μ)-invariant for L, for $i = 1, 2$. If $x \in E_1 \backslash \{0\}$ we know that $\lim_n \|P_{1,n}x - x\| = \|P_1 x - x\| = 0$, so defining $x_n := P_{1,n}x \in E_{1,n}$, we get $x_n \to x$. As $0 = P_{2,n}x_n \to P_2 x$, we know $x \in \ker P_2$ and therefore $x \notin E_2$. This argument shows that $E_1 \cap E_2 = \{0\}$, hence L is not ϕ-Erg because it admits two (L, μ)-invariant subspaces. ■

Corollary 3.7.10. *Given μ over M_k with $\#\operatorname{supp}\mu > 1$, $\mathcal{B}^\phi_\mu(M_k)$ is open, dense and, therefore, generic.*

3.8 Some examples

In this section, we present several examples. The main one is Example 3.8.5 that considers a quantum channel which is a kind of version of a Markov chain. We can show in expression (3.8) that the entropy of this channel coincides with the entropy of the associated stationary Markov Process. This is a piece of clear evidence that our definition is a natural extension of the classical concept of entropy. In [BKL21a] it is shown that the entropy of this channel is related to one of the Lyapunov exponents of the associated time evolution process which are described in sections 3.5 and 3.6.

Example 3.8.1. Let $V_{2n} = c \cdot \begin{pmatrix} \frac{1}{2n} & 0 \\ 0 & 0 \end{pmatrix}$ and $V_{2n-1} = d \cdot \begin{pmatrix} 0 & \frac{1}{2n-1} \\ 0 & 0 \end{pmatrix}$, for all $n \geqslant 1$ (with constants c and d to be defined). Then,

$$V_{2n}^{\dagger} V_{2n} = \frac{c^2}{(2n)^2} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } V_{2n-1}^{\dagger} V_{2n-1} = \frac{d^2}{(2n-1)^2} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Setting $L = I$ (the identity map $v \mapsto v$) and $\mu = \sum_{n=1}^{\infty} \delta_{V_n}$, we have

$$\int_{M_k} L(v)^{\dagger} L(v)\, d\mu(v) = \sum_{n=1}^{\infty} V_n^{\dagger} V_n$$

$$= c^2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \sum_{n=1}^{\infty} \frac{1}{(2n)^2} + d^2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}.$$

Choosing

$$c = \left(\sum_{n=1}^{\infty} \frac{1}{(2n)^2} \right)^{-1/2} \text{ and } d = \left(\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \right)^{-1/2},$$

we get $\int L(v)^{\dagger} L(v)\, d\mu(v) = \mathrm{Id}$. Now, notice that

$$\int \|L(v)\|\ d\mu(v) = c \cdot \sum_{n=1}^{\infty} \frac{1}{2n} + d \cdot \sum_{n=1}^{\infty} \frac{1}{2n-1} = \infty,$$

whereas $\|L(v)\| \leqslant max\{c, d\} < \infty$, for all $v \in supp(\mu)$. Even when the last integral is not finite, the limitation on the norm above should produce an invariant probability for the kernel, according to Theorem 3.5.2. To show this will be our goal. Before that, we will compute the action of the quantum channel (in order to clear out what is the fixed density).

For a general density $\rho = \begin{pmatrix} \rho_1 & \rho_2 \\ \rho_3 & \rho_4 \end{pmatrix}$, we have

$$V_{2n}\rho V_{2n}^{\dagger} = \frac{c^2}{(2n)^2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \rho_1 & \rho_2 \\ \rho_3 & \rho_4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{c^2}{(2n)^2}\begin{pmatrix} \rho_1 & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$\begin{aligned} V_{2n-1}\rho V_{2n}^{\dagger} &= \frac{d^2}{(2n-1)^2}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \rho_1 & \rho_2 \\ \rho_3 & \rho_4 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ &= \frac{d^2}{(2n-1)^2}\begin{pmatrix} \rho_4 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

That is,

$$\begin{aligned} \phi_L(\rho) = \sum_{n=1}^{\infty}\left(\frac{c^2}{(2n)^2}\rho_1 + \frac{d^2}{(2n-1)^2}\rho_4\right)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} &= (\rho_1+\rho_4)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ &= \mathrm{tr}\,(\rho)\cdot |\, e_1 \rangle\langle e_1 \,| = |\, e_1 \rangle\langle e_1 \,|. \end{aligned}$$

This ϕ_L is not irreducible but it is an interesting example. It is a case where the invariant probability is unique as we will see soon.

Clearly, the only fixed point for ϕ_L is $\rho_{inv} = |\, e_1 \rangle\langle e_1 \,|$. What we should expect for invariant probabilities over $P(\mathbb{C}^k)$? As the fixed point is itself a projection and the proposition 3.6.2 says it is an average of projections around any invariant probability, the only option is a probability concentrated in $\hat{e}_1$, which is $\nu = \delta_{\hat{e}_1}$. Let's check that it is the case.

For a general probability ν over $P(\mathbb{C}^k)$ and a Borel set $B \subset P(\mathbb{C}^k)$, we have

$$\nu\Pi_L(B) = \int_{M_k}\int_{P(\mathbb{C}^k)} \mathbf{1}_B(L(v)\cdot\hat{x})\,\|L(v)x\|_{HS}^2\, d\mu(v)d\nu(\hat{x})$$

$$= \int_{P(\mathbb{C}^k)}\sum_{n=1}^{\infty}\left[\mathbf{1}_B(V_{2n}\cdot\hat{x})\,\|V_{2n}x\|_{HS}^2 + \mathbf{1}_B(V_{2n-1}\cdot\hat{x})\,\|V_{2n-1}x\|_{HS}^2\right] d\nu(\hat{x}).$$

Notice that $V_{2n}\cdot\hat{x} = \hat{e}_1$ for $\hat{x} \neq \hat{e}_2$ and $V_{2n-1}\cdot\hat{x} = \hat{e}_1$ for $\hat{x} \neq \hat{e}_1$, whereas $V_{2n}e_1 = V_{2n-1}e_2 = 0$. Also, for a representative $x = (x_1, x_2)$ of norm 1, we got $(|\, x \rangle\langle x \,|)_{ij} = x_i\overline{x_j}$. So,

$$\mathrm{tr}\,(V_{2n}\,|\, x \rangle\langle x \,|\, V_{2n}^{\dagger}) = \frac{c^2}{(2n)^2}\cdot(|\, x \rangle\langle x \,|)_{11} = \frac{c^2}{(2n)^2}|x_1|^2,$$

and

$$\operatorname{tr}(V_{2n-1}\,|\,x\,\rangle\langle\,x\,|\,V_{2n-1}^{\dagger}) = \frac{d^2}{(2n-1)^2}\cdot(|\,x\,\rangle\langle\,x\,|)_{22} = \frac{d^2}{(2n-1)^2}|x_2|^2.$$

Then,

$$\begin{aligned}
\nu\Pi_L(B) &= \int_{P(\mathbb{C}^k)} \sum_{n=1}^{\infty} \mathbf{1}_B(\hat{e}_1)\left[\frac{c^2}{(2n)^2}|x_1|^2 + \frac{d^2}{(2n-1)^2}|x_2|^2\right] d\nu(\hat{x}) \\
&= \int_{P(\mathbb{C}^k)} \mathbf{1}_B(\hat{e}_1)(|x_1|^2 + |x_2|^2) d\nu(\hat{x}) \\
&= \int_{P(\mathbb{C}^k)} \mathbf{1}_B(\hat{e}_1) d\nu(\hat{x}) \\
&= \mathbf{1}_B(\hat{e}_1).
\end{aligned}$$

We conclude that if $\nu\Pi_L = \nu$, then $\nu = \delta_{\hat{e}_1}$. We also get a bonus: the invariant probability is unique.

To illustrate Proposition 3.6.2 (under the irreducible condition) we write down the following example.

Example 3.8.2. The next example is somehow related to Example 3.8.5. Let's define

$$V_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } V_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

These two matrices generate the same elements which we will consider in Example 3.8.5, since for $\mu = \delta_{V_1} + \delta_{V_2}$,

$$\phi_I(\rho) = V_1\rho V_1^{\dagger} + V_2\rho V_2^{\dagger} = |\,e_1\,\rangle\langle\,e_1\,|.$$

Also, we get that ϕ_I is not irreducible. Wanting to fix this issue, we introduce

$$V_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ and } V_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Notice that these two matrices generates another channel ψ that maps every density ρ into $|\,e_2\,\rangle\langle\,e_2\,|$. So, it is also not irreducible. Now, redefining $\mu = \frac{1}{2}\sum_{i=1}^{4}\delta_{V_i}$, we get that

$$\phi_I(\rho) = \frac{1}{2}\sum_{i=1}^{4} V_i \rho V_i^\dagger = \frac{1}{2}(|\, e_1 \rangle\langle e_1 \,| + |\, e_2 \rangle\langle e_2 \,|) = \frac{1}{2}\mathrm{Id}.$$

In this case, μ is a measure and not a probability.

We compute the products

$$V_1^\dagger V_1 = V_1,\ V_2^\dagger V_2 = V_4,$$

and

$$V_3^\dagger V_3 = V_1 \text{ and } V_4^\dagger V_4 = V_4.$$

In this way,

$$\phi_I^*(\mathrm{Id}) = \frac{1}{2}\sum_{i=1}^{4} V_i^\dagger V_i = V_1 + V_4 = \mathrm{Id},$$

and ϕ_I is stochastic. As $\mathrm{Id} > 0$, we get that $(I+\phi)(\rho) = \rho+\phi(\rho) = \rho+\mathrm{Id} > 0$, and so ϕ is irreducible. Clearly, $\rho_{inv} = \frac{1}{2}\,\mathrm{Id}$.

Now, for a general ν over $P(\mathbb{C}^k)$ and a Borel set $B \subset P(\mathbb{C}^k)$, we get

$$\begin{aligned}\nu\Pi_I(B) &= \int_{P(\mathbb{C}^k)}\int_{M_k} \mathbf{1}_B(L(v)\cdot\hat{x})\,\|L(v)x\|_{HS}^2\ d\mu(v)d\nu(\hat{x})\\ &= \int_{P(\mathbb{C}^k)}\sum_{i=1}^{4}\frac{1}{2}\mathbf{1}_B(V_i\cdot\hat{x})\,\|V_i x\|_{HS}^2\ d\nu(\hat{x}).\end{aligned}$$

Remember that

$$V_1\,|\, x \rangle\langle x \,|\, V_1^\dagger = \begin{pmatrix} |x_1|^2 & 0 \\ 0 & 0 \end{pmatrix},\quad V_2\,|\, x \rangle\langle x \,|\, V_2^\dagger = \begin{pmatrix} |x_2|^2 & 0 \\ 0 & 0 \end{pmatrix},$$

$$V_3\,|\, x \rangle\langle x \,|\, V_3^\dagger = \begin{pmatrix} 0 & 0 \\ 0 & |x_1|^2 \end{pmatrix} \text{ and } V_4\,|\, x \rangle\langle x \,|\, V_4^\dagger = \begin{pmatrix} 0 & 0 \\ 0 & |x_2|^2 \end{pmatrix}.$$

So,

$$\nu\Pi_I(B) = \frac{1}{2}\int_{P(\mathbb{C}^k)} [\mathbf{1}_B(V_1\cdot\hat{x})+\mathbf{1}_B(V_3\cdot\hat{x})]|x_1|^2+[\mathbf{1}_B(V_2\cdot\hat{x})+\mathbf{1}_B(V_4\cdot\hat{x})]|x_2|^2\,d\nu(\hat{x})$$

$$= \frac{1}{2}\int_{P(\mathbb{C}^k)} [\mathbf{1}_B(\hat{e}_1) + \mathbf{1}_B(\hat{e}_2)]|x_1|^2 + [\mathbf{1}_B(\hat{e}_1) + \mathbf{1}_B(\hat{e}_2)]|x_2|^2\,d\nu(\hat{x})$$

$$= \frac{1}{2}\int_{P(\mathbb{C}^k)} \mathbf{1}_B(\hat{e}_1) + \mathbf{1}_B(\hat{e}_2)\,d\nu(\hat{x})$$

$$
\begin{aligned}
&= \frac{1}{2}\mathbf{1}_B(\hat{e}_1) + \mathbf{1}_B(\hat{e}_2) \\
&= \frac{1}{2}\,\delta_{\hat{e}_1}(B) + \frac{1}{2}\,\delta_{\hat{e}_2}(B).
\end{aligned}
$$

We conclude that if $\nu = \nu\Pi_I$, then $\nu = \frac{1}{2}\,\delta_{\hat{e}_1} + \frac{1}{2}\,\delta_{\hat{e}_2}$. Note that (see the concept of barycenter in Section 3.6)

$$
\int_{P(\mathbb{C}^k)} \pi_x \, d\nu(\hat{x}) = \frac{1}{2}\,\pi_{e_1} + \frac{1}{2}\,\pi_{e_2} = \frac{1}{2}\mathrm{Id} = \rho_{inv}.
$$

Example 3.8.3 (L is a C^*-automorphism)**.** Suppose that μ over M_k satisfies the below conditions:

- $\displaystyle\int_{M_k} v^\dagger v \; d\mu(v) = \mathrm{Id}$; and
- $\displaystyle\int_{M_k} \|v\|^2 \, d\mu(v) < \infty$, where $\|\cdot\|$ is the Hilbert-Schmidt norm.

Take an unitary matrix $U \in M_k$ and define $L(v) = UvU^\dagger$. Note that $\left\|UvU^\dagger\right\|^2 = \mathrm{tr}\,(UvU^\dagger) = \mathrm{tr}\,(v) = \|v\|^2$. Moreover,

$$
\begin{aligned}
\int_{M_k} L(v)^\dagger L(v) \; d\mu(v) &= \int_{M_k} Uv^\dagger U^\dagger UvU^\dagger \; d\mu(v) \\
&= U \int_{M_k} v^\dagger v \; d\mu(v) U^\dagger \\
&= \mathrm{Id}.
\end{aligned}
$$

Remark 3.8.4. The operators of the form $L(v) = UvU^\dagger$, for an unitary U, are the C^*-automorphisms of M_k (see Section 1.4 in [Arv98]).

In the next example, we adapt the reasoning of an Example 4 in [BLLC10] to the present setting.

We will show that for a certain μ and L (and, quantum channel) the value we get here for the entropy is equal to the classical entropy of a Markov Chain (when the state space is finite).

Example 3.8.5 (The Markov model in quantum information)**.** Suppose that $P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$ is a irreducible (in the classical sense for a Markov chain) column stochastic matrix. Define μ over M_2 by

$$
\mu = \sum_{i=1}^{4} \delta_{V_i},
$$

where the matrices V_i are

$$V_1 = \begin{pmatrix} \sqrt{p_{00}} & 0 \\ 0 & 0 \end{pmatrix}, \; V_2 = \begin{pmatrix} 0 & \sqrt{p_{01}} \\ 0 & 0 \end{pmatrix},$$

$$V_3 = \begin{pmatrix} 0 & 0 \\ \sqrt{p_{10}} & 0 \end{pmatrix} \text{ and } V_4 = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{p_{11}} \end{pmatrix}.$$

We take $L = I$ and $\phi_I = \phi_L$, in order to get the quantum channel

$$\phi(\rho) = \sum_1^4 V_i \rho V_i^\dagger,$$

whose dual is

$$\phi^*(\rho) = \sum_1^4 V_i^\dagger \rho V_i.$$

Note that

$$V_1^\dagger V_1 = \begin{pmatrix} p_{00} & 0 \\ 0 & 0 \end{pmatrix}, \quad V_2^\dagger V_2 = \begin{pmatrix} 0 & 0 \\ 0 & p_{01} \end{pmatrix}$$

$$V_3^\dagger V_3 = \begin{pmatrix} p_{10} & 0 \\ 0 & 0 \end{pmatrix} \text{ and } V_4^\dagger V_4 = \begin{pmatrix} 0 & 0 \\ 0 & p_{11} \end{pmatrix}, \tag{3.5}$$

that is,

$$\phi^*(\mathrm{Id}_2) = \begin{pmatrix} p_{00} + p_{10} & 0 \\ 0 & p_{01} + p_{11} \end{pmatrix} = \mathrm{Id}_2$$

The channel ϕ is stochastic. We claim that the channel is irreducible (later we will exhibit the associated invariant density operator ρ). Consider first the positive operator

$$\rho = \begin{pmatrix} \rho_1 & \rho_2 \\ \rho_3 & \rho_4 \end{pmatrix}$$

where $\rho_1, \rho_2 \in \mathbb{R}$ and $\rho_3 = \overline{\rho_2}$ (in order to get that $\rho \geqslant 0$)

The $V_i \rho V_i^\dagger$ are given by:

$$\rho^1 := V_1 \rho V_1^\dagger = \begin{pmatrix} p_{00}\rho_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \rho^2 := V_2 \rho V_2^\dagger = \begin{pmatrix} p_{01}\rho_4 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\rho^3 := V_3 \rho V_3^\dagger = \begin{pmatrix} 0 & 0 \\ 0 & p_{10}\rho_1 \end{pmatrix} \text{ and } \rho^4 := V_4 \rho V_4^\dagger = \begin{pmatrix} 0 & 0 \\ 0 & p_{11}\rho_4 \end{pmatrix} \tag{3.6}$$

It follows that

$$\phi(\rho) = \begin{pmatrix} p_{00}\rho_1 + p_{01}\rho_4 & 0 \\ 0 & p_{10}\rho_1 + p_{11}\rho_4 \end{pmatrix}.$$

In the diagonal one can find the classical action on vectors of the Markov Chain described by P.

In the same way for $v = (v_1, v_2) \in \mathbb{C}^2$, we get

$$\langle v \mid \phi(\rho)v \rangle = (p_{00}\rho_1 + p_{01}\rho_4)|v_1|^2 + (p_{10}\rho_1 + p_{11}\rho_4)|v_2|^2 \geqslant 0.$$

Moreover, the equality only happens when

$$p_{00}\rho_1 + p_{01}\rho_4 = p_{10}\rho_1 + p_{11}\rho_4 = 0.$$

From this we get $\rho_1 = \rho_4 = 0$, because $p_{ij} \geqslant 0$.

In this case, we get $\rho = 0$.

This means that , $\rho \neq 0, \rho \geqslant 0 \Rightarrow \phi(\rho) > 0$, and, finally, we get that ϕ is positive improving. From this, it follows that ϕ is irreducible.

Now, we will look for the invariant density matrix. Assuming $\rho_1 + \rho_4 = 1$, we observe that $\phi(\rho) = \rho \Rightarrow \rho_2 = \rho_3 = 0$, and

$$\begin{cases} \rho_1 = p_{00}\rho_1 + p_{01}\rho_4 \\ \rho_4 = p_{10}\rho_1 + p_{11}\rho_4. \end{cases} \tag{3.7}$$

We get

$$(1 - p_{00})\rho_1 = p_{01}\rho_4 = p_{01}(1 - \rho_1) = p_{01} - p_{01}\rho_1$$

$$\Rightarrow (1 - p_{00} + p_{01})\rho_1 = p_{01}.$$

As P is irreducible, it follows that $0 < p_{ij} < 1$ e $1 - p_{00} + p_{01} > 0$. That is,

$$\rho_1 = \frac{p_{01}}{1 - p_{00} + p_{01}} \quad \text{and} \quad \rho_4 = \frac{1 - p_{00}}{1 - p_{00} + p_{01}}.$$

An invariant density matrix is

$$\rho = \begin{pmatrix} \dfrac{p_{01}}{1 - p_{00} + p_{01}} & 0 \\ 0 & \dfrac{1 - p_{00}}{1 - p_{00} + p_{01}} \end{pmatrix}.$$

Note that $\pi = (\rho_1, \rho_4) \in \mathbb{R}^2$ is the vector of probability which is invariant for the stochastic matrix P (see (3.7)).

Now, we will estimate the entropy of the quantum channel ϕ. Using (3.6) in the expression $\mathrm{tr}\,(V_jV_i\rho V_i^\dagger V_j^\dagger)$ we get

$$\begin{cases} \mathrm{tr}\,(V_1\rho^iV_1^\dagger) = p_{00}(\rho^i)_1 \\ \mathrm{tr}\,(V_2\rho^iV_2^\dagger) = p_{01}(\rho^i)_4 \\ \mathrm{tr}\,(V_3\rho^iV_3^\dagger) = p_{10}(\rho^i)_1 \\ \mathrm{tr}\,(V_4\rho^iV_4^\dagger) = p_{11}(\rho^i)_4 \end{cases}$$

For example,

$$\mathrm{tr}\,(V_3V_1\rho V_1^\dagger V_3^\dagger) = \mathrm{tr}\,(V_3\rho^1V_3^\dagger) = p_{10}(\rho^1)_1 = p_{10}p_{00}\rho_1.$$

From this we get the table.

$\mathrm{tr}\,(V_jV_i\rho V_i^\dagger V_j^\dagger)$	i 1	2	3	4
j				
1	$p_{00}^2\rho_1$	$p_{00}p_{01}\rho_4$	0	0
2	0	0	$p_{01}p_{10}\rho_1$	$p_{01}p_{11}\rho_4$
3	$p_{00}p_{10}\rho_1$	$p_{10}p_{01}\rho_4$	0	0
4	0	0	$p_{11}p_{10}\rho_1$	$p_{11}^2\rho_4$
$\mathrm{tr}\,(V_i\rho V_i^\dagger)$	$p_{00}\rho_1$	$p_{01}\rho_4$	$p_{10}\rho_1$	$p_{11}\rho_4$

The entropy we defined in the text is given by

$$h_\mu(L) = -\int_{M_k\times M_k} \mathrm{tr}\,(L(v)\rho L(v)^\dagger)P(v,w)\log(P(v,w))d\mu(v)d\mu(w),$$

where $P(v,w) = \frac{\mathrm{tr}\,(L(w)L(v)\rho L(v)^\dagger L(w)^\dagger)}{\mathrm{tr}\,(L(v)\rho L(v)^\dagger)}$.
We assumed before that $L = I$ and $\mu = \sum_i \delta_{V_i}$. Then, we finally get,

$$h_\mu(I) = -\sum_{i=1}^4\sum_{j=1}^4 \mathrm{tr}\,(V_jV_i\rho V_i^\dagger V_j^\dagger)\cdot\log\left(\frac{\mathrm{tr}\,(V_jV_i\rho V_i^\dagger V_j^\dagger)}{\mathrm{tr}\,(V_i\rho V_i^\dagger)}\right)$$

$$= -\big[p_{00}^2\rho_1 log(p_{00}) + p_{00}p_{10}\rho_1\log(p_{10}) + p_{00}p_{01}\rho_4\log(p_{00}) + p_{10}p_{01}\rho_4\log(p_{10})$$
$$+p_{01}p_{10}\rho_1\log(p_{01}) + p_{11}p_{10}\rho_1\log(p_{11}) + p_{01}p_{11}\rho_4\log(p_{01}) + p_{11}^2\rho_4\log(p_{11})\big]$$

$$= -\big[p_{00}\log(p_{00})(p_{00}\rho_1 + p_{01}\rho_4) + p_{10}\log(p_{10})(p_{00}\rho_1 + p_{01}\rho_4)$$
$$+p_{01}\log(p_{01})(p_{10}\rho_1 + p_{11}\rho_4) + p_{11}\log(p_{11})(p_{10}\rho_1 + p_{11}\rho_4)\big]$$

$$= -p_{00}\log(p_{00})\rho_1 - p_{10}\log(p_{10})\rho_1 - p_{01}\log(p_{01})\rho_4 - p_{11}\log(p_{11})\rho_4$$

$$= -p_{00}\log(p_{00})\pi_0 - p_{10}\log(p_{10})\pi_0 - p_{01}\log(p_{01})\pi_1 - p_{11}\log(p_{11})\pi_1 =$$

$$-\sum_{i,j=0}^{1} \pi_j p_{ij}\log(p_{ij}).$$

Therefore,

$$h_\mu(I) = -\sum_{i,j=0}^{1} \pi_j p_{ij}\log(p_{ij}). \tag{3.8}$$

The last expression is the value of the classical Shannon-Kolmogorov entropy of the stationary Markov Process associated to the line stochastic matrix $P = (p_{ij})_{i,j=0,1}$ (see [Spi72] and [PY98]).

The entropy is positive because the a priori μ is a measure (of mass equal to 4) and not a probability.

Now, let's look at the kernel Π_L and find an invariant probability. For a given probability ν in $P(\mathbb{C}^k)$ and a Borel set $B \subset P(\mathbb{C}^k)$, we have

$$\nu\Pi_L(B) = \int_{P(\mathbb{C}^k)}\int_{M_k} \mathbf{1}_B(L(v)\cdot\hat{x})\,\|L(v)x\|_{HS}^2\; d\mu(v)d\nu(\hat{x}),$$

which means

$$\nu\Pi_L(B) = \int_{P(\mathbb{C}^k)}\sum_{i=1}^{4} \mathbf{1}_B(V_i\cdot\hat{x})\,\|V_i\,x\|_{HS}^2\; d\nu(\hat{x}).$$

Note that

$$V_1\cdot\hat{x} = \hat{e}_1,\ \text{if } \hat{x}\neq\hat{e}_2;\quad V_2\cdot\hat{x} = \hat{e}_1,\ \text{if } \hat{x}\neq\hat{e}_1;$$

$$V_3\cdot\hat{x} = \hat{e}_2,\ \text{if } \hat{x}\neq\hat{e}_2;\quad V_4\cdot\hat{x} = \hat{e}_2,\ \text{if } \hat{x}\neq\hat{e}_1$$

$$\text{and } V_1(e_2) = V_2(e_1) = V_3(e_2) = V_4(e_1) = 0.$$

It follows that

$$\nu\Pi_L(B) = \int_{P(\mathbb{C}^k)} \mathbf{1}_B(\hat{e}_1)\,[\|V_1x\| + \|V_2x\|] + \mathbf{1}_B(\hat{e}_2)\,[\|V_3x\| + \|V_4x\|]\,d\nu(\hat{x}).$$

Now, we compute

$$\operatorname{tr}(V_1\,|\,x\rangle\langle x\,|\,V_1^\dagger) = p_{00}\,|x_1|^2,$$

$$\operatorname{tr}(V_2 \,|\, x \rangle\langle x \,|\, V_2^\dagger) = p_{01}\,|x_2|^2,$$
$$\operatorname{tr}(V_3 \,|\, x \rangle\langle x \,|\, V_3^\dagger) = p_{10}\,|x_1|^2$$
$$\textit{and}\ \operatorname{tr}(V_4 \,|\, x \rangle\langle x \,|\, V_4^\dagger) = p_{11}\,|x_2|^2.$$

In this way, we get

$$\nu\Pi_L(B) = \int_{P(\mathbb{C}^k)} \mathbf{1}_B(\hat{e}_1)\,(p_{00}\,|x_1|^2 + p_{01}\,|x_2|^2) + \mathbf{1}_B(\hat{e}_2)\,(p_{10}\,|x_1|^2 + p_{11}\,|x_2|^2)\,d\nu(\hat{x}).$$

From the last expression, we conclude that $\nu\Pi_L$ has support in the set $\{\hat{e}_1, \hat{e}_2\}$.

In this way, if $\nu = \nu\Pi_L$, then it has to be equal to $\alpha \cdot \delta_{\hat{e}_1} + \beta \cdot \delta_{\hat{e}_2}$, with constants $\alpha, \beta \geqslant 0$, such that, $\alpha + \beta = 1$. As we know the expression for ρ_{inv}, we can go further:

$$\rho_{inv} = \int_{P(\mathbb{C}^k)} \pi_x\,d\nu(\hat{x}) = \alpha \cdot \pi_{e_1} + \beta \cdot \pi_{e_2}.$$

As

$$\rho_{inv} = \begin{pmatrix} \dfrac{p_{01}}{1 - p_{00} + p_{01}} & 0 \\ 0 & \dfrac{1 - p_{00}}{1 - p_{00} + p_{01}} \end{pmatrix},$$

we get that $\alpha = \dfrac{p_{01}}{1 - p_{00} + p_{01}}$ and $\beta = \dfrac{1 - p_{00}}{1 - p_{00} + p_{01}}$.

In order to finish our example, we write down the invariant probability

$$\nu = \frac{p_{01}}{1 - p_{00} + p_{01}} \cdot \delta_{\hat{e}_1} + \frac{1 - p_{00}}{1 - p_{00} + p_{01}} \cdot \delta_{\hat{e}_2} = \pi_1\,\delta_{\hat{e}_1} + \pi_2\,\delta_{\hat{e}_2},$$

and we point out that the two constants are no more no less then the entries of the invariant probability vector $\pi = (\pi_1, \pi_2)$ for the Markov chain with transitions $P = (p_{ij})_{i,j=1,2}$.

In this way, the concept of entropy we considered before in Section 3.4 is a natural generalization of the classical Kolmogorov-Shannon entropy and the process $X_n, n \in \mathbb{N}$, of Section 3.5 is a natural generalization of the classical Markov Chain process.

Example 3.8.6. Consider a measure μ with support on the set

$$\{\begin{pmatrix} x & -y \\ y & x \end{pmatrix} \mid x, y \in \mathbb{R}\} \subset M_2,$$

such that has density $f(x,y) = \frac{1}{4\pi}\, e^{-\frac{(x^2+y^2)}{2}}$ (see also (9) in [EW07])

Taking $L = I$ we get that $\rho_0 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ satisfies $\phi_I(\rho_0) = \rho_0$.

Indeed the channel is given by

$$\rho = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \to \phi_I(\rho) =$$

$$\int\int \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \frac{1}{4\pi}\, e^{-\frac{(x^2+y^2)}{2}}\, dx\, dy = \begin{pmatrix} 1/2 & \frac{b-c}{2} \\ \frac{c-b}{2} & 1/2 \end{pmatrix}.$$

Notice that although $\begin{pmatrix} 1/2 & b \\ -b & 1/2 \end{pmatrix}$ is a fixed point of ϕ_I, it is not a density unless $b = 0$. Thus, ρ_0 is the only eigendensity.

Given a probability ν on $P(\mathbb{C}^k)$ the expression for the kernel is

$$\nu\Pi_L(S) = \int_{P(\mathbb{C}^k)} \Pi_L(\hat{w}, S)\; d\nu(\hat{w}) =$$

$$\int_{P(\mathbb{C}^k)\times M_k} \mathbf{1}_S(L(v)\cdot \hat{w})\, \|L(v)w\|^2\; d\nu(\hat{w})\, d\mu(v) =$$

$$\int_{P(\mathbb{C}^k)\times M_k} \mathbf{1}_S \widehat{\begin{pmatrix} v_1\, w_1 - v_2\, w_2 \\ v_2\, w_1 + v_1\, w_2 \end{pmatrix}}\; (v_1^2 + v_2^2)\, \frac{1}{4\pi}\, e^{-\frac{(v_1^2+v_2^2)}{2}}\, dv_1\, dv_2\, d\nu(\hat{w}).$$

Now, we will estimate the entropy (which will be negative).

Using the fixed density operator $\rho_0 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ we get (according to Section 3.4)

$$P(v,w) = \frac{\operatorname{tr}(wv\rho_0 v^\dagger w^\dagger)}{\operatorname{tr}(v\rho_0 v^\dagger)}.$$

We denote

$$w = \begin{pmatrix} w_1 & -w_2 \\ w_2 & w_1 \end{pmatrix} \text{ and } v = \begin{pmatrix} v_1 & -v_2 \\ v_2 & v_1 \end{pmatrix},$$

and we get

$$
\begin{aligned}
\operatorname{tr}(v\rho_0 v^\dagger) &= \frac{1}{2}\operatorname{tr}\left(\begin{pmatrix} v_1 & -v_2 \\ v_2 & v_1 \end{pmatrix}\begin{pmatrix} v_1 & v_2 \\ -v_2 & v_1 \end{pmatrix}\right) \\
&= \frac{1}{2}\operatorname{tr}\begin{pmatrix} v_1^2+v_2^2 & 0 \\ 0 & v_1^2+v_2^2 \end{pmatrix} = v_1^2+v_2^2
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{tr}(wv\rho_0 v^\dagger w^\dagger) &= \frac{1}{2}\operatorname{tr}\left(w\begin{pmatrix} v_1 & -v_2 \\ v_2 & v_1 \end{pmatrix}\begin{pmatrix} v_1 & v_2 \\ -v_2 & v_1 \end{pmatrix}w^\dagger\right) \\
&= (v_1^2+v_2^2)\operatorname{tr}(ww^\dagger) = (v_1^2+v_2^2)(w_1^2+w_2^2).
\end{aligned}
$$

Thus, we get the following expression for the entropy (remember that $\int_0^\infty x^3 e^{-\frac{x^2}{2}}dx = 2$):

$$
\begin{aligned}
h_\mu(L) &= -\frac{1}{16\pi^2}\int (v_1^2+v_2^2)(w_1^2+w_2^2)\log(w_1^2+w_2^2)e^{-\frac{v_1^2+v_2^2}{2}}e^{-\frac{w_1^2+w_2^2}{2}}dv_1dv_2dw_1dw_2 \\
&= -\frac{1}{4}\int_0^\infty\int_0^\infty r_v^3 r_w^3 \log(r_w^2)e^{-\frac{r_v^2}{2}}e^{-\frac{r_w^2}{2}}dr_v dr_w \\
&= -\frac{1}{4}\int_0^\infty\left[\int_0^\infty r_v^3 e^{-\frac{r_v^2}{2}}dr_v\right]r_w^3\log(r_w^2)e^{-\frac{r_w^2}{2}}dr_w \\
&= -\frac{1}{2}\int_0^\infty r_w^3\log(r_w^2)e^{-\frac{r_w^2}{2}}dr_w \\
&= -\int_0^\infty r_w^3\log(r_w)e^{-\frac{r_w^2}{2}}dr_w \\
&\approx -1.11593
\end{aligned}
$$

We used polar coordinates above.

3.9 Conclusion and relations with other works

We introduce a concept of entropy and pressure (definitions depending on an *a priori* probability μ). For a given $H : M_k \to M_k$ (which plays the role of an Hamiltonian, or a Liouvillian) we define a version of the Ruelle operator $\phi_H : M_k \to M_k$, via the expression:

$$
\rho \to \phi_H(\rho) = \int_{M_k} H(v)\rho H(v)^\dagger \, d\mu(v).
$$

After that, we presented a type of Ruelle Theorem: a variational principle of pressure related to an eigenvalue problem for the Ruelle operator (see Theorem 3.4.8). The entropy and the Ruelle operator are linked via the *a priori* probability in a natural and fundamental way.

The definition of entropy considered here is not based on the point of view of dynamical partitions. It is a kind of generalization of Rokhlin Formula which says the entropy of an σ-invariant probability ν is $H(\nu) = -\int \log J d\nu$, where J is the Jacobian (a dynamical version of Radon-Nikodym derivative). Note that this entropy is not relative but absolute. Results in [LMMS15] - for the classical (not quantum) Thermodynamic Formalism theory - include the case where the alphabet M (a compact metric space) is uncountable. We did not use the results of [LMMS15] we just mentioned it to say that we followed similar reasoning.

A common procedure in Statistical Mechanics (for the one-dimensional lattice $M^{\mathbb{N}}$ or $M^{\mathbb{Z}}$) is to define entropy by considering first a finite box of size, let's say n, and then take the limit on the size of the box: the thermodynamic limit. The probability on the finite box M^n has no dynamical content. On the limit, when $n \to \infty$, it may have dynamical content (where the dynamics of shift corresponds to translation in the lattice $M^{\mathbb{N}}$ or $M^{\mathbb{Z}}$). We say in this case that the entropy was obtained via finite partitions. In this setting, probabilities maximizing pressure are obtained in a similar way, like via the limit $\frac{e^{-H} dP}{\int e^{-H} dP}$, $n \to \infty$, where the Hamiltonian H is in some way defined on each box of size n. The procedure is different in Thermodynamic Formalism, where you work primarily with the Shannon–Kolmogorov entropy on the lattice $M^{\mathbb{N}}$ or $M^{\mathbb{Z}}$ (which has dynamical content) for getting shift invariant probabilities that maximize pressure. This entropy can be estimated by a version of the Rokhlin Formula (see [LMMS15]). The Ruelle operator also played an important role in our definition of entropy. Both concepts are linked in a natural and fundamental way (see [LMMS15], or section 4 in [BKL21a] for the classical thermodynamic formalism case).

In [BKL21a] the authors show a relation of the entropy presented here with Lyapunov exponents, and this is a clear indication of its dynamical nature.

Below we will present some clarifications on which directions our work is related to relevant issues in the area related to quantum entropy.

First of all, is needed to say that the von Neumann entropy, which is given by - trace $(\rho \log \rho)$, in the same way as the expressions $-\sum_{i=1}^{d} p_i \log p_i$, or $\int \log f(x) f(x) dx$, where f is positive and $\int f(x) dx = 1$, are not exactly dynamical entropies (at least from our point of view).

Quantum entropies with dynamical content were considered in a large number of papers and books for several decades. We believe our point of view does not coincide exactly (as far as we know) with the quite important results on the topic we describe next.

In [Ara73] and [Ara69] H. Haraki considers the relative entropy which can be defined for arbitrary normal states on a von Neumann algebra. As it is a relative entropy is different from ours.

A very well know version is the dynamical entropy of C^*-algebras and von Neumann algebras of A. Connes, H. Narnhofer, and W. Thirring (see [CNT87]); as far we understand is in "some sense based" on the principle of dynamic partitions.

L. Accardi, A. Souissi and E. Soueidy in [ASS20] consider a Quantum version of Markov Chains which is in "some sense" based' on the principle of dynamic partitions. It is different from ours.

R. Alicki and M. Fannes in [AF01] considers the concept of quantum dynamical entropy from different points of view: section 12 considers entropy production; section 13.1 consider the case of the quantum cat map; section 13.2 consider noncommutative Lyapunov exponents and the Ruelle inequality (the dynamics are associated with the continuous-time semigroup generated by the Laplacian in a compact Riemannian manifold); section 13.3 is devoted to quasi-free fermionic dynamics. All of them are different from ours.

The setting of [Sło03] which considers iterated function systems and Markov operators is the point of view closer to our work. But this reference does not consider the variational principle of pressure neither a version of the Ruelle operator. Results in [BLLC10], [BLLC11a] and [BLLC11b] addressed these topics and they were generalized here.

The book [Pet07] consider the relative *von Neumann entropy* in Quantum information with a view to some applications like the Quantum Stein Lemma, Quantum Chernoff bounds, and Quantum Fisher information.

T. Sagawa in [Sag21] consider the relative entropy of von Neumann and questions related to the second law of Thermodynamics and majorization: what happens with the value of the entropy of a density matrix after the iteration by a quantum channel? The book [Pet07] addresses preliminarily the question of majorization when a matrix is applied on a finite probability ([LR22] consider a similar problem considering the iteration of the dual of the Ruelle operator and not a matrix). Maybe a future work could be to analyze majorization under the context of the present paper.

C. Pinzari, Y. Watatani, and K. Yonetani in [PY$^+$00] consider entropy and a variational principle of entropy from the point of view of C^*-algebras. A version of the Perron-Frobenius theorem was used as an important tool for analyzing KMS states for some interesting examples arising from subshifts in symbolic dynamics. The relationship between the Voiculescu topological entropy and the topological entropy of the associated subshift is studied. In the case of the Cuntz-Krieger algebras, explicit construction of the state of maximal entropy was done. We understood that the space of symbols (the alphabet) considered in [PY+00] is finite. Our results correspond to the case where the alphabet (in some sense the support of the *a priori* probability μ) can be uncountable.

In [KP02] the variational principle of pressure is considered by D. Kerr and C. Pinzari. They introduce a notion of pressure for a selfadjoint element in a C^*-algebra, adapting Voiculescu's formulation of topological entropy for a nuclear C^*-algebra (see [Voi95] and [Stø02]). The variational inequality holds for the Connes-Narnhofer-Thirring entropy. They also introduce the concept of local state approximation entropy which is different from our definition of entropy.

I. Nechita and C. Pellegrini addressed questions related to generic properties for quantum channels. In [NP12] the authors show that for a fixed density matrix $\beta : \mathbb{C}^n \to \mathbb{C}^n$, the existence of a set of full measure for the Haar measure, on the set of unitary operator $U : \mathbb{C}^n \otimes \mathbb{C}^n \to \mathbb{C}^n \otimes \mathbb{C}^n$, satisfying the property that for the associated quantum channel $Q \to \Phi(Q) = Tr_2(U(q \otimes \beta)U^*)$ there exists a unique fixed point. In [LS15] the authors show that, in fact, there exists an open and dense set of unitary operators U with such property.

A final remark: our main theorems considered the case of the C^*-algebra of matrices M_k and a natural question is if our proofs can be implemented for a general C^*-algebra? Several results for completely positive maps that were used here are also known in a more general scope. This eventual extension would involve several issues that by their nature would be much more complex; in its generality would encompass - in a sense - the classical thermodynamic formalism for potentials that depends on an infinite number of coordinates. The main eigenfunction for the Ruelle operator of a continuous potential may not exist; the existence requires the use of the Holder regularity of the potential. For the Markov case, the Perron Theorem provides similar results without further hypotheses due to the fact that a potential that depends on two coordinates is automatic of Holder class. We leave the question related to the general C^*-algebra for future work.

Bibliography

[ACR18] D. Aguiar, L. Cioletti, and R. Ruviaro. A variational principle for the specific entropy for symbolic systems with uncountable alphabets. *Mathematische Nachrichten*, 291(17-18):2506–2515, 2018.

[AF94] Robert Alicki and Mark Fannes. Defining quantum dynamical entropy. *letters in mathematical physics*, 32(1):75–82, 1994.

[AF01] Robert Alicki and Mark Fannes. *Quantum dynamical systems.* Oxford university press; Oxford, 2001.

[AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savare. Gradient flows in metric spaces and in the wasserstein space of probability measures, 2005.

[Ara69] Huzihiro Araki. Gibbs states of a one dimensional quantum lattice. *Communications in Mathematical Physics*, 14(2):120–157, 1969.

[Ara73] Huzihiro Araki. Relative hamiltonian for faithful normal states of a von neumann algebra. *Publications of the Research Institute for Mathematical Sciences*, 9(1):165–209, 1973.

[Arv98] William Arveson. An invitation to c*-algebras, 2nd printing. *Graduate Text in Mathematics, Springer*, 1998.

[ASS20] Luigi Accardi, Abdessatar Souissi, and El Gheteb Soueidy. Quantum markov chains: a unification approach. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 23(02):2050016, 2020.

[AW19] George Androulakis and Duncan Wright. On the nonlinearity of quantum dynamical entropy. *Journal of Mathematical Physics*, 60(5):053504, 2019.

[BCL+11] Alexandre Baraviera, Leandro Cioletti, Artur O. Lopes, Joana Mohr, and Rafael Rigão Souza. On the general one-dimensional xy model: positive and zero temperature, selection and non-selection. *Reviews in Mathematical Physics*, 23(10):1063–1113, 2011.

[BD85] Michael F. Barnsley and Stephen Demko. Iterated function systems and the global construction of fractals. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 399(1817):243–275, 1985.

[BDEG88] Michael F. Barnsley, Stephen G. Demko, John H. Elton, and Jeffrey S. Geronimo. Invariant measures for markov processes arising from iterated function systems with place-dependent probabilities. In *Annales de l'IHP Probabilités et statistiques*, volume 24, pages 367–394, 1988.

[Ben03] Fabio Benatti. *Classical and quantum entropies: information and dynamics*. Princeton University Press, 2003.

[Ben09] Fabio Benatti. *Dynamics, information and complexity in quantum systems*. Springer Science & Business Media, 2009.

[BFPP19] Tristan Benoist, Martin Fraas, Yan Pautrat, and Clément Pellegrini. Invariant measure for quantum trajectories. *Probability Theory and Related Fields*, 174(1):307–334, 2019.

[BKL21a] Jader E. Brasil, Josué Knorst, and Artur O. Lopes. Lyapunov exponents for quantum channels: an entropy formula and generic properties. *Journal of Dynamical Systems and Geometric Theories*, 19(2):155–187, 2021.

[BKL21b] Jader E. Brasil, Josué Knorst, and Artur O. Lopes. Thermodynamic formalism for quantum channels: Entropy, pressure, gibbs channels and generic properties. *Communications in Contemporary Mathematics*, page 2150090, 2021.

[BLLC10] Alexandre Baraviera, Carlos Felipe Lardizabal, Artur O. Lopes, and M. Terra Cunha. A thermodynamic formalism for density matrices in quantum information. *Applied Mathematics Research Express*, 2010(1):63–118, 2010.

[BLLC11a] Alexandre Baraviera, Carlos Felipe Lardizabal, Artur O. Lopes, and M. Terra Cunha. A dynamical point of view of quantum information: entropy and pressure. In *Dynamics, Games and Science I*, pages 81–121. Springer, 2011.

[BLLC11b] Alexandre Baraviera, Carlos Felipe Lardizabal, Artur O. Lopes, and M. Terra Cunha Cunha. Quantum stochastic processes, quantum iterated function systems and entropy. *São Paulo Journal of Mathematical Sciences*, 5(1):53–87, 2011.

[BLMM21] Jader E. Brasil, Artur O. Lopes, Jairo K. Mengue, and Carlos G. Moreira. Quantum spin probabilities at positive temperature are hölder gibbs probabilities. *Communications in Contemporary Mathematics*, 23(01):1950050, 2021.

[BOS22] Jader E. Brasil, Elismar R. Oliveira, and Rafael Rigão Souza. Thermodynamic formalism for general iterated function systems with measures. *preprint*, 2022.

[BP13] Vasile Berinde and Madalina Pacurar. The role of the pompeiu-hausdorff metric in fixed point theory. *Creative Mathematics and Informatics*, 22(2):143–150, 2013.

[CA06] D. Charalambos and Border Aliprantis. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag Berlin and Heidelberg GmbH & Company KG, 3 edition, 2006.

[CNT87] Alain Connes, Heide Narnhofer, and W. Thirring. Dynamical entropy ofc* algebras and von neumann algebras. *Communications in Mathematical Physics*, 112(4):691–719, 1987.

[CO17] L. Cioletti and Elismar R. Oliveira. Thermodynamic formalism for iterated function systems with weights, 2017.

[CO19] L. Cioletti and Elismar R. Oliveira. Applications of variable discounting dynamic programming to iterated function systems and related problems. *Nonlinearity*, 32(3):853, 2019.

[DM78] Claude Dellacherie and Paul-André Meyer. Probabilities and potential, vol. 29 of north-holland mathematics studies, 1978.

[Dum13] Dan Dumitru. Attractors of infinite iterated function systems containing contraction type functions. *An. Ştiinţ. Univ. Al. I. Cuza Iaşi, Mat. NS*, 59:281–298, 2013.

[Elt87] John H. Elton. An ergodic theorem for iterated maps. *Ergodic Theory and Dynamical Systems*, 7(4):481–488, 1987.

[Eva78] D. Evans. R. h0egh-krohn, spectral properties of positive maps on c*-algebras. *J. London Math. Soc*, 17:345–355, 1978.

[EW07] J. Eisert and M. M. Wolf. Quantum information with continuous variables of atoms and light, 2007.

[FL99] Ai Hua Fan and Ka-Sing Lau. Iterated function system and ruelle operator. *Journal of mathematical analysis and applications*, 231(2):319–344, 1999.

[GL08] E. Garibaldi and A. O. Lopes. On the aubry–mather theory for symbolic dynamics. *Ergodic Theory and Dynamical Systems*, 28(3):791–815, 2008.

[GLJ05] Gertruda Gwóźdź-Lukawska and Jacek Jachymski. The hutchinson-barnsley theory for infinite iterated function systems. *Bulletin of the Australian Mathematical Society*, 72(3):441–454, 2005.

[GS$^+$21] Aman Gupta, Cyril Shaju, et al. A study of the international stock market behavior during covid-19 pandemic using a driven iterated function system. *Computational economics*, pages 1–12, 2021.

[HMU02] Pawel Hanus, R. Mauldin, and Mariusz Urbański. Thermodynamic formalism and multifractal analysis of conformal infinite iterated function systems. *Acta Mathematica Hungarica*, 96(1-2):27–98, 2002.

[Hut81] John E. Hutchinson. Fractals and self similarity. *Indiana University Mathematics Journal*, 30(5):713–747, 1981.

[JP19] Vinayak Jagadish and Francesco Petruccione. An invitation to quantum channels. *arXiv preprint arXiv:1902.00909*, 2019.

[Käe04] Antti Käenmäki. On natural invariant measures on generalised iterated function systems. *Ann. Acad. Sci. Fenn. Math*, 29(2):419–458, 2004.

[KOW99] A. Kossakowski, M. Ohya, and N. Watanabe. Quantum dynamical entropy for completely positive map. *Infinite dimensional analysis, quantum probability and related topics*, 2(02):267–282, 1999.

[KP02] David Kerr and Claudia Pinzari. Noncommutative pressure and the variational principle in cuntz–krieger-type c*-algebras. *Journal of Functional Analysis*, 188(1):156–215, 2002.

[Lid19] Daniel A. Lidar. Lecture notes on the theory of open quantum systems. *arXiv preprint arXiv:1902.00967*, 2019.

[LMMM18] Artur O. Lopes, Jairo K. Mengue, Joana Mohr, and Carlos G. Moreira. Large deviations for quantum spin probabilities at temperature zero. *Stochastics and Dynamics*, 18(06):1850044, 2018.

[LMMS15] Artur O. Lopes, Jairo K. Mengue, Joana Mohr, and Rafael R. Souza. Entropy and variational principle for one-dimensional lattice systems with a general a priori probability: positive and zero temperature. *Ergodic Theory and Dynamical Systems*, 35(6):1925–1961, 2015.

[LO09] Artur O. Lopes and Elismar R. Oliveira. Entropy and variational principles for holonomic probabilities of ifs. *Discrete & Continuous Dynamical Systems*, 23(3):937–955, 2009.

[Lop11] Artur O. Lopes. Thermodynamic formalism, maximizing probabilities and large deviations. `http://mat.ufrgs.br/~alopes/pub3/notesformteherm.pdf`, 2011.

[LR22] Artur O. Lopes and R. Ruggiero. Nonequilibrium in thermodynamic formalism: The second law, gases and information geometry. *Qualitative Theory of Dynamical Systems*, 21(1):1–44, 2022.

[LS15] Artur O. Lopes and Marcos Sebastiani. Generic properties for random repeated quantum iterations. *Quantum Studies: Mathematics and Foundations*, 2(4):389–402, 2015.

[ŁŻS03] Artur Łoziński, Karol Życzkowski, and Wojciech Słomczyński. Quantum iterated function systems. *Physical Review E*, 68(4):046110, 2003.

[MU00] R. D. Mauldin and M. Urbański. Parabolic iterated function systems. *Ergodic Theory and Dynamical Systems*, 20(5):1423–1447, 2000.

[NP12] Ion Nechita and Clément Pellegrini. Random repeated quantum interactions and random invariant states. *Probability Theory and Related Fields*, 152(1):299–320, 2012.

[OP04] Masanori Ohya and Dénes Petz. *Quantum entropy and its use.* Springer Science & Business Media, 2004.

[Pet07] Dénes Petz. *Quantum information theory and quantum statistics.* Springer Science & Business Media, 2007.

[PP90] William Parry and Mark Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*, 187(188):1–268, 1990.

[PY98] Mark Pollicott and Michiko Yuri. *Dynamical systems and ergodic theory.* Number 40. Cambridge University Press, 1998.

[PY$^+$00] Claudia Pinzari, K. Yonetani, et al. Kms states, entropy and the variational principle in full c*-dynamical systems. *Communications in Mathematical Physics*, 213(2):331–379, 2000.

[Rue67] David Ruelle. A variational formulation of equilibrium statistical mechanics and the gibbs phase rule. *Communications in Mathematical Physics*, 5(5):324–329, 1967.

[Rue68] David Ruelle. Statistical mechanics of a one-dimensional lattice gas. *Communications in Mathematical Physics*, 9(4):267–278, 1968.

[Sag21] Takahiro Sagawa. *Entropy, Divergence, and Majorization in classical and quantum thermodynamics.* Springer Verlag, Singapor, 2021.

[Sch00] Robert Schrader. Perron-frobenius theory for positive maps on trace ideals. *arXiv preprint math-ph/0007020*, 2000.

[Sło03] Wojciech Słomczyński. *Dynamical entropy, Markov operators, and iterated function systems.* Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 2003.

[Spi72] Frank Spitzer. A variational characterization of finite markov chains. *The Annals of Mathematical Statistics*, 43(1):303–307, 1972.

[SS17] Wojciech Słomczyński and Anna Szczepanek. Quantum dynamical entropy, chaotic unitaries and complex hadamard matrices. *IEEE Transactions on Information Theory*, 63(12):7821–7831, 2017.

[SSU01] Károly Simon, Boris Solomyak, and Mariusz Urbański. Invariant measures for parabolic ifs with overlaps and random continued fractions. *Transactions of the American Mathematical Society*, 353(12):5145–5164, 2001.

[Ste02] Örjan Stenflo. Uniqueness of invariant measures for place-dependent random iterations of functions. In *Fractals in Multimedia*, pages 13–32. Springer, 2002.

[Stø02] Erling Størmer. A survey of noncommutative dynamical entropy. In *Classification of Nuclear C*-Algebras. Entropy in Operator Algebras*, pages 147–198. Springer, 2002.

[SŻ94] Wojciech Sl/omczyński and Karol Życzkowski. Quantum chaos: an entropy approach. *Journal of Mathematical Physics*, 35(11):5674–5700, 1994.

[Voi95] Dan Voiculescu. Dynamical approximation entropies and topological entropy in operator algebras. *Communications in mathematical physics*, 170(2):249–281, 1995.

[Wal75] Peter Walters. Ruelle's operator theorem and g-measures. *Transactions of the American Mathematical Society*, 214:375–387, 1975.

[Wol12] Michael M. Wolf. Quantum channels & operations: Guided tour. `https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf`, 2012.